Title: Improved methods for identifying the operational determinants of a bank's capital ratio

Journal: Applied Economics

Author no. 1: Anna P. I. Vong

ORCID no. 0009-0000-6659-0371

Faculty of Business Administration, University of Macau,

Avenida da Universidade E22 Taipa, Macau SAR, China, Tel. +853 8822-4667

E-mail: annavong@um.edu.mo

Author no. 2 (corresponding): Duarte Trigueiros

ORCID no. 0000-0002-3619-4615

ISTAR, ISCTE, University Institute of Lisbon,

Av. Forças Armadas, 1649-026 Lisboa, Portugal, Tel. +351 210 464 023

E-mail: duarte.trigueiros@iscte-iul.pt

Acknowledgement: This research is supported by the Portuguese "Fundação para a Ciência e a Tecnologia, I.P." (FCT) [ISTAR Projects UIDB/04466/2020 and UIDP/04466/2020].

Improved methods for identifying the operational determinants of a bank's capital ratio

Abstract

Published research using econometric models to identify the determinants of bank capital ratios has produced inconsistent results. This is partly due to the failure of model formulations to distinguish between operational and managerial effects. This paper explains how the use of bank leverage can separate these effects, how to prevent financial ratios from undermining model interpretability, and how to identify and avoid ratio-related biases. The application of these improvements is tested on a panel of East Asian retail bank data from 2004 to 2014, covering China, Hong Kong, Indonesia, India, Japan, Singapore, Malaysia, the Philippines, and Thailand. Conclusions are drawn from quasi-experimental designs comparing two-way GMM models before and after implementation of the improvements. The results show that the improved models identify the operational determinants of capital ratios and avoid simultaneity and omitted variable bias and ratio-induced opacity in the results. The use of complementary data segments helps to interpret the results and identify paradigmatic cases for a better understanding of the relationship between regulatory capital and risk.

Keywords: Capital Ratio; Basel Leverage Ratio; Financial Ratio; Ratio-Induced distortion.

JEL codes: C18, C23, C52, G21, G28.

1. Introduction

Banks are required to maintain a minimum level of capital commensurate with the risks they take. In particular, the ratio of available funds to an estimate of the bank's risk, known as the capital ratio, must not fall below a certain level.[1]

Bank managers would find it useful to know which factors arising from the bank's risk exposures lead to increases or decreases in capital ratios. Such knowledge would allow

[1] Basel regulations (Barth *et al.*, 2013). From 2018 onwards, the third round of regulations is in place.

them to focus on best practices and not have to act in an ad hoc, reactive manner by hastily increasing the bank's capital in response to unexpected increases in risk exposures (Cohen and Scatigna, 2016). However, after a few years of attention, the search for the determinants of capital ratios seems to have stalled without clarifying what these factors are. In light of recent developments in the relationship between bank leverage and capital ratios, this paper proposes an improved methodology that will give new impetus to this area of research and allow the applied use of capital ratio models.

Studies attempting to identify the determinants of capital ratios (e.g., Rime, 2001; Kleff and Weber, 2008; Brewer *et al*., 2008; Fonseca and Gonzalez, 2010; Bateni *et al*., 2014; Aktas *et al*., 2015; Klepczarek, 2015; Shingjergjin and Hyseni, 2015) provide conflicting results.[2] Most of these studies have focused on the Total Capital ratio, also known as the Capital Adequacy ratio (CAR), which was originally considered as the main indicator of a bank's ability to withstand losses (Jeff, 1990), rather than the Tier 1 (CR) or Common Equity Tier 1 (CET1) capital ratios, which are the most widely used today.[3]

More recent regulation and published research have highlighted the relationship between CR and the Tier 1 Leverage Ratio (LR),[4] which is easier to calculate than CR but also capable of assessing a bank's likelihood of surviving (Estrella et al., 2000). CR and LR can both be used to evaluate risk (Jarrow, 2013), and a leverage ratio restriction may help limit the risk of bank runs (Dermine, 2015), leading to more truthful risk reporting on the part of banks (Blum, 2008). Rime (2001), Nilssona *et al.* (2014), Hasan *et al.* (2015), and Barth and Seckinger (2018) also agree on the relevance of LR for bank risk assessment. The mounting awareness of the relationship between CR and LR contrasts with a lack of interest in the role that LR can play in the identification of CR determinants.

This paper shows that LR is necessary to validly identify the operational determinants

[2] As part of the discussion of the results of this paper, a critical description of this research is included.
[3] Capital ratios differ in their numerator: Total (CAR), Tier 1 (CR), Common Equity Tier 1 (CET1) Capital.
[4] LR, known in banking circles as Leverage Ratio, has the same numerator as CR but the denominator is Assets.

of CR and to avoid simultaneity and omitted variable bias. It also uncovers constraints that can distort the meaning of coefficients of financial ratios in regression formulations and explains how to mitigate the degrading effect of ratios on the interpretability of models. After providing sufficient reasons for the failure of previous attempts to identify CR determinants, the paper proposes improvements and empirically tests their value using quasi-experimental designs to compare GMM[5] models that ignore and apply the improved methods.

The illustrations and empirical tests presented use a data panel of East Asian retail banks from 2004 to 2014, the only extended period with comparable figures on regulation. The jurisdictions covered are China (CN), Hong Kong (HK), Indonesia (ID), India (IN), Japan (JP), Singapore (SG), Malaysia (MA), the Philippines (PH), and Thailand (TH).

Some of these jurisdictions are fast-growing economies, while others are stable and rich; some are huge, while others are tiny; some are at the forefront of implementing advanced risk assessment and management approaches, while others are just beginning to implement Basel II;[6] some follow government guidelines strictly, while others are allowed more discretion. Using data from jurisdictions where economic and banking characteristics are complementary makes it easier to interpret the results and to identify paradigmatic cases to improve understanding of the relationship between bank capital and risk.

2. Data description and illustration of poorly understood barriers to CR modelling

After a brief description of the dataset used, this section shows the existence of constraints that can distort models that use financial ratios as predictors, explains how to preserve the interpretability of the same models, and shows how to avoid simultaneity and omitted variable biases in CR-determining models without having to use systems of equations.

BankScope (Duprey and Mathias, 2016) provides a 2004-2014 panel of 24 indicators,

[5] Generalized Method of Moments (Arellano and Bond, 1991),
[6] The second round of regulations, Basel II, came into force between 2006 and 2010. The Basel rules allow banks of a certain size to use advanced approaches, rather than standard approaches, to estimate the risks they want to offset. These banks also use these added capabilities to implement advanced risk management tools.

to which the World Bank adds 5 jurisdiction-specific economic factors. Table 1 lists these variables, which include bank ratios, economic factors, and a proxy for the bank size. Table 1 also shows abbreviations and pooled means and standard deviations before and after the 2008 financial crisis.

Table 1

The 595 original banks of the panel were used in the preliminary statistics in Table 1 and in the 2010-2011 pooled regressions shown in Table 2, but only the 254 banks that had adopted Basel II by 2006 were included in the CR determinant models in Tables 3 and 4.

Capital ratios increased globally in the years following the 2008 crisis (Cohen and Scatigna, 2016), but apart from Singapore, East Asian retail banks reduced them. Leverage ratios were stable over the period and bank size moved in line with GDP, including in Japan, which entered a recession in 2010. Except for India, consumer prices experienced a shock in 2008, with both Japan and China experiencing deflation after the 2008 financial crisis. Profitability was lower in Japan than in other jurisdictions and worsened after the crisis, with shocks in ROA and ROE and a reduction in reserves. Indian banks' ratios experienced a shock in 2013, not 2008.

Observed correlations between variables are not stable, depending on jurisdiction and period, except for profitability ratios which have two orthogonal sources of variability, with ROE and Cost to Income (CI) on one side and ROA with other efficiency ratios on the other.

There are clear differences between developed and developing economies. The ratio of loans to assets was stable in developed economies but increased in developing economies with an upward shock in 2008. CR, LR, ROA and ROE were higher in developing economies than in developed ones. Shocks to ROE occurred in developed economies only. The correlation between CR and LR (Cathcart et al., 2015) decreased over the period in developed economies but was high and steady in developing economies with a downward shock in 2008.

The left-hand side of Table 2 lists the partitions of the dataset that take advantage of the diversity of the region to compare models under complementary economic and banking characteristics. Each partition provides two complementary data segments. The partition by GDP *per capita*, for example, compares banks in developed and developing economies, while the partition by GDP compares banks in larger and smaller economies, and there are also partitions that use individual bank characteristics to segment data.

Table 2

2.1 Ratios with common numerators in the same formulation

This section now discusses the difficulties involved in modelling CR using ratios, which have so far been poorly understood. Ratios are a tool of financial analysis, conveying scale-free information. They are also used in econometric models as dependent and independent variables (also referred to in this paper as "predictors") but some authors have criticised this use based on their innate correlation, faulty control of scale, or the simultaneity that may arise when the predicted variable, being a ratio, carries more than one effect. Other authors have dismissed these concerns, arguing that if models using ratios are correctly treated, they are not necessarily faulty. Wiseman (2009) reviews this debate.

The question of which disagreement is most enduring, concerns whether ratios with common components will distort the models in which they are included. It is a well-known fact that ratios with common numerators, for example, ROA and ROE, are correlated even if the components are strictly independent. This has led Kronmal (1993) to claim that when such ratios are used in regressions, estimates reflect spurious associations. Firebaugh and Gibbs (1985) contend that such correlation is an integral part of the model.

Here, it is shown that regression coefficients may suffer distortions when two ratios with common numerators are used as predictors in the same formulation. This is no small matter since almost half of the papers cited in relation to CR determinant prediction include ratios with common numerators. The same is true for models used, for example, by Beaver *et al*. (1997) and Ou and Penman (1989), to name just a few, well-known cases.

Besides listing data partitions, Table 2 shows in columns 4 to 6 the results of running three pooled regressions (years 2010-2011) in which ROA and ROE predict CR, first separately (rows 1 and 2 in each segment) and then together (row 3). Table 2 therefore shows three regressions per segment: one with ROA, another with ROE and the third with ROA and ROE. The latter is a case of two ratios with common numerators in the same formulation.

Instead of regression parameters, which are useless in this case, the table displays *t*-statistics (coefficients divided by standard errors). Irrespective of the unit of measurement, *t*-statistics show the relative magnitude and direction of the effects of ROA and ROE on CR, being therefore suited for comparisons. The logarithmic (hereafter "log") transformation is applied to CR and the previous year CR (CR_{t-1}) is included as independent variable.[7]

In some of the segments, labelled in column 7 as a "match", the R-squared and the *t*-statistics of regressions in which ROA alone explains CR (columns 4 and 5, row 1) are extremely low but increase sharply after adding a significant ROE to the regression (row 3). Such increases suggest that ROA and ROE interact, but the usual type of interaction, namely ROA-ROE correlation, is unable to provide a credible explanation for the *t*-statistic of ROA after ROE is included. Indeed, increases of such magnitude would call for a strong ROA-ROE correlation but since ROE alone significantly explains CR (row 2), any strong correlation that might exist between ROA and ROE would inevitably show in regressions in which ROA alone explains CR, which is not the case.

The origin of interactions becomes apparent when the regressions that use ROA and ROE together are performed using log-transformed ROA and ROE, as shown in columns 8 and 9. In logs, ratios are subtractions, and, for those segments labelled as "matches", the

---

[7] The highly skewed distribution of CR requires the use of log-transformed CR to avoid influential cases and to control for heteroscedasticity in error terms. The inclusion of CR_{t-1} aims to reduce model misspecification.

modelling algorithm estimates coefficient values of log ROA and log ROE that bring the respective t-statistics to near symmetry. For simplicity, let us assume unit standard errors so that t-statistics equal regression coefficients. Then, the functional form of these regressions is

$$log\,CR_t = a + b_0\,log\,CR_{t-1} + b_1\,log\frac{NI}{avg\,A} + b_2\,log\frac{NI}{avg\,E} + u,$$

where NI is Net Income, A is Assets, E is Equity, *avg* is year-begin and year-end average, $NI\ /\ avg\ A$ is ROA, and $NI\ /\ avg\ E$ is ROE. Since symmetric t-statistics in this case means $b_1 \approx -b_2 \approx b$, the functional form is, in fact,

$$log\,CR_t = a + b_0\,log\,CR_{t-1} + b\,log\frac{avg\,E}{avg\,A} + u$$

where NI is no longer present, CR is explained by the log of the ratio of average Equity to average Assets, and the functional form has one less parameter. Table 2 shows that, for the matching segments, the t-statistics of regressions in which the logs of ROA and ROE predict CR (columns 8 and 9, row 3) approximate to the strong association between the log of the ratio of average Equity to average Assets and CR (column 10).

Symmetric t-statistics in regression coefficients denote two predictors with the same effect on prediction, but with opposite directions. Therefore, the observed symmetries are a successful attempt, on the part of the modelling algorithm, to discard the variability of NI in the two identical numerators of ROA and ROE, balancing one against the other. Although this balancing is made possible by the fact that, in logs, the effects of the numerators of ROA and ROE are additive and therefore separable from the effects of denominators, it is not obvious how the modelling algorithm separates them. Symmetries in no way follow from a simple manipulation of the functional form, in which there are four effects to be modelled (two numerators plus two denominators), but only two available parameters, b_1 and b_2.

Rearranging the functional form as

$$log\,CR_t = a + b_0\,log\,CR_{t-1} + log\,NI\,(b_1 + b_2) - b_1\,log\,avg\,A - b_2\,log\,avg\,E + u$$

it is apparent that the variability available to model CR has two terms. One is NI-related,

$$log\, NI\, (b_1 + b_2)$$

and the other is NI-unrelated,

$$b_0\, log\, CR_{t-1} - b_1\, log\, avg\, A - b_2\, log\, avg\, E.$$

Hence if log NI is strictly independent of $log\, CR_t$, and its variability is not negligible (as is the case with matching segments), then the modelling algorithm will have to make $b_1 + b_2$ equal to zero. In other words, every time log NI has no explanatory power over $log\, CR_t$ but carries sizeable variability that must be dealt with, the modelling algorithm will not be free to find the optimal solution in any locus of the b_1, b_2 space, being constrained to find the solution in the line $b_1 = -\, b_2$ using one degree of freedom (parameter) less. Symmetric regression coefficients therefore follow from constraint $b_1 + b_2 = 0$. This is not surprising, as algorithms seek to maximize explained variability, hence the variability of log NI is disregarded when it plays no role in explaining CR but, since log NI is a component of the variability of two predictors, each with its own regression coefficient, the algorithm must balance one coefficient against the other. The rearranged functional form shows that the algorithm can do this, and how it is achieved.

The term "constraint" is used here for lack of a better name, but it is a gradual, not a yes or no effect. In fact, if the correlation of log NI to $log\, CR_t$ is small but not zero and the variability of log NI is not negligible, then $b_1 + b_2$ is small but not exactly zero. However, negligible variability of log NI rules out constraints, no matter how small the correlation of log NI to $log\, CR_t$ may be. With this in mind and noting that the roles of ROA and ROE in the generation of constraints may swap, it is easy to understand what the algorithm is trying to achieve in the non-matching segments in Table 2.

Constraints arise from the presence of too many parameters in the functional form compared to the sources of variability being modelled. For example, in some of the regressions in Table 2, ROA should not have been included as a predictor. Therefore, to limit constraints, ratios with common components should not be included in the same regression, but if common components are unavoidable, then the number of ratios should be reduced without sacrificing useful variability. This being the case, the belief that adding inert predictors to the functional form is harmless should be set aside when it comes to ratios.

So far it has been explained how the modelling algorithm makes symmetric coefficients appear in transformed ratios with common numerators. The distortions observed in the *t*-statistics of the untransformed ratios reflect the symmetries observed in the transformed ratios. In transformed ratios, the algorithm achieves full balancing of effects, whereas in untransformed ratios, the algorithm achieves partial balancing because the relationship between the numerator and the denominator does not lend itself to exact separation of effects. The origin of distorted coefficient values of untransformed ratios in regressions would therefore be found in approximations to symmetry.

2.2 Ratio-induced opacity of regression formulations

Another poorly understood difficulty that can become acute in regressions that use many ratios, as is the case with CR determinant models, is ratio-induced opacity of the results. Any approach to ratio-induced opacity should be based on a clear understanding of how ratios work, so that opacity is identified as such and not as some other characteristic.

A ratio is a scale-free observation because the numerator and denominator reflect scale (bank size) that is cancelled out when the ratio is formed. However, a division removes any effect that appears on its two factors, as long as that effect is multiplicative rather than additive,[8] as is the case with reported accounting numbers or market prices. Multiplicative variables are those in which distributions are preserved when they are multiplied or divided,

[8] Since early, economists have noticed that random processes where the multiplications of probabilities play a major role, lead to a type of randomness where effects are proportionate, and observations behave exponentially (Aitchison and Brown, 1957; Singh and Whittington, 1968; Ijiri and Simon, 1977).

not when they are added or subtracted. While the simplest additive formulation is $x = \mu + u$ (x explained as expected μ plus deviate u), the multiplicative equivalent would be $w = w_0 v$ where each realisation of w is explained as constant level w_0 multiplied by random factor v.

Consider the following two types of ratio component: the deflator, which contains the effect of scale and nothing more and the active component, which includes a scale-free effect in addition to scale. If s means scale, $v = s$ for the deflator type and $v = sf$ for the active component type, with f meaning a scale-free financial feature, e.g., liquidity or profitability. Total Assets is a typical example of a deflator, as it is supposed to reflect the effect of scale and nothing more.

Ratios in which one of the components, typically the denominator, is a deflator, apportion one source of variability to prediction, namely the scale-free feature of the other component. ROA, for example, apportions scale-free Net Income to the modelling. As the denominator is inert with respect to prediction, the statistical behaviour of ratios such as ROA is the same as that of any other multiplicative random variable.

Ratios in which each component is active, having its own feature in addition to scale, are not like any other variable because they bring more than one effect into the modelling, namely the two main effects (the features of the numerator and denominator), plus their interaction, which is the ratio itself. If f_n is the feature of the numerator and f_d is the feature of the denominator, then the two types of ratios just mentioned are formally described as

$$\frac{sf_{\mathrm{n}}}{s} = f_{\mathrm{n}} \quad \text{and} \quad \frac{sf_{\mathrm{n}}}{sf_{\mathrm{d}}} = f_{\mathrm{n/d}},$$

constant levels apart. Ratios with deflators as denominators belong to the f_n type and ratios with active denominators belong to the $f_{\mathrm{n/d}}$ type.

Ratios of the f_n type will not create constraints because the denominator's scale is cancelled by the numerator's scale. Only ratios with common f_{d} or f_{n} can create constraints.

Moreover, ratios of the $f_{\mathrm{n/d}}$ type will increase the opacity of regression formulations, firstly because the feature in the denominator, f_{d}, introduces nonlinearity into the estimate, and also because common denominators, not just numerators, will now be able to create constraints, which will distort coefficients directly or make estimation unstable as the algorithm is free to select the solution from a line of optimal loci, not just one locus. Therefore, to limit opacity, ratios of the type f_n should be preferred.

2.3 Simultaneity and misspecification in CR-predicting regressions

Lastly, this section turns to the predicted variable, CR. To comply with regulation, bank managers act on the numerator of CR, which is Tier 1 Capital, to balance adverse changes in the denominator, Risk-Weighted Assets (RWA) that reflect risks associated with bank operations and cannot adjust in the short term. Therefore, there are two active, interrelated components in CR, one management-made and the other related to operational risk.[9]

CR modelling faces a simultaneity problem in that the same predictors must explain two effects, one of which (changes in Tier 1 Capital) is partly the result of actions taken in response to the other (changes in Risk-Weighted Assets). Besides, the modelling of CR also faces a misspecification issue because what managers want to discover by using CR-determinant models is not the complete list of those determinant factors, but only those operational factors that can help them avoid having to act on Tier 1 Capital.

Both difficulties would be solved if the numerator of CR, Tier 1 capital, was removed from the model, but an econometric model, being a complement to enquiry must provide as required. What is needed here is a ratio capable of taking management-induced variability away from the model. If *C* is Tier 1 capital, A is assets, and x_1, x_2 ... are CR predictors, the inclusion of $C/A =$ LR among predictors will do just this, accounting for and, at the same time, assessing the effect of changes in Tier 1 capital on CR. Consider the functional form,

[9] Bartlett and Partnoy (2020) offer a critical review of the literature on the use of ratios as predicted variables.

$$\frac{C}{RWA} = a + b_1 \frac{C}{A} + b_2 x_2 + b_3 x_3 + \cdots + u$$

For C to cancel out on either side of this formulation, C must add, not multiply, other variables. The use of log-transformed CR and LR would achieve this. In practice, the highly skewed distribution of CR already requires log-transforming, and it matters little whether LR, where the numerator is bounded by the denominator, is log-transformed or not.

Once *C* is controlled, simultaneity is no longer an issue and the x_1, x_2 ... can be chosen to explain the operational risk component of CR, offering managers what they want to know. Thus, the inclusion of LR among CR predictors extricates operating influences from others, with the extra advantage that it also annuls any differences that may exist in the definition of regulatory capital among the various jurisdictions, making for more robust models.

3. The experiment

The improvements that should be made to avoid constraints and simultaneity and to increase the interpretability of models incorporating ratios are summarised below:

i. Avoid using in the same formulation ratios with common, active components.
ii. Use the ratios required to cover the sources of variability to model, no more.
iii. Avoid using ratios with active denominators.
iv. Include as predictor a ratio that explains the feature of the dependent variable causing simultaneity, use independent variables that explain the other feature.

This section empirically compares models with and without these improvements. The bank data is from years 2006 to 2014[10] and the panel is limited to the 254 banks that, at the time, followed Basel II regulation, which restricts the scope of the models but not the validity of the tests, which are based on comparisons rather than individual results.

Partitions are circumscribed to the most relevant three, namely,

[10] The initial year is 2006, not 2004, as two years are engaged by first- and second-order terms.

- By GDP *per capita* as per 2006, with segments being the developed (HK, JP, SG) and developing economies (CN, ID, IN, MA, PH, TH).
- By GDP as per 2006, with segments being the smaller (HK, ID, IN, MA, PH, SG, TH) and larger economies (CN, JP).
- By average bank size, with smaller and larger banks separated by the median DIM.

3.1 Description of models

GMM estimation is used for CR modelling. Two-way differencing cancels unobserved bank and year effects, and the dynamic term CR_{t-1} accounts for CR persistency.[11] Differencing, which is advisable when including a dynamic term, is also ideal when comparing results. All formulations incorporate a proxy for scale (DIM), and GDP, GDG, and INFL to capture jurisdiction-specific variability that would otherwise add to the error term.

Tables 3 and 4 describe the models obtained by, respectively, ignoring and applying the improvements to the same banks, period, and dependent variable. For each model, the tables provide the list of predictors, coefficient values, and z-statistics with significance levels. Tests of over-identifying restrictions in instruments, serial correlation and the Chi Square statistics for predictors and year-dummies are also given. The coefficient values of year-dummies are not shown. Ratios are log-transformed when skewed with positive values only.

Tables 3 and 4

To detect constraints, instruments that avoid the common ratio component are fitted to the ratio in question, and the ratio is then replaced by its fitted values in the original model. If the fitted coefficient differs in sign from the original coefficient, a constraint is assumed. Constraints and interpretability issues can be observed in GMM as well as in other modelling tools, in which, though, the respective estimation assumptions might not have been met.

[11] GMM addresses estimation issues posed by the dynamic term when the period is small. Endogeneity issues are addressed using lagged variables as instruments. Heteroscedastic data make it advisable to use robust covariance matrix estimation (Windmeijer, 2005). See e.g., Croissant and Millo (2014) for details and references.

Models that ignore improvements (Table 3) include the variables listed in Table 1 that show traces of significance in predicting CR, except LR. Some of these are ratios with common components. Models that apply improvements (Table 4) use 8 ratios, namely,

- The Tier 1 Leverage ratio (LR) which is the year-end Tier 1 Capital to Assets ratio.
- Year-end Loan Loss Reserves to Gross Loans ratio (LLR).
- Year-end Loans to Deposits ratio (LDS).
- Year-end Loans to Assets ratio (LA).
- Year-end Impaired Loans to Gross Loans ratio (NPL).
- The new ratio of yearly Loan Loss Provisions to Average Earning Assets (PRO, not to be found in Table 1, computed from LLP and NIM).
- Yearly operating efficiency, with two ratios, Net Interest Margin (NIM) and Other Operating Income (OOI, not to be found in Table 1, computed from ORA, NIM and IRA). The denominators of NIM and OOI are average (year-begin and year-end) Earning Assets. The two numerators add to Yearly Total Operating Income.

Therefore, in models applying improvements, equity items are omitted, and operating income items replace Net Income. These omissions cannot lead to missing or inverse effects because LR takes into account the effect of Tier 1 Capital and Net Income cannot affect the operations of the year. Two ratios, LR and LA, have deflators in the denominator. The others require active denominators to make the model meaningful, namely Gross Loans deflates loan related items and Average Earning Assets deflates annual flows. Two ratios, LA and LDS, have a common numerator (Gross Loans), but it would not make sense to exclude one.

3.2 Results and discussion

In the models applying the improvements (Table 4), the reduction in the parameters made available for modelling seems to have inhibited constrained coefficients in LA and LDS. LLR, NPL and INFL are not significant anywhere and operating ratios only explain CR in developed economies. Significant CR_{t-1} appear in certain partitions only, which is also the case for LR and LA.

In particular, for smaller banks and banks in small and developing economies, i.e. banks with a lower standing in the region, CR_{t-1} and LA are significant, but LR and the year dummies are not. In large and developed economies, i.e. banks with a higher standing in the region, the opposite is true: LR and the year dummies are significant, but CR_{t-1} and LA are not. The segment with the bigger banks adds significant CR_{t-1} to the characteristics of higher-standing banks. LDS is significant for bigger banks and large economies. Efficiency ratios are significant for developed economies only.

The significance of LR stems from events affecting Tier 1 capital. It is therefore plausible to conclude from Table 4 that management action is more likely in banks with a higher standing, which are subject to greater scrutiny and take action to keep CR in line. In banks with lower standing, LR is less significant because managers can wait for the annual results to play out as usual, especially since thrifty customers and trouble-free finances are common in the less vibrant East Asian economies of the time (Jones and Zeitz, 2017). In turn, the significance of CR_{t-1} indicates a more persistent CR. For banks with a lower standing in the region, the standard approaches to risk assessment are more prevalent than the advanced approaches (Barth et al., 2013), and risk charges are therefore selected from a small list of assets. RWA, therefore, will be less diverse over time and across assets. When diversity is low, variability is limited and RWA has less to explain. Since, as seen, Tier 1 capital is also less diversified in this case, CR_{t-1} should explain a large part of CR variability.

Given this, the models in Table 4 plausibly suggest that lower standing banks are likely to use less advanced approaches to risk assessment, where LA conveys the required information, while other effects, whether ratios or year dummies, are not significant. Conversely, higher standing banks, where the use of advanced approaches is more common, should have higher volatility and lower CR persistence. In this case, LDS, together with operating efficiency and other ratios, provides the necessary information to predict CR.

In summary, Table 4 shows a clear division between banks in large and developed economies on the one hand and small banks and banks in small and developing economies on the other, the former with high LR, low CR_{t-1} and low LA and the latter with the opposite. This split explains the models almost entirely and is also plausible.

In terms of the direction of the factors influencing CR, bank efficiency increases CR, while loans, loans to deposits and loan loss provisions decrease CR. Bank size and GDP can either increase or decrease CR. Given the diversity and inconsistency of published results, it is reassuring that ours are clear and intuitively consistent.

With minor adjustments, the improved models can be used to find the determinants of CAR, CET1, not just CR, or with different banking regulatory frameworks, namely Basel III. For banks from the same jurisdiction XOA can be used in place of GDP, GDG and INFL.

Models that ignore improvements (Table 3) replicate the undesirable features of published formulations of capital ratio prediction, namely the large number of ratios, the use of ratios with common components, and the poor separation between management-made and operational effects. As a result, these models suffer from constraints, namely in the NPL-ILE and ROA-ROE pairs, while the two risk-related ratios LA and LDS, also with a common numerator, are not significant in any segment, which is puzzling. At least one of these ratios should be significant as LA and LDS are closely related to bank risk. CR is also related to XOA in smaller banks and to ROA, NIM and REP in developed economies. CR_{t-1} is highly significant for all the segments which, as seen above, is misleading. Clearly, it is the inclusion of LR that has led to this distinctiveness of CR_{t-1} and to the completeness and plausibility of the models that apply improvements. Therefore, the models in Table 3 where LR should be present but is not, are poorly specified.

The following lines put the improvements presented here and their usefulness into perspective, briefly reviewing the most common approaches and methodologies used to identify the drivers of bank capital.

Rime (2001) avoids simultaneity in models by using a system of simultaneous equations that considers the numerator and denominator of the capital ratio separately. Other examples of authors who estimate size-adjusted bank capital rather than the capital ratio are Kleff and Weber (2008), Brewer et al. (2008) and Fonseca and González (2010). These authors use dynamic modelling or GMM, and the denominator of the capital ratio, risk-weighted assets, is added to models as a predictor. Although this avoids simultaneity, the models say little about the factors that determine a bank's risk.

Brewer et al. (2008) and Fonseca and González (2010) use heterogeneous samples, including banks from around the world, to test whether public policy, regulation and other characteristics of different jurisdictions significantly affect bank capital. Additional predictors are included in the formulations to characterise jurisdictions. As a result, the number of predictors is typically large, and results are difficult to interpret. This opacity is exacerbated by the fact that risk-related factors cannot be separated from management-related factors.

Other publications also mentioned replicate the financial analysis approach, where each bank ratio is included as a predictor to capture a specific feature (Bateni *et al*., 2014; Aktas *et al*., 2015; Klepczarek, 2015; Shingjergjin and Hyseni, 2015). As a result, these models are prone to simultaneity and constraining problems. For example, Klepczarek (2015) finds that ROA and ROE affect bank capital in opposite directions, while Shingjergji and Hyseni (2015) find similar behaviour for LA and LDS. Other conflicting results are found when comparing Bateni *et al*. (2014) with Klepczarek (2015) on the ratio of Deposits to Assets. Aktas *et al*. (2015) and Shingjergji and Hyseni (2015) use proxies for bank leverage in their models, but do not use LR. Therefore, in addition to suffering from misspecification bias, the results are ambiguous as it is not clear whether the observed changes in the dependent variable are due to managerial action or operational risk factors.

In none of the cited publications is there any attempt to divide the observations into complementary segments of economic or banking relevance, which leads to a superficial interpretation of the results. For example, the common finding that large banks tend to have lower capital ratios (Klepczarek, 2015; Shingjergji and Hyseni, 2015; Aktas *et al*., 2015) could have been made more precise by distinguishing between higher and lower standing banks. The fact that authors such as Shingjergji and Hyseni (2015) find no correlation between bank capital and profitability is consistent with the findings reported here for lower standing banks. The difference, however, is that for these authors such a finding is puzzling due to missing context, whereas in the present paper the same finding is illuminating.

In summary, apart from the awareness of the distortions and biases that ratios can introduce in models, the two interrelated features that distinguish the methodology presented here from that of other publications are the use of LR and data segments with complementary economic and banking characteristics. The first makes it possible to predict the determinants of banks' operational risk while avoiding simultaneity, misspecification, and ambiguity. The second adds context to results that would otherwise be difficult to understand.

4. Conclusion

The paper has presented improved methods for identifying the operational determinants of a bank's capital ratio, starting with a discussion of the difficulties inherent in using bank ratios as predictors. The discussion has helped to integrate the reality of bank ratios into the context of econometric modelling and to address issues that we believe have been poorly understood.

The paper has made it clear that when predictors are ratios with common numerators, regression coefficients may model something different from what was intended. Constraints may also make models unstable because trivial differences in the data may lead to algorithms selecting one pair of coefficients rather than another. Common denominators that, rather than being purely scale deflators, play an active role in modelling, can also lead to constraints. Models subject to this type of distortion are those that use many ratios as predictors, as is the case, for example, in the search for the factors that determine bank risk, the direction of changes in future earnings, or bank efficiency. Since constraints are caused by over-specification of the functional form, it is recommended that the number of predictors should be kept to a minimum.

Ratio-induced opacity was also discussed. Ratios with active denominators were identified as capable of increasing the complexity of model results, and it was recommended that when selecting predictors, preference should be given to ratios in which the denominators perform scale deflation and nothing more.

The question of how to conduct a more focused examination of the factors that influence capital ratios was then explored. It was shown that the inclusion of the Basel leverage ratio as a predictor in models leads to the separation of operational influences from others, providing what managers are interested in and eliminating differences that may exist in the definition of regulatory capital across jurisdictions.

Using the leverage ratio and data segments with complementary economic and banking characteristics, it was possible to identify two paradigm groups of banks, one with a lower standing in the region, where capital ratios have a trend and little remains to be explained, and the other with a higher standing, where capital ratios have little memory of the past but leverage and other ratios have relevant information about the present. In addition to establishing the role of leverage ratios in explaining capital ratios, this finding provides a starting point for future modelling and adds context to the interpretation of the results.

The dataset used for illustrative purposes is stable regarding regulatory requirements, which is something hard to find nowadays. The improvements described and tested here apply equally to data based on Basel II and Basel III banking regulations, but Basel III data would not have been stable in terms of regulation.[12] Conclusions were drawn from quasi-experimental designs, that is, not from models in isolation, but from the comparison of models that used standard and improved methods for the same data.

The improvements aimed at obtaining reliable and interpretable models are difficult to reconcile with the need to cover the main sources of variability in the predicted variable. This conflict between reliability and comprehensiveness stems from the need to avoid the difficulties inherent in using ratios as dependent and independent variables. A major benefit of this paper is to make modellers aware of these difficulties and how they can be mitigated.

5. References

Aitchison, J., Brown, J., 1957. The Lognormal Distribution (Cambridge University Press, UK).

Aktas, R., Acikalin, S., Bakin, B., Celik, B., 2015. The determinants of banks' capital adequacy ratio: some evidence from south-eastern European countries, *Journal of Economics and Behavioural Studies*, 7 (1), 79-88.

Arellano, M., Bond, S., 1991. Some tests of specification for panel data: Monte Carlo evidence application to employment equations. *Review of Economic Studies*, 58, 227–297.

Bartlett, R., Partnoy, F., 2020, The ratio problem, working paper SSRN 3605606.

Barth, A., Seckinger, C., 2018. Capital regulation with heterogeneous banks –Unintended consequences of a too strict leverage ratio, *Journal of Banking and Finance*, 88, 455-465.

Barth, J., Caprio, J., Levine, R., 2013. Bank regulation and supervision in 180 countries from 1999 to 2011. NBER Working Paper No. w18733.

Bateni, L., Vakiliford, H., Asghari, F., 2014. The influential factors on capital adequacy ratio in Iranian banks, *International Journal of Economics and Finance*, 6 (11), 108-116.

[12] See e.g., https://www.investopedia.com/terms/b/basell-iii.asp on the controversy surrounding Basel III.

Beaver, W., McAnally, M., Stinson, C., 1997, The information content of earnings and prices: a simultaneous equations approach, *Journal of Accounting and Economics*, 23, 53-81.

Blum, J. 2008. Why Basel II may need a leverage ratio restriction, *Journal of Banking and Finance*, 32 (8), 1699-1707.

Brewer, E., Kaufman, G., Wall, L. 2008. Bank capital ratios across countries: why do they vary? *Journal of Financial Services Research*, 4 (December), 177-201.

Cathcart, L., El-jahel, U. and R. Jabbour, 2015. Can regulators allow banks to set their own capital ratios? *Journal of Banking and Finance*, 53, 112-123.

Cohen, B. and M. Scatigna, 2016. Banks and capital requirements: Channels of adjustment, *Journal of Banking and Finance*, 69, 556-569.

Croissant, Y., Millo, G., 2008. Panel data econometrics in R: the plm package. *Journal of Statistical Software*, 27 (2).

Dermine, J. 2015. Basel III leverage ratio requirement and the probability of bank runs. *Journal of Banking and Finance*, 53, 266-277.

Duprey, T., Mathias, L. 2016. BankScope dataset getting started. Available at SSRN 2191449.

Estrella, A., Park, S., Peristiani, S. 2000. Capital ratios as predictors of bank failure, *FRBNY Economic Policy Review*, 6 (2), 33-52.

Fonseca, A., González, F. 2010. How bank capital buffers vary across countries: the influence of cost of deposits, market power and bank regulation, *Journal of Banking and Finance*, 34, 892-902.

Firebaugh, G., Gibbs, J. 1985. User's Guide to ratio variables. *American Sociological Review*, 50 (5), 713-722.

Ijiri, I., Simon, H. 1977. Skewed distributions and the size of business firms (North-Holland, Amsterdam).

Hasan, I., Siddique, I., Sun, X., 2015. Monitoring the invisible hand of market discipline: capital adequacy revisited. *Journal of Banking and Finance*, 50, 475-492.

Jarrow, R. 2013. A leverage ratio rule for capital adequacy. *Journal of Banking and Finance*, 37, 973-976.

Jeff, L., 1990. Capital adequacy: the benchmark of the 1990's. *Bankers Magazine*, 173 (1), 14-18.

Jones, E., Zeitz, A., 2017. The limits of globalising Basel banking standards. *Journal of Financial Regulation*, 3, 89-124.

Kleff, V., Weber, M., 2008. How do banks determine capital? Evidence from Germany, *German Economic Review*, 9 (3), 354-372.

Klepczarek, E., 2015. Determinants of European banks' capital adequacy. *Comparative Economic Research*, 18 (4), 81-96.

Kronmal, R., 1993. Spurious correlation and the fallacy of the ratio standard revisited, *Journal of the Royal Statistical Society*, Series A, 156(3), 379-392.

Nilssona, V., Nordströmb, J., Bredmar, K., 2014. Need for liquidity and the capital structure of Swedish banks following the financial crisis. *International Journal of Finance and Banking Studies*, 3 (2), 10-27.

Ou, J., and Penman, S., 1989. Financial statement analysis and the prediction of stock percentage change in prices. *Journal of Accounting and Economics*, 11, 295- 330.

Rime B., 2001. Capital requirements and bank behaviours: empirical evidence for Switzerland. *Journal of Banking and Finance*, 25 (4), 789-805.

Shingjergjin, A. and Hyseni M., 2015. The determinants of the capital adequacy ratio in the Albanian banking system during 2007 – 2014. *International Journal of Economics, Commerce and Management*, Vol. III, Issue 1, Jan 2015, p. 1-10.

Singh, A., Whittington, G. 1968. Growth, Profitability and Valuation (Cambridge University Press, UK).

Windmeijer, F., 2005. A finite sample correction for the variance of linear efficient two-steps GMM estimators. *Journal of Econometrics*, 126, 25-51.

Wiseman, R., 2009, On the use and misuse of ratios in strategic management research, Bergh, D.D. and Ketchen, D.J. (Ed.) Research Methodology in Strategy and Management, Vol. 5, pp. 75-110 (Emerald Group Publishing Limited, Leeds). https://doi.org/10.1108/S1479-8387(2009)0000005004, ISSN: 1479-8387.[13]

[13] Robert M. Wiseman, Eli Broad Graduate School of Management, Michigan State University, East Lansing, MI.

Table 1: abbreviations and description of variables. Not included is previous year CR ($\log CR_{(t-1)}$). In the description of variables, the term 'av' denotes the period-begin and period-end average. The sign '%' indicates that ratio values are percentages, that is, multiplied by 100. CAR, CR, LR and RWR values are as per Basel II. There are missing values for any of the years of the period. Negative Equity cases are excluded from computations and modelling.

		2004-2008		2009-2014	
Abbreviation	Variable Description	Mean	Standard Deviation	Mean	Standard Deviation
CAR	Tier 1 + 2 Capital to Risk-Weighted Assets %	15	10	15	8.2
CR	Tier 1 Capital to Risk-Weighted Assets %	12	10	12	7.9
LR	Tier 1 Capital to Assets (total) %	6.7	4.7	6.9	4.75
RWR	Risk-Weighted Assets to Assets (total) %	0.6	0.11	0.6	0.21
DIM	Logarithm basis 10 of Assets (total, th. USD)	6.9	0.79	7.2	0.75
ROA	Net Income to av. Assets (total) %	0.81	1.8	0.71	1.9
ROE	Net Income to av. Equity (total) %	8.8	16	8.4	11
NIM	Net Interest Margin %, as Net Interest Revenue to av. Earning Assets	2.9	2.6	2.7	4.1
IRA	Net Interest Revenue to av. Assets %	2.6	2.4	2.3	1.4
ORA	Other Operating Income to av. Assets %	0.76	0.83	0.79	0.77
XOA	Non-Interest Expenses to av. Assets %	2.2	1.4	2.1	2.1
REP	Recurring earning power % as stable income to av. Assets	1.4	1.3	1.4	1.2
CI	Cost to Income % as overhead costs to Net Interest Revenue Plus Other Operating Income	59	32	59	23
EA	Equity to Assets (total) %, inverse of equity multiplier	7.7	5.1	8	4.7
EL	Equity to Net Loans %	17	23	17	26
EDS	Equity to Deposits, Money Market, S-T Funding %	10	8.7	10	13
ELB	Equity to Liabilities (total) %	9.1	7.3	9.1	10
LLR	Loan Loss Reserves to Gross Loans %	2.9	3.1	2.1	2.5
LLP	Loan Loss Provisions to Net Interest Revenue %	12	43	13	25
NPL	Impaired Loans to Gross Loans %	4.9	4.6	3.3	3.7
ILE	Impaired Loans to Equity %	49	57	31	45
LA	Net loans to Assets (total) %	57	15	58	14
LDS	Net loans to Deposits, Money Market, S-T Funding %	67	21	68	21
LDB	Net Loans to Total Deposits and Borrowing %	65	15	66	16
GDP	Logarithm basis 10 of jurisdiction GDP (th. USD)	12	0.6	12	0.5
GDG	Jurisdiction GDP annual growth rate %	5.4	3.5	3.6	3.9
CDP	Jurisdiction bank capital as a percentage of GDP %	7	2.5	7.2	2.8
PPC	Jurisdiction GDP *per capita* (th. USD).	18	17	19	17
INFL	Jurisdiction Consumer Price Index (annual) %	2.9	3.5	3.4	3.6

Table 2: Partitions of the dataset into segments. The number of banks per segment is displayed in column 1. Each segment shows three pooled 2010-2011 regressions, one (labelled in column 3 as '1') in which $log\ CR$ is explained by $log\ CR_{t-1}$ and ROA, another (labelled in column 3 as '2') in which $log\ CR$ is explained by $log\ CR_{t-1}$ and ROE, and finally the third regression (labelled in column 3 as '3') in which $log\ CR$ is explained by $log\ CR_{t-1}$, ROA, and ROE. Column 4 shows the incremental R-Square, and columns 5-6 show *t*-statistics (coefficients divided by standard errors). Column 7 signals segments that match the conditions stated in the paper for a constraint to occur, having an incremental R-squared of the ROA regression below 0.007 and an incremental R-squared of ROE regression above 0.012. Columns 8-9 show *t*-statistics of three regressions in which $log\ CR$ is first explained by $log\ CR_{t-1}$ and log ROA, then by $log\ CR_{t-1}$ and log ROE, and finally by $log\ CR_{t-1}$, log ROA, and log ROE. Column 10 shows *t*-statistics of regressions in which $log\ CR_{t-1}$ and $log\ (E/A)$, explain $log\ CR$. It is verified that all the matching cases exhibit nearly symmetric *t*-statistics which are, in module, like the *t*-statistics of $log\ (E/A)$. In column 7, the case labelled (1) is a match but has a significant ROA, and in the case labelled (2), the roles of ROA and ROE are reversed. The significance level is in asterisk notation: '***' is p < 0.000, '**' is p < 0.01, '*' is p < 0.05 '.' is p < 0.1. p < 0.05 if |*t*| > 1.96 approximately. Outliers ($CR > 80$) and cases with zero and negative average Equity are excluded.

1. Partition	2. Segment	3. Row and predictor	Not Transformed						Transformed				10. $log\ (avg\ E/avg\ A)\ t$	
			4. Inc. R Sq.	5. ROA *t*		6. ROE *t*		7. Note	8. $log\ ROA\ t$		9. $log\ ROE\ t$			
No partition (330 banks)	Not segmented	1 ROA	0.003	-0.822					0.615					
		2 ROE	0.031			-4.996	**	**match**			-3.192	*		
		3 ROA and ROE	0.095	3.627	*	-6.148	**		7.854	***	-8.501	***	7.335	**
By GDP *per capita* of jurisdiction (204 and 126 banks)	HK, JP, KR, SG, TW (developed)	1 ROA	0.162	-3.681	*				-1.595	.				
		2 ROE	0.029			-11.244	***				-2.306	*		
		3 ROA and ROE	0.031	3.359	**	-11.079	***		1.346	.	-2.136	*	0.838	
	CN, IN, ID, MA, PH, TH (developing)	1 ROA	0.000	0.383					0.488					
		2 ROE	0.017			-3.244	*	**match**			-3.180	*		
		3 ROA and ROE	0.181	4.055	**	-5.207	**		6.178	**	-6.983	**	6.727	**
By GDP of jurisdiction (199 and 131 banks)	HK, ID, MA, PH, SG, TH, TW (smaller)	1 ROA	0.000	-0.637					0.340					
		2 ROE	0.048			-2.921	*	**match**			-2.474	*		
		3 ROA and ROE	0.076	2.117	*	-3.561	*		5.950	**	-6.482	**	6.391	**
	CN, IN, JP, KR (larger)	1 ROA	0.004	-2.694	*				0.642					
		2 ROE	0.030			-5.292	**	**match** (1)			-0.064			
		3 ROA and ROE	0.109	1.239		-4.686	**		2.084	*	-1.983	*	1.793	.
By extreme decile of individual bank assets (34 and 33 banks)	Lowest SIZE decile (smallest banks)	1 ROA	0.000	-0.897					-0.163					
		2 ROE	0.028			-3.204	*	**match**			-2.600	*		
		3 ROA and ROE	0.042	1.361		-3.364	*		7.849	***	-8.469	***	7.530	**
	Highest SIZE decile (biggest banks)	1 ROA	0.157	-3.722	*				-2.537	*				
		2 ROE	0.273			-2.901	*				-2.267	*		
		3 ROA and ROE	0.287	-2.298	*	0.440			-1.118		0.107		-1.283	.
By Basel II adoption year in individual bank (182 and 148 banks)	Early adoption (before 2009)	1 ROA	0.026	1.558	.				0.889					
		2 ROE	0.000			-1.959	*	**match** (2)			-1.445	.		
		3 ROA and ROE	0.066	3.840	*	-4.023	*		5.227	**	-5.356	**	4.478	**
	Late adoption (2009 and after)	1 ROA	0.000	-2.503	*				0.575					
		2 ROE	0.120			-5.153	**	**match**			-0.598			
		3 ROA and ROE	0.217	1.777	.	-4.814	**		3.034	*	-3.039	*	3.627	*

Table 3: GMM models which ignore improvements, explaining the log of the Tier 1 Capital to Risk-Weighted Assets ratio (CR) for 3 partitions of the East Asian dataset, 6 models in total. The table shows abbreviations of predictors, coefficient estimates, coefficients divided by standard errors (z), significance asterisk notation ('***' is p < 0.000; '**' is p < 0.01; '*' is p < 0.05 '.' is p < 0.1). The decimal logarithm is denoted 'log' and the logarithm of CR lagged by 1 year is $\log CR_{t-1}$ The table also shows model statistics. Some constraining distortions are highlighted. Only the 254 banks covering the 2006-2014 period under Basel II regulation are included. Outliers (CR > 80) or cases with zero and negative average Equity are excluded.

Partition	By GDP *per capita* of jurisdictions						By GDP of jurisdiction						By DIM (bank size)					
Segment	Developed economies			Developing economies			Small economies			Large economies			Below median			Above median		
Predictor	Estimate	z		Estimate	z		Estimate	z		Estimate	z		Estimate	z		Estimate	z	
$log\ CR_{(t-1)}$	0.3929	3.15	**	0.3553	5.45	***	0.4612	4.29	***	0.2425	4.03	***	0.4231	6.83	***	0.3544	5.15	***
ROA	0.1148	2.66	**	-0.0922	$\boxed{-2.36}$	*	-0.0517	-1.31		0.0592	1.08		-0.0345	-0.54		-0.0202	-0.42	
ROE	-0.0029	-1.81	.	0.0062	$\boxed{2.24}$	*	0.0057	2.40	*	-0.0016	-0.79		0.0028	0.59		0.0012	0.57	
NIM	0.0051	3.69	***	0.0179	0.88		0.0008	0.65		0.0458	1.61		0.0017	1.71	.	0.0131	0.53	
XOA	-0.0576	-1.92	.	-0.0235	-1.25		-0.0367	-1.82	.	0.0225	0.50		-0.0486	-2.62	**	0.0118	0.35	
REP	-0.0949	-2.97	**	0.0082	0.32		-0.0046	-0.19		-0.0334	-0.92		0.0157	0.59		-0.0316	-1.05	
EL	0.0008	0.36		0.0032	4.59	***	0.0036	4.82	***	-0.0003	-0.31		0.0033	4.19	***	-0.0029	-1.84	.
EDS	-0.0623	-1.91	.	0.0055	4.69	***	0.0056	3.62	***	-0.0401	-0.65		0.0048	2.93	**	0.0056	0.15	
ELB	0.1033	3.13	**	-0.0054	-6.88	***	-0.0052	-6.29	***	0.0894	1.42		-0.0045	-3.87	***	0.0441	1.01	
log LLR	0.0589	1.98	*	-0.0011	-0.04		0.0397	1.63		0.0354	0.73		0.0013	0.05		0.0768	2.90	**
LLP	0.0004	1.74	.	-0.0006	-2.37	*	-0.0003	-0.84		-0.0003	-1.03		0.0011	0.08		-0.0007	-2.01	*
log NPL	0.2457	$\boxed{2.77}$	**	0.4601	$\boxed{3.56}$	***	0.3671	$\boxed{2.74}$	**	0.2721	$\boxed{2.69}$	**	0.3929	$\boxed{2.96}$	**	0.2203	$\boxed{2.08}$	*
log ILE	-0.2729	$\boxed{-3.11}$	**	-0.4963	$\boxed{-3.73}$	***	-0.3992	$\boxed{-2.88}$	**	-0.3164	$\boxed{-3.51}$	***	-0.4222	$\boxed{-3.04}$	**	-0.2591	$\boxed{-2.46}$	*
LA	-0.0112	-1.68	.	-0.0026	-0.93		-0.0022	-0.67		-0.0041	-0.42		0.0002	0.06		-0.0011	-0.14	
LDS	0.0065	1.16		0.0006	0.52		0.0011	0.73		0.0031	0.35		0.0005	0.36		-0.0024	-0.40	
DIM	0.2915	3.04	**	-0.1209	-1.72	.	-0.0583	-0.67		0.0862	0.84		-0.0084	-0.12		0.2748	3.22	**
GDP	-0.2332	-2.02	*	0.0124	0.08		0.1553	1.10		-0.2552	-2.23	*	-0.0422	-0.46		-0.3633	-3.80	***
GDG	0.0073	0.99		0.0011	0.32		0.0003	0.08		-0.0021	-0.62		0.0041	0.84		-0.0013	-0.45	
INFL	-0.0165	-1.51		0.0018	0.53		0.0001	0.01		-0.0112	-1.34		-0.0031	-0.63		-0.0017	-0.47	
No. observations	128 banks, 9 years			126 banks, 9 years			125 banks, 9 years			129 banks, 9 years			126 banks, 9 years			128 banks, 9 years		
Sargan chisq (16)	23.1			14.9			16.8			23.9			21.6			27.2 *		
Autocorrelation 1st	-2.97 **			-4.93 ***			-4.37 ***			-3.43 **			-4.69 ***			-2.60 **		
Autocorrelation 2nd	-0.867			-0.433			-0.675			-0.965			-0.661			-0.928		
Wald coeff. chisq (13)	810.4 ***			866.8 ***			632.6 ***			259.3 ***			585.5 ***			340.5 ***		
Wald dum. chisq (9)	57.4 ***			25.1 **			24.7 **			87.8 ***			24.9 **			63.0 ***		

Table 4: GMM models which apply improvements, explaining the log of the Tier 1 Capital to Risk-Weighted Assets ratio (CR) for 3 partitions of the East Asian dataset, 6 models in total. The table shows abbreviations of predictors, coefficient estimates, coefficients divided by standard errors (z), significance in asterisk notation ('***' is p < 0.000; '**' is p < 0.01; '*' is p < 0.05 '.' is p < 0.1). The decimal logarithm is denoted 'log' and the log of CR lagged by 1 year is $\log CR_{(t-1)}$. The table also shows model statistics. Only the 254 banks covering the 2006-2014 period under Basel II regulation are included. Outliers (CR > 80) or cases with zero and negative average Equity are excluded.

Partition	By GDP *per capita* of jurisdictions						By GDP (size) of jurisdiction						By DIM (bank size)					
Segment	Developed economies			Developing economies			Small economies			Large economies			Below median			Above median		
Predictor	Estimate	z		Estimate	z		Estimate	z		Estimate	z		Estimate	z		Estimate	z	
$\log CR_{(t-1)}$	-0.0274	-0.14		0.4873	5.79	***	0.6099	5.64	***	0.1181	1.82	.	0.4757	5.14	***	0.3196	3.54	***
LR	0.0935	7.66	***	0.0219	0.95		0.0203	0.95		0.1374	9.27	***	0.0258	1.15		0.1163	12.0	***
NIM	0.0237	2.82	**	0.0044	0.26		-0.0071	-0.81		-0.0036	-0.10		-0.0028	-0.28		-0.0001	-0.00	
OOI	0.0391	2.54	*	-0.0231	-1.12		-0.0142	-0.95		0.0109	0.41		-0.0068	-0.39		-0.0155	-0.85	
log LLR	-0.0241	-1.01		-0.0257	-0.80		0.0032	0.08		-0.0019	-0.09		-0.0254	-0.76		0.0073	0.33	
PRO	-0.0545	-2.43	*	-0.0186	-0.60		-0.0558	-1.36		-0.0245	-0.79		-0.0522	-1.68	.	-0.0055	-0.15	
log NPL	-0.0043	-0.31		-0.0061	-0.37		-0.0062	-0.35		-0.0015	-0.06		-0.0128	-0.75		0.0059	0.35	
LA	-0.0047	-0.98		-0.0126	-2.86	**	-0.0131	-3.52	***	0.0039	0.75		-0.0112	-3.15	**	-0.0008	-0.31	
LDS	-0.0042	-1.15		0.0015	1.01		0.0021	1.41		-0.0082	-2.14	*	0.0015	0.98		-0.0067	-2.97	**
DIM	0.1441	1.17		-0.2796	-1.81	.	-0.1983	-1.33		0.0653	0.89		-0.1473	-1.15		0.1336	2.15	*
GDP	-0.3034	-2.91	**	0.3301	1.51		0.4214	2.15	*	-0.2044	-2.11	*	0.1182	0.82		-0.2091	-3.39	***
GDG	0.0086	1.87	.	-0.0037	-0.94		-0.0042	-1.34		-0.0023	-0.85		0.0025	0.62		-0.0009	-0.47	
INFL	-0.0121	-1.14		0.0018	0.53		0.0028	0.92		-0.0033	-0.44		-0.0054	-1.33		-0.0018	-0.73	
No. observations	128 banks, 9 years			126 banks, 9 years			125 banks, 9 years			129 banks, 9 years			126 banks, 9 years			128 banks, 9 years		
Sargan chisq (16)	22.3			11.7			10.8			23.4			14.2			19.8		
Autocorrelation 1st	-0.55			-4.37 ***			-4.45 ***			-2.43 *			-3.70 ***			-4.22 ***		
Autocorrelation 2nd	-0.85			-0.21			0.27			-1.54			-0.49			0.15 ***		
Wald coeff. chisq (13)	355.6 ***			410.9 ***			366.1 ***			249.6 ***			289.1 ***			238.8 ***		
Wald dum. chisq (9)	111.2 ***			9.4			18.3 *			93.5 ***			19.8 *			55.3 ***		