



EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

Luca Mingarelli, Beatrice Ravanetti,
Tamarah Shakir, Jonas Wendelborn

Dawn of the (half) dead: the twisted
world of zombie identification

No 2743 / October 2022

Disclaimer: This paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Abstract

Since the term was first coined in studies on the 1990s Japanese crisis, the concept of zombification has been investigated and revived repeatedly when concerns arise about credit misallocation and stagnating productivity growth in an economy. The starting point for these studies nearly always involves trying to identify the so-called ‘zombie’ firms. This has led in the past years to a proliferation of different definitions and identification methodologies. We survey the most prominent definitions, discussing advantages and limitations of each. We also undertake a comparison of methodologies on a common dataset for euro area firms from 2004-2019, with the exercise revealing limited overlap and low comparability in the firms identified by several prominent studies. In response, we introduce a formalisation of zombie-classifications which helps to make order in the growing number of variations and identification methodologies. Moreover, this formalisation also helps extending the concept of binary identification to that of fuzzy zombie-identification. In particular, we introduce a general procedure to turn arbitrary binary classifications into fuzzy ones showing it successfully increases consistency between zombie definitions.

Keywords: Zombie firms, Vulnerable firms.

JEL Codes: L25, D22, D24, C55, O40.

Non-technical summary

The term *zombie firm*, originated in the context of the Japanese economic stagnation started in the early 1990s and has been used to indicate companies which would exit normally functioning markets but manage nonetheless to survive, typically relying on subsidised credit. Since then, the concept has reemerged for other economies where concerns arise around credit misallocation or risks of stagnating productivity, this includes debates in the euro area in the wake of the sovereign debt crisis in the 2010s and in the wake of the COVID-19 pandemic in 2020.

However, despite the growing amount of research on the topic, the methodologies available to detect this phenomenon are still imperfect and prone to misclassification. Moreover, with time, the concept has been made more evanescent, being often used interchangeably for vulnerable or financially weak firm, especially in the public debate. In this work we seek to establish conceptual order on the one hand, and provide a comparison and consistency assessment of some of the most prominent zombie identification methodologies on the other.

First, we introduce a framework useful to formalise binary zombie-identifications, i.e. identification methods that classify firms as either being a zombie or not. Employing firm level and firm-bank level data for euro area companies we compare the classifications of different methodologies. We find that while some show similar trends in time, all methods identify sets of firms with little or no overlap. This is concerning as, despite being all called *zombies*, these firms actually represent distinct subsets of the economy. Hence, comparing empirical results across methodologies is not possible, warranting additional care in drawing policy considerations.

We then introduce a general procedure to extend existing zombie-identifications from binary to fuzzy, meaning that we associate continuous scores between zero and unity to each firm, loosely quantifying a *degree of zombieness*. Crucially, our procedure does not require any additional data requirement and is associated only with limited additional computational complexity. Most importantly, we show that employing this fuzzy generalisation

increases the consistency among methods. Therefore, we argue this approach should be preferred over binary classifications as it can make empirical results and analyses more comparable and better able to inform policy making.

1 Introduction

In the early 1990s, the Japanese economy experienced a period of sustained stagnation and weak productivity following the burst of an asset price and credit bubble. In this period – which would become known as the ‘Lost Decade’ – lax banking supervision and government encouragement in extending more credit as a route out of stagnation, contributed to the creation of incentives for credit institutions to forebear on or ‘evergreen’ their bad debt so as to avoid recognising it as non-performing, which would have pushed them closer to their minimum requirements of capital. As a consequence, low-quality and unprofitable borrowers were able to maintain access to cheap subsidised credit, allowing them to survive where they would have otherwise been taken over by more profitable firms or exited the market. These non-viable firms became known in the economic literature as *zombies* (Caballero et al., 2008), a term first coined by Kane (1987).

The market distortions associated with the presence of such firms, which range from decreased competitiveness, to market congestion, chronic weak productivity, and misallocation of private and public resources, raised the concern that such a phenomenon could materialise also in other economies. In Europe, the very low interest rate environment and accommodative credit conditions present in the euro area since 2010 have prompted arguments that its economy may also have seen an increase in zombie firms, particularly in the context of an extended period of low productivity in many euro area economies (Acharya et al., 2019). The theme also reemerged in the immediate wake of the COVID-19 pandemic: with the corporate sector suffering from severe disruptions in its activities, the large-scale public support schemes faced the difficult challenge of providing an efficient allocation of resources and, therefore, distinguishing between temporarily illiquid firms and structurally insolvent ones (Laeven et al., 2020). While policy measures mitigated the economic impact of the pandemic and prevented waves of bankruptcies (Gourinchas et al., 2020), policy makers and observers alike are left with the concern that funding non-viable firms may have further slowed down the long-term recovery with the unintended consequences of inflating the zombification problem (Helmersson et al., 2021).

The potential ramifications of this phenomenon in terms of efficient allocation of resources, productivity, and long term growth have therefore sparked a strand of literature attempting to both capture the presence and share of such firms in the economy, as well as to quantify their causes and effects. However, while the early methodology put forward by Caballero et al. (2008) to identify zombie firms was aimed at a specific phenomenon – the extension of subsidised credit to undeserving borrowers – many identification methodologies that followed had to resort to proxies, mainly because of stringent data requirements necessary for a more robust assessment. This has resulted in many scholars employing their own definition of zombie firms in function of the data available to them. Proxies searching for inability of a firm to cover interest payments, checks on operating incomes, and persistent lack of profitability may have also been a contributing factor to the increasing evanescence of the concept of zombie firms, which is often times confused with that of financially weak or vulnerable firm.

An additional source of confusion is brought into the picture by the fact that different authors have often restricted their analysis, once again mostly due to limited data availability, to distinct geographical jurisdictions, economic sectors, time periods, or types of firms (publicly listed versus private companies). This makes comparative exercises of estimates and empirical analysis challenging. However, these comparative exercises are fundamental to understand the degree of cross-method consistency and eventually to inform policy making in the design of better public support measures.

In this spirit, this work seeks to establish more order in the space of zombie-identifications by surveying the most prominent methods in the literature and comparing the sample of firms each identifies as zombies. This is done both by replicating, to the best of our ability, these methodologies on a common dataset for euro area firms for the period between 2004 and 2019. Starting from the observation that most identification methodologies use a binary classification, i.e. classify firms either as a zombie or a non-zombie, we introduce a formal framework to generalise these binary definitions. This framework is then used to generalise binary identifications to fuzzy ones, meaning classifications that allow for a *degree of zombieness*. This is done by turning an indicator function confined on the

Boolean domain $\{0,1\}$ into a membership function mapping on the real interval $[0,1]$. We argue the proposed fuzzy approach is preferable as it accounts for the bluriness inherent in the zombie -concept as opposed to black-or-white classifications, and most importantly increases the consistency across methods in terms of the subsets of firms identified, albeit at the cost of decreased conservativeness.

Using joint *largest identifiable subsets* (LIS) and thereby making results comparable across identifications, we find that the main binary methodologies classify as zombies different subsets of the economy. This result, in turn, raises questions regarding the comparability of such identifications and their ability to provide generalised conclusions about the relative importance and implications of the presence of zombie firms in the economy. In part, this limited overlap in identified zombie firms is due to ultimately arbitrary classification thresholds used in binary identifications, which is mitigated by our proposed fuzzy approach.

The remainder of this work is structured as follows. In Section 2, we survey some of the most prominent works on corporate zombification, discussing for each advantages as well as limitations, both conceptual and in terms of data requirements. We then present a collection of datasets we have available to replicate some of these zombie-identification methodologies and introduce a formal framework to describe all zombie-identifications (Section 3). Cross-method consistency is discussed. Section 4 presents a formal procedure to generalise the identifications discussed before to fuzzy identifications, providing evidence that this approach increases consistency among different identifications. Finally, Section 5 draws conclusions and final considerations. For simplicity, we shall use certain conventions for symbols and variables throughout this work: the reader can refer to Table A.2 in the Appendix for their definitions.

2 A survey of binary identifications

After the term *zombie* was used to describe the surge of structurally insolvent Japanese borrowers kept afloat by subsidised lending in the 1990s (Caballero et al., 2008), a long

stream of literature followed trying to better characterise the phenomenon, and searching for its presence in other economies. In 1990s Japan, as documented by Peek and Rosengren (2005) and Hoshi (2006), the practice of credit evergreening was widespread¹ and driven by wrong incentives rooted in lax bank regulation and supervision policies, which made possible to turn early indirect evidence (Sekine et al., 2003) into quantitative characterisation of the phenomenon. The detection of this phenomenon in other economies however, has proved a hard task, giving rise to a zoo of different methods for the identification of zombies, and at times changing the meaning authors attach to the term, blurring the boundary between zombies and vulnerable firms.

In this Section, we provide qualitative comparisons of some of the most prominent methodologies for the identification of zombie firms. First, we survey different approaches put forward to identify zombie firms, discussing the conceptual and technical advantages they bring as well as their limitations. Different methodologies have been employed to estimate the relevance of zombification for different economies, but often being limited to short and different windows of time, different sectors and geographical jurisdictions, types of firms, and datasets.

The possibility that a zombification of the economy such as the one observed in Japan in the early '90s would come about in other economies as well has been a growing concern in the past decade, giving rise to a stream of literature attempting to estimate the risks of such an event materialising. The necessity to identify non-viable firms has lead research to find clever proxies and rationales to classify companies as belonging to this category. However, a major constraint on further developments has often been the lack of available data with sufficient granularity. In order to detect the issue originally identified in Japan one would ideally need information on the granular debt structure of individual companies including associated costs and crucially ad hoc information on interest paid. This highly granular and sensitive information is unfortunately hard to come by, and, when available, has often the drawback of having limited sample sizes - this has been the main factor

¹Up to 35% of Japanese firms have been estimated (Caballero et al., 2008) to receive subsidised credit in the period 1995 to 2001.

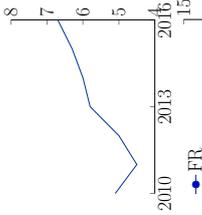
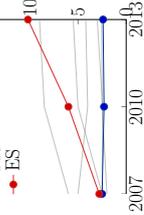
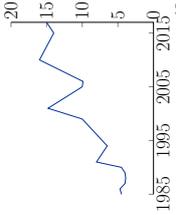
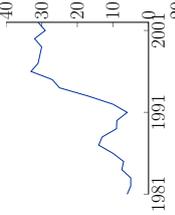
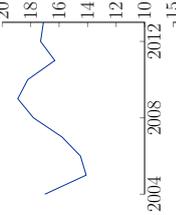
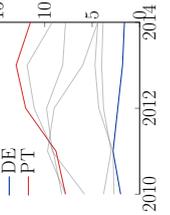
contributing to the development of alternative identifications based on companies' balance sheet data or market data attempting to capture persistent financial weakness.

One can therefore distinguish different zombie identification methodologies into those based on *firm level* data and those requiring *firm-bank level* information. As some methodologies are built on market data, identifications can be further divided into those restricted to *listed* companies and those applicable more broadly also to *non-listed* ones. In Table 1 the most prominent and widely employed definitions are summarised, distinguishing between data type (firm level versus firm-bank level), and firm type (listed versus non-listed). There, the coverage of different geographical regions and sectors considered in each original work is also detailed. Finally, the zombie share as quantified by each method is presented for the time window considered in each work. As mentioned, many more methods have been put forward in the literature, both as new methods in their own right but mostly as variations around those presented in Table 1, and a more comprehensive list can be found in appendix in Table A.1. The focus of this work is, however, on the most prominent methods listed in Table 1.

The most notable definition is that of Caballero et al. (2008) (CHK08), classifying as zombies those firms kept artificially alive by subsidised credit, defined as firms' actual interest payment on debt being below a hypothetical lower bound interest payment expected for the most creditworthy borrowers, and thus captured by a negative interest rate gap² $G^{IR} < 0$. This approach provides a concrete method to identify a specific phenomenon (that is the existence of lending at lower than expected rates) while also avoiding the hard-wiring by construction of correlations with specific market segments' growth and profitability that affect identifications based exclusively on operating characteristics. At the same time, however, this method has the drawback of requiring rarely available data. Indeed, the reference lower rate associated with most creditworthy borrowers is hypothetical in nature: while this may be sensible when considering a sample of listed firms only,

²In Caballero et al. (2008) firm-specific short-term and long-term bank loans, and total outstanding bonds are used to determine the hypothetical interest paid to be compared with the actual interest paid by each firm. The reader can refer to equation (7) in Definition 5 for an explicit definition of the interest rate gap G^{IR} as put forward by Caballero et al. (2008).

Table 1: Summary of the most prominent methodologies to identify zombie firms

Definition	Label	Dataset	Data type	Firm type	Coverage	# firms [k]	Identification	Zombie share [%]
Acharya et al. (2020)	ACEE20	Amadeus	Firm-level	All	AT, BE, DE, DK, ES, FI, FR, IT, PL, PT, SE, SK, ex. Utilities and financial sector	1168	$IC < M_{\text{cty-sec}}(IC)$, $L > M_{\text{cty-sec}}(L)$, and $G^{IR} < 0$	
Adalet McGowan et al. (2018)	AAM18	Orbis	Firm-level	All	BE, FI, FR, IT, KR, SI, ES, SE, GB (2003-2013) + AT, DE, LU, PT for 2013, ex. NACE2 01-09, 64-66, 84-99	1012	For 3 years: $(IC)_2 < 1$ and $T > 10$	
Banerjee and Hofmann (2020)	BH20	Worldscope	Firm-level	Listed	AT, BE, CA, CH, DE, DK, ES, FR, GB, IT, JP, NL, SE, US	32	For 2 years: $IC < 1$, and $q^T < M_{\text{sec}}(q^T)$	
Caballero et al. (2008)	CHK08	Nikkei Needs Corporate Financial	Firm-bank level	Listed	JP	2.2	$G^{IR} < 0$	
Schivardi et al. (2017)	SST17	Cerved	Firm-bank level	Limited liability	IT, ex. Agriculture and financial sector	242	$(RoA)_3 < CoC$, $L > L^*$	
Storz et al. (2017)	SKSW17	Amadeus	Firm-level	Non-listed SMEs	DE, ES, FR, GR, IE, PT, SI, ex. NACE2 01-09, 64-66, 84, 99	423	For 2 years: $RoA < 0$, $I^N < 0$, and $S^C < 5\%$	

Note: We denote by $\langle x \rangle$ the mean of x , and by $\langle x \rangle_y$ its y -years moving average. The last column reproduces the share of zombie time series as presented in each work; when a country breakdown is provided only the two most extreme observations are labelled, with the others drawn in gray. For works reporting firm-year observations a per-year average is presented in this table. For an overview of how the individual variables are defined see Table A.2 in the Appendix. CoC : cost of capital, IC : interest coverage ratio, T : firm age, q^T : Tobin's q , RoA : return on assets, S^C : debt servicing capacity, I^N : net investments, L : leverage, G^{IR} : interest rate gap.

it might not be as suitable when considering small and medium enterprises, whose rates received often differ significantly from those of large listed ones.

Among the most noteworthy and widely employed definitions constructed on firm level data instead is the work by Adalet McGowan et al. (2018) (AAM18), where zombies are identified as firms reporting an interest coverage ratio less than one, $IC < 1$, for at least three consecutive years³ and with an age of at least ten years, $T \geq 10$. The rationale is that the survival of firms despite low interest coverage ratios can potentially represent an indicator of subsidies extending beyond explicit subsidised credit and capturing for example non-performing loans, government guarantees, and weak insolvency regimes. In addition, IC is argued to be easily comparable across countries and preferable to parameters such as negative profits as less endogenous to productivity. At the same time, excluding younger firms prevents misidentifying start-ups with current weak performance despite their long-term potential and expected future profits.

However, while relatively simple to implement and going beyond the sample limitation to listed firms, this approach may have the unintended effect of shifting the meaning of the term *zombie* from that of ‘firms receiving subsidies’ to that of ‘vulnerable firm’. Indeed, quite to the contrary of the original meaning, low interest coverage ratios could be associated with higher interest payments rather than lower rates, as it was originally pointed out in Storz et al. (2017) (SKSW17). We call this the *IC critique*. In addition, the age requirement introduced by Adalet McGowan et al. (2018) is sometimes criticized as it does not clearly rule out why a young firm could not equally be considered a zombie (Banerjee and Hofmann, 2020). In order to address the IC critique, a profitability-oriented identification is proposed by SKSW17: here, zombies are firms reporting for at least two consecutive years negative returns on assets, $RoA < 0$, negative net investments, $N^I < 0$, and debt servicing capacity below five percent, $S^C < 5\%$. Within this framework, zombies are therefore identified as those unprofitable firms unable to invest past their depreciation value, with the negative investment requirement excluding start-ups whose yet unrealised profits may be driven by current large investment, and the constraint on debt servicing capacity providing a better indicator for highly indebted firms and

³The methodology was originally introduced in the Bank of Korea’s Financial Stability Report (Bank of Korea, 2013) to identify vulnerable firms, but we attribute the method to Adalet McGowan et al. (2018) as they were the first to employ it as a proxy to identify zombie firms.

potentially for subsidised credit than the interest coverage ratio of AAM18. The two consecutive years requirement aims to prevent misclassification due to fluctuations in the business cycle.

Another criticism against the early approach of CHK08, is that it may fail to provide a reliable zombie measure due to the influence of exogenous factors. In particular, the identification might be biased with respect to long-standing credit relationships — which may affect the levels of interest rates used to perform such classification — and more generally under-perform within an environment of exceptionally low interest rates, such as those observed in more recent years in developed economies (Fukuda and Nakamura, 2011; Banerjee and Hofmann, 2020). In an attempt to address these limitations, Acharya et al. (2020) (ACEE20)⁴ complement the subsidised credit criterion with a quality criterion: low interest coverage ratio and high leverage (below country-sector-year median for the interest coverage ratio, $IC < M_{\text{cty-sec}}(IC)$, and above for leverage, $L > M_{\text{cty-sec}}(L)$). Similarly, Fukuda and Nakamura (2011) extended the identification of CHK08 with a *profitability criterion* whereby zombies are such that reported before-tax profits (EBIT) are lower than a hypothetical risk-free interest payment, and an *evergreening criterion* accounting for high leverage and increased debt levels (cf. Table A.1).

Because wide-spread granular reporting of firms balance sheets and financial ratios came about only in recent years, analysis based on this data often falls short with respect to the time dimension. The ability to quantify zombification of economies across a wider time span is necessary to analyse the phenomenon over several decades, and multiple business cycles, which in turn is crucial for solid statistical inference and ex-post policy assessment. Better coverage of historical data is often available for publicly listed firms, which brings the added benefit of market equity prices, providing a proxy for the perceived potential for growth of firms (Banerjee and Hofmann, 2020) (BH20). The latter feature in particular has been introduced in Banerjee and Hofmann (2018, 2020), where zombies are defined as firms reporting for two consecutive years an interest coverage ratio below unity, and Tobin's q below sector median, $q^T < M_{\text{sec}}(q^T)$. This approach, which can be

⁴An earlier step in this direction was presented also in Acharya et al. (2019), also cf. Table A.1.

seen as an extension of previous work by AAM18, is still subject to the same criticism which applies to using interest coverage for identification of firms that might be receiving subsidised credit, therefore potentially identifying vulnerable firms rather than zombies per se. Nonetheless, it undoubtedly improves by introducing a more forward-looking metric of zombies' performance, namely, checking for expected low future profits (reflected in lower than median q^T in each sector) and by substantially extending the time span over which the identification can be conducted. However, the increased historical coverage comes at the notable cost of restricting the analysis to publicly listed firms, whose sample is unlikely to be representative of the broader population of firms across countries. First, any economy is mostly composed of small and medium-sized enterprises. Second, different countries have significantly different propensity to list, making any cross-country analysis challenging.

A different approach is taken by Schivardi et al. (2017) (SST17), where zombie firms are characterised as those whose expected marginal return of capital is below the market cost of capital after risk adjustment. SST17 link banks' characteristics to the prevalence of zombie firms and use both profitability and default risk indicators in their identification. A zombie firm must be unprofitable (having 3-year moving average ROA lower than the cost of capital incurred by the safest firms in the sample) and highly indebted (with leverage exceeding a time-invariant threshold). The authors also apply an alternative definition of zombies, substituting the condition on ROA with a measure of interest coverage ratio (calculated as EBITDA over interest expenses). This criterion is able to pick a stricter subset of zombie firms than identified using the threshold on ROA.

Different identifications are therefore characterised by a wide range of advantages and limitations. These are summarised in Table 2. Some of these depend on data availability: information on publicly listed firms for example is more widely available than for non-listed companies, and balance sheet information is easier to come by than firm-bank level relationships. Others instead are structural and depend on the specific issue the authors have in mind. As the range of different methodologies to identify zombie firms grows, it becomes increasingly desirable to be able to compare them, as the risk

Table 2: Advantages and limitations of different zombie identification methodologies

Method	Advantages	Limitations
ACEE20	Adds quality criterion	Subject to IC critique
AAM18	Simplest to compute Wide firm type and geographical coverage	Subject to IC critique Low historical coverage Potentially capturing vulnerability rather than zombieness
BH20	Wide historical coverage	Subject to IC critique Limited to listed firms Potentially capturing vulnerability rather than zombieness
CHK08	Simple criterion More explicit detection of subsidised credit Avoids hardwired sectoral correlations	Data hard to come by Potentially biased with respect to exogenous factors Limited to listed firms
SST17	Explicit firm-bank data Conceptually appealing approach to compare a firm's profitability with a measure of cost of capital	Unclear definition of Z-score
SKSW17	Goes beyond IC critique	Low historical coverage

Note: ACEE20: Acharya et al. (2020), AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SST17: Schivardi et al. (2017), SKSW17: Storz et al. (2017).

of identifying disjoint subsets of the economy grows as well. These cross-methodology heterogeneities, however, make such a direct comparison hard, as in addition to different conceptual frameworks underlying the work of each scholar, the empirical evidence provided is restricted to specific (short) time windows, different economies and sectors, and different firm types. Diverse definitions and identifications lead therefore to somewhat different results, eventually leaving the conceptual debate on the size, real effects, and policy implications of the zombie phenomenon open to misinterpretation. For example, while there is clearly value in analysing both the growth of vulnerable firms as well as the zombification of companies through subsidised credit, we believe it is important to be aware of the distinction and maintain a separation between the two abstract notions. The risk is otherwise to confound causes and consequences of events such as the recent COVID-19 crisis by identifying as zombies firms which are in fact viable (Laeven et al., 2020).

3 A conceptual framework for binary identifications

Constructing a dataset by gathering information on firms' balance sheets and financial ratios from the Orbis database and other ECB internal data sources, allows to replicate different zombie definitions on a common subset of firms and over a common window of time. This section will present such a dataset and seek to provide clarity on properties of different identification methodologies and the extent to which they are consistent with each other. Furthermore, a common framework is provided within which all these methodologies can be formalised: this step will be useful in Section 4 when we will extend these definitions from binary to fuzzy ones.

3.1 Data

To conduct our comparative analysis, we use firm-level data from Bureau van Dijk's Orbis database. The Orbis database collects information on firms' balance sheets, cash flows, activity status and ownership structures for millions of firms globally, mainly from local data providers and firm registries. As such, Orbis is the largest publicly available firm-level dataset and has also been used in a number of studies on zombie firms (e.g. Adalet McGowan et al. (2018); Storz et al. (2017); Acharya et al. (2019, 2020))⁵.

While Orbis has the advantage that it covers a large number of firms from many different countries and across different firm sizes and makes available data from financial statements in a roughly consistent and comparable format, it also has a number of caveats. Essentially, data collection is done on a best-effort basis, which means that there are many gaps in the dataset, in the form of missing values for individual variables, but also in the time-coverage of single firms (i.e. there can be gaps in between years of reports available for a firms), which has implications for the number of firms to which individual zombie identification methods can be applied. Further, Orbis has a reporting lag of around two years and coverage varies significantly over time as well as countries, depending on the underlying data source. Thus, samples from Orbis are not necessarily representa-

⁵Some of these studies use the Amadeus database, which is essentially the European subset of Orbis.

tive of a countries' firm population and can be biased in different directions⁶. However, Kalemlı-Özcan et al. (2015) show that for a number of countries roughly representative firm samples can be constructed from Orbis data.

For our analysis we look at annual reports of firms from the 19 euro area countries.⁷ Due to comparatively lower coverage in earlier years and the two year lag in Orbis data, we limit our analysis to the period between 2004 and 2019.

Following mainly Storz et al. (2017) and Kalemlı-Özcan et al. (2015), we perform a number of steps to clean and prepare the data for analysis. First, we correct the year to which a firm's report is attributed. Since the accounting years of firms can differ from the calendar year, we attribute reports with a closing date before the 1st of June to the previous year. Second, as most studies on zombie firms do, we limit our analysis to non-financial corporations (NFCs). To this end, we exclude firms based on their NACE Rev. 2 industry classification from the financial and insurance activities (NACE divisions 64, 65, and 66), the public sector (NACE division 84), activities of households (NACE divisions 97 and 98) and extraterritorial organisations (NACE division 99). Given the structural differences to other firms, we additionally exclude firms from the primary sector (NACE divisions starting with 0). Third, we only consider firms that are still active according to Orbis. Fourth, we look at the highest level of consolidation available. Accordingly we limit our analysis to reports with Orbis consolidation codes C1 (consolidated statement of a mother company where no unconsolidated companion is reported in Orbis), C2 (consolidated statement of a mother company where an unconsolidated companion is reported in Orbis) and U1 (unconsolidated statement of a company with no consolidated companion in Orbis)⁸. Fifth, we only consider firms for which the reported balance sheets are consistent. To this end, we require firms to report positive total assets and non-negative debt and we check that the sum of equity and debt (i.e. total liabilities) does not deviate

⁶See Bajgar et al. (2020) for an assessment of how representative Orbis data is of the firm population of several OECD countries.

⁷We use the composition of the euro area as of 2015: Austria, Belgium, Cyprus, Estonia, Finland, France, Germany, Greece, Italy, Ireland, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, Portugal, Slovakia, Slovenia and Spain. We take this composition as fixed throughout the sample period and abstract from the fact that some of these countries adopted the euro only during the sample period.

⁸See also Bureau van Dijk (2017).

more than an percent from the firm's total assets. Finally, we drop remaining duplicates and, to facilitate identification methods which require several consecutive years, we interpolate key variables for single missing years in between two existing firm accounts by taking the arithmetic mean of the respective variables from the adjacent existing accounts. See Table 3 for descriptive statistics on key variables used in the different identification methods.

To facilitate the replication of specific identification methods, we augment the data obtained from Orbis with other data sources where needed. Specifically, the identification method by BH20 requires data on market capitalisations to compute Tobin's q . Since the coverage of market capitalisations is rather poor in Orbis for years prior to 2011, we use the firms' equity ISINs reported in Orbis to fill the missing values drawing on data from Bloomberg and the ECB's centralised securities database (CSDB).

For their identification of zombie firms in Japan, CHK08 employ so called *prime rates* to determine the minimum amount of interest a firm would have been charged on its bank loans in a given year. These prime lending rates are compiled and published by the Bank of Japan and comprise the minimum short- and long-term lending rates adopted by the city banks. To the best of our knowledge there is no corresponding interest rate published for the euro area. One could use average interest rates charged of NFCs, which are published by the ECB⁹, but these would contravene CHK08's idea to provide the most conservative estimate of what interest a firm would need to pay on its bank loans under the most favourable conditions in a given year. To get the best estimate of prime rates for the euro area available to us, we use the individual MFI Interest Rate Statistics (iMIR). This dataset contains information on interest rates charged by individual MFIs by different breakdowns of loan maturity, loan purpose and counter-party sector. The data is collected by euro area national central banks on a best effort basis and covers around 300 credit institutions from the euro area, but has a limited time series coverage, starting only in 2007. Since the data is collected on a monthly basis, we first take the yearly average of interest rates charged on loans to NFCs for each credit institution and

⁹See MFI Interest Rate Statistics (MIR Statistics) in the ECB's Statistical Data Warehouse

then select the minimum across euro area credit institutions for every given year. We exclude observations that have the value zero as yearly average as there are indications that there are a number of cases where reported zeros should actually be missing values. In taking the minimum across euro area banks, we assume that any euro area firm from the Orbis sample would have been able to obtain a loan from any euro area bank at the most favourable conditions; we believe that this assumption does clearly not hold in reality, but serves the conservative spirit adopted in CHK08. The only drawback of using this dataset is that it is not publicly available and hence results cannot be reproduced by researchers outside the Eurosystem.

Data on coupon rates on convertible corporate bonds as used by CHK08 as a proxy for interest paid on bonds outstanding are obtained from the CSDB.

Table 3: Descriptive statistics

	Firms [k]	10%	25%	50%	75%	90%
Total assets [€M]	3580.2	0.014	0.064	0.271	0.990	3.319
Leverage	3580.2	0.08	0.29	0.62	0.9	1.2
<i>RoA</i>	2766.4	-11.75	-0.74	2.13	9.29	22.45
Productivity	1476.1	1.08	1.24	1.4	1.57	1.75
<i>IC</i>	1538.0	-27.44	0.17	5.34	37.75	267.43
<i>S^C</i>	2392.9	-0.12	0.02	0.13	0.4	1.15
<i>I^N</i>	3304.4	-0.09	-0.03	0.0	0.01	0.11
<i>T</i>	3580.2	2.0	5.0	11.0	21.0	31.0
<i>q^T</i>	2.7	0.76	0.95	1.2	1.8	3.02

Note: descriptive statistics are shown for reporting year 2019 and across euro area countries. Productivity indicates total factor productivity calculated by means of Solow residuals. *RoA*: return on assets, *IC*: interest coverage ratio. *S^C*: debt servicing capacity, *I^N*: net investments, *T*: firm age, *q^T*: Tobin's q. For variable definitions see Table A.2 in the Appendix

3.2 Cross-methodology analysis: a quantitative comparison

As we have discussed, while scholars have been conducting a wide range of research on the risks and causes of corporates' zombification, it remains challenging to put these results together. On the one hand, this is due to the vast heterogeneity in terms of methods used to identify such firms. But aside from the conceptual differences between various identifications, a further difficulty arising when attempting to draw a comparison is found

in the fact that different analyses are conducted on different subsets of the economy, in specific geographies, sectors, and selecting different firm types. The reason is found in the data each identification requires.

In this section we shall therefore try to provide a quantitative comparison among the most prominent zombie identifications described earlier in Table 1. In order to mitigate the difficulties discussed above, the comparison is performed on the same dataset, selecting the relevant variables. Because the coverage of different firm characteristics still varies within our dataset, and because some methods naturally restrict the sample of firms to a specific subset, we introduce the notion of *largest identifiable subset* (LIS), that is the subset of firms on which a given identification (or set of identifications) can be applied^{10,11}.

Before continuing with the quantitative comparison of the share of zombie firms identified by different methods, it is useful to formalise what constitutes a zombie-identification into a framework that can be applied across different methods.¹² Thus, Definition 1 provides a formalisation of the concept of crisp zombie-identification (or binary¹³, that is classifying a firm as either a zombie or not) in terms of a vector of constrained firm characteristics \mathcal{V} .

Definition 1 *A crisp zombie-identification is a vector $\mathcal{V}_y^{(i)}$ of firm i characteristics in year y , together with an indicator function $Z_{i,y} : \mathcal{V}_y^{(i)} \rightarrow \{0, 1\}$ taking each firms' vector of characteristics onto the Boolean domain. In most cases this function can be expressed*

¹⁰As an example, the LIS of AAM18 is composed of all firms in our dataset reporting operating income and interest expenses (which are required to compute a firm's interest coverage ratio IC) in a given reporting year as well as the previous two years (to fulfill the three consecutive year condition). Conversely, firms reporting missing values for one of these variables in either of the three years are not part of the LIS.

¹¹When performing a firm level comparison over identifications associated with different data coverage, it is useful to restrict the comparison to the joint LIS of all identifications considered.

¹²As we shall see in more detail in Section 4, this formalisation will also help to make a direct extension to fuzzy identifications.

¹³We prefer to use the term *crisp* to binary as it more directly relates to the terminology used in fuzzy set theory, therefore making this definition more consistent with what will be discussed in Section 4 when Definition 1 will be extended to incorporate a notion of fuzziness.

in the form

$$Z_{i,y} = \left(\prod_{x_y \in \mathcal{V}_y^{(i)}} k(x_y) \right)^{\frac{1}{|\mathcal{V}|}}, \quad (1)$$

with the kernel $k(x_y) \equiv \mathbb{1}_{x_y < 0}$ determining the thresholds of interest for each firm's characteristic in vector $\mathcal{V}_y^{(i)}$ relevant for identification in year y . Therefore, a firm is called a zombie (or non-viable) in year y if $Z_{i,y} = 1$, and healthy (or viable) if $Z_{i,y} = 0$.^{14,15} We refer to $n = |\mathcal{V}|$ as the dimensionality of the zombie-identification.

The functional form (1) can incorporate the requirement of Y -consecutive years into the vector $\mathcal{V}_y^{(i)}$ by including the corresponding historical reported values of characteristic x : $(x_{y-(Y-1)}, \dots, x_{y-1}, x_y)$.¹⁶ One can alternatively make such requirement explicit by rewriting (1) as

$$Z_{i,y} = \left(\prod_{w=0}^{Y-1} \prod_{x_y \in \mathcal{V}_y^{(i)}} k(x_y) \right)^{\frac{1}{Y|\mathcal{V}|}}, \quad (2)$$

where the vector $\mathcal{V}_y^{(i)}$ contains exclusively the characteristics of firm i reported in year y , as the multiple-years requirement is explicitly accounted for. As we will discuss, this functional form suggests a natural way to extend the identification beyond the Boolean domain, also allowing for different weights on each requirement, or previous years. The functional form of the indicator function $Z_{i,y}$ in (1) accounts only for identifications in the form of logical conjunctions of requirements on different firms characteristics.¹⁷ While it is straightforward to write a more general form allowing for logical disjunctions,¹⁸ almost all zombie-identifications known to us can be expressed as in (1).¹⁹ Therefore, all

¹⁴As we will discuss, the exponent here is superfluous; however, it will help generalising this functional form to fuzzy identification.

¹⁵Notice that while $k_w(x_y)$ accounts for conditions of the form $x_y < 0$, this does not prevent one to consider more general conditions. As an example, a condition of the form $x_y > c$ would be accounted for by the entry $c - x_y$ in $\mathcal{V}_y^{(i)}$.

¹⁶As an example, an identification requiring two consecutive years of a given firm characteristic x would have $\mathcal{V}_y^{(i)} = (x_{y-1}, x_y)$.

¹⁷E.g. $C_1 < 0$ AND $C_2 < 0$ for some firm characteristics C_1 and C_2 .

¹⁸E.g. $C_1 < 0$ OR $C_2 < 0$ for some firm characteristics C_1 and C_2 . The indicator function would in this case take the form $\bigvee_{x \in \mathcal{V}_y^{(i)}} k(x)$.

¹⁹One exception is the definition of Andrews and Petroulakis (2019). This was not explicitly considered in this work as it consists of a modification of the identification by Storz et al. (2017). In particular,

identifications we consider here are uniquely determined by a specific characteristic vector \mathcal{V} , sharing the same indicator function $Z_{i,y}$ presented in (1).

One can now consider a concise formal definition of the identifications considered so far exclusively in terms of the characteristics entering the identification vector \mathcal{V} : these are presented in the following Definitions 2 to 7. For an overview of how the single variables are defined, see Table A.2 in the Appendix.

Definition 2 (ACEE20) *The zombie-identification ACEE20 (Acharya et al., 2020) is uniquely defined by the characteristics vector*

$$\mathcal{V}_y = (\langle IC_y \rangle_2 - M_{cty-sec}(\langle IC_y \rangle_2), M_{cty-sec}(L_y) - L_y, G_y^{IR}), \quad (3)$$

where $\langle IC_y \rangle_2$, L_y , and G_y^{IR} are the firm's 2-years moving average interest coverage ratio, leverage, and interest rate gap in year y respectively, and $M_{cty-sec}$ denotes the median functions applied at country-sector level. Here, the interest rate gap, G_y^{IR} , is defined as the difference between the interest rate paid by a firm and the median of interest rates paid by AAA-rated firms in the same year, where interest rates are determined as the ratio of its interest expenses relative to the sum of its outstanding loans, credit, and bonds and the AAA rating is inferred from firms' interest-coverage ratio being above 12.5%.

Definition 3 (AAM18) *The zombie-identification AAM18 (Adalet McGowan et al., 2018) is uniquely defined by the characteristics vector*

$$\mathcal{V}_y = (IC_{y-2} - 1, IC_{y-1} - 1, IC_y - 1, 10 - T), \quad (4)$$

with T the age of the firm.

Definition 4 (BH20) *The zombie-identification BH20 (Banerjee and Hofmann, 2020)*

Andrews and Petroulakis (2019) define a zombie as a firm reporting for three consecutive years (i) low debt servicing capacity (below 5% as in Storz et al. (2017)) and (ii) either negative returns on assets or negative net investment. This definition can be accommodated in the framework described above as hinted in the previous footnote.

is uniquely defined by the characteristics vector

$$\mathcal{V}_y = (IC_{y-1} - 1, IC_y - 1, q_{y-1}^T - M_{sec}(q_{y-1}^T), q_y^T - M_{sec}(q_y^T)). \quad (5)$$

Here, q_y^T is the firm's Tobin's q in year y , and M_{sec} the median function applied at sectoral level.

Definition 5 (CHK08) The zombie-identification CHK08 (Caballero et al., 2008) is uniquely defined by the characteristics vector

$$\mathcal{V}_y^{(i)} = (G_y^{IR}), \quad (6)$$

where for each firm the interest rate gap G_y^{IR} in year y is defined as

$$G_y^{IR} = \frac{R_y - rs_{y-1}BS_{y-1} - \langle rl_y \rangle_5 BL_{y-1} - rcb_{\min \text{ over last 5 years}, y} \text{Bonds}_{y-1}}{BS_{y-1} + BL_{y-1} + \text{Bonds}_{y-1} + CP_{y-1}}. \quad (7)$$

Here, R_y is the interest payment, and BS_y , BL_y , Bonds_y , and CP_y , denote the amount of short-term bank loans, long-term bank loans²⁰, total bonds outstanding²¹, and the amount of outstanding commercial paper respectively in year y . Similarly, rs_y , rl_y , $rcb_{\min \text{ over last 5 years}}$ are the average short-term prime rate, long-term prime rate, and minimum observed coupon rate on any convertible corporate bond issued in the last five years.

Definition 6 (SST17) The zombie-identification SST17 (Schivardi et al., 2017) is uniquely defined by the characteristics vector

$$\mathcal{V}_y = (\langle RoA_y \rangle_3 - CoC_y, M(L_y) - L_y). \quad (8)$$

Here, CoC_y and L_y are the firm's cost of capital and leverage in year y , while $\langle RoA_y \rangle_3$ is the three-years moving average of returns on assets computed as EBITDA over total

²⁰Short-term bank loans refers to loans with maturity below one year, and similarly long-term bank loans to maturities greater than one year.

²¹Total bonds outstanding includes convertible bonds and warrant-attached bonds.

assets:

$$\langle RoA_y \rangle_3 = \frac{1}{3} \sum_{w=0}^2 \frac{EBITDA_{y-w}}{TA_{y-w}}. \quad (9)$$

Definition 7 (SKSW17) *The zombie-identification SKSW17 (Storz et al., 2017) is uniquely defined by the characteristics vector*

$$\mathcal{V}_y = (RoA_{y-1}, RoA_y, S_{y-1}^C - 5\%, S_y^C - 5\%, I_{y-1}^N, I_y^N), \quad (10)$$

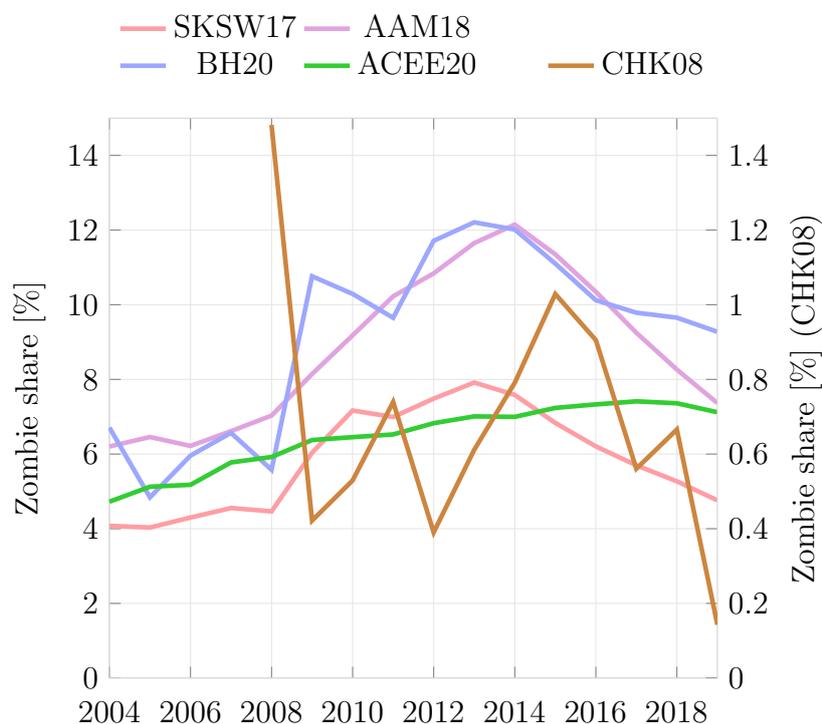
with S_y^C denoting debt servicing capacity, defined as EBITDA over financial debt, and I^N investment, defined as the net change in total fixed assets relative to the previous year. Here, return on assets is computed as net income over total assets.

Replicating the estimation of the share of zombies as identified per Definitions 2 to 7 above, we obtain the shares presented in Figure 1.²² The identification is conducted on each method's LIS, for all firms within the euro area and over a time window of sixteen years.²³ The trends and levels of AAM18 and BH20 are remarkably similar, which is reassuring, and possibly to be expected given that both methods base their identification on low interest coverage ratios. This however also speaks in favour of the method by Adalet McGowan et al. (2018) as it is simpler and less restrictive (can be applied also to non-listed firms) than BH20. In particular, our results suggest the main driver of BH20 remains the condition $IC < 1$. A similar trend can be observed for the method by SKSW17, although levels are lower given the more conservative nature of this identi-

²²We are able to replicate all identifications except for that of Schivardi et al. (2017) (SST17). This has to do with a number of issues we encounter when attempting to replicate SST17, but mostly two stand out. First, in determining their cost of capital variable CoC the authors make use of a subset of creditworthy firms they label *Safest Firms* which they say is determined as those firms with Altman-Z score of either 1 or 2. The score used by SST17 is, however, a credit score provided by Cerved, the vendor from which they obtain their data, and not the classical Altman-Z score in the sense of (Altman, 1968), which uses a different scale. The underlying methodology is likely similar, but the exact methodology used is not disclosed by Cerved. As the Cerved scores are not available to us, we try replicating the results using traditional Altman-Z scores in the sense of (Altman, 1968), also varying the threshold for creditworthiness based on different percentiles of the Z-score distribution, but find the results are highly sensitive to variations in this threshold. We also find results are highly sensitive to variations in the threshold for leverage. Therefore we do not report additional analysis on SST17 as we are unable to replicate the identification.

²³The exception is CHK08 for which we do not have data on interest rates prior to 2008.

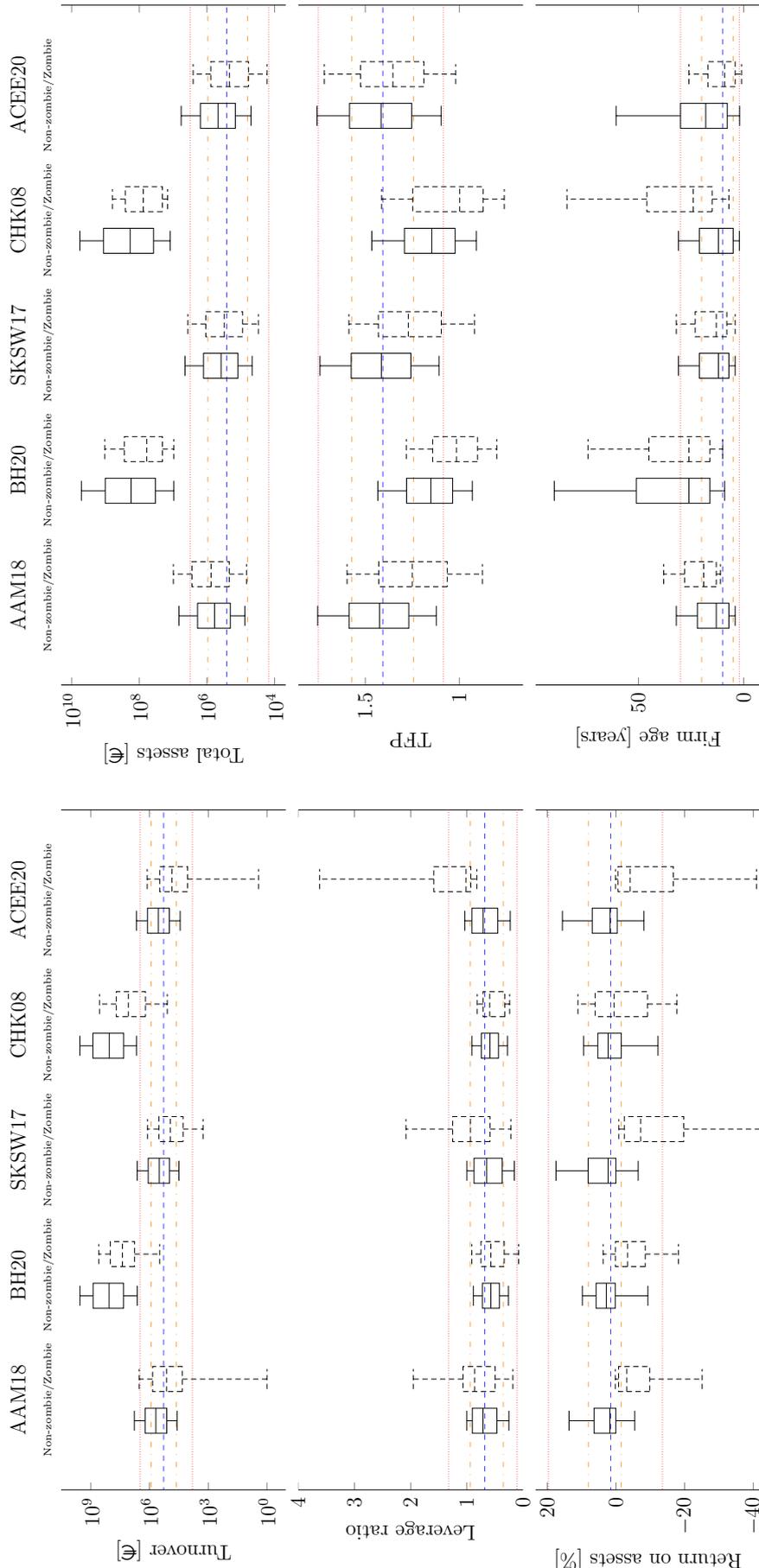
Figure 1: Share of zombie firms



Note: each methodology is applied on its associated LIS. All methods are applied on the window of time spanning the years from 2004 to 2019, with the exception of CHK08 for which we do not have sufficient data before 2008. ACEE20: Acharya et al. (2020), AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SKSW17: Storz et al. (2017).

fication. While AAM18, BH20, and SKSW17 all describe an increasing share of zombies in the aftermath of the global financial crisis, but declining since the euro area sovereign debt crisis, identification ACEE20, despite including a similar condition on interest coverage, depicts instead a different story with shares of zombies pretty constantly increasing over the whole time window considered in our analysis. On the other hand, the trend resulting from applying the definition of CHK08 can also be observed to be somewhat similar to that of AAM18, BH20, and SKSW17, although the large degree of volatility makes it difficult to make a conclusive statement on the extent of correlation between the two. This is likely also a function of the low share of firms identified which reduces the statistical robustness of the estimate. This fact actually applies more broadly across all methodologies, making apparent the importance of coverage for robust estimates, as higher volatility is associated with smaller LIS.

Figure 2: Firm characteristics by zombie status for reference year 2018



Note: boxplots with solid lines show the distributions for non-zombie firms, boxplots with dashed lines for zombie firms. Whiskers denote the 10th and 90th percentiles. The dashed lines show the distributions for the entire Orbis sample after applying the filters lined out in Section 3.1. The dashed blue line denotes the median, the dashed yellow lines the interquartile range and the dotted red lines the 10th and 90th percentiles. TFP: total factor productivity. TFP is calculated by means of Solow residuals. ACEE20: Acharya et al. (2020), AAM18: Azalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SKSW17: Storz et al. (2017).

Observing that most methods are somewhat consistent in trend and levels is reassuring but not sufficient to determine to which extent different identifications actually relate to the same subset of firms in the economy. In order to shed more light on the compatibility between different definitions, we first look at how firms identified by the different methods as zombies and non-zombies compare across a set of firm characteristics and secondly at the overlaps in firm level identifiers, comparing the actual groups of firms identified as zombies.

Figure 2 shows how different firm characteristics are distributed across non-zombie firms (solid boxplots) and zombie firms (dashed boxplots) for each of the five methods. From these distributions we can see a number of commonalities, but also important differences between zombies and non-zombies as well as different identifications, which can mostly be explained by the different criteria used by each identification. Across all methods zombies tend to have lower turnover than non-zombies and tend to be less productive in terms of total factor productivity (TFP)²⁴. Zombie firms also tend to have lower and often negative returns on assets than their peers. In the case of SKSW17 this relationship is hardwired in the identification criteria and it is similarly expected for identifications that use a criterion on interest coverage ratios²⁵; the only method for which this relationship is less distinct is CHK08 which neither has a criterion on return on assets nor on interest coverage ratios. Zombie firms tend to be smaller in terms of total assets, except for AAM18, which might be due to the ten year age requirement. Due to this requirement, firms identified as zombies by AAM18 are by definition older than their non-zombie peers. This relationship is similarly clear cut only for CHK08 and for ACEE20 zombie firms are actually younger than non-zombies. This finding underlines the discussion on whether an age requirement is actually a good proxy to include in zombie identifications. For AAM18, SKSW17 and ACEE20, zombie firms tend also to be higher leveraged; in the case of ACEE20 this relationship is explicitly included in the identification criteria.

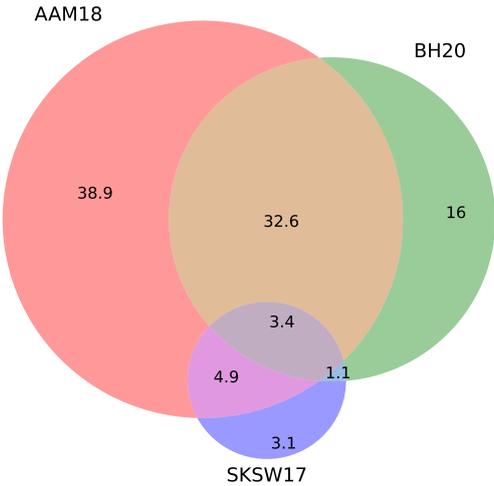
Given that they can only be applied to listed firms, BH20 and CHK08 have structurally different LIS than the other three methods. This is apparent across many of the firm characteristics shown and explains differences to the median firms identified by other

²⁴TFP is computed by means of Solow residuals.

²⁵Note the close accounting relationship between an interest coverage ratio below one and a negative return on assets.

methods. For example, these listed firms are much larger in terms of total assets, have higher turnover, but have lower TFP. They have similar median returns on assets, but the differences in leverage between zombies and non-zombies are much smaller than for firms in the other samples.

Figure 3: Overlap between zombie identifications

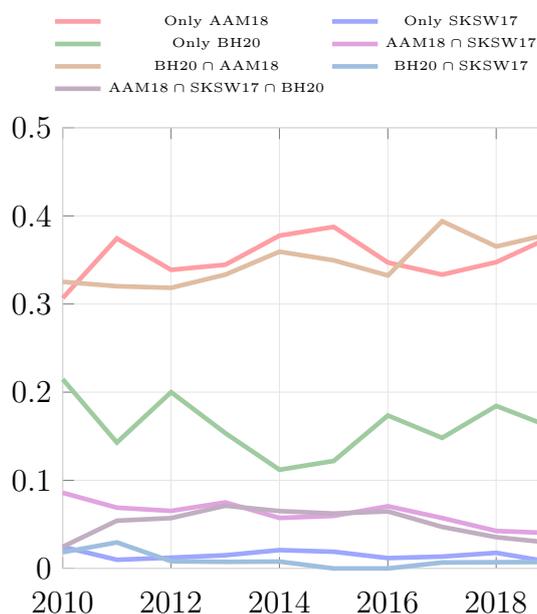


Note: the Venn diagram represents the overlaps between a selection of three zombie-identifications: AAM18, BH20, and SKSW17. Data refers to reporting year 2019. Identification is performed over the joint LIS of the three methods. Numbers labelling each area refer to the share of firms with respect to all firms identified as zombie by any of the three methods. AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), SKSW17: Storz et al. (2017).

Figure 3 presents a Venn diagram for a selection of zombie-identifications (AAM18, BH20, SKSW17). To make results comparable across methods, a common sample of firms corresponding to the joint LIS of all three methods is used. The Venn diagram shows firms jointly identified as zombies by different methods are relatively few. Firms identified as zombies by both AAM18 and BH20 represent 68% of those identified by BH20 alone, 36% of all firms identified as zombies by all three methods jointly, and 37% of those jointly identified by AAM18 and BH20, 45% of those identified by AAM18 alone.²⁶ When considering SKSW17 and BH20 instead, this number falls to 36% of firms identified as zombies with respect to those identified by SKSW17 alone (9% of those identified by BH20 alone,

²⁶Notice that when comparing the intersection of two sets taking as a reference only the largest of the two, the maximum overlap is constrained by the size of the smallest set.

Figure 4: Overlap fractions over time



Note: the overlap fractions of different intersections of the identifications AAM18, BH20, SKSW17 remain relatively constant in time. Each identification is computed over the associated LIS. Colours correspond to those for each area presented in Figure 3. AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), SKSW17: Storz et al. (2017).

7% of those jointly identified by SKSW17 and BH20, and 4.5% of those identified by all methods). These figures are somewhat better but still low when comparing AAM18 with SKSW17, where we find 66% of firms identified as zombies by both methods with respect to those identified by the most restrictive one SKSW17 (11% with respect to BH20, and 10% with respect to all methods).

These findings are relatively constant in time as shown in Figure 4. Here the evolution of the overlapping areas depicted in Figure 3 are tracked across time (colours are maintained across the two graphics), showing both constancy in terms of magnitude, as well as in terms of relative ordering.

Finally, a more complete analysis of the overlaps of different identifications is provided in Table 4. Here, the fraction of overlapping zombies is presented both in relative terms and with respect to the complete set of jointly identified firms. First, the overlap between each pair of definitions is reported with respect to the size of the smallest set. That is, for

any two sets of zombies A and B identified by two different methods, the *relative overlap* is $|A \cap B| / \min\{|A|, |B|\}$. Secondly, the fraction of zombies jointly identified by each pair of definitions is also reported with respect to the number of zombies identified by both. That is, for any two sets of zombies A and B identified by two different methods, the *overlap fraction* is the share $|A \cap B| / |A \cup B|$.

Table 4: Pairwise overlaps between zombie identifications

	Relative overlap [%]	Overlap fraction [%]
AAM18 \cap BH20	67.66	35.51
AAM18 \cap SKSW17	81.82	8.26
AAM18 \cap ACEE20	33.96	5.08
AAM18 \cap CHK08	33.33	0.38
BH20 \cap SKSW17	40.91	5.45
BH20 \cap ACEE20	7.14	1.54
BH20 \cap CHK08	0.0	0.0
SKSW17 \cap ACEE20	0.0	0.0
SKSW17 \cap CHK08	0.0	0.0
ACEE20 \cap CHK08	0.0	0.0

Note: *relative Overlap* refers to the fraction of firms jointly identified as zombies by each pair of methodologies with respect to the number of firms captured by the most restrictive methodology, that is $|A \cap B| / \min\{|A|, |B|\}$. *Overlap Fraction* refers instead to the fraction of firms jointly identified by both methods with respect to all firms captures by either of the two, that is $|A \cap B| / |A \cup B|$. Data refers to reporting year 2019. ACEE20: Acharya et al. (2020), AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SKSW17: Storz et al. (2017).

Most overlapping shares presented in Table 4 are worryingly small, with some being null. While this can be expected when considering the overlaps with CHK08 whose corresponding estimated absolute level of zombie share is very low, it is perhaps less so when considering the overlap with ACEE20. These results raise concerns regarding the ability to generalise empirical findings obtained under a specific zombie definition and associated LIS to others and most importantly to the level of whole economies. Moreover, as we have briefly discussed before, this leaves open questions concerning the ability of some of these methods to capture phenomena which can be attributed to the actual zombification of firms, meaning they receive subsidised credit, as opposed to them being simply financially weak and less profitable.

Most of the limitations we have discussed so far exist on both a conceptual and structural

level. As such it may perhaps be possible to fully address them only through better data, which is, however, often not available to most scholars. However, as we shall argue in the next section, a compromise can be found by acknowledging the fuzzy nature of this phenomenon and extending the binary zombie-identifications considered so far to fuzzy ones, which we show partially addresses these issues.

4 Beyond binary identifications

The discussion has so far considered exclusively dichotomic classifications of zombie firms. This is what we have referred to as *crisp zombie-identification* which, as defined in (1), lives on the Boolean domain. The concept of a zombie firm however is not clear-cut, so that one might find it desirable to associate each firm with a *degree of zombieness*, rather than a purely binary classification, in order to allow for finer analysis of the population of firms, of their characteristics, and weaknesses. Moreover, the thresholds used for identification, despite the efforts of the researchers to determine them empirically and provide sensible definitions, are nonetheless to an extent arbitrary. This provides a further motive to base the analysis of a firm's viability on fuzzy, non-binary variables. Crisp identifications are therefore prone to misclassification of non-zombies, both conceptually and methodologically. In this spirit, we shall now seek to enrich these crisp definitions by extending the domain from the Booleans to the real unit interval, so as to obtain fuzzy identifications which are more robust to misclassification.

Although most analysis found in the literature employ crisp zombie-identifications, there have been some attempts in the past to extract fuzzy variables capable of describing the degree of a firm's viability. One example is provided by the work of SST17, where the authors employ principal component analysis to reduce the dimensionality of the identification space to one by selecting the first principal component. The application of principal component analysis in this context, however, presents a number of deficiencies, most notably the violation of fundamental assumptions underlying the derivation of results concerning principal components. The linearity assumption is almost always

violated in the variables used for identification: their mutual relations often show strong non-linear behaviour instead (see e.g. Figure B.2); the principal components are not orthogonal, which is a reflection of the non-gaussianity of the associated copula; the distributions of the variables used for identification are fat-tailed which can lead to spurious results; finally, larger variance in one variable is not necessarily associated with greater relevance in identifying zombies, which is instead what principal component analysis would assume (Shlens, 2014; Taleb, 2020). This warrants additional care in applying principal component analysis to the creation of fuzzy identifications of zombie firms.

A different approach, first employed in the zombie literature by CHK08 and also adopted by Acharya et al. (2019), relies on fuzzy set theory. The authors define the set of zombie firms as a fuzzy set, whose elements can have a degree of membership rather than obeying a principle of bivalence. Practically, this is done by turning an indicator function confined on the Boolean domain $\{0, 1\}$ into a linear indicator with domain spanning the real unit interval $[0, 1]$. An example of such an indicator function is provided in Figure 5. The functional form of the indicator function might be seen as an additional degree of arbitrariness, but the main reason why this approach has not been employed elsewhere is most likely to be found in the fact that CHK08 employ a one-variable zombie identification, which is straightforward to generalise to a fuzzy one. On the contrary, most of the literature relies on higher-dimensional crisp identifications, which has so far prevented the application of a generalised fuzzy-theoretic approach.

Solutions to the limitations of both approaches introduced above exist. In the remainder of this section however, we shall focus our attention on generalising the latter approach as it is the most intuitive and simplest to generalise. The following Definition 8 therefore provides a generalisation of Definition 1.²⁷

Definition 8 *A fuzzy zombie-identification is a vector $\mathcal{V}_y^{(i)}$ of firm i characteristics in year y , together with a membership function $\mathcal{Z}_{i,y} : \mathcal{V}_y^{(i)} \rightarrow [0, 1]$ taking each firm's vector of characteristics onto the unit interval. The membership function $\mathcal{Z}_{i,y}$ generalises (1) to*

²⁷An application of this approach to introduce a fuzzy zombie measure on the method by Storz et al. (2017) was presented by us already in Helmersson et al. (2021), although with less technical detail.

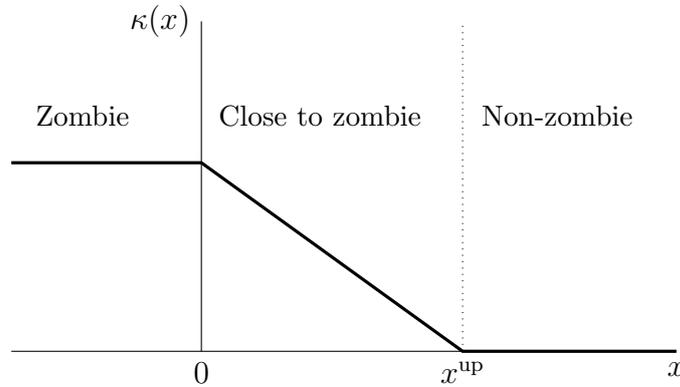
$$\mathcal{Z}_{i,y} = \left(\prod_{x \in \mathcal{V}_y^{(i)}} \kappa(x) \right)^{\frac{1}{|\mathcal{V}|}}, \quad (11)$$

where the kernel κ is monotonically decreasing over $[0, \infty[$ and non increasing over $]-\infty, 0]$. One simple functional form which we will consider here is the piece-wise linear form defined by

$$\kappa(x) = \mathbb{1}_{x < 0} + \frac{x^{\text{up}} - x}{x^{\text{up}}} \mathbb{1}_{0 \leq x < x^{\text{up}}} \quad (12)$$

for some upper threshold x^{up} , determining the thresholds of interest for each firm's characteristic in vector $\mathcal{V}_y^{(i)}$. Therefore, a firm is called a zombie (or non-viable) in year y if $\mathcal{Z}_{i,y} = 1$, healthy (or viable) if $\mathcal{Z}_{i,y} = 0$, and close-to-zombie (or quasi-zombie) of degree $\mathcal{Z}_{i,y}$ if $0 < \mathcal{Z}_{i,y} < 1$. We refer to $n = |\mathcal{V}|$ as the dimensionality of the zombie-identification.

Figure 5: Piece-wise linear membership function



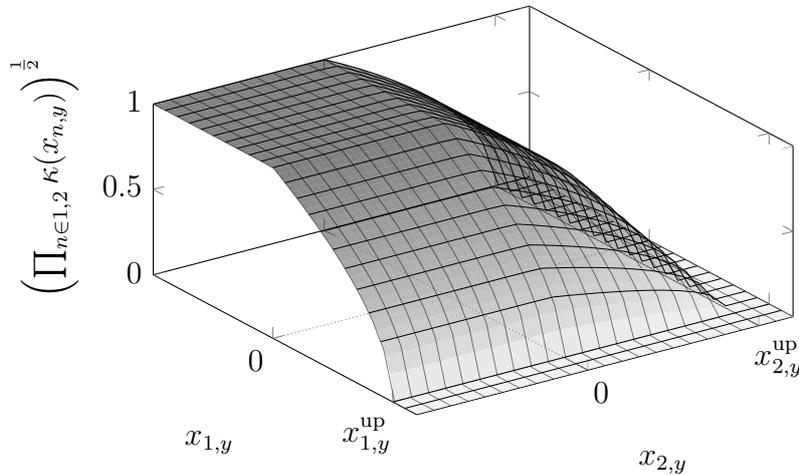
Note: the figure shows the piece-wise linear kernel $\kappa(x)$ from equation (12). The boundary is assumed to be linearly decreasing on the interval $[0, x^{\text{up}}]$. In the case of a one-dimensional identification such as that of CHK08, the three regions in this way determined, identify zombie, non-zombie, and close-to-zombie firms as labelled. Clearly, this does not necessarily have to be the case for a n -dimensional identification with $n > 1$.

The choice for the functional form of κ is clearly arbitrary, but one desirable property is for κ to be monotonically decreasing on the positive real line. This property ensures a lower score $\mathcal{Z}_{i,y}$ to firms further away from the origin. Further requiring convexity on $[0, \infty[$ can be of help to give more prominence to quasi-zombies closer to the origin, but a

piece-wise linear form as in (12), and depicted in Figure 5, provides the simplest and most intuitive option. For each variable one therefore defines an upper threshold above which the firm’s characteristic is deemed healthy with high certainty, such as the median $x^{\text{up}} = \bar{x}$, and a lower threshold which, as before, is incorporated by rescaling each element of $\mathcal{V}_y^{(i)}$ appropriately. Moreover, as for (1), the membership function (11) already incorporates the requirement of Y -consecutive years into the vector $\mathcal{V}_y^{(i)}$ by including the corresponding historical reported values of characteristic x : $(x_{y-(Y-1)}, \dots, x_{y-1}, x_y)$.

Effectively, the membership function (11) is constructed as a geometric mean of individual membership functions (the kernels $\kappa(x)$) applied to each firm’s characteristic used for identification, which is the most appropriate choice of average in this context. A depiction of the membership function $\mathcal{Z}_{i,y}$ is given in Figure 6 for the case of a two-dimensional vector of characteristics $|\mathcal{V}| = 2$. One clear consequence of that is that no firm will be classified as zombie or close to zombie if at least one of its characteristics used for identification is above the sample upper threshold x^{up} .²⁸

Figure 6: Two-dimensional membership function



Note: membership function $\mathcal{Z}_{i,y}$ for a fuzzy two-dimensional zombie set.

Through Definition 8 it is therefore possible to generalise any crisp zombie identification to a fuzzy one. We apply this approach to generalise Definitions 3, 5, 7, 2, and 5 (AAM18,

²⁸Relaxing this condition is possible simply by making a different choice for κ ; one suitable example could be an exponential functional form, or a power-law reflecting the distribution of the parameter under consideration.

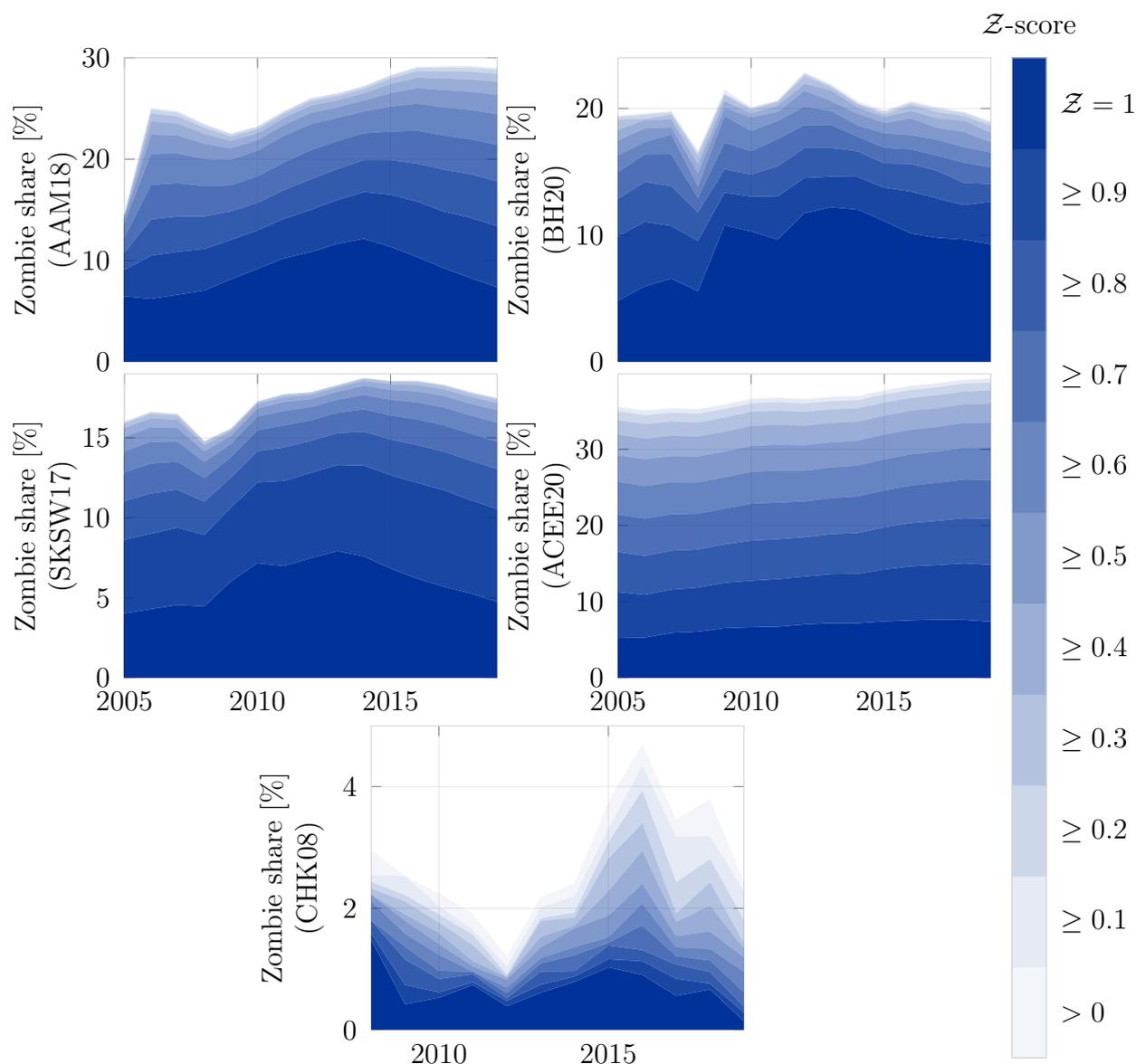
BH20, SKSW17, ACEE20, CHK08) to fuzzy zombie identifications, with $x^{\text{up}} = \bar{x}$ for all methods²⁹ except for Definition 5 (CHK08) where the upper threshold defined by the authors themselves is used instead (that is the upper threshold $x_{\text{CB08}}^{\text{up}} = 75\text{bp}$ for the interest rate gap). Note that this procedure leaves untouched the share of zombies identified by the crisp identification and simply adds close to zombie firms with an associated decreasing value of \mathcal{Z} . That is, for any identification one has $\sum_i \mathcal{Z}_{i,y} = \sum_i \mathbb{1}_{\mathcal{Z}_{i,y}=1}$. The resulting fuzzy distributions are presented in Figure 7.

Because the very concept of zombie firms is fuzzy, any binary definition faces the difficulty of setting a specific (arbitrary) threshold, therefore cutting out possibly ambiguous firms. The fuzzy approach improves on this limitation by assigning a degree of belonging to the set of zombies, therefore improving our ability to capture such non-viable firms. The additional distributional dimension introduced by the fuzzy metric provides additional information when analysing trends and the evolution of the population of zombie and quasi-zombie firms. As an example, the decline in share of zombies ($\mathcal{Z} = 1$) leaves open questions regarding how firms who have exited such a status are fairing. A binary black-and-white approach might suggest the idea that firms have recovered simply because they are not strictly classified as zombies any longer, but their performance might have actually only slightly improved, just by the amount necessary to lift them up from the zombie status. This is evidenced when looking at Figure 7, showing the recent decline of zombies, has not always been met by a similar decline in quasi-zombies ($\mathcal{Z} > 0$).

Regarding the fuzzy version of CHK08, there are two observations sticking out opposed to other methods. First, the distribution of firms with positive scores is not generally tilted towards higher scores (below 1) and second, the increase of firms with positive scores from 2014 onward is much steeper than for the firms with $\mathcal{Z} = 1$. Both features

²⁹We generally use time varying medians and in cases where the lower thresholds (from the binary identification methods) are set on a more granular level (e.g. sector-year or sector-country-year level) we adopt the same level of granularity to determine the upper thresholds. Also, if a median should happen to lie below the lower threshold, we instead determine the upper threshold on the sub-sample of firms for which the respective variable is above the lower threshold. Notice that the choice of median for x^{up} with definitions based on a single identification variable such as Definition 5 by Caballero et al. (2008), will by construction imply 50% of firms with $\mathcal{Z}_{i,y} > 0$. For this reason, in computing this specific fuzzy identification we stick with the threshold $x_{\text{CB08}}^{\text{up}} = 75\text{bp}$ proposed by the authors.

Figure 7: Distributions of fuzzy zombie scores

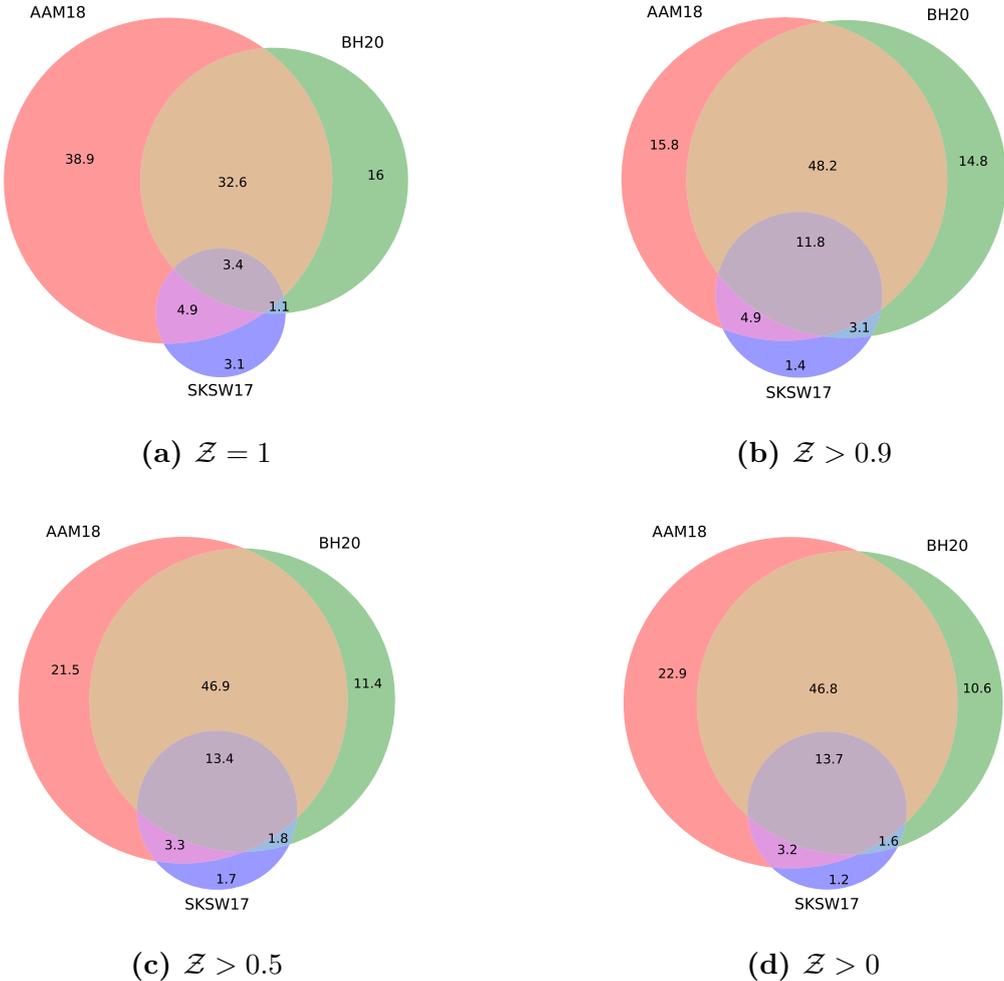


Note: distributions of fuzzy zombie scores respectively for fuzzy generalisation of Definitions 3 (AAM18), 5 (BH20), 7 (SKSW17), 2 (ACEE20), and 5 (CHK08). ACEE20: Acharya et al. (2020), AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SKSW17: Storz et al. (2017).

can be attributed to the fixed upper threshold and shifts in the underlying distribution of the interest rate gap. The first observation relates to the fact that the interest rate gap distribution is not concentrated around zero as firms usually pay a higher rate than the prime rate and there are relatively fewer firms close to an interest rate gap of zero. The

second observation can be explained by a narrowing of the distribution of interest rates paid by firms as the overall level of rates declined during this period, dominantly driven by monetary policy. Accordingly, as the upper threshold is not moved, a compression of the range of rates paid automatically shifts firms closer to the zero interest rate gap and hence increases the share of firms receiving positive CHK08 zombie scores.

Figure 8: Overlaps between zombie identifications for different fuzzy scores



Note: accounting for close-to-zombie firms through fuzzification progressively increases the agreement between different identification methodologies. Here, identification is performed for a selection of three methods (AAM18, BH20, SKSW17) on their joint LIS. Data refers to reporting year 2019. Numbers labelling each area refer to the fraction (in percentages) of firms identified as zombies with respect to zombies identified by either of the three methods on their joint LIS. AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), SKSW17: Storz et al. (2017).

Most importantly, introducing fuzziness brings increased consistency across different def-

initions. It is worth noting this may come at the cost of an increase in errors of the first kind, while reducing type II errors. Both variations are however mitigated by being associated with the fuzzy score \mathcal{Z} . Therefore, turning an identification from crisp to fuzzy has the drawback of making it less conservative, potentially including firms which are not actual zombies. This drawback is mitigated by associating decreasing importance to such first kind errors, which is reasonable to expect further away from the thresholds mandated by each zombie identification. Conversely, by reducing false negatives, that is reducing the likelihood of an identification missing potentially non-viable firms, increases consistency across identifications, which may be desirable especially when comparing empirical extrapolations and hence to generalise.

In Figure 8 the overlaps of identified firms is shown for a selection of three different methods (AAM18, BH20, SKSW17) where the shares are computed over their joint LIS, making evident the increased consistency both pairwise as well as in terms of zombies jointly identified by all three methods³⁰. A more complete analysis of all pairwise combinations is reported in Table 5. The fraction of overlapping zombies is presented both in relative terms and with respect to the complete set of jointly identified firms for both binary definitions as well as for their fuzzy extension, as done previously in Section 2. As expected, the share of jointly identified zombies increases when moving to fuzzy identification methods³¹.

Finally, the fuzzy analysis also allows for a comprehensive assessment of how firms evolve over time between zombie, quasi-zombie and non-zombie buckets. Figure 9 illustrates the probabilities of transitioning from one state to another for zombies, quasi-zombies and non-zombies according to the SKSW17 methodology in the period 2014-2015. Most recovering zombies or quasi-zombies are shown to turn to a healthy status while those that did not experience an improvement, saw their situation deteriorating, progressing towards a higher level of zombieness. Further, firms with a high fuzzy score ($\mathcal{Z} \geq 0.9$)

³⁰We restrict the representation to three methods as it is not possible to draw general n -venn diagrams for $n > 3$.

³¹Since CHK08 and BH20 can be applied only to a smaller sample of firms and especially CHK08 identifies only very few firms as zombies overall, the overlaps with other methods can in these cases be driven by only a few data points and accordingly should be interpreted cautiously.

Table 5: Pairwise overlaps between zombie identifications for different fuzzy scores

	Relative overlap [%]		Overlap fraction [%]		Mean productivity	
	$\mathcal{Z} = 1$	$\mathcal{Z} > 0$	$\mathcal{Z} = 1$	$\mathcal{Z} > 0$	$\mathcal{Z} = 1$	$\mathcal{Z} > 0$
AAM18 \cap BH20	67.66	81.69	35.51	36.33	0.97	1.07
AAM18 \cap SKSW17	81.82	87.69	8.26	17.78	0.93	1.06
AAM18 \cap ACEE20	33.96	38.33	5.08	17.78	1.12	1.11
AAM18 \cap CHK08	33.33	17.39	0.38	1.0	1.78	1.24
BH20 \cap SKSW17	40.91	52.17	5.45	18.27	0.99	1.03
BH20 \cap ACEE20	7.14	25.62	1.54	12.0	1.5	1.14
BH20 \cap CHK08	0.0	4.0	0.0	0.46	—	0.99
SKSW17 \cap ACEE20	0.0	28.12	0.0	7.33	—	1.1
SKSW17 \cap CHK08	0.0	6.45	0.0	1.29	—	1.08
ACEE20 \cap CHK08	0.0	30.23	0.0	2.52	—	1.07

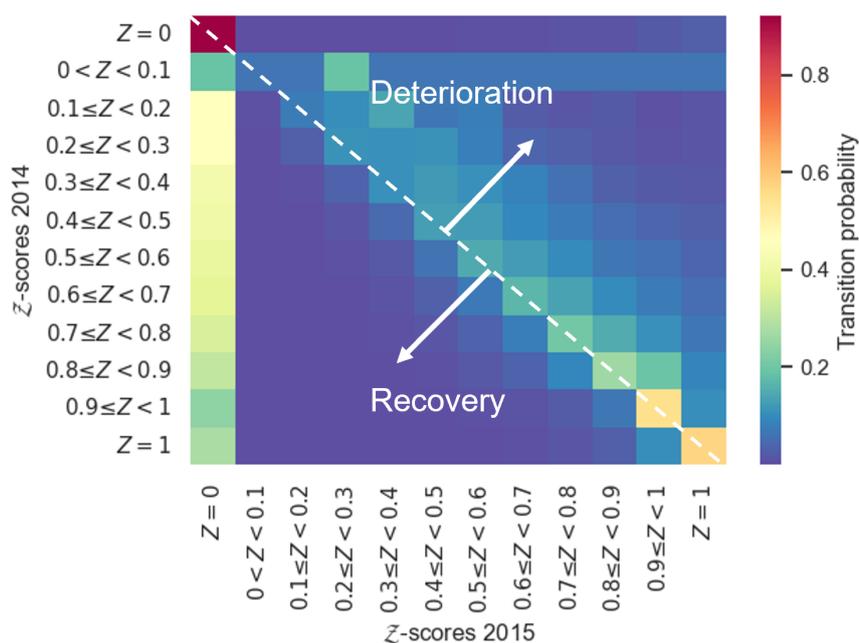
Note: *relative Overlap* refers to the fraction of firms jointly identified as zombies and close-to-zombie ($\mathcal{Z} > 0$) by each two methodologies with respect to the number of firms captured by the most restrictive methodology, that is $|A \cap B| / \min\{|A|, |B|\}$. *Overlap Fraction* refers instead to the fraction of firms jointly identified by both methods with respect to all firms captures by either of the two, that is $|A \cap B| / |A \cup B|$. Mean productivity refers to average total factor productivity among firms in the corresponding intersection. Results are shown for the reporting year 2019. ACEE20: Acharya et al. (2020), AAM18: Adalet McGowan et al. (2018), BH20: Banerjee and Hofmann (2020), CHK08: Caballero et al. (2008), SKSW17: Storz et al. (2017).

were more likely to remain in their (quasi-)zombie status while firms with lower scores had a higher probability to recover. We plan to extend on these transition dynamics in future research.

5 Conclusions

Firstly introduced in the context of Japan’s ‘Lost Decades’, the analysis of zombie firms has gained importance also in Europe, where extremely accommodative credit conditions and subdued economic growth following the sovereign debt crisis have hinted at an expanding zombification problem in the real economy. During the COVID-19 pandemic, extensive support measures have been put in place to curb the pandemic-induced economic fallout, and the risk of such measures having indirectly assisted non-viable firms supported the zombification phenomenon to acquire even more relevance in the public debate.

Figure 9: Transition matrix



Note: the transition matrix depicts the probability of transitioning from one state to another, following the fuzzy approach for the SKSW17 methodology in the period 2014-2015. Most recovering zombies or quasi-zombies are shown to turn to a healthy status while those that did not experience an improvement, saw their situation deteriorating, progressing towards a higher level of zombieness.

In this context, we survey the main approaches in the literature, highlighting their heterogeneity both at a conceptual and quantitative level. We provide a general framework formalising the concept of crisp zombie-identifications and examine the shares and trends of zombie firms in the economy, comparing the five main methodologies and their respective LIS as well as their joint LIS over a time span of sixteen years. We shed light on their different properties and the level of consistency between approaches. Although the shares and trends of zombies are somewhat consistent between methods, we find that these are identified using different subsets of the population of firms.

Acknowledging important data limitations faced by scholars in the identification design, we expand the literature by providing a generalised fuzzy-theoretic approach. Such methodology presents the threefold contribution of providing a more nuanced measure of zombieness, decreasing the arbitrariness in choosing thresholds, and adding information on the evolution in the shares and trends of zombie and quasi-zombie firms. Interestingly,

while most binary methods show a decreasing trend in the share of zombies in recent years, the results inferred from the fuzzy approach suggest that this decrease has not been met by a similar reduction in quasi-zombie firms. Most importantly, analysing the overlaps of zombie firms as identified by different methods, we find that including fuzziness increases consistency and allows for a better comparison between different identifications.

Despite increased consistency between methods, these findings show that a mismatch between the conceptual notion of a zombie and the different methodologies for its actual identification exists. While our findings do not confute the important results achieved in the literature on zombie firms, they suggest that cross-methodology heterogeneity makes it difficult to draw clear conclusions on the extent of the zombification problem in the economy. Any discussion on corporate zombification and generalisation of empirical findings should be conducted cautiously, take into account different identifications and their respective limitations.

Accordingly, it is also difficult to design an optimal policy mix to tackle zombification. Indeed, it might be the best option for policy makers not to focus particularly at zombie firms at all. The zombification problem may be best addressed by general economic policies that stimulate innovation, productivity and growth (Bindseil and Schaaf, 2020) as well as ensuring that the banking system is well capitalised and not incentivised to evergreen loans. Similarly, by reducing the cost of market exit, a reformation of insolvency regimes could help preventing the emergence of zombie firms (McGowan et al., 2017; Becker and Ivashina, 2021).

References

- V. V. Acharya, T. Eisert, C. Eufinger, and C. Hirsch. Whatever it takes: The real effects of unconventional monetary policy. *The Review of Financial Studies*, 32(9):3366–3411, 2019.
- V. V. Acharya, M. Crosignani, T. Eisert, and C. Eufinger. Zombie credit and (dis-)

- inflation: evidence from Europe. Working Paper 27158, National Bureau of Economic Research, 2020.
- M. Adalet McGowan, D. Andrews, and V. Millot. The walking dead? Zombie firms and productivity performance in OECD countries. *Economic Policy*, 33(96):685–736, 2018.
- E. I. Altman. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23(4):589–609, 1968.
- D. Andrews and F. Petroulakis. Breaking the Shackles: Zombie Firms, Weak Banks and Depressed Restructuring in Europe. Working Paper Series 2240, European Central Bank, 2019.
- M. Bajgar, G. Berlingieri, S. Calligaris, C. Criscuolo, and J. Timmis. Coverage and representativeness of Orbis data. OECD Science, Technology and Industry Working Papers 2020/06, OECD Publishing, 2020.
- R. Banerjee and B. Hofmann. The rise of zombie firms: causes and consequences. *BIS Quarterly Review September*, 2018.
- R. N. Banerjee and B. Hofmann. Corporate zombies: Anatomy and life cycle. BIS Working Papers 882, Bank for International Settlements, 2020.
- Bank of Korea. Financial stability report. 2013.
- F. Bargagli Stoffi, M. Riccaboni, and A. Rungi. Machine learning for zombie hunting. firms’ failures and financial constraints. Available at SSRN: <https://ssrn.com/abstract=3588410>, 2020.
- B. Becker and V. Ivashina. Corporate Insolvency Rules and Zombie Lending. In *Proceedings of the ECB Forum on Central Banking 2021*, pages 32–52. European Central Bank, 2021.
- U. Bindseil and J. Schaaf. Zombification is a real, not a monetary phenomenon: Exorcising the bogeyman of low interest rates. *VoxEU*, 2020. URL

<https://voxeu.org/content/zombification-real-not-monetary-phenomenon-exorcising-bogeyman-low-interest-rates>.

Bureau van Dijk. *Orbis - Data Guide*, 2017.

R. J. Caballero, T. Hoshi, and A. K. Kashyap. Zombie lending and depressed restructuring in Japan. *American economic review*, 98(5):1943–77, 2008.

S. Fukuda and J. Nakamura. Why Did ‘Zombie’ Firms Recover in Japan? *The World Economy*, 34:1124–1137, 2011.

P.-O. Gourinchas, S. Kalemli-Özcan, V. Penciakova, and N. Sander. COVID-19 and SME Failures. NBER Working Papers 27877, National Bureau of Economic Research, Inc, 2020.

I. Hallak, P. Harasztosi, and S. Schich. Fear the Walking Dead? Incidence and Effects of Zombie Firms in Europe. JRC Technical Report, Publications Office of the European Union, 2018.

T. Helmersson, L. Mingarelli, B. Mosk, A. Pietsch, B. Ravanetti, T. Shakir, and J. Wendelborn. Corporate zombification: post-pandemic risks in the euro area. *Financial stability review*, European Central Bank, May 2021.

T. Hoshi. Economics of the living dead. *The Japanese Economic Review*, 57(1):30–49, 2006.

K. Imai. A panel study of zombie SMEs in Japan: Identification, borrowing and investment behavior. *Journal of the Japanese and International Economies*, 39(C):91–107, 2016.

S. Kalemli-Özcan, B. Sorensen, C. Villegas-Sanchez, V. Volosovych, and S. Yesiltas. How to construct nationally representative firm level data from the Orbis global database: New facts and aggregate implications. Working Paper 21558, National Bureau of Economic Research, 2015.

- E. Kane. Dangers of capital forbearance: The case of the FSLIC and ‘Zombie’ S&Ls. *Contemporary Economic Policy*, 5(1):77–83, 1987.
- L. Laeven, G. Schepens, and I. Schnabel. Zombification in Europe in times of pandemic. *VoxEU*, 2020. URL <https://voxeu.org/article/zombification-europe-times-pandemic>.
- M. A. McGowan, D. Andrews, and V. Millot. Insolvency regimes, zombie firms and capital reallocation. Oecd economics department working papers, OECD Publishing, 2017.
- S. Nurmi, J. Vanhala, and M. Virén. The life and death of zombies – evidence from government subsidies to firms. Research Discussion Papers 8/2020, Bank of Finland, 2020.
- J. Peek and E. Rosengren. Unnatural Selection: Perverse Incentives and the Misallocation of Credit in Japan. *The American Economic Review*, 2005.
- F. Schivardi, E. Sette, and G. Tabellini. Credit misallocation during the European financial crisis. Temi di Discussione (Working Paper) 1139, Bank of Italy, 2017.
- T. Sekine, K. Kobayashi, and Y. Saita. Forbearance Lending: The Case of Japanese Firms. *Monetary and Economic Studies*, 21(2):69–92, 2003.
- J. Shlens. A tutorial on principal component analysis. *arXiv preprint arXiv:1404.1100*, 2014.
- M. Storz, M. Koetter, R. Setzer, and A. Westphal. Do we want these two to tango? On zombie firms and stressed banks in Europe. Working Paper Series 2104, European Central Bank, 2017.
- N. N. Taleb. Statistical consequences of fat tails: Real world preasymptotics, epistemology, and applications. *arXiv preprint arXiv:2001.10488*, 2020.
- S. Urionabarrenetxea, J. D. Garcia-Merino, L. San-Jose, and J. L. Retolaza. Living with zombie companies: do we know where the threat lies? *European Management Journal*, 36(3):408–420, 2018.

C. Zhang, Y. Chen, and H. Zhou. Zombie Firms and Soft Budget Constraints in the Chinese Stock Market. *Asian Economic Journal*, 34(1):51–77, 2020.

A Appendix

Table A.1: Identifications and their variations

Identification	\mathcal{V}_y
Acharya et al. (2020)	$IC_y - M_{\text{cty-sec}}(IC_y), G_y^{IR}$
Adalet McGowan et al. (2018)	$IC_{y-w} - 1$ for $w \in \{0, 1, 2\}$ and $10 - T$
Nurmi et al. (2020)	$IC_{y-w} - 1$ for $w \in \{0, 1, 2\}$
Banerjee and Hofmann (2018)	$IC_{y-w} - 1$ for $w \in \{0, 1, 2\}$, $10 - T$, and $q_{y-w}^T - M_{\text{SEC}}(q_{y-w}^T)$
Hallak et al. (2018)	$IC_{y-w} - 1$ for $w \in \{0, 1, 2, 3, 4\}$ and $10 - T$
Banerjee and Hofmann (2020)	$IC_{y-w} - 1, q_{y-w}^T - M_{\text{SEC}}(q_{y-w}^T)$ for $w \in \{0, 1\}$
Bargagli Stoffi et al. (2020)	above 9th decile of predictions of failure for at least three years.
Caballero et al. (2008)	G_y^{IR}
Fukuda and Nakamura (2011)	$G_y^{IR}, EBIT_{y-}$ risk-free $E_y^I, 1/2 * A_{(y-w)} - D_y$ for $w \in \{0, 1\}$, increased borrowing in y
Zhang et al. (2020)	$G_y^{IR}, EBIT_{y-} - NRGL_{y-}$ risk-free $E_y^I, 1/2 * A_{(y-w)} - D_y$ for $w \in \{0, 1\}$, increased borrowing in y
Imai (2016)	$G_y^{IR}, \sum_{m=0}^T (EBIT_{t-m} - \text{risk-free } E_{t-m}^I), 1/2 * A_{(y-w)} - D_y$ for $w \in \{0, 1\}$, increased borrowing in y
Acharya et al. (2019)	$G_y^{IR}, \text{rating} < BB$, and constant syndicate composition
Schivardi et al. (2017)	$\langle \text{RoA}_y \rangle_3 - \text{CoC}, M(L_y) - L_y$
Storz et al. (2017)	$\text{RoA}_{y-w}, S_{y-w}^C - 5\%, I_{y-w}^N$ for $w \in \{0, 1\}$
Andrews and Petroulakis (2019)	$\text{RoA}_{y-w} \vee I_{y-w}^N, S_{y-w}^C - 5\%$ for $w \in \{0, 1, 2\}$
Urionabarrenetxea et al. (2018)	$EZ_{Index} = \sum_{i=1}^4 k_i I_{i, \text{standardized}}$ where $I = \left\{ \frac{\text{Neg. } E_i}{\sum_{i=1}^N \text{Neg. } E_i}, -\frac{\text{Neg. } E_i}{A_i}, 1 - \frac{A_i}{D_i}, 1 - \frac{CF_i}{D_i} \right\}$ and k is the proportion of variance explained by I

Note: main identifications are marked in bold, followed by associated variations. Urionabarrenetxea et al. (2018) built a composite indicator aimed at identifying ‘extreme zombie’ firms and adopted a static and dynamic approach. The table reports the static approach only. We denote by $\langle x \rangle$ the mean of x , and by $\langle x \rangle_y$ its y -years moving average. CoC : cost of capital, IC : interest coverage ratio, T : firm age, q^T : Tobin’s q , RoA : return on assets, S^C : debt servicing capacity, I^N : net investments, L : leverage, G^{IR} : interest rate gap, A : total assets, D : total external debt, $NRGL$: non-recurring gains and losses, EZ_{Index} : extreme zombie index, E : equity, CF : cash flow.

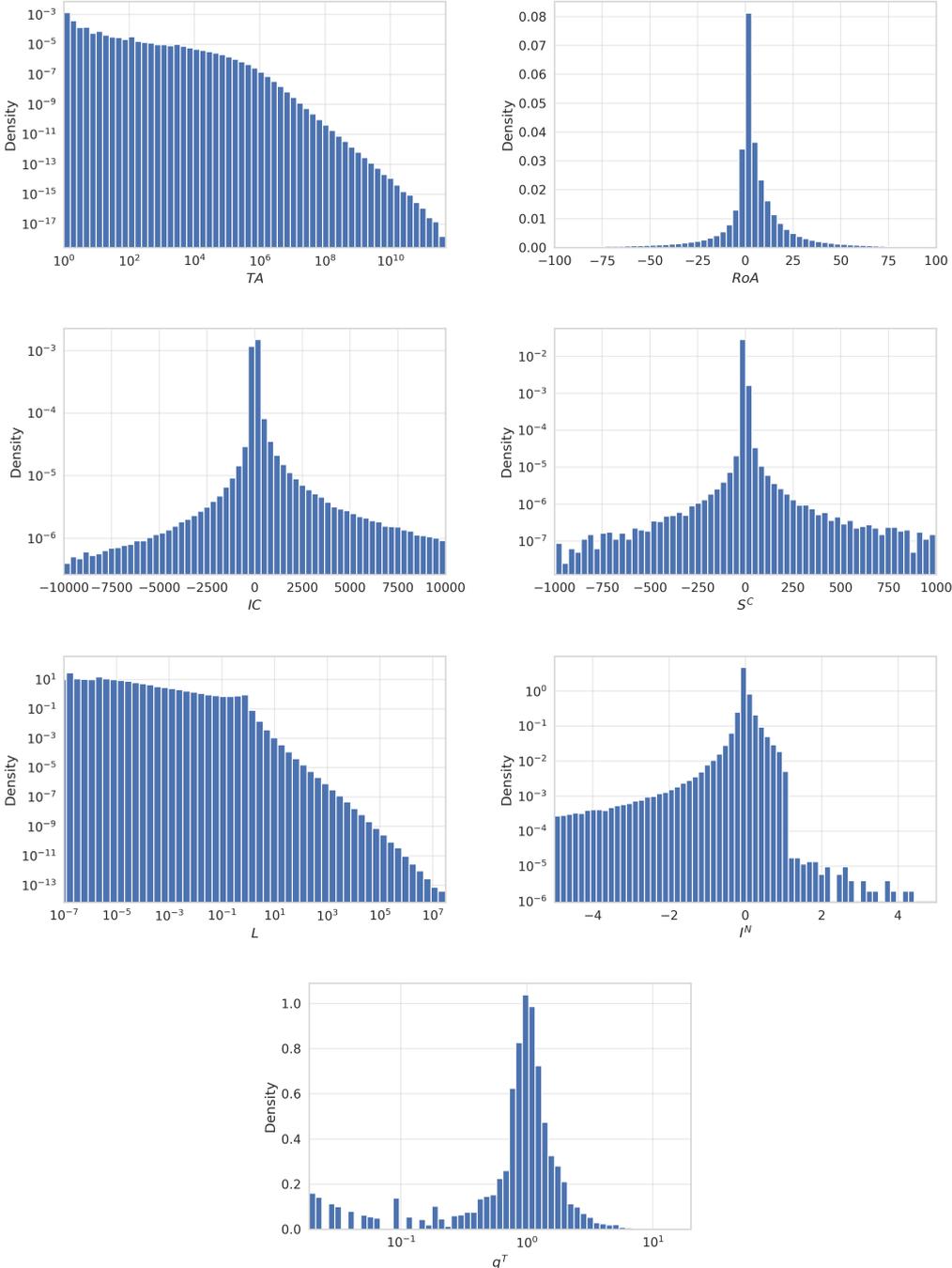
Table A.2: Variable definitions in original papers and this paper's implementation

Method	Variable	Original definition	Present definition
AAM18	IC	Operating income/interest expenses	EBIT/interest paid
	T	-	year - year of incorporation
BH20	IC	EBIT/interest payments	EBIT/interest paid
	q^T	(Market cap + current liabilities + non-current liabilities)/total assets	Same
SKSW17	RoA	Net income/total assets	Same
	S^C	EBITDA/financial debt	EBITDA/(current + non-current liabilities)
	I^N	Net change in total fixed assets relative to previous year	Fixed assets _{t} - fixed assets _{$t-1$}
ACEE20	IC	EBIT/interest expense	EBIT/interest paid
	L	(loans + short-term credit + long-term debt)/total assets.	(Current + non-current liabilities - provisions)/total assets
	G^{IR}	$IR - IR$ of AAA-rated firms	Same
	IR	Interest expenses/(outstanding loans + credit + bonds)	Interest paid/(current + non-current liabilities - provisions)
AAA-rated firms	BS	Short-term bank loans (< 1 year)	Current loans and overdrafts
	BL	Long-term bank loans (> 1 year)	Bank loans + current portion of long-term debt
Bonds		Total bonds outstanding (including convertible bonds and warrant-attached bonds)	Other short-term debt + debentures and convertible debt + other long-term interest bearing debt
	CP	Commercial paper outstanding	0
CHK8	R	Interest payments	Interest paid
	r_s	Average short-term prime rate	Min. interest rate for NFCs among EA MFIs with maturity ≤ 1 year
	r_l	Average long-term prime rate	Min. interest rate for NFCs among EA MFIs with maturity > 5 years
	rcb	Min. observed coupon rate on any conv. corp. bond issued in the last 5 years	Same
	G^{IR}	equation (7)	equation (7)

Note: IC : interest coverage ratio, T : firm age, q^T : Tobin's q , RoA : return on assets, S^C : debt servicing capacity, I^N : net investments, L : leverage, G^{IR} : interest rate gap, IR : interest rate.

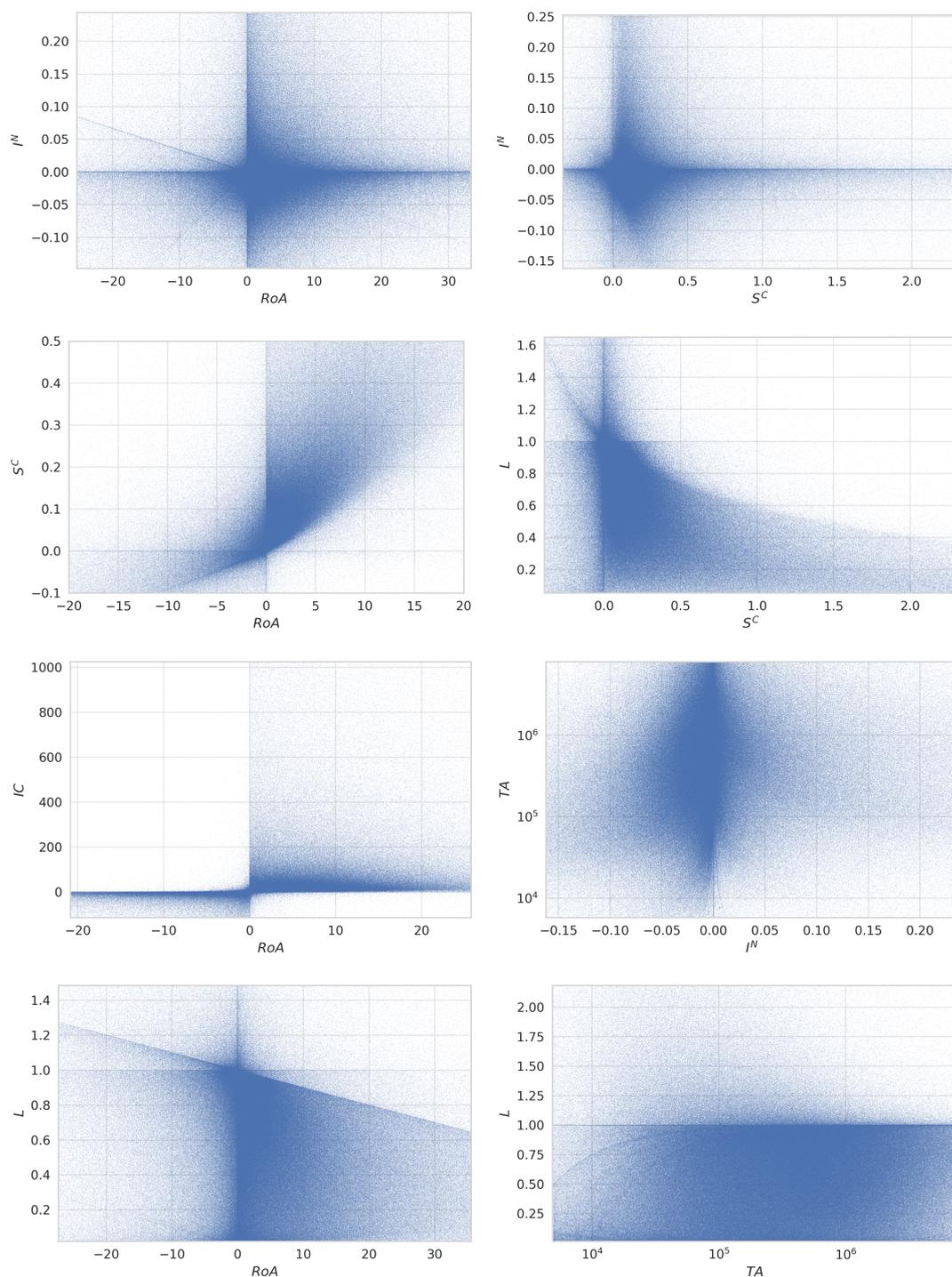
B Further descriptives

Figure B.1: Densities of some of the relevant variables from Table A.2.



Note: the bivariate distributions reveal a strong non-linear dependency structure. Axis restricted to regions of interest. TA : total assets, RoA : return on assets, IC : interest coverage ratio, SC : debt servicing capacity, L : leverage, I^N : net investments, q^T : Tobin's q.

Figure B.2: Bivariate distributions of some of the relevant variables from Table A.2



Note: the bivariate distributions reveal a strong non-linear dependency structure. Axis restricted to regions of interest. IC : interest coverage ratio, TA : total assets, RoA : return on assets, S^C : debt servicing capacity, I^N : net investments, L : leverage.

Acknowledgements

We are grateful to Benjamin Hartung, Paloma Lopez-Garcia, Giulio Nicoletti, Ralph Setzer, Mika Tujula and Peter Welz for useful comments and discussions during early stages of this work. We would further like to thank Takeo Hoshi, Ivan Huljak and other participants at the Virtual Peer Workshop on Non-Viable ('Zombie') Companies at the Joint Vienna Institute for fruitful discussions and valuable feedback.

The views expressed here are those of the authors and do not necessarily represent the views of the European Central Bank or the European Bank for Reconstruction and Development. All errors are our own.

Luca Mingarelli

European Central Bank, Frankfurt am Main, Germany; email: Luca.Mingarelli@ecb.europa.eu

Beatrice Ravanetti

European Bank for Reconstruction and Development, London, United Kingdom; email: ravanetb@ebrd.com

Tamarah Shakir

European Central Bank, Frankfurt am Main, Germany; email: Tamarah.Shakir@ecb.europa.eu

Jonas Wendelborn

European Central Bank, Frankfurt am Main, Germany; email: Jonas.Wendelborn@ecb.europa.eu

© European Central Bank, 2022

Postal address 60640 Frankfurt am Main, Germany

Telephone +49 69 1344 0

Website www.ecb.europa.eu

All rights reserved. Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from www.ecb.europa.eu, from the [Social Science Research Network electronic library](#) or from [RePEc: Research Papers in Economics](#). Information on all of the papers published in the ECB Working Paper Series can be found on the [ECB's website](#).

PDF

ISBN 978-92-899-5391-7

ISSN 1725-2806

doi:10.2866/836398

QB-AR-22-108-EN-N