Nonstationary Z-score measures

Mare, Davide Salvatore and Moreira, Fernando and Rossi, Roberto

Oce of the Chief Economist, Europe and Central Asia Region, World Bank Group, USA, Business School, Credit Research Centre, The University of Edinburgh, UK, Business School, The University of Edinburgh, UK

11 November 2015

Nonstationary Z-score measures

Davide Salvatore Mare*[1,2], Fernando Moreira[2], and Roberto Rossi†[3]

[1]Office of the Chief Economist, Europe and Central Asia Region, World Bank Group, USA
[2]Business School, Credit Research Centre, The University of Edinburgh, UK
[3]Business School, The University of Edinburgh, UK

November 11, 2015

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 the degree of 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 while estimating the overall risk of bank insolvency.

JEL-Classification: C20, C60, G21

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

*The findings, interpretations, and conclusions expressed in this work are entirely those of the authors. They do not represent the views of the World Bank Group.

†Corresponding author: University of Edinburgh Business School, 29 Buccleuch Place, EH89JU, 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 credit institutions or of the overall sector. Banking financial stability 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], 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. The Z-Score is one of the most broadly used accounting measures in the literature for estimating the overall bank solvency because it combines together information on performance (return on assets indicator), leverage (equity to assets indicator) and risk (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. In the banking literature we do not have a unique indication on how to construct the Z-Score. A recent paper by Lepetit and Strobel [2013] surveys the literature to show different approaches to estimate the Z-Score. It also proposes an adjustment that brings the best results in terms of the root mean squared error (RMSE) criterion compared to existing approaches.

The goal of our work is to obtain reliable estimates of this measure of financial stability. To achieve this, we extend the study of Lepetit and Strobel [2013] and we contribute to the literature on Z-score estimators as follows:

- We introduce bias reduction strategies to improve effectiveness of estimators in Lepetit and Strobel [2013].
- We introduce novel estimators whose aim is to effectively capture nonstationary stochastic returns; for these estimators, 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 credit 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] as well as of our novel ones.
- We carry out an extensive empirical study that contrast results obtained with the aforementioned estimators 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.

The remainder of the paper is structured as follows. Section 2 presents a brief overview of the relevant literature. Section 3 defines the Z-Score whilst Section 4 summarises the current methods for computing the Z-Score and 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. Section 10 illustrates two case studies that contrast results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva. Section 11 concludes.

2 Measures of bank stability

A large literature employs accounting information, market information or a combination of the two sources to compute bank stability measures. The Z-Score, discussed in next section, is by far the most widely used accounting ratio in the literature Boyd et al. [2006], Mercieca et al. [2007], Laeven and Levine [2009], Fiordelisi and Mare [2014], Bolton et al. [2015]. Other accounting ratios, such as non-performing loans to total loans or the level of provisioning Houston et al. [2010], Fiordelisi et al. [2011], are also used to capture different risk dimensions although the focus is on specific narrower aspects such as credit risk, operational risk, liquidity risk and market risk. Market-based measures include observable market prices (equity prices, debt prices, credit default swap spreads, bond yield spreads) that allow to capture forward-looking information by incorporating market expectations. Moreover, structural models are populated using both accounting and market information assuming a theoretical framework for formalising the occurrence of bank default.

Equity Beta and Equity Return volatility Stever [2007], Fahlenbrach et al. [2012] are market-based measures computed using equity prices. Beta captures systematic risk or the risk of an investment that cannot be diversified. It is estimated using asset pricing models, such as the capital asset pricing model Sharpe [1964], Lintner [1969] or the three-factor model of Fama and French [1993]. The higher the value of this indicator the higher the instability of a bank. Equity return volatility captures both systematic and non-systematic risk and it is computed as the standard deviation of bank's equity returns over a certain time period. Higher values indicate higher fluctuation in stock returns hence higher risk.

Insolvency risk may be estimated implicitly looking at market prices and specifically bond yield spreads and credit default swaps (CDS). These prices are used to estimate the risk-neutral probabilities that include systematic risk. Both measures draw on the risk-free rate to compute the present value of the expected losses. The bond yield spread is calculated as the difference between the interest rate on a certain bank bond and the risk-free rate. The price of the risky bond should reflect the expected risk of the security and it captures the expected loss to investors. In a CDS the seller of the contract accepts to compensate the buyer of the CDS in case the bank defaults on the bond. In return for assuming the credit risk on the bank, the seller receives a stipulated premium from the buyer of the CDS.

Two standard measures of firm level risk are Value-at-Risk (VaR) and Expected-Shortfall (ES). These seek to measure the potential loss incurred by the firm as a whole in an extreme event at a prefixed confidence level. VaR expresses the maximum loss given a certain level of confidence over a certain time horizon. The expected shortfall, see e.g. Acharya et al. [2010], Ellul and Yerramilli [2013], is defined as the expected loss conditional on returns being less than some α-percentile. Specifically, ES is the average negative return over the $\alpha\%$ worst return days for a bank's stock price.

Structural models provide a framework for estimating default probabilities. The most widely used is the Merton's model [Merton, 1973] to estimate a distance-to-default indicator (DD). The bank is in technical default (insolvent) when the market value of its assets is below the market value of its liabilities. Since the DD is a market-based measure of distress, it contains expectations of market participants and it is forward-looking, making it a popular measure in bank insolvency prediction.

3 The Z-Score

The Z-Score is employed by researchers and by multilateral organisations — for instance, it is included among the indicators of The Global Financial Development Database (World Bank) — to

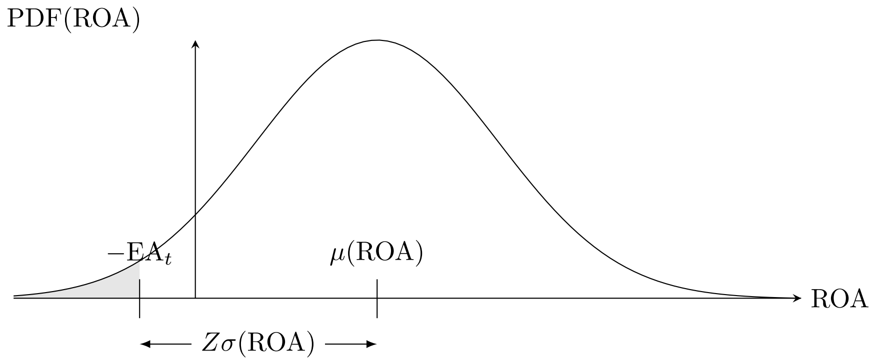

Figure 1: This figure summarizes the information provided by the Z-Score. When the only random variable is ROA, we seek for the point in the distribution where capital is depleted. We are then able to infer the probability of bankruptcy. Higher values of the Z-Score denote higher stability because the distance to default is larger.

estimate the overall degree of solvency in banking. It is built using the return on asset ratio (ROA) augmented by the equity-to-asset ratio (EA) all divided by a measure of variability in returns — often, the standard deviation of ROA. The rationale is that the lower the capital base the higher is the likelihood of bankruptcy. Moreover, higher variability in returns also increases the probability of bankruptcy. In a cross-sectional setting, the standard approach to estimate the Z-Score for an individual bank (or a set of banks) is as follows:

$$Z = \frac{-\mathrm{EA} - \mu(\mathrm{ROA})}{\sigma(\mathrm{ROA})} \tag{1}$$

where μ denotes the expected value and σ denotes the standard deviation of the ROA. In the literature, it is customary to revert sign so to obtain a positive Z-Score, i.e. $(\mu(\mathrm{ROA}) + \mathrm{EA})/\sigma(\mathrm{ROA})$. In the rest of this work we will adopt this convention.

The Z-Score combines balance sheet and income statement information to provide a measure of bank soundness. It indicates the number of standard deviations by which returns have to diminish in order to deplete the equity of a bank or a banking system. A higher Z-Score implies a higher degree of solvency and therefore it gives a direct measure of bank stability. Figure 1 illustrates the rationale for using the Z-Score as a measure of the overall bank stability.

An important element of Equation 1 is establishing whether we have one or more random variables involved in the formula. The common approach in the literature is to consider EA deterministic hence ROA is the only random variable Boyd and Graham [1986]. This is not a necessary assumption if we consider EA to be just the upper bound of the probability of insolvency.

When we introduce the time dimension and work in a cross-sectional stochastic setting, there are few options that can be exploited. Lepetit and Strobel [2013] suggest that ROA, EA and

σ(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. Nevertheless, the ambiguity on how to compute the different components of the Z-Score has favoured the development of different approaches that bring surprisingly different results in terms of linear dependence of the different measures; see for instance Table 3 in Lepetit and Strobel [2013]. These have led to a number of inconsistencies. First, if we consider the last period value for ROA and we compute σ(ROA) over the whole sample horizon, it is not clear what random variable is being estimated. 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 measure of variability could be affected, along with the value of the mean. Finally, since annual data is often employed for the computation of the different elementary components of the Z-Score, the information set available varies at different points in time and this is likely to affect the accuracy of the computations.

4 Existing approaches to compute the Z-Score

Consider the elementary information in Table 1 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. The literature proposes five main approaches to compute the Z-Score for a given bank Lepetit and Strobel [2013]. These are reported below as $Z_1,\dots,Z_5$.

We shall consider first

$$Z_1 \equiv \frac{\mu(\text{ROA}_T) + \mu(\text{EA}_T)}{\sigma(\text{ROA}_T)} \tag{2}$$

where $T = t-2,\dots,t$. According to this measure realisations EA_t also come from a random variable whose sample mean is estimated from past observations. We can imagine that in certain settings equity over asset at time t should be modeled as a random variable. However, if the model comprises this random variable, then the measure should also take into account the variability associated with it.

The second measure is a rather straightforward implementation of the classic Z-Score discussed in Section 3.

$$Z_2 \equiv \frac{\mu(\mathrm{ROA}_T) + \mathrm{EA}_t}{\sigma(\mathrm{ROA}_T)} \tag{3}$$

where $T = t-2, \ldots, t$. 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, the reliability of this measure is low because of statistical bias in the estimation of the standard deviation of ROA.

As mentioned in Section 3, in the third measure

$$Z_3 \equiv \frac{\mathrm{ROA}_t + \mathrm{EA}_t}{\sigma(\mathrm{ROA}_T)} \tag{4}$$

where $T = 1, \ldots, t$, the model considers only the last period value for ROA, while it computes $\sigma(\mathrm{ROA})$ 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.

To compute the fourth approach to calculate the Z-Score, we need to define the instantaneous standard deviation as

$$\sigma_{\mathrm{inst}}(\mathrm{ROA}_T) \equiv |\mathrm{ROA}_t - \mu(\mathrm{ROA}_T)| \tag{5}$$

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

$$Z_4 \equiv \frac{\mathrm{ROA}_t + \mathrm{EA}_t}{\sigma_{\mathrm{inst}}(\mathrm{ROA}_T)} \tag{6}$$

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

Finally,

$$Z_5 \equiv \frac{\mu(\mathrm{ROA}_T) + \mathrm{EA}_t}{\sigma(\mathrm{ROA}_T)} \tag{7}$$

where $T = 1, \ldots, t$, 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 associated with 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.

5 Eliminating bias in estimators Z_2 and Z_5

Estimators Z_2 and Z_5 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 generalise these estimators by introducing estimator

$$Z_6^k \equiv \frac{\mu(\mathrm{ROA}_T) + \mathrm{EA}_t}{\sigma(\mathrm{ROA}_T)} \tag{8}$$

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

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. It is therefore worthwhile to develop an unbiased variant of this estimator.

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 $\bar{\chi}(k-1)$ 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 $\sigma = \bar{s}\bar{\chi}(k-1)/\sqrt{k-1}$. The unbiased variant of Z_6^k is then

$$\bar{Z}_6^k \equiv \frac{\mu(\mathrm{ROA}_T) + \mathrm{EA}_t}{\sigma(\mathrm{ROA}_T)\bar{\chi}(k-1)/\sqrt{k-1}} \tag{9}$$

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

$$\hat{Z}_6^k \equiv \frac{\mu(\mathrm{ROA}_T) + \mathrm{EA}_t}{(1 + 1/(4k))\sigma(\mathrm{ROA}_T)} \tag{10}$$

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

It is well-known that, although $\sigma(\mathrm{ROA}_T)$ 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 or by means of distribution dependent approximate correction factors.

6 A dynamic estimator for nonstationary ROA

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

$n := t - k - 1;\ \ d := \{\};\ \ x := \{\};\ \ k := 1;$
for $i \leq n + 1$ **do**
  fit a trend line $f(y) : a + by$ with intercept a and slope b to the time series $r_i, \ldots, r_{i+k-1}$;
  $x := x \bigcup \{f(i + (k-1)/2)\}$;
  $d := d \bigcup \{r_{i+(k-1)/2} - f(i + (w-1)/2)\}$;
end
$m :=$ Mean(x);
$s :=$ StandardDeviation(d);
$\bar{\tau} := (1 + 1/(4(n+1)))s/m$;
if $|\bar{\tau} f(t)| \leq \epsilon$ **then**
  standard deviation forecast very close to zero, i.e. smaller than ϵ;
  $\bar{s} := s\bar{\chi}(n+1)/\sqrt{n}$;
  $z := -(EA + f(t))/\bar{s}$;
else
  $z := -(EA + f(t))/(\bar{\tau} f(t))$;
end

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 potentially 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

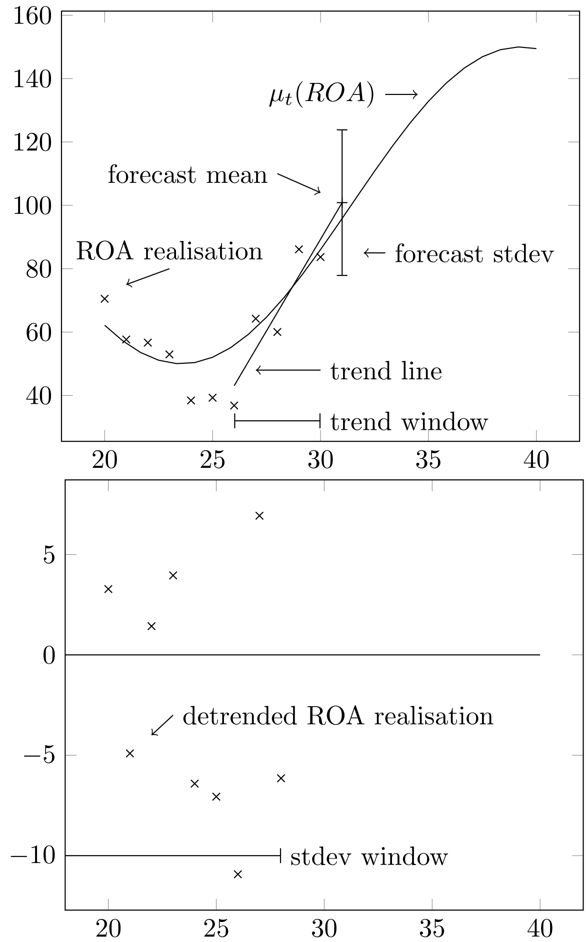

Figure 2: Dynamic estimator for nonstationary ROA

unobserved constant coefficient of variation τ, where $\tau = \sigma_t(ROA)/\mu_t(ROA)$. 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 $i = 1, \ldots, n+1$ consider time window $i, \ldots, i+k-1$ (line 6); fit a trend line $f(y) : a+by$ — note that a nonlinear regression is also possible — to ROA observations 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 $n+1, \ldots, n+k$, 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.

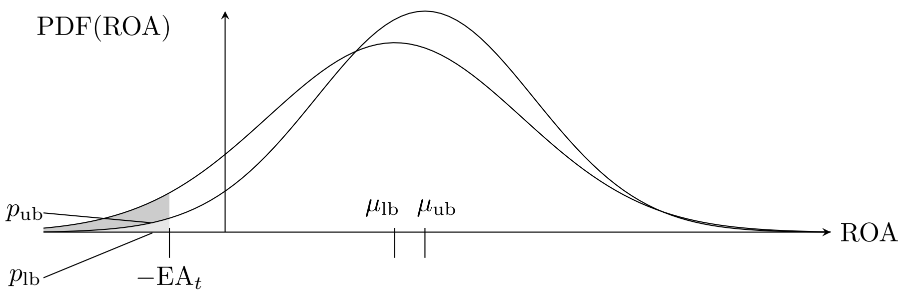

Figure 3: Graphical representation of the confidence-based Z_6^k

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

We shall assume without loss of generality normally distributed ROA. 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

$$\left(m - t_{k-1}\left(\frac{1+\alpha}{2}\right)\frac{1}{\sqrt{k}}s, m + t_{k-1}\left(\frac{1+\alpha}{2}\right)\frac{1}{\sqrt{k}}s\right) = (\mu_{\mathrm{lb}}, \mu_{\mathrm{ub}})$$

$$\left(\frac{(k-1)s^2}{\chi^2_{k-1}(\frac{1+\alpha}{2})}, \frac{(k-1)s^2}{\chi^2_{k-1}(\frac{1-\alpha}{2})}\right) = (\sigma^2_{\mathrm{lb}}, \sigma^2_{\mathrm{ub}})$$

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

$$Z^k_{6,\mathrm{lb}} \equiv \frac{-\mathrm{EA}_T - \mu_{\mathrm{lb}}}{\sigma_{\mathrm{ub}}} \tag{11}$$

$$Z^k_{6,\mathrm{ub}} \equiv \frac{-\mathrm{EA}_T - \mu_{\mathrm{ub}}}{\sigma_{\mathrm{lb}}} \tag{12}$$

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 $p_{\text{lb}} = \Phi(Z5_{\text{lb}})$ and $p_{\text{ub}} = \Phi(Z5_{\text{ub}})$; 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 $(n + k - 3(n/k))$

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

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

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. It should be noted that the coefficient of variation is, by definition, a positive value. We can address this issue by taking the absolute value $|s/m|$ of s/m in the above expressions and by not allowing intervals spanning over negative values. However, a more elegant but slightly more complicated solution for normally distributed ROA can be obtained by adopting the approach 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.

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

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

where t_{k-2}^{-1} is the inverse t distribution with $k - 2$ degrees of freedom. $Z_{6,\text{lb}}^k$ and $Z_{6,\text{ub}}^k$ 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.

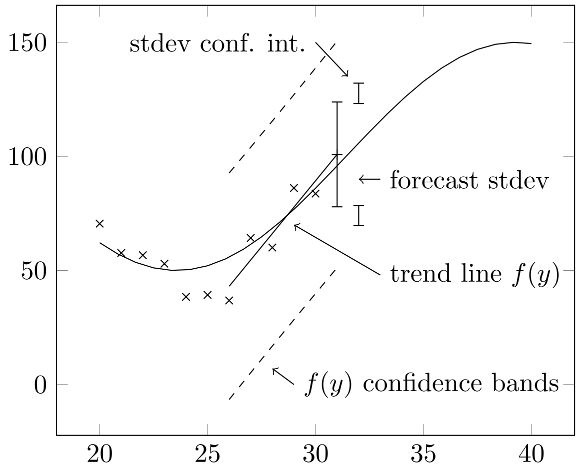

Figure 4: Confidence interval analysis

8 Numerical analysis

In this numerical study we test the effectiveness of Z_6^k, $\bar{Z}_6^k$ and Z_7^k. To do so, we employ the six stochastic processes illustrated in Table 2 and Figure 5. We assume that ROA_t in each period $t = 1, \ldots, 50$ 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. 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 $t = 21, \ldots, 50$, to estimate the Z-score. Experiments are carried out under a common random number settings; periods $1, \ldots, 20$ 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. In practice, 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*, as introduced in Prestwich et al. [2014]. In our experiments, EA is fixed to 10; error measures used to compare different Z-score measures, namely mean error (ME), mean absolute error (MAE) and root mean squared error (RMSE), are computed for periods $21, \ldots, 50$ in the forecasting horizon over the 300 realisations considered for each time series.

Our numerical study confirms the effectiveness of the bias reduction strategy discussed in Section 5 and of the dynamic estimator Z_7^k discussed in Section 6. In Tables 3, 4 and 5 we report

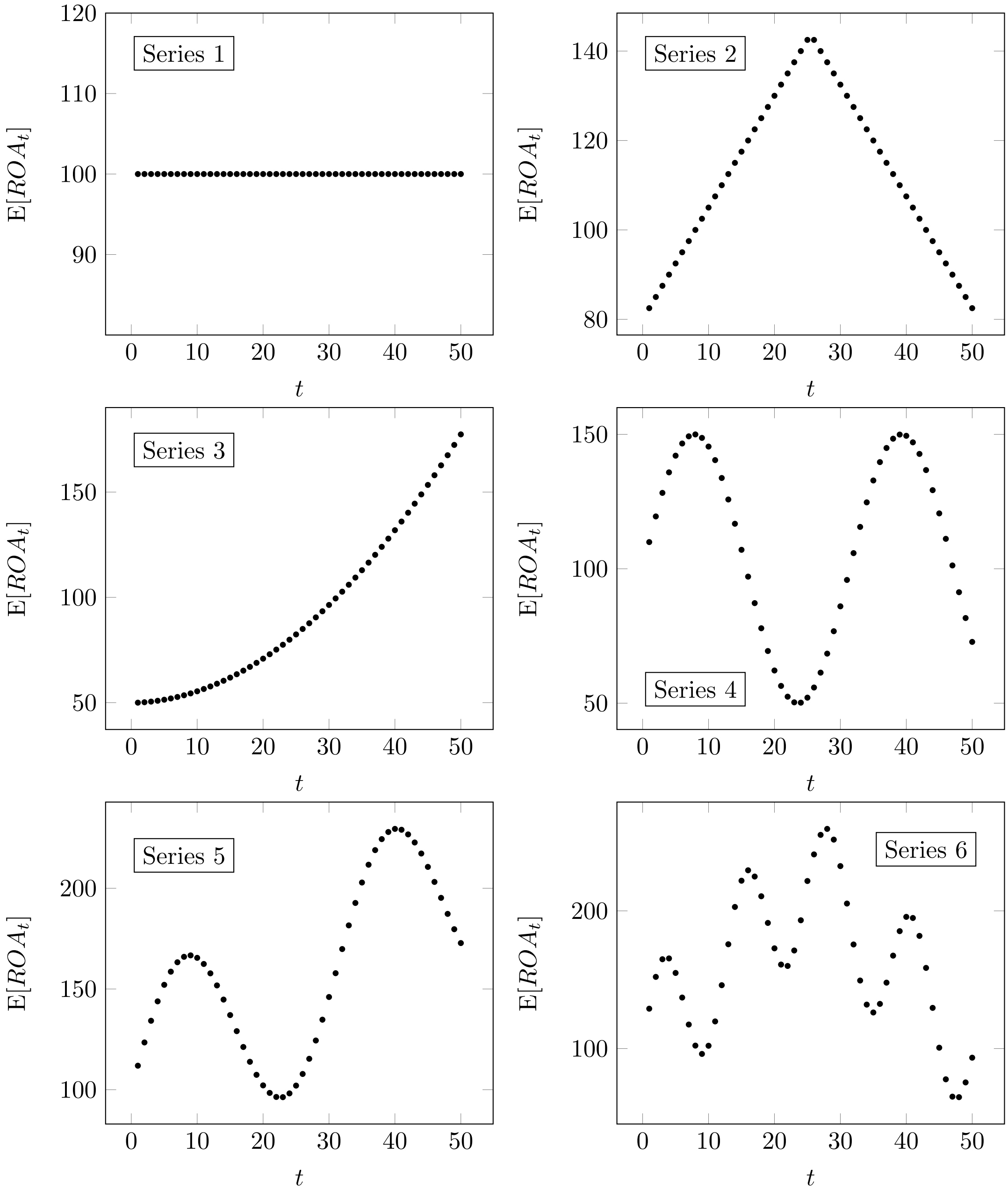

Figure 5: The six series of E[ROA_t] employed in our numerical study

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

Table 2: Expected ROA patterns in our empirical study

		Z_6^k			$\bar{Z}_6^k$			Z_7^k		
τ	Series\\k	3	5	t	3	5	t	3	5	7
0.1	1	9.13	2.91	0.28	6.84	2.08	0.19	2.73	1.46	0.99
	2	8.26	2.16	-4.92	6.08	1.38	-4.97	2.58	1.31	0.85
	3	7.91	1.47	-6.46	5.77	0.73	-6.49	2.39	1.15	0.75
	4	4.10	-1.85	-7.93	2.36	-2.41	-7.96	2.24	0.88	0.09
	5	6.70	0.23	-6.13	4.73	-0.42	-6.17	2.38	1.09	0.45
	6	1.78	-3.44	-6.94	0.36	-3.87	-6.97	2.08	-0.32	-2.44
0.25	1	3.66	1.17	0.11	2.74	0.83	0.07	1.17	0.62	0.41
	2	3.34	1.09	-0.61	2.46	0.76	-0.64	1.11	0.56	0.36
	3	3.64	1.09	-1.14	2.73	0.76	-1.17	1.02	0.48	0.31
	4	2.99	0.59	-1.82	2.14	0.28	-1.84	0.97	0.44	0.22
	5	3.32	0.89	-1.02	2.46	0.58	-1.05	1.00	0.48	0.28
	6	2.58	0.14	-1.37	1.80	-0.12	-1.39	0.97	0.39	0.01
0.5	1	3.66	1.17	0.11	2.74	0.83	0.07	1.17	0.62	0.41
	2	3.34	1.09	-0.61	2.46	0.76	-0.64	1.11	0.56	0.36
	3	3.64	1.09	-1.14	2.73	0.76	-1.17	1.02	0.48	0.31
	4	2.99	0.59	-1.82	2.14	0.28	-1.84	0.97	0.44	0.22
	5	3.32	0.89	-1.02	2.46	0.58	-1.05	1.00	0.48	0.28
	6	2.58	0.14	-1.37	1.80	-0.12	-1.39	0.97	0.39	0.01
		Z_2		Z_5						

Table 3: ME for Z_6^k, $\bar{Z}_6^k$ and Z_7^k

ME, MAE and RMSE 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. 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 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, MAE and RMSE).

As variability increases, $\tau = 0.5$, $\bar{Z}_6^k$ begins to be more effective than Z_7^k. We believe that this behavior 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

		Z_6^k			$\bar{Z}_6^k$			Z_7^k		
τ	Series\k	3	5	t	3	5	t	3	5	7
0.1	1	11.12	4.66	1.12	9.64	4.30	1.11	2.97	1.96	1.64
	2	10.39	4.31	4.92	9.04	4.05	4.97	2.88	1.88	1.59
	3	10.13	3.97	6.46	8.83	3.79	6.49	2.78	1.86	1.62
	4	7.75	4.25	7.93	7.06	4.37	7.96	2.75	1.82	1.51
	5	9.22	3.73	6.13	8.13	3.69	6.17	2.72	1.77	1.42
	6	7.01	4.68	6.94	6.64	4.89	6.97	2.50	1.24	2.47
0.25	1	4.50	1.91	0.46	3.91	1.76	0.46	1.32	0.82	0.68
	2	4.19	1.86	0.69	3.64	1.72	0.71	1.28	0.79	0.66
	3	4.48	1.85	1.22	3.89	1.72	1.24	1.21	0.77	0.67
	4	3.94	1.65	1.82	3.44	1.58	1.84	1.24	0.81	0.70
	5	4.17	1.71	1.03	3.63	1.60	1.06	1.17	0.75	0.63
	6	3.66	1.53	1.37	3.23	1.51	1.39	1.24	0.78	0.63
0.5	1	2.34	1.02	0.26	2.04	0.95	0.26	0.83	0.56	0.45
	2	2.23	1.01	0.26	1.94	0.93	0.26	0.79	0.54	0.44
	3	2.30	1.01	0.32	2.00	0.94	0.32	0.78	0.52	0.42
	4	2.30	0.97	0.43	2.00	0.91	0.45	0.82	0.58	0.49
	5	2.20	0.97	0.27	1.92	0.90	0.27	0.71	0.49	0.40
	6	2.13	0.89	0.34	1.86	0.84	0.35	0.73	0.56	0.48
		Z_2		Z_5						

Table 4: MAE for Z_6^k, $\bar{Z}_6^k$ and Z_7^k

		Z_6^k			$\bar{Z}_6^k$			Z_7^k		
τ	Series\k	3	5	t	3	5	t	3	5	7
0.1	1	45.06	8.35	1.48	39.70	7.65	1.45	3.85	2.61	2.23
	2	30.14	8.69	4.95	26.40	8.03	5.00	3.76	2.52	2.16
	3	40.93	6.50	6.60	36.06	6.00	6.63	3.70	2.52	2.21
	4	18.77	5.76	7.95	16.40	5.67	7.97	3.74	2.48	2.02
	5	48.13	5.73	6.17	42.51	5.40	6.20	3.63	2.39	1.92
	6	20.78	5.57	6.96	18.35	5.66	6.99	3.30	1.56	2.64
0.25	1	18.47	3.40	0.61	16.28	3.11	0.60	1.75	1.10	0.93
	2	10.60	3.39	0.78	9.25	3.11	0.80	1.71	1.07	0.90
	3	24.70	3.53	1.34	21.82	3.25	1.36	1.65	1.05	0.92
	4	10.58	2.80	1.84	9.25	2.59	1.86	1.73	1.13	0.99
	5	12.73	3.08	1.11	11.17	2.84	1.13	1.59	1.03	0.87
	6	14.26	2.41	1.42	12.56	2.27	1.44	1.70	1.15	0.95
0.5	1	9.69	1.79	0.34	8.55	1.65	0.34	1.27	0.92	0.77
	2	6.94	1.76	0.33	6.09	1.61	0.33	1.22	0.90	0.71
	3	8.72	1.77	0.39	7.68	1.62	0.39	1.22	0.87	0.68
	4	6.85	1.72	0.48	6.01	1.59	0.50	1.33	1.01	0.84
	5	6.56	1.65	0.33	5.76	1.51	0.33	1.10	0.80	0.64
	6	8.61	1.43	0.39	7.59	1.32	0.40	1.18	0.98	0.87
		Z_2		Z_5						

Table 5: RMSE for Z_6^k, $\bar{Z}_6^k$ and Z_7^k

the best strategy is simply to ignore trends and revert to simpler measures such as $\bar{Z}_6^k$ to prevent overfitting.

Finally, an interesting result is the fact that estimators Z_6^k and $\bar{Z}_6^k$ present a considerable bias when the underlying stochastic process features trends and low variability; this can be observed by contrasting ME and MAE. Z_7^k, in contrast, features a much lower bias.

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 6: Distribution of banks by country and specialization

9 Empirical study

In this section we present a comparative study of how the different banks are classified in the deciles of the distribution of $\bar{Z}_6^k$ 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. We retrieve data from BvD Bankscope covering the period 2005–2013. We select all types of depository institutions (commercial banks, savings banks and cooperative banks) that operate in G20 countries. We also exclude 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 6. In Figure 6, we pool together the data for all banks and all years and classify each observation in the decile of the distribution of each measure. We then relate each numerical value corresponding to the decile of the distribution by comparing the assessment for each measure and compute the difference in the deciles. In 20% of the cases, all three methods 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). Z_5 and 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. Hence we conclude that in general these measures produce different results. We further investigate this matter in the context of two case studies.

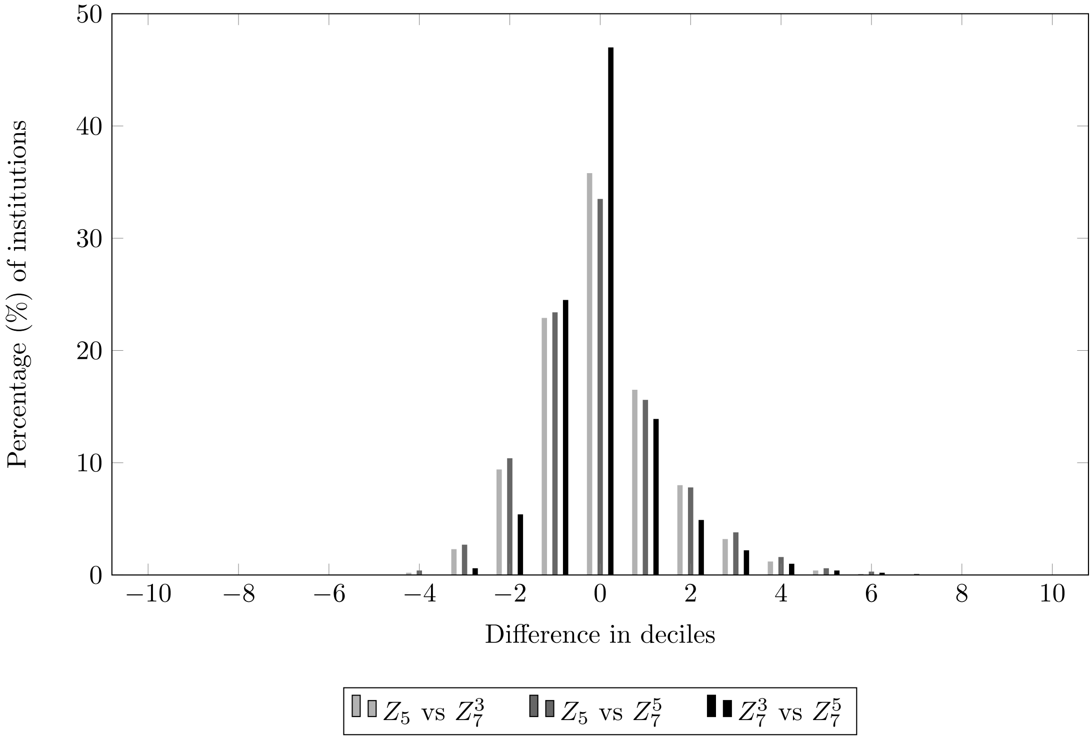

Figure 6: Comparison among the classification rankings of different methods for computing the Z-Score

10 Case studies

In this section 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. 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 for the two aforementioned institutions (Figure 7). 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 8, 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.

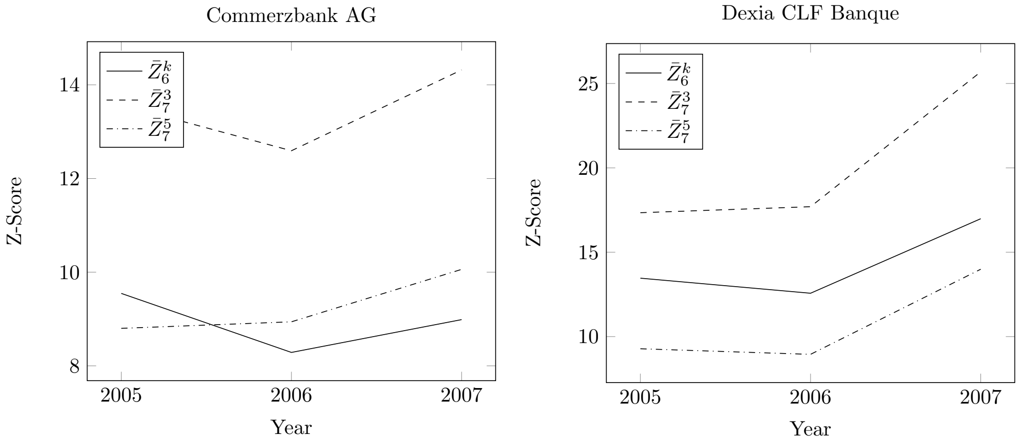

Figure 7: Performance of Z-score measures in the context of two real-world scenarios (2005-2007)

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 7 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 8). 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 preceeding the distress. In this case, Figure 7 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 8) and clearly indicate a situation of financial distress.

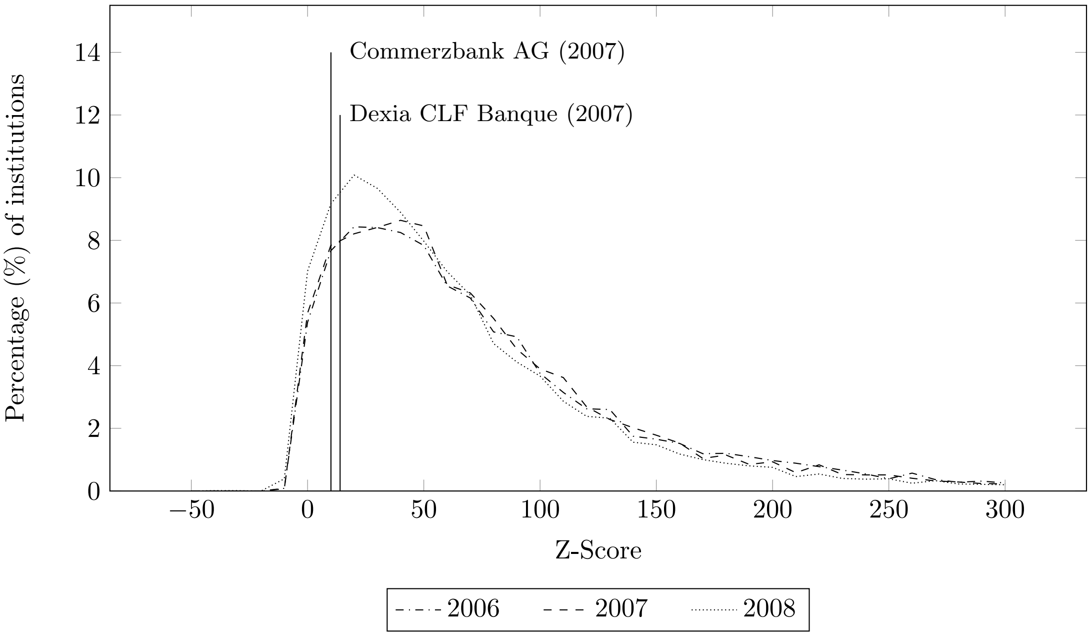

Figure 8: Relative position of Commerzbank AG and Dexia CLF Banque in 2007 with respect to all other banks in the sample. The position is computed using $\bar{Z}_7^5$, but other measures do not produce a substantially different rank.

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 9 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.

11 Conclusions

In this work we focused on the issue of determining reliable estimates of the Z-Score, a popular 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

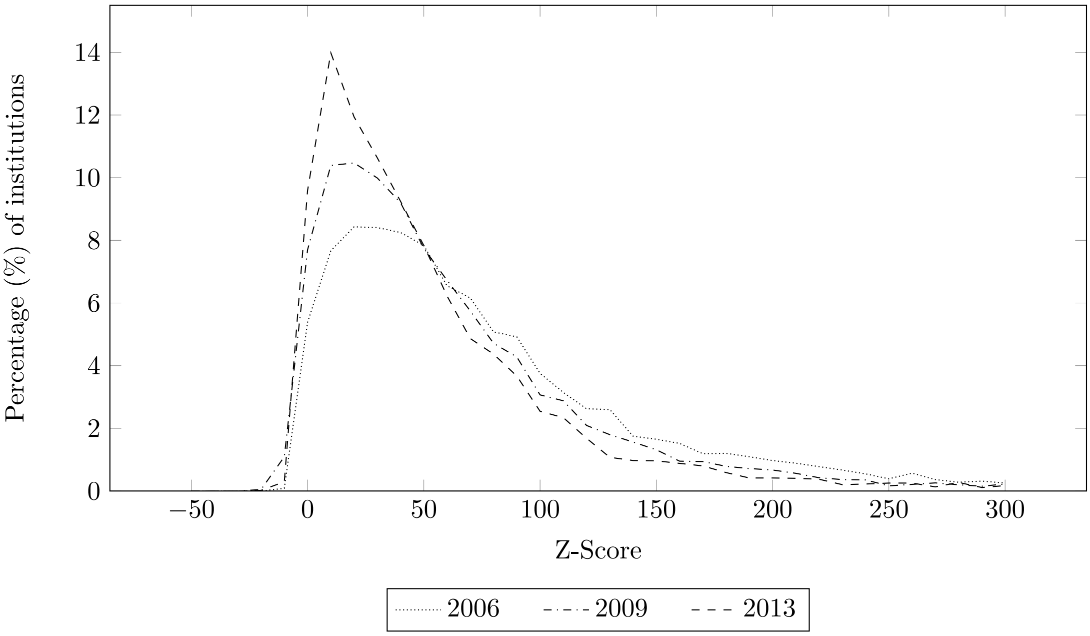

Figure 9: Distribution of bank Z-Score ($\bar{Z}_7^5$) in our sample; evolution between 2006 and 2013

of our novel ones; we carried out an extensive empirical study that contrast results obtained with the aforementioned estimators on over ten thousand banks; and we contrasted results obtained by using Z-score estimators against business news on the banking sector obtained from Factiva.

Acknowledgements

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

References

V. V. Acharya, L. H. Pedersen, T. Philippon, and M. Richardson. Measuring systemic risk. Technical report, New York University, 2010.

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.

G. Edward Miller. Asymptotic test statistics for coefficients of variation. *Communications in Statistics - Theory and Methods*, 20(10):3351–3363, Jan. 1991.

A. Ellul and V. Yerramilli. Stronger risk controls, lower risk: Evidence from U.S. bank holding companies. *Journal of Finance*, 68(5):1757–1803, 2013.

R. Fahlenbrach, R. Prilmeier, and R. M. Stulz. This time is the same: Using bank performance in 1998 to explain bank performance during the recent financial crisis. *The Journal of Finance*, 67 (6):2139–2185, Dec. 2012.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1):3–56, February 1993.

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.

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.

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.

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.

J. Lintner. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets: A Reply. *The Review of Economics and Statistics*, 51(2):222–24, May 1969.

S. Mercieca, K. Schaeck, and S. Wolfe. Small european banks: Benefits from diversification? *Journal of Banking & Finance*, 31(7):1975–1998, July 2007.

R. C. Merton. On the pricing of corporate debt: the risk structure of interest rates. Working papers 684-73., Massachusetts Institute of Technology (MIT), Sloan School of Management, 1973.

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 newsvendor problem under binomial, poisson and exponential demand. *European Journal of Operational Research*, 239(3):674–684, Dec. 2014.

N. Salkind. *Encyclopedia of Research Design*. SAGE Publications, Inc., 2455 Teller Road, Thousand Oaks California 91320 United States, 2010.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3):425–442, 1964.

R. Stever. Bank size, credit and the sources of bank market risk. Technical Report 238, BIS, 2007.