



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Nonstationary Z-score measures

Citation for published version:

Mare, D, Moreira, F & Rossi, R 2017, 'Nonstationary Z-score measures', *European Journal of Operational Research*, vol. 260, no. 1, pp. 348-358. <https://doi.org/10.1016/j.ejor.2016.12.001>

Digital Object Identifier (DOI):

[10.1016/j.ejor.2016.12.001](https://doi.org/10.1016/j.ejor.2016.12.001)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

European Journal of Operational Research

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Nonstationary Z-score measures

Davide Salvatore Mare¹, Fernando Moreira¹, and Roberto Rossi*²

¹Business School, Credit Research Centre, The University of Edinburgh, UK

²Business School, The University of Edinburgh, UK

September 7, 2016

Abstract

In this work we develop advanced techniques for measuring bank insolvency risk. More specifically, we contribute to the existing body of research on the Z-Score. We develop bias reduction strategies for state-of-the-art Z-Score measures in the literature. We introduce novel estimators whose aim is to effectively capture nonstationary returns; for these estimators, as well as for existing ones in the literature, we discuss analytical confidence regions. We exploit moment-based error measures to assess the effectiveness of these estimators. We carry out an extensive empirical study that contrasts state-of-the-art estimators to our novel ones on over ten thousand banks. Finally, we contrast results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva. Our work has important implications for researchers and practitioners. First, accounting for nonstationarity in returns yields a more accurate quantification of the degree of solvency. Second, our measure allows researchers to factor in the degree of uncertainty in the estimation due to the availability of data.

JEL-Classification: C20, C60, G21

Keywords: bank stability; prudential regulation; insolvency risk; financial distress; Z-Score

*Corresponding author: University of Edinburgh Business School, 29 Buccleuch Place, EH89JS, Edinburgh, UK.
Phone: +44 (0)1316515077, Fax: +44 (0)1316513197, E-mail: roberto.rossi@ed.ac.uk.

1 Introduction

The measurement of financial stability in banking aims at assessing the degree of solvency of individual financial institutions or of the overall sector. Financial stability in banking has been investigated in relation to a broad variety of determinants such as corporate governance [Laeven and Levine, 2009], competition [Fiordelisi and Mare, 2014], efficiency [Fiordelisi et al., 2011], the diversification strategy of shareholders [García-Kuhnert et al., 2013], and creditor rights and information sharing [Houston et al., 2010]. It is paramount both under the regulatory and supervisory perspectives because it drives policy choices to assure the resilience and the functional working of the banking sector, along with optimal social welfare and economic growth.

Bank financial instability is proportional to the likelihood that creditors of a bank are not repaid partially or in full. This comes to be true when financial losses (expected and unexpected) are not covered with provisions or capital and the value of the assets is not sufficient to repay in full debt obligations. In practice, assessment of bank's insolvency risk should capture both the variability in revenues and the buffers — both in terms of reserves and equity — to absorb financial losses [Hannan and Hanweck, 1988, Boyd and Runkle, 1993, Delis et al., 2014].

While the Operational Research (OR) literature has assessed bank stability by means of methods such as Neural Networks, Support Vector Machine and Data Envelopment Analysis (see, for example Kumar and Ravi [2007], Demyanyk and Hasan [2010], Ioannidis et al. [2010]), many studies in the empirical banking literature have employed approaches based on accounting ratios which seem to be more appealing to practitioners due to the simplicity and ready availability of data for these techniques compared to more sophisticated methods. Accounting ratios, such as non-performing loans to total loans or the level of provisioning [Houston et al., 2010, Fiordelisi et al., 2011], are employed to capture different risk dimensions although the focus is on specific narrower aspects such as credit risk, operational risk, liquidity risk and market risk. A more comprehensive accounting ratio, the Z-Score, is by far the most widely used in the literature for estimating the overall bank solvency [Boyd et al., 2006, Mercieca et al., 2007, Laeven and Levine, 2009, Fiordelisi and Mare, 2014, Bolton et al., 2015]. It is also used by multilateral organisations — for instance, it is included among the indicators of The Global Financial Development Database (World Bank) — to proxy for financial stability for the overall banking sector. This ratio combines together information on performance (for instance, return on assets indicator), leverage (equity to assets indicator) and risk (for instance, standard deviation of return on assets). A bank can therefore be classified as being less stable, or closer to insolvency, if it shows lower performance, it is less capitalized or it has a higher degree of variation in returns.

Constructing Z-Score measures for cross-sectional analysis is fairly straightforward; however, there is little consensus in the literature on how to construct a Z-Score measure when stochastic returns are nonstationary. Lepetit and Strobel [2013] survey existing studies and contrast a number of approaches to accomplish this task. Nevertheless, existing approaches to compute the Z-Score entail biases and are not designed to fully capture nonstationarity of returns. The goal of our work is to propose a rigorous approach to reduce estimation bias, capture nonstationary returns, and account for estimation errors while constructing a Z-Score measure. To achieve this, we build on the study of Lepetit and Strobel [2013] and contribute to the Z-Score literature as follows:

- We introduce bias reduction strategies consistent to the theoretical premises of the Z-Score to improve effectiveness of estimators in [Lepetit and Strobel, 2013].
- We introduce a novel estimator whose aim is to effectively capture nonstationary stochastic returns; for this estimator, as well as for existing ones in the literature, we discuss analytical confidence regions. Due to the small sample size typically employed, we argue that accounting for estimation errors is important to obtain consistent ranking of financial institutions according to their overall risk of solvency.
- For the first time in the literature, we exploit moment-based error measures — a novel tool for ranking forecasters recently introduced in [Prestwich et al., 2014] — to assess the effectiveness of existing estimators from [Lepetit and Strobel, 2013] and of our novel ones; these estimators are tested on random variates sampled from a selection of stochastic processes featuring characteristics, such as trends and seasonalities, commonly investigated in time series analysis.
- We carry out an extensive empirical study that contrasts results obtained with the aforementioned estimators on a large dataset from the banking sector covering the period 2005–2013.
- Finally, to identify actual financial distress events, we contrast results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva.

The remainder of the paper is structured as follows. Section 2 defines the Z-Score. Section 3 summarises the current methods for computing the Z-Score; Section 4 outlines some important limitations. Section 5 illustrates how to remove bias from existing methods for computing the Z-Score. Section 6 introduces our novel nonstationary estimators, while Section 7 presents analytical confidence regions for these estimators. Section 8 puts to the test our novel estimators against existing methods for computing the Z-Score. Section 9 presents an empirical study based on data from BvD Bankscope, covering the period 2005–2013. The study investigates ranking discrepancies

observed between our novel estimators and existing ones in the literature; it also illustrates two case studies that contrast results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva. Section 10 draws conclusions.

2 The Z-Score

The theoretical foundations of the Z-Score date back to the work of Roy [1952]. In the literature it is common to define bank insolvency as a state in which the sum of the equity-to-asset ratio (EA) and the return on asset ratio (ROA) is less or equal to zero, i.e. $EA + ROA \leq 0$. In presence of uncertainty, it is customary to consider EA deterministic [Boyd and Graham, 1986] and model ROA as a random variable with finite mean $\mu(ROA)$ and variance $\sigma^2(ROA)$. Following the discussion outlined in [Hannan and Hanweck, 1988, Boyd and Runkle, 1993] one may exploit Bienaymé-Chebyshev inequality to determine an upper bound on the probability of insolvency as follows

$$\Pr\{ROA \leq -EA\} \leq Z^2 \quad (1)$$

where

$$Z \equiv \frac{EA + \mu(ROA)}{\sigma(ROA)} > 0 \quad (2)$$

denotes the “Z-Score,” a value that is inversely proportional to an upper bound on the probability of insolvency $\Pr\{ROA \leq -EA\}$ even for weak distributional assumptions [Strobel, 2011, Lepetit and Strobel, 2015].

It is interesting to observe that for the special case in which ROA is normally distributed with mean $\mu(ROA)$ and standard deviation $\sigma(ROA)$, the Z-Score admits an intuitive interpretation: it represents the number of standard deviations by which returns have to diminish in order to deplete the equity of a bank or a banking system. More formally,

$$\Pr\{ROA \leq -EA\} = \Phi\left(\frac{-EA - \mu(ROA)}{\sigma(ROA)}\right) = 1 - \Phi\left(\frac{EA + \mu(ROA)}{\sigma(ROA)}\right),$$

where Φ is the standard normal cumulative distribution function and the second equality holds because the normal distribution is symmetric. Figure 1 illustrates the rationale for using the Z-Score as a measure of the overall bank stability. It is clear that a higher Z-Score implies a higher degree of solvency and therefore it gives a direct measure of bank stability.

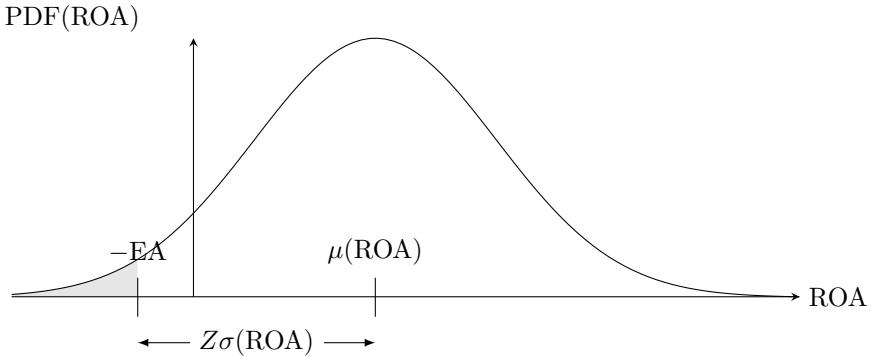


Figure 1: This figure illustrates the Z-Score for normally distributed returns. If ROA is less or equal to $-EA$ we observe a default, the probability of which is represented in gray. Higher values of the Z-Score denote higher stability because the distance to default is larger.

3 Existing approaches to compute the Z-Score

As discussed in [Lepetit and Strobel, 2013] the implementation of Z-Score measures for cross-sectional analysis does not raise major questions; however, when stochastic returns are nonstationary constructing a Z-Score is less straightforward. When we introduce the time dimension there are few strategies that can be adopted, which we shall next summarise. In Table 1, we report the elementary information for the calculation of the time-varying Z-Score.

| Notation | Description |
|------------------|---|
| t | time index |
| NOPAT_t | net operating profit after taxes at time t |
| TOTA_t | total assets at time t |
| TE_t | total equity at time t |
| ROA_t | return on asset at time t , defined as $\text{NOPAT}_t/\text{TOTA}_t$ |
| EA_t | equity over asset at time t , defined as $\text{TE}_t/\text{TOTA}_t$ |

Table 1: Elementary information for the calculation of time-varying Z-Score

Consider a random variable ρ with unknown distribution representing the stochastic process that generates realisations ROA_t . We are given a set ROA_T of realisations sampled from ρ for T periods, i.e. $T \equiv \{t-k, \dots, t\}$, where t denotes the most recent time period. Let $\mu(\text{ROA}_T)$ represent the sample mean of these realisations and $\sigma(\text{ROA}_T)$ their sample standard deviation.

Lepetit and Strobel [2013] survey the empirical banking literature and propose a new approach for the estimation of the Z-Score. The authors suggest that in the literature there are five main approaches to compute the Z-Score for a given bank; these are reported below as Z_1, \dots, Z_5 .

We shall consider first

$$Z_1 \equiv \frac{\mu(\text{ROA}_T) + \mu(\text{EA}_T)}{\sigma(\text{ROA}_T)} \quad (3)$$

where $T = t-2, \dots, t$. According to this measure, originally discussed in [Boyd et al., 2006] section

III.A, realisations EA_t also come from a random variable whose sample mean is estimated from past observations. In certain settings, equity over asset at time t should be modeled as a random variable; however, if one decides to model EA as a random variable, then the Z-score should also account for the variability associated with this random variable, namely $\sigma(EA)$.

The second measure [Yeyati and Micco, 2007] is a rather straightforward implementation of the classic Z-Score discussed in Section 2.

$$Z_2 \equiv \frac{\mu(\text{ROA}_T) + EA_t}{\sigma(\text{ROA}_T)} \quad (4)$$

where $T = t - 2, \dots, t$. Now equity over asset is a known value, while mean and standard deviation of ROA are estimated from past realisations observed in the last three periods. Unfortunately, as we shall see in the next section, when sample size is small the reliability of this measure is low because of statistical bias in the estimation of the standard deviation of ROA.

The third measure is introduced in [Hesse and Cihák, 2007]

$$Z_3 \equiv \frac{\text{ROA}_t + EA_t}{\sigma(\text{ROA}_T)} \quad (5)$$

where $T = 1, \dots, t$; it represents a model that considers only the last period value for ROA, while it computes $\sigma(\text{ROA})$ over the whole sample horizon. In this case, it is not clear what random variable is being estimated and no clear judgement can be made on the statistical properties of this estimator.

The fourth measure, originally discussed in [Boyd et al., 2006] section III.B, is based on the concept of “instantaneous standard deviation,” which is defined as follows

$$\sigma_{\text{inst}}(\text{ROA}_T) \equiv |\text{ROA}_t - \mu(\text{ROA}_T)| \quad (6)$$

where $T = 1, \dots, t$ and $|x|$ denotes the absolute value of x . Z_4 can then be computed as

$$Z_4 \equiv \frac{\text{ROA}_t + EA_t}{\sigma_{\text{inst}}(\text{ROA}_T)} \quad (7)$$

where $T = 1, \dots, t$. This measure features drawbacks similar to those discussed for Z_3 .

Finally,

$$Z_5 \equiv \frac{\mu(\text{ROA}_T) + EA_t}{\sigma(\text{ROA}_T)} \quad (8)$$

where $T = 1, \dots, t$, introduced in [Lepetit and Strobel, 2013], is essentially a modified version of Z_2 in which sample mean and sample standard deviation are computed over the whole sample

horizon as opposed to the last three periods. Due to the conceptual problems we highlighted for estimators Z_1 , Z_3 and Z_4 , in the rest of this work we shall concentrate on developing enhanced versions of estimators Z_2 and Z_5 .

4 Limitations to existing approaches

When translating the theoretical premises of the Z-Score into the practical challenge of quantifying bank insolvency, a number of issues arises. As Roy [1952] posits, the estimates of the mean and standard deviation of the distribution of gross returns do have sampling distributions. Since these moments are estimated using past information, an important element to obtain consistent estimates is to ascertain the sampling distributions of the respective estimators. Another issue with the computation of the Z-Score is that it does not capture the short-term fluctuations of bank risk because the variance component is computed using information from a number of periods in the past [Delis et al., 2014]. Along the same lines, Tsionas [2016] suggests that the computation of the Z-Score may be faulty if the variance of returns is not properly estimated. This is particularly important as the original formulation of the Z-Score does not provide an approach to properly estimate the variance of returns.

The theoretical literature on the Z-Score is very thin and it focuses mainly on the distributional assumption imposed on bank's returns. Strobel [2011] provides justification for the Z-Score when bank's distribution of returns is unimodal. In this case, we can relax the assumption of finiteness of variance and obtain a substantially improved upper bound of the probability of insolvency. Lepetit and Strobel [2015] further provide evidence that imposing a distributional assumption on bank's return is not fundamental for the probabilistic foundation of the Z-Score. The authors propose a refined insolvency probability bound that is more effective than the traditional insolvency probability bound in proxying a bank's probability of insolvency. This refined measure is particularly effective in quantifying the probability of insolvency for banks with very low solvency.

As discussed in [Lepetit and Strobel, 2013], the implementation of Z-Score measures for cross-sectional analysis is largely uncontroversial; however, the construction of Z-Score measures under nonstationary returns is less straightforward. In this setting, ROA becomes a nonstationary stochastic process ρ_t with time-varying mean $\mu_t(\text{ROA})$ and standard deviation $\sigma_t(\text{ROA})$ for time period t ; EA could be modelled as a time independent safety buffer for the level of return, or it could be represented by a random variable whose parameters may also be estimated. The ambiguity on how to compute the different components of the Z-Score has favoured the development

of different approaches; for instance, Lepetit and Strobel [2013] suggest that EA,¹ $\mu_t(\text{ROA})$ and $\sigma_t(\text{ROA})$ can be computed using either the current period values or values over a set of time periods; e.g., rolling time window using three or five observations. These approaches bring surprisingly different results in terms of linear dependence of the different measures (see for instance Table 3 in [Lepetit and Strobel, 2013]).

As far as we are aware, the theoretical literature on the Z-Score does not discuss bias reduction strategies for the estimation of the mean and standard deviation of bank returns but this is paramount to consistently compute individual bank solvency. Furthermore, from a theoretical perspective, the ambiguity on how to compute the different components of the Z-Score may lead to difficulties. First, if we estimate $\mu_t(\text{ROA})$ according to Z_3 by using only the last period value and we estimate $\sigma_t(\text{ROA})$ over the whole sample horizon, it becomes unclear what the values estimated actually represent, since the stochastic process we are estimating has not been clearly defined and past observations have not been deseasonalised in line with the structure of this nonstationary process. Second, higher returns may be associated with higher variance (heteroscedasticity) and no existing methods reflect the degree of estimation error associated with available data. Third, considering different lengths in the time series of data of each individual bank and then comparing the resultant values of the Z-Score may lead to inconsistencies, as the sample size may affect accuracy and precision of our estimation for the mean and the standard deviation of returns [Delis et al., 2014].

In the following sections we will try to address each of the points highlighted in the previous paragraph: in Section 5 we discuss how to eliminate bias in estimators Z_2 and Z_5 and we introduce the unbiased estimator Z_6^k ; in Section 6 we introduce a novel estimator Z_7^k able to capture the structure of an underlying nonstationary stochastic process for the ROA; finally, in Section 7 we discuss how to compute confidence regions for Z_7^k in order to capture uncertainty of our measurement.

5 Unbiased variants of Z_2 and Z_5

Estimators Z_2 and Z_5 discussed in [Lepetit and Strobel, 2013] are essentially the same estimator, with the only difference that Z_5 employs the whole set of past realisations to estimate mean and standard deviation of ROA, while Z_2 only employs the last three realisations. For ease of exposition, we introduce a more compact notation parameterised by the length k of the time window used to

¹Or $\mu_t(\text{EA})$, if EA is modelled as a random variable

estimate mean and standard deviation of ROA:

$$Z_6^k \equiv \frac{\mu(\text{ROA}_T) + \text{EA}_t}{\sigma(\text{ROA}_T)} \quad (9)$$

where $T = t - k + 1, \dots, t$.

It is well-known in statistics that the sample standard deviation is a biased estimator of a random variable standard deviation, see Bolch [1968]. The great advantage of Z_6^k — and consequently of estimators Z_2 and Z_5 — is its simplicity and intuitive nature; unfortunately, Z_6^k is biased; it is therefore worthwhile to develop an unbiased variant of this estimator. A similar discussion in the context of inventory theory has recently been carried out by [Prak et al., 2016, p.5].

Unfortunately, there exists no estimator of the standard deviation that is unbiased and distribution independent — note that Bessel's correction does not yield an unbiased estimator of standard deviation. However, if we assume normally distributed ROA, to correct the bias we can exploit Cochran's theorem, which implies that the square of $\sqrt{n-1}s/\sigma$, where s is the sample standard deviation and σ is the actual standard deviation, has chi distribution with $n-1$ degrees of freedom. Let $\bar{\chi}(n-1)$ denote the expected value of a chi distribution with $n-1$ degrees of freedom and \bar{s} denote the expected value of the sample standard deviation; it follows that $\sigma = \bar{s}\bar{\chi}(n-1)/\sqrt{n-1}$. The unbiased variant of Z_6^k is then

$$\bar{Z}_6^k \equiv \frac{\mu(\text{ROA}_T) + \text{EA}_t}{\sigma(\text{ROA}_T)\bar{\chi}(k-1)/\sqrt{k-1}} \quad (10)$$

where $T = t - k + 1, \dots, t$. A simpler approximation can be obtained by exploiting the correction factor for the estimator of the coefficient of variation of a normally distributed random variable [Salkind, 2010],

$$\hat{Z}_6^k \equiv \frac{\mu(\text{ROA}_T) + \text{EA}_t}{(1 + 1/(4k))\sigma(\text{ROA}_T)} \quad (11)$$

where $T = t - k + 1, \dots, t$.

It is well-known that, although $\sigma(\text{ROA}_T)$ is biased, it performs better than the corrected estimator in terms of the mean squared error criterion, see e.g. [Johnson and Wichern, 2007]; however, since in \bar{Z}_6^k the estimator of the standard deviation appears in the denominator, as we will demonstrate in our experiments, this represents for us an advantage. If ROA is not normally distributed, bias can be reduced via bootstrapping [Efron, 1979] or by means of distribution dependent approximate correction factors.

Algorithm 1: Computing Z_7^k

Data: r : an array of ROA realisations; t : the period for which we aim to estimate the Z-score; EA: equity over asset at time t ; k : the time window (in periods) used for trend estimation, an odd number greater than one.

Result: z : the estimated Z-score at period t

```
1  $n := t - k - 1$ ;  $d := \{\}$ ;  $x := \{\}$ ;  $k := 1$ ;
2 for  $i \leq n + 1$  do
3   | fit a trend line  $f(y) : a + by$  with intercept  $a$  and slope  $b$  to the time series  $r_i, \dots, r_{i+k-1}$ ;
4   |  $x := x \cup \{f(i + (k-1)/2)\}$  ;
5   |  $d := d \cup \{r_{i+(k-1)/2} - f(i + (w-1)/2)\}$  ;
6 end
7  $m := \text{Mean}(x)$ ;
8  $s := \text{StandardDeviation}(d)$ ;
9  $\bar{\tau} := (1 + 1/(4(n+1)))s/m$ ;
10 if  $|\bar{\tau}f(t)| \leq \epsilon$  then
11   | standard deviation forecast very close to zero, i.e. smaller than  $\epsilon$ ;
12   |  $\bar{s} := s\bar{\chi}(n+1)/\sqrt{n}$ ;
13   |  $z := -(EA + f(t))/\bar{s}$ ;
14 else
15   |  $z := -(EA + f(t))/(\bar{\tau}f(t))$ ;
16 end
```

6 A dynamic estimator for nonstationary ROA

Despite being simple and intuitive, the key assumption underpinning Z_6^k and its unbiased variant is that there is a stationary stochastic process generating ROA realisations — or a process whose mean and standard deviation change slowly over time. In fact, these measures are essentially based on moving averages and standard deviations and may therefore fail to properly capture the structure of an underlying nonstationary stochastic process for the ROA. Setting a low value of k as in Z_2 partially addresses this problem by reducing the size of the window of past observations that are used to estimate mean and standard deviation of ROA. Unfortunately, if the underlying process features trends and it is heteroskedastic, estimates of mean and standard deviation may lag behind the actual stochastic process.

To deal with a heteroskedastic ROA, in this section we introduce a new method that operates under the assumption that the stochastic process associated with the ROA is nonstationary with unobserved time dependent mean $\mu_t(ROA)$ and unobserved constant coefficient of variation τ , where $\tau \equiv |\sigma_t(ROA)/\mu_t(ROA)|$, where $|x|$ denotes the absolute value of x . The measure we propose, which we shall name Z_7^k , is essentially an heteroskedastic extension of \bar{Z}_6^k , which can be computed as shown in Algorithm 1. Let r be a time series of ROA realisations, stored in an array; t be the time period for which we aim to compute a Z-Score; EA the equity over asset at time t ; k the size (in periods) of a rolling time window, where k is odd and greater than one. For each period $i = 1, \dots, n + 1$ consider time window $i, \dots, i + k - 1$ (line 6); fit a trend line $f(y) : a + by$ to

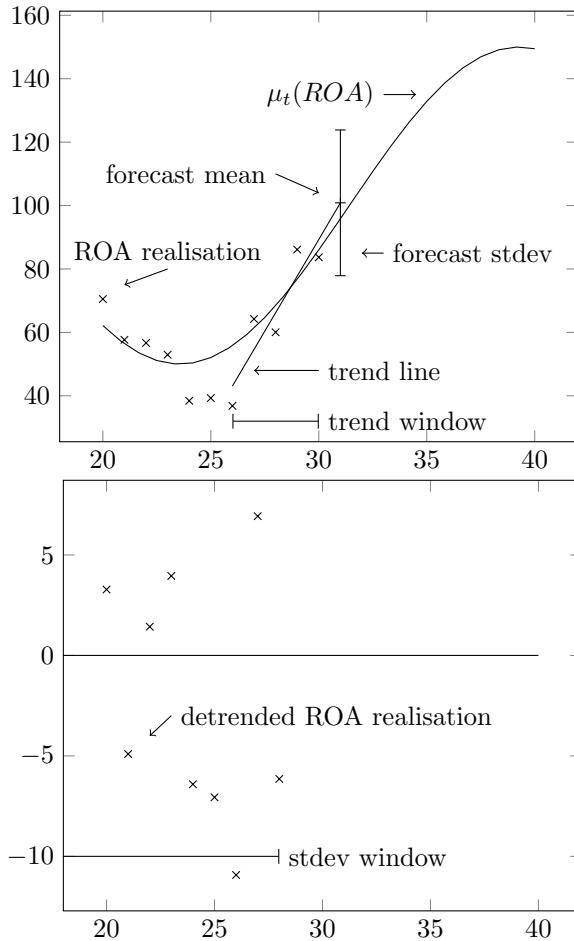


Figure 2: Dynamic estimator for nonstationary ROA

ROA observations² within this time window (line 3); and use this trend line to detrend the ROA realisation at period $i + (k - 1)/2$. Maintain a record of detrended ROA realisations (line 5) and associated estimates of mean ROA values (line 4); estimate $\bar{\tau}$ using these values (line 9). By using the trend line obtained for time window $n+1, \dots, n+k$, forecast mean ROA in period t as $f(t)$ and ROA standard deviation in period t as $f(t)\bar{\tau}$ (line 15). If, however, $|\bar{\tau}f(t)| \leq \epsilon$, where ϵ is a small number, use a more conservative homoscedastic strategy in which the Z-score is computed from the bias-adjusted sample standard deviation \bar{s} as shown in line 13. A graphical representation of the approach is shown in Figure 2.

7 Confidence regions for Z-score estimators

In this section, we discuss how to construct confidence intervals around Z_6^k and confidence regions around Z_7^k in such a way as to account for the weight of evidence at hand; these may be employed

²It should be noted that a nonlinear regression is also possible; however, it is well-known that if the process under scrutiny is Gaussian then the best predictor at an unobserved location is a linear function of the observed values.

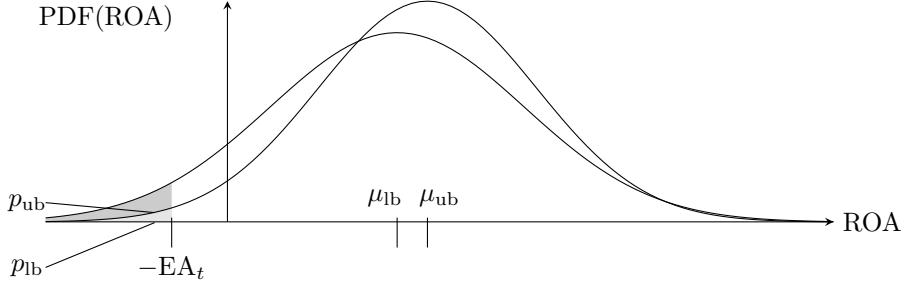


Figure 3: Graphical representation of the confidence-based Z_6^k

to carry out classical statistical analysis or, if we are interested in the newsvendor-like problem of determining capital requirements for institutions via confidence-based reasoning [Rossi et al., 2014].

We operate under normally distributed ROA; confidence regions for other distributions may not be available in closed form, for this reason one may need to resort to numerical techniques such as bootstrapping [Efron, 1979]. Recall that ROA_t is the return on asset observed at time t ; consider k periods and let the sample mean of the ROA be m and its sample standard deviation be s . We compute α confidence intervals around the ROA sample mean and sample standard deviation according to established formulae for the normal distribution parameters

$$\begin{aligned} \left(m - t_{k-1} \left(\frac{1+\alpha}{2} \right) \frac{1}{\sqrt{k}} s, m + t_{k-1} \left(\frac{1+\alpha}{2} \right) \frac{1}{\sqrt{k}} s \right) &= (\mu_{lb}, \mu_{ub}) \\ \left(\frac{(k-1)s^2}{\chi_{k-1}^2(\frac{1+\alpha}{2})}, \frac{(k-1)s^2}{\chi_{k-1}^2(\frac{1-\alpha}{2})} \right) &= (\sigma_{lb}^2, \sigma_{ub}^2) \end{aligned}$$

where $t_{k-1}(\cdot)$ is the inverse Student's t distribution with $k-1$ degrees of freedom; and $\chi_{k-1}^2(\cdot)$ is the inverse χ^2 distribution with $k-1$ degrees of freedom.

We construct confidence intervals around Z_6^k as we argue that data availability is a key element whilst drawing comparisons between point estimates. The new confidence based Z_6^k has the following lower (lb) and upper (ub) bounds:

$$Z_{6,lb}^k \equiv \frac{-\text{EA}_T - \mu_{lb}}{\sigma_{ub}} \quad (12)$$

$$Z_{6,ub}^k \equiv \frac{-\text{EA}_T - \mu_{ub}}{\sigma_{lb}} \quad (13)$$

The graphical representation is reported in Figure 3. In this figure, p_{lb} and p_{ub} are lower and upper bounds, respectively, for the default probability. That is $p_{lb} = \Phi(Z_{5,lb}^k)$ and $p_{ub} = \Phi(Z_{5,ub}^k)$;

where $\Phi(\cdot)$ is the standard normal cumulative distribution function.

Finally, we apply confidence based reasoning to the computation of Z_7^k . In Algorithm 1 consider the set d of detrended realisations obtained at line 5 and the set x of means obtained at line 4; let s be the standard deviation of d and m the mean of x .

Confidence intervals around the coefficient of variation τ can be constructed by using a variant of the approach in Edward Miller [1991], which accounts for the reduced number of degrees of freedom ($n + k - 3(n/k)$)

$$\tau_{lb} \equiv \frac{s}{m} - \Phi^{-1}\left(\frac{1+\alpha}{2}\right) \sqrt{(n+k-3(n/k))^{-1} \left(\frac{s}{m}\right)^2 \left(0.5 + \left(\frac{s}{m}\right)^2\right)}$$

$$\tau_{ub} \equiv \frac{s}{m} + \Phi^{-1}\left(\frac{1+\alpha}{2}\right) \sqrt{(n+k-3(n/k))^{-1} \left(\frac{s}{m}\right)^2 \left(0.5 + \left(\frac{s}{m}\right)^2\right)}$$

where $\Phi^{-1}(\cdot)$ is the inverse standard normal cumulative distribution function. The rationale behind the number $(n + k - 3(n/k))$ is the following: $n + k$ is the total number of ROA realisations, 3 degrees of freedom (slope, intercept, and mean used in the detrending step) are lost every time we fit a trend over a time window that does not overlap with any other time window; this happens n/k times, since there are n/k of such time windows. By definition, the coefficient of variation is a positive value, we therefore take the absolute value $|s/m|$ of s/m in the above expressions in order to prevent intervals from spanning over negative values; a more elegant but slightly more complicated solution for normally distributed ROA is discussed in [Koopmans et al., 1964].

Confidence bands around the trendline $f(y) : a + by$ can be constructed using standard approaches in linear regression, i.e.

$$f(n+k)_{lb} \equiv f(n+k) - t_{k-2}^{-1}\left(\frac{1+\alpha}{2}\right) \sqrt{\frac{1}{k} + \frac{(k/2+1)^2}{\sum_{i=n}^{n+k} (i-n-k/2)^2}} \sqrt{\frac{1}{k-2} \sum_{i=1}^{n+1} d_i^2}$$

$$f(n+k)_{ub} \equiv f(n+k) + t_{k-2}^{-1}\left(\frac{1+\alpha}{2}\right) \sqrt{\frac{1}{k} + \frac{(k/2+1)^2}{\sum_{i=n}^{n+k} (i-n-k/2)^2}} \sqrt{\frac{1}{k-2} \sum_{i=1}^{n+1} d_i^2}$$

where t_{k-2}^{-1} is the inverse t distribution with $k - 2$ degrees of freedom. $Z_{6,lb}^k$ and $Z_{6,ub}^k$ can be immediately constructed using the results just presented. In Figure 4 we construct, for a simple numerical example, confidence bands for the trendline as well as confidence intervals around the standard deviation.

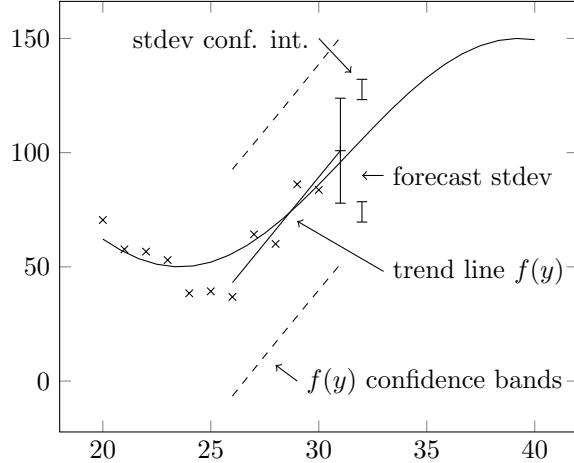


Figure 4: Confidence interval analysis

8 Numerical analysis

In this numerical study we contrast the effectiveness of Z_6^k , \bar{Z}_6^k and Z_7^k by relying on two widely used performance measures from the forecasting literature: the mean percentage error (MPE) and the mean absolute percentage error (MAPE). To do so, we employ the six stochastic processes illustrated in Table 2 and Figure 5; these processes range from stationary returns (series 1), peak returns (series 2), exponential returns (series 3), seasonal returns (series 4), a return process featuring a combination of an upward trend and a seasonal effect (series 5), and a complex process featuring a combination of a life cycle trend plus a seasonal effect (series 6); our aim is to capture a comprehensive selection of features, such as trends and seasonalities, that one typically tries to capture while performing time series analysis. We assume that ROA_t in each period $t = 1, \dots, 50$ is a normally distributed random variable with mean $\mu_t(ROA)$ and standard deviation $\tau\mu_t(ROA)$, where τ is the coefficient of variation, which in our experimental design takes values 0.1, 0.25 and 0.5. It is worth remarking that modeling the ROA as a Gaussian process is essentially equivalent to stating that forecast errors are normally distributed around a mean value $\mu_t(ROA)$ and with standard deviation $\tau\mu_t(ROA)$; the assumption that forecast errors are normally distributed around a given expected value is widely adopted in the forecasting literature, e.g. ordinary linear regression, ARIMA, etc.; we shall provide further evidence of the soundness of this assumption in the context of our study in Section 9. However, the reader should keep in mind that the Z-Score represents a value that is inversely proportional to an upper bound on the probability of insolvency even under weak distributional assumptions [Strobel, 2011, Lepetit and Strobel, 2015], and that the the Z-Score measures presented in Sections 3 and 6 are free of distributional assumptions. We leave the investigation of the effectiveness of Z_6^k , \bar{Z}_6^k and Z_7^k under alternative distributional assumptions as

future work motivated by strong improvements that we are about to discuss for the Gaussian case.

We generate 300 random realisations of each series, for a total of 5400 series. We then apply Z_6^k , \bar{Z}_6^k and Z_7^k at each period $t = 21, \dots, 50$, to estimate the Z-score. Experiments are carried out under a common random number settings; periods $1, \dots, 20$ are kept as “warm up” periods and the Z-score for these periods is not forecasted. The actual Z-score for each period can be obtained analytically from the stochastic process that generated the ROA realisations. In our study the forecasting error is given by the difference between the actual Z-score at a given period and the one estimated by a given estimator. Essentially, this means we measure forecasting errors against parameters of the underlying stochastic process that generated the data — in particular mean and standard deviation used in the computation of the Z-score. Since we do not measure forecasting errors against ROA realisations, we effectively employ mean-based error measures, or we should rather say *moment-based error measures*, a novel performance measurement framework recently introduced in Prestwich et al. [2014]. In our experiments, EA is fixed to 10; error measures used to compare different Z-score measures, namely MPE and MAPE are computed for periods $21, \dots, 50$ in the forecasting horizon over the 300 realisations considered for each time series.

| Series | Analytical expression |
|--------|--|
| 1 | $E[ROA_t] = 100$ |
| 2 | $E[ROA_t] = \begin{cases} 80 + 2.5t & \text{if } 1 \leq t \leq 25 \\ 142.5 - (t - 26) & \text{if } 26 \leq t \leq 50 \end{cases}$ |
| 3 | $E[ROA_t] = \begin{cases} 50 & \text{if } t = 1 \\ E[ROA_{t-1}] + 0.1t & \text{if } 2 \leq t \leq 50 \end{cases}$ |
| 4 | $E[ROA_t] = 100 + 50\sin(0.2t)$ |
| 5 | $E[ROA_t] = 100 + 50\sin(0.2t) + 2t$ |
| 6 | $E[ROA_t] = \begin{cases} 100 + 50\sin(0.5t) + 5t & \text{if } 1 \leq t \leq 25 \\ 100 + 50\sin(0.5t) + 125 - 5t & \text{if } 26 \leq t \leq 50 \end{cases}$ |

Table 2: Expected ROA patterns in our empirical study

In Tables 3 and 4 we report MPE, and MAPE for the various estimators derived from the literature, i.e. Z_6^k , their unbiased variants \bar{Z}_6^k , and for our novel estimator Z_7^k . We did not include a table with the root mean squared percentage error (RMSPE) since this performance measure was consistent with the previous two.

Our numerical study confirms the effectiveness of the bias reduction strategy discussed in Section 5 — especially for short time windows of $k = 3$ and $k = 5$ periods. The average MAPE reduction achieved by \bar{Z}_6^t over Z_6^k is 12% when $k = 3$, 4.8% when $k = 5$; however, it is negligible when $k = t$.

The dynamic estimator Z_7^k discussed in Section 6 overwhelmingly outperforms other measures in 12 scenarios out of 18 according to MPE and in 10 scenarios out of 18 according to MAPE. The average MAPE reduction achieved by Z_7^k over \bar{Z}_6^t is 64% when $k = 3$, 50% when $k = 5$ and, 12%

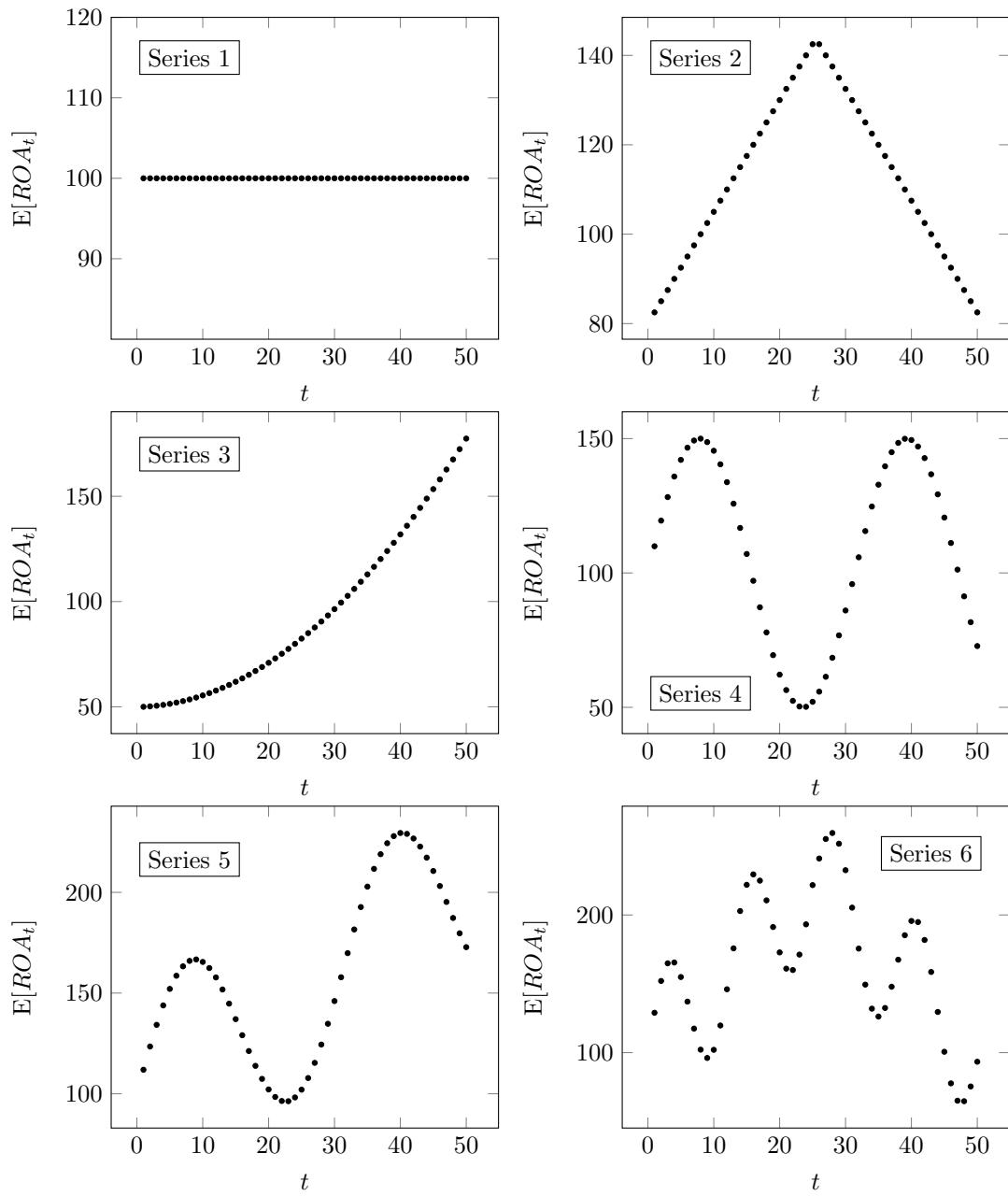


Figure 5: The six series of $E[ROA_t]$ employed in our numerical study

| τ | Series\k | Z_6^k | | | \bar{Z}_6^k | | | Z_7^k | | |
|--------|----------|---------|-------------|------------|---------------|------|-------------|---------|-------------|-------------|
| | | 3 | 5 | t | 3 | 5 | t | 3 | 5 | 7 |
| 0.1 | 1 | -83 | -26 | -2.5 | -62 | -19 | <u>-1.7</u> | -25 | -13 | -9.0 |
| | 2 | -76 | -20 | 45 | -56 | -13 | 46 | -24 | -12 | <u>-7.7</u> |
| | 3 | -73 | -13 | 59 | -53 | -6.6 | 60 | -22 | -10 | <u>-6.8</u> |
| | 4 | -37 | 16 | 71 | -21 | 21 | 71 | -20 | <u>-7.7</u> | <u>-0.6</u> |
| | 5 | -63 | <u>-2.2</u> | 58 | -44 | 3.9 | 58 | -22 | -10 | -4.1 |
| | 6 | -17 | 32 | 65 | -4.1 | 36 | 65 | -19 | <u>3.1</u> | 23 |
| 0.25 | 1 | -83 | -26 | -2.5 | -62 | -19 | <u>-1.6</u> | -27 | -14 | -9.3 |
| | 2 | -77 | -25 | 14 | -57 | -17 | 15 | -25 | -13 | <u>-8.3</u> |
| | 3 | -84 | -25 | 26 | -63 | -17 | 27 | -23 | -11 | <u>-7.0</u> |
| | 4 | -67 | -13 | 41 | -48 | -6.5 | 41 | -22 | -10 | <u>-4.7</u> |
| | 5 | -78 | -21 | 24 | -58 | -14 | 25 | -23 | -11 | <u>-6.4</u> |
| | 6 | -61 | -3.8 | 32 | -42 | 2.4 | 32 | -23 | -9.0 | <u>-0.1</u> |
| 0.5 | 1 | -83 | -27 | -2.4 | -63 | -19 | <u>-1.5</u> | -23 | -15 | -10 |
| | 2 | -79 | -26 | <u>3.5</u> | -59 | -18 | 4.2 | -23 | -13 | -9.2 |
| | 3 | -82 | -27 | <u>7.9</u> | -62 | -19 | 8.7 | -20 | -13 | -8.2 |
| | 4 | -80 | -22 | 19 | -59 | -15 | 20 | -18 | -10 | <u>-5.9</u> |
| | 5 | -80 | -25 | 8.0 | -60 | -17 | 8.7 | -21 | -12 | <u>-7.3</u> |
| | 6 | -75 | -17 | 13 | -55 | -10 | 14 | -19 | -9.2 | <u>-4.5</u> |
| | | Z_2 | Z_5 | | | | | | | |

Table 3: MPE for Z_6^k , \bar{Z}_6^k and Z_7^k ; underlined figures identify the best performing approach(es) in each row.

| τ | Series\k | Z_6^k | | | \bar{Z}_6^k | | | Z_7^k | | |
|--------|----------|---------|-------|-----------|---------------|----|-----------|---------|-----------|-----------|
| | | 3 | 5 | t | 3 | 5 | t | 3 | 5 | 7 |
| 0.1 | 1 | 101 | 42 | <u>10</u> | 88 | 39 | <u>10</u> | 27 | 18 | 15 |
| | 2 | 95 | 40 | 45 | 83 | 37 | 46 | 26 | 17 | <u>15</u> |
| | 3 | 93 | 36 | 59 | 81 | 35 | 60 | 25 | 17 | <u>15</u> |
| | 4 | 70 | 38 | 71 | 64 | 39 | 71 | 25 | 16 | <u>13</u> |
| | 5 | 87 | 35 | 58 | 77 | 35 | 58 | 26 | 17 | <u>13</u> |
| | 6 | 66 | 43 | 65 | 62 | 45 | 65 | 23 | <u>12</u> | 23 |
| 0.25 | 1 | 102 | 43 | 11 | 89 | 40 | <u>10</u> | 30 | 19 | 15 |
| | 2 | 96 | 43 | 16 | 84 | 40 | 16 | 29 | 18 | <u>15</u> |
| | 3 | 103 | 42 | 28 | 89 | 39 | 29 | 28 | 18 | <u>15</u> |
| | 4 | 89 | 37 | 41 | 77 | 35 | 41 | 28 | 18 | <u>16</u> |
| | 5 | 98 | 40 | 24 | 85 | 38 | 25 | 28 | 18 | <u>15</u> |
| | 6 | 86 | 36 | 32 | 76 | 35 | 32 | 29 | 18 | <u>15</u> |
| 0.5 | 1 | 106 | 47 | <u>12</u> | 93 | 43 | <u>12</u> | 38 | 25 | 21 |
| | 2 | 102 | 46 | <u>12</u> | 89 | 43 | <u>12</u> | 36 | 25 | 20 |
| | 3 | 105 | 47 | <u>15</u> | 92 | 43 | <u>15</u> | 35 | 24 | 19 |
| | 4 | 103 | 44 | <u>19</u> | 90 | 41 | 20 | 36 | 26 | 22 |
| | 5 | 104 | 45 | <u>13</u> | 90 | 42 | <u>13</u> | 33 | 23 | 19 |
| | 6 | 99 | 41 | <u>16</u> | 87 | 39 | <u>16</u> | 34 | 26 | 22 |
| | | Z_2 | Z_5 | | | | | | | |

Table 4: MAPE for Z_6^k , \bar{Z}_6^k and Z_7^k ; underlined figures identify the best performing approach(es) in each row.

when $k = t$. It is immediately apparent that Z_6^3 (i.e. Z_2 from Lepetit and Strobel [2013]) and Z_6^5 , including their unbiased variants, perform poorly. Z_6^t (i.e. Z_5 from Lepetit and Strobel [2013]) is instead competitive, especially in its unbiased variant \bar{Z}_6^t .

As expected, estimators Z_6^k and \bar{Z}_6^k are very effective in dealing with a stationary pattern (series

| Country | Commercial | Cooperative | Savings | Total |
|--------------------|------------|-------------|---------|--------|
| Argentina | 277 | 16 | 8 | 301 |
| Austria | 352 | 493 | 381 | 1,226 |
| Brazil | 400 | 9 | 0 | 409 |
| Canada | 55 | 4 | 0 | 59 |
| China | 226 | 4 | 1 | 231 |
| France | 556 | 190 | 118 | 864 |
| Germany | 627 | 5,634 | 3,479 | 9,740 |
| India | 358 | 24 | 0 | 382 |
| Indonesia | 276 | 0 | 0 | 276 |
| Italy | 156 | 1,080 | 94 | 1,330 |
| Japan | 994 | 3,815 | 9 | 4,818 |
| Mexico | 182 | 0 | 2 | 184 |
| Republic Of Korea | 0 | 2 | 6 | 8 |
| Russian Federation | 2,222 | 0 | 16 | 2,238 |
| Saudi Arabia | 81 | 0 | 0 | 81 |
| South Africa | 38 | 0 | 0 | 38 |
| Spain | 53 | 124 | 25 | 202 |
| Turkey | 96 | 0 | 0 | 96 |
| United Kingdom | 433 | 0 | 0 | 433 |
| USA | 29,543 | 55 | 1,030 | 30,628 |
| Total | 36,925 | 11,450 | 5,169 | 53,544 |

Table 5: Distribution of banks by country and specialization

1); however, Z_7^k is also competitive in this setting. If the underlying stochastic process features trends and low/medium variability $\tau \in \{0.1, 0.25\}$, estimators Z_6^k and \bar{Z}_6^k are generally inferior to Z_7^k across the board (ME and MAE).

As variability increases, $\tau = 0.5$, \bar{Z}_6^k begins to be more effective than Z_7^k . We believe that this behaviour ought to be expected. If variability is very high, trying to capture underlying trends becomes a difficult task. In this setting, although Z_7^k remains quite competitive, it appears that the best strategy is simply to ignore trends and revert to simpler measures such as \bar{Z}_6^k to prevent overfitting.

9 Empirical study

The goal of this empirical study is twofold: first, we aim to show that the assumption of normality of ROA, upon which the analysis carried out in the previous section is based, is supported by available data; second, we aim to show that our novel dynamic measures discussed in Section 6, Z_7^k , produce results that are substantially different from those obtained with other state-of-the-art measures from the literature, i.e. \bar{Z}_6^k . We illustrate this latter result in general, by carrying out a comprehensive analysis over a large dataset of banks, as well as by means of two case studies that focus on specific institutions that faced financial distress during the 2007–2008 financial crisis.

We retrieved data from BvD Bankscope covering the period 2005–2013 and we selected all

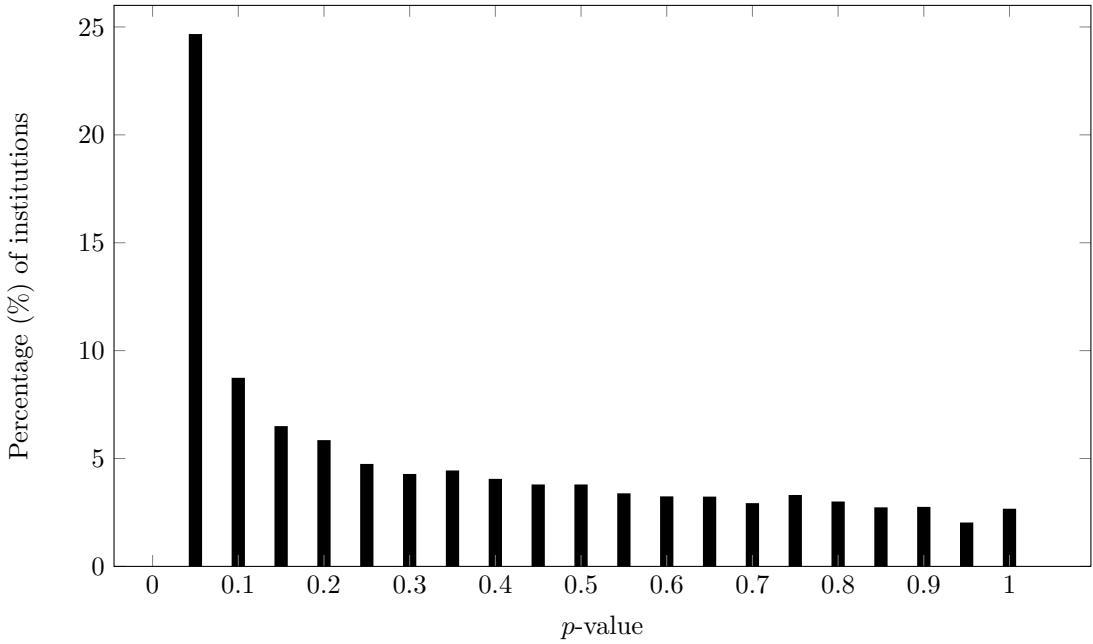


Figure 6: Frequency plot of p -values obtained via Shapiro-Wilk test of normality for the institutions in our sample; ROA observations for each institution cover the period 2005–2013.

types of depository institutions (commercial banks, savings banks and cooperative banks) that operate in G20 countries. We also excluded all institutions where data was not available for one of the following accounting items: total assets, total equity and pre-tax profit. The total number of observations per country and type of credit institution appears in Table 5.

The analysis carried out in Section 8 focused on normally distributed ROA. In order to validate this assumption, we carried out a Shapiro-Wilk test³ of normality for ROA observations we retrieved from BvD Bankscope; more precisely, we computed a p -value for each bank in the sample that features 9 observations in the period 2005–2013; our findings are illustrated in Fig. 6. The hypothesis of normality fails to be rejected for a considerable number of banks: over 75% of the banks at 5% significance, and over 85% of the banks at 1% significance in our sample. However, the fact that this hypothesis is rejected for a non-negligible proportion of banks should motivate future studies about the effectiveness of the measures here discussed under alternative distributional assumptions.

We now present a comparative study of how the different banks are classified by different measures; namely, \bar{Z}_6^k with k equal to the whole sample available, \bar{Z}_7^3 and \bar{Z}_7^5 . The aim is to contrast assessment of the overall risk of bankruptcy among different measures and show that, in a real-world scenario, these measures produce substantially different results. To obtain Figure 7 we proceeded as follows. For every given year we ranked all banks on the basis of their Z-score

³It is easy to verify via Monte-Carlo simulation that Shapiro-Wilk test provides good power for a sample size of 9 and significance values of 0.05 or 0.01.

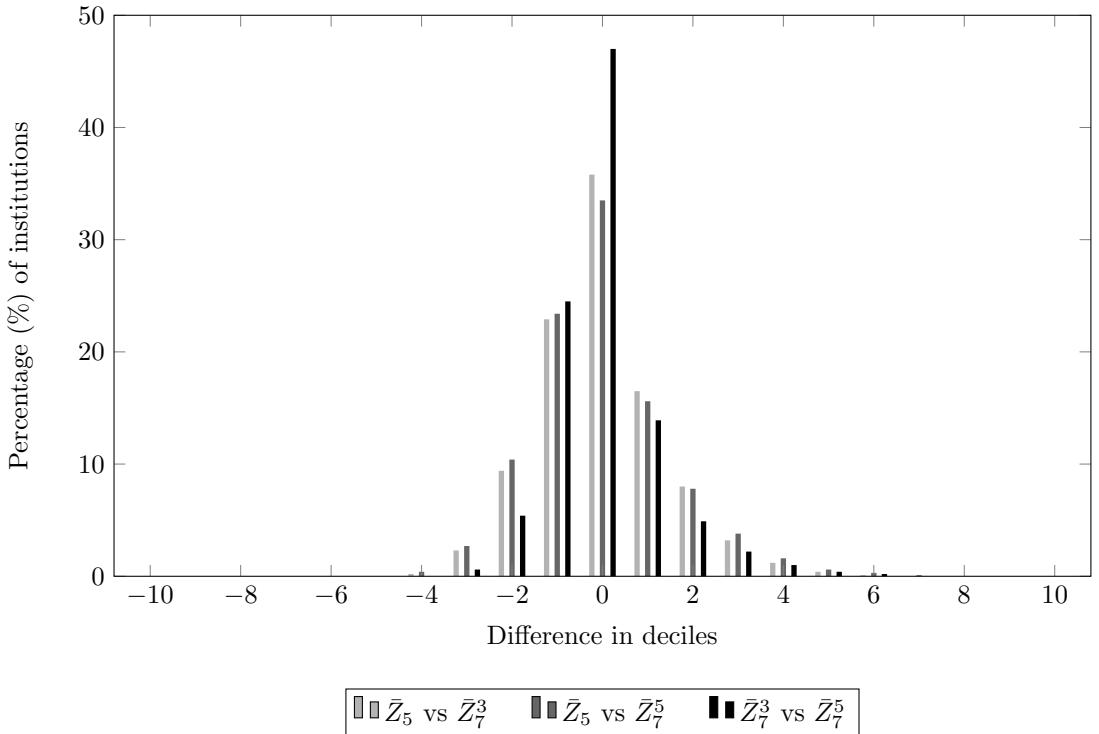


Figure 7: Comparison among the classification rankings of different methods for computing the Z-Score

and then we split this ranking distribution into deciles. We then computed the difference in deciles obtained for every given pair of measures. In 20% of the cases, all three measures lead to the same classification, i.e. banks are assigned to the same decile of the distribution. \bar{Z}_7^3 and \bar{Z}_7^5 show the highest level of correspondence (47% of the cases). \bar{Z}_6^k and \bar{Z}_7^5 agree only in slightly more than 30% of the cases, while in approximately 65% of the cases they differ by \pm one decile or more. In 30% of the cases the difference is of two deciles or more, and in 10% of the cases it is of three deciles or more; \bar{Z}_6^k and \bar{Z}_7^3 display a pattern similar to \bar{Z}_6^k and \bar{Z}_7^5 . Hence we conclude that in general \bar{Z}_6^k and \bar{Z}_7 produce different results. We further investigate this matter in the context of two case studies.

We operationalise the Z-score measures presented in Section 5 (\bar{Z}_6^k with k equal to the whole sample available) and in Section 6 (\bar{Z}_7^k with k equal to 3 and 5; therefore \bar{Z}_7^3 and \bar{Z}_7^5 , respectively) in the context of two real-world scenarios. To support this analysis, we carried out an extensive search for events that indicate financial distress in banks. We used Factiva news database to collect data on such events between 1997 and 2011. Among all events retrieved from the database only two refer to financial institutions (Commerzbank AG and Dexia CLF Banque) for which we have sufficient data in our numerical data set described in Section 9. We use these data to estimate three different Z-score measures (\bar{Z}_6^k , \bar{Z}_7^3 and \bar{Z}_7^5) until the year before the distress events observed

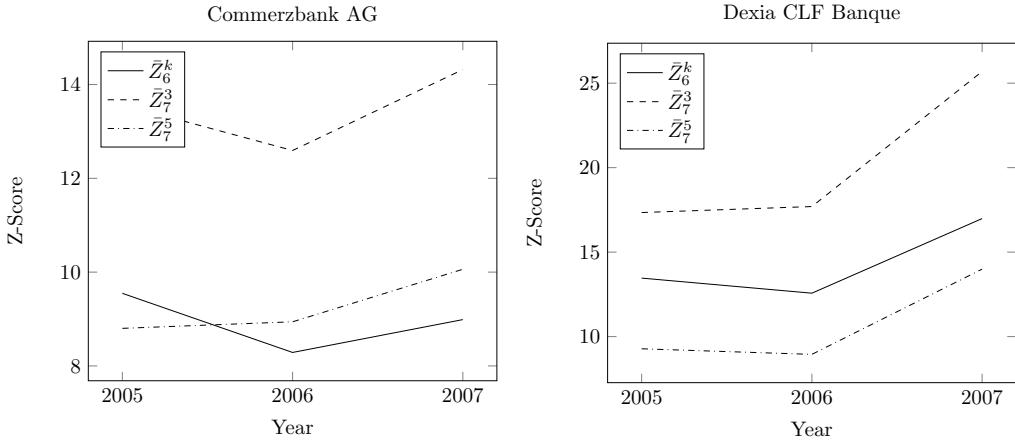


Figure 8: Performance of Z-score measures in the context of two real-world scenarios (2005-2007)

for the two aforementioned institutions (Figure 8). The level of distress in financial institutions is inversely proportional to the Z-Score value [see Lepetit and Strobel, 2013]. The reader should keep in mind that a low value of the Z-score is associated with high levels of financial distress. As shown in Figure 9, in 2007 only approximately 7.5% of the banks in the sample feature lower Z-Scores (i.e. higher levels of financial distress) than Commerzbank AG and Dexia CLF Banque. For reference, the figure also reports the evolution of the distribution of bank Z-Score in our sample between 2006 and 2008; in particular, it is apparent the increase of the level of financial distress in 2008.

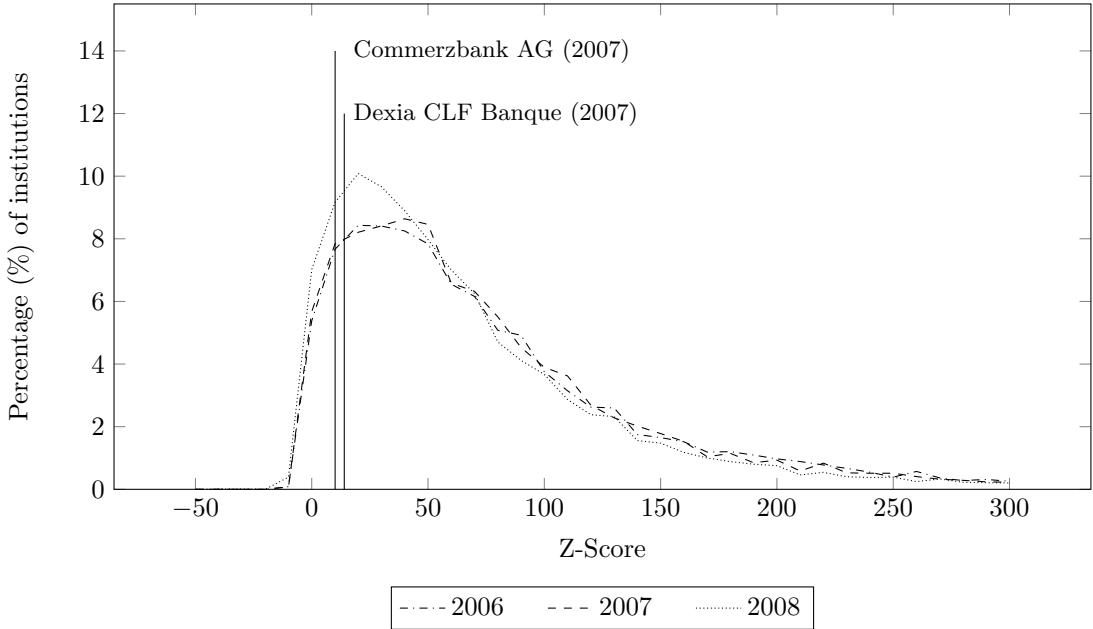


Figure 9: Relative position of Commerzbank AG and Dexia CLF Banque in 2007 with respect to all other banks in the sample. The position is computed using \bar{Z}_7^5 , but other measures do not produce a substantially different rank.

Commerzbank AG. On the 2nd of November 2008, it was announced that Commerzbank AG, the second-biggest German bank, would receive a government rescue of around 19 billion euros. According to the bank's chief executive, the need of a bailout was due to the abrupt rise in capital requirements demanded by supervisory authorities, rating agencies and the capital markets after the financial crisis. As other German banks were facing the same regulatory changes but many of them did not rely on new capital injections by the national government, we would expect that a good Z-score measure would capture the unusual capital depletion at Commerzbank. Given the inverse relationship between Z-score and distress level mentioned above, it is expected that Commerzbank Z-score should be low immediately before 2008. Figure 8 shows that \bar{Z}_6^k is the measure that predicts the highest level of financial distress, while \bar{Z}_7^3 appears to be the less accurate among the three measures. However, all three measures (\bar{Z}_6^k , \bar{Z}_7^3 and \bar{Z}_7^5) are fairly low in comparison with other banks in the sample (Figure 9). This indicates a significant level of distress between 2005 and 2007, which can be seen as an early warning signal on Commerzbank's financial situation.

Dexia CLF Banque. On the 30th of September 2008 Dexia Bank received a 6.4 billion euro bailout from France, Belgium and Luxembourg. In the weeks following the bankruptcy of the American investment bank Lehman Brothers, rumours on the weak financial situation of Dexia spread in the European market and its shares plunged by nearly 30% on the day before the bailout was announced. As in the previous example, we would expect a low Z-score reading for Dexia in the years preceding the distress. In this case, Figure 8 shows that \bar{Z}_7^5 is the measure that predicts the highest level of financial distress in 2007; once more \bar{Z}_7^3 appears to be the less accurate among the three measures. However, as in the previous case study, all three measures (\bar{Z}_6^k , \bar{Z}_7^3 and \bar{Z}_7^5) are low in comparison with other banks in the sample (Figure 9) and clearly indicate a situation of financial distress.

While studying the two financial distress events here discussed, we observed a steady increase in the the level of financial distress of institutions between 2007 and 2008. This ought to be expected, as in this period we were approaching the 2008 financial crisis. However, in Figure 10 we now show the evolution, between 2006 and 2013, of the distribution of bank Z-Score (\bar{Z}_7^5) in our sample. This figure appears to suggest that the steady increase in the the level of financial distress of institutions in our sample goes beyond the 2008 financial crisis and appears to span over the whole period 2006 - 2013. We feel that it is out of the scope of this paper to cross-validate this result by using alternative indicators and to discuss the implications of this finding; however, we believe this preliminary result deserves further investigation.

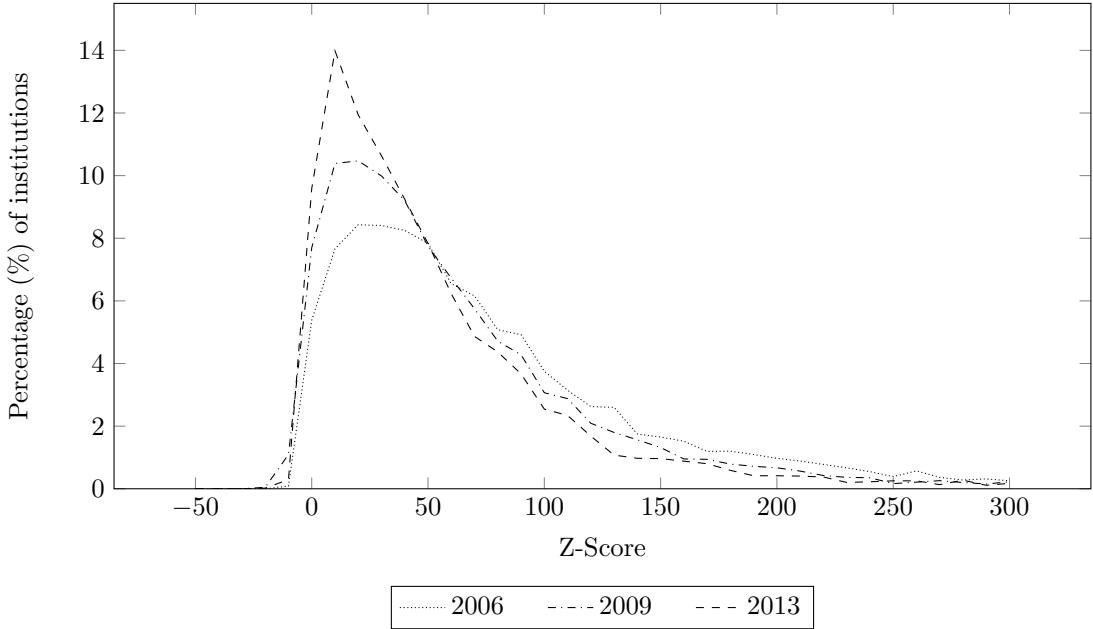


Figure 10: Distribution of bank Z-Score (\bar{Z}_7^5) in our sample; evolution between 2006 and 2013

10 Conclusions

In this work we focused on the issue of determining reliable estimates of the Z-Score, a widely used measure of financial stability. To achieve this, we extended the study of Lepetit and Strobel [2013] by introducing bias reduction strategies to improve effectiveness of their estimators. We also introduced a number of novel estimators whose aim is to effectively capture nonstationary stochastic returns; for these estimators, as well as for existing ones in the literature, we discussed analytical confidence regions. For the first time in the literature, we exploited moment-based error measures to assess the effectiveness of existing estimators from Lepetit and Strobel [2013] as well as of our novel ones; we carried out an extensive empirical study that contrasts results obtained with the aforementioned estimators on over ten thousand banks. We also contrasted results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva. Our results confirm the effectiveness of our bias reduction strategy (\bar{Z}_6^t) especially for short time windows (average MAPE reduction up to 12%). They also show that our novel dynamic estimator Z_7^k overwhelmingly outperforms (average MAPE reduction up to 64%) existing measures when return variability is medium or low, while it remains competitive with other existing approaches when return variability is high.

Acknowledgements

The authors wish to thank Tom Archibald for advice and support, and Angelos Karapappas for research assistance.

References

- B. W. Bolch. The teacher's corner: More on unbiased estimation of the standard deviation. *The American Statistician*, 22(3):27, June 1968.
- P. Bolton, H. Mehran, and J. Shapiro. Executive compensation and risk taking. *Review of Finance*, 19(6):2139–2181, Jan. 2015.
- J. Boyd, D. Nicoló, and A. G., Jalal. Bank risk-taking and competition revisited: new theory and new evidence. Technical Report WP/06/297, International Monetary Fund, Washington, D.C, 2006.
- J. H. Boyd and S. L. Graham. Risk, regulation, and bank holding company expansion into non-banking. *Quarterly Review*, 10(Spr):2–17, 1986.
- J. H. Boyd and D. E. Runkle. Size and performance of banking firms. *Journal of Monetary Economics*, 31(1):47–67, Feb. 1993.
- M. D. Delis, I. Hasan, and P. Kazakis. Bank regulations and income inequality: Empirical evidence. *Review of Finance*, 18(5):1811–1846, Aug. 2014.
- Y. Demanyuk and I. Hasan. Financial crises and bank failures: A review of prediction methods. *Omega*, 38(5):315–324, Oct. 2010.
- G. Edward Miller. Asymptotic test statistics for coefficients of variation. *Communications in Statistics - Theory and Methods*, 20(10):3351–3363, Jan. 1991.
- B. Efron. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics*, 7(1):1–26, 1979.
- F. Fiordelisi and D. S. Mare. Competition and financial stability in European cooperative banks. *Journal of International Money and Finance*, 45(C):1–16, 2014.
- F. Fiordelisi, D. Marques-Ibanez, and P. Molyneux. Efficiency and risk in European banking. *Journal of Banking & Finance*, 35(5):1315–1326, May 2011.

- Y. García-Kuhnert, M.-T. Marchica, and R. Mura. Shareholder diversification, bank Risk-Taking, and capital allocation efficiency. *SSRN Electronic Journal*, 2013.
- T. H. Hannan and G. A. Hanweck. Bank insolvency risk and the market for large certificates of deposit. *Journal of Money, Credit and Banking*, 20(2):203–211, 1988.
- H. Hesse and M. Cihák. Cooperative Banks and Financial Stability. IMF Working Papers 07/2, International Monetary Fund, 2007.
- J. F. Houston, C. Lin, P. Lin, and Y. Ma. Creditor rights, information sharing, and bank risk taking. *Journal of Financial Economics*, 96(3):485–512, June 2010.
- C. Ioannidis, F. Pasiouras, and C. Zopounidis. Assessing bank soundness with classification techniques. *Omega*, 38(5):345–357, Oct. 2010.
- R. A. Johnson and D. W. Wichern. *Applied Multivariate Statistical Analysis (6th Edition)*. Pearson, 6 edition, Apr. 2007.
- L. H. Koopmans, D. B. Owen, and J. I. Rosenblatt. Confidence intervals for the coefficient of variation for the normal and log normal distributions. *Biometrika*, 51(1/2):pp. 25–32, 1964.
- P. Kumar and V. Ravi. Bankruptcy prediction in banks and firms via statistical and intelligent techniques - a review. *European Journal of Operational Research*, 180(1):1–28, July 2007.
- L. Laeven and R. Levine. Bank governance, regulation and risk taking. *Journal of Financial Economics*, 93(2):259–275, August 2009.
- L. Lepetit and F. Strobel. Bank insolvency risk and time-varying z-score measures. *Journal of International Financial Markets, Institutions and Money*, 25:73–87, July 2013.
- L. Lepetit and F. Strobel. Bank insolvency risk and z-score measures: A refinement. *Finance Research Letters*, 13:214–224, May 2015.
- S. Mercieca, K. Schaeck, and S. Wolfe. Small european banks: Benefits from diversification? *Journal of Banking & Finance*, 31(7):1975–1998, July 2007.
- D. Prak, R. Teunter, and A. Syntetos. On the calculation of safety stocks when demand is forecasted. *European Journal of Operational Research*, June 2016.
- S. Prestwich, R. Rossi, S. Armagan Tarim, and B. Hnich. Mean-based error measures for intermittent demand forecasting. *International Journal of Production Research*, 52(22):6782–6791, May 2014.

- R. Rossi, S. Prestwich, S. A. Tarim, and B. Hnich. Confidence-based optimisation for the newsven-dor problem under binomial, poisson and exponential demand. *European Journal of Operational Research*, 239(3):674–684, Dec. 2014.
- A. D. Roy. Safety first and the holding of assets. *Econometrica*, 20(3):431–449, 1952.
- N. Salkind. *Encyclopedia of Research Design*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States, 2010.
- F. Strobel. Bank insolvency risk and z-score measures with unimodal returns. *Applied Economics Letters*, 18(17):1683–1685, Nov. 2011.
- M. G. Tsionas. Parameters measuring bank risk and their estimation. *European Journal of Oper-ational Research*, 250(1):291–304, Apr. 2016.
- E. L. Yeyati and A. Micco. Concentration and foreign penetration in Latin American banking sectors: Impact on competition and risk. *Journal of Banking & Finance*, 31(6):1633–1647, June 2007.