Asset correlations and procyclical impact
Kung-Cheng Ho, Jiun-Lin Chen and Shih-Cheng Lee

Rating momentum in the macroeconomic stress testing and scenario analysis of credit risk
Jimmy Skoglund and Wei Chen

A model combination approach to developing robust models for credit risk stress testing: an application to a stressed economy
Georgios Papadopoulos

Point-in-time probability of default term structure models for multiperiod scenario loss projection
Bill Huajian Yang

ADDENDUM

Addendum to “A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration”
Torsten Pyttlik, Mark Rubtsov and Alexander Petrov

For all subscription queries, please call:
UK/Europe: +44 (0) 207 316 9300
USA: +1 646 736 1850   ROW: +852 3411 4828
Risk.net in numbers

200+
New articles & technical papers (each month)

140,000
Users (each month)

370,000
Page views (each month)

58,000+ articles stretching back 20 years

6,500+
on Asset Management

6,900+
on Commodities

£21,000+
on Derivatives

19,400+
on Regulation

19,600+
on Risk Management

See what you’re missing
Visit the world’s leading source of exclusive in-depth news & analysis on risk management, derivatives and complex finance now.
The Journal of Risk Model Validation
EDITORIAL BOARD

Editor-in-Chief
STEVE SATCHELL  University of Cambridge

Associate Editors

MOHAN BHATIA  Wipro Ltd
STEFAN BLOCHWITZ  Deutsche Bundesbank
J. L. BREEDEN  Prescient Models LLC
WEI CHEN  SAS Institute Inc.
GEORGE CHRISTODOUKAKIS  Manchester Business School
MARCELO CRUZ
KLAUS DUELLMANN  European Central Bank
DOUGLAS DWYER  Moody’s Analytics
CHRISTOPHER C. FINGER  Federal Reserve Board
DAVID LI  Mitsubishi Bank

CHRISTIAN MEYER  DZ BANK AG
PETER MIU  McMaster University
BOGIE OZDEMIR  Sun Life Financial Group
PETER QuELL  DZ BANK AG
DANIEL RÖSCH  University of Hannover
HARALD SCHEULE  University of Technology, Sydney
ROGER STEIN  Moody’s Investors Service
ALLAN TIMMERMANN  University of California, San Diego
BILL HUAJIAN YANG  Royal Bank of Canada

SUBSCRIPTIONS

The Journal of Risk Model Validation (Print ISSN 1753-9579 | Online ISSN 1753-9587) is published quarterly by Infopro Digital Limited, Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK. Subscriptions are available on an annual basis, and the rates are set out in the table below.

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>Europe</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk.net Journals</td>
<td>£2045</td>
<td>€2895</td>
<td>$3225</td>
</tr>
<tr>
<td>Individual Journal (print)</td>
<td>£775</td>
<td>€1075</td>
<td>$1265</td>
</tr>
<tr>
<td>Risk.net Premium</td>
<td>£2895</td>
<td>€4145</td>
<td>$4545</td>
</tr>
</tbody>
</table>

Academic discounts are available. Please enquire by using one of the contact methods below.

All prices include postage. All subscription orders, single/back issues orders, and changes of address should be sent to:

UK & Europe Office: Infopro Digital, Haymarket House, 28-29 Haymarket, London, SW1Y 4RX, UK. Tel: +44 207 316 9300

US & Canada Office: Infopro Digital, 55 Broad Street, Floor 22, New York, NY, 10005. Tel: +1 646 736 1850; Fax: +1 646-390-6612

Asia & Pacific Office: Infopro Digital, Unit 1704-05, Berkshire House, Taikoo Place, 25 Westlands Road, Quarry Bay, Hong Kong. Tel: +852 3411 4828 Fax: +852 3411 4811 Website: www.risk.net/journal E-mail: info@risk.net

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
Manuscripts and research papers submitted for consideration must be original work that is not simultaneously under review for publication in another journal or other publication outlets. All articles submitted for consideration should follow strict academic standards in both theoretical content and empirical results. Articles should be of interest to a broad audience of sophisticated practitioners and academics.

Submitted papers should follow Webster’s New Collegiate Dictionary for spelling, and The Chicago Manual of Style for punctuation and other points of style, apart from a few minor exceptions that can be found at www.risk.net/journal. Papers should be submitted electronically via email to: journals@incisivemedia.com. Please clearly indicate which journal you are submitting to.

You must submit two versions of your paper; a single \LaTeX{} version and a PDF file. \LaTeX{} files need to have an explicitly coded bibliography included. All files must be clearly named and saved by author name and date of submission. All figures and tables must be included in the main PDF document and also submitted as separate editable files and be clearly numbered.

All papers should include a title page as a separate document, and the full names, affiliations and email addresses of all authors should be included. A concise and factual abstract of between 150 and 200 words is required and it should be included in the main document. Five or six keywords should be included after the abstract. Submitted papers must also include an Acknowledgements section and a Declaration of Interest section. Authors should declare any funding for the article or conflicts of interest. Citations in the text must be written as (John 1999; Paul 2003; Peter and Paul 2000) or (John et al 1993; Peter 2000).

The number of figures and tables included in a paper should be kept to a minimum. Figures and tables must be included in the main PDF document and also submitted as separate individual editable files. Figures will appear in color online, but will be printed in black and white. Footnotes should be used sparingly. If footnotes are required then these should be included at the end of the page and should be no more than two sentences. Appendixes will be published online as supplementary material.

Before submitting a paper, authors should consult the full author guidelines at:

http://www.risk.net/static/risk-journals-submission-guidelines

Queries may also be sent to:

The Journal of Risk Model Validation, Infopro Digital,
Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK
Tel: +44 (0)20 7004 7531; Fax: +44 (0)20 7484 9758
E-mail: journals@incisivemedia.com
The Journal of Risk Model Validation

The journal

As monetary institutions rely heavily on economic and financial models for a wide array of applications, model validation has become progressively inventive within the field of risk. The Journal of Risk Model Validation focuses on the implementation and validation of risk models, and it aims to provide a greater understanding of key issues including the empirical evaluation of existing models, pitfalls in model validation and the development of new methods. We also publish papers on backtesting. Our main field of application is in credit risk modeling but we are happy to consider any issues of risk model validation for any financial asset class.

The Journal of Risk Model Validation considers submissions in the form of research papers on, but not limited to, the following topics.

- Empirical model evaluation studies.
- Backtesting studies.
- Stress-testing studies.
- New methods of model validation/backtesting/stress testing.
- Best practices in model development, deployment, production and maintenance.
- Pitfalls in model validation techniques (all types of risk, forecasting, pricing and rating).

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
## CONTENTS

**Letter from the Editor-in-Chief**

**RESEARCH PAPERS**

*Asset correlations and procyclical impact*

Kung-Cheng Ho, Jiun-Lin Chen and Shih-Cheng Lee

1

*Rating momentum in the macroeconomic stress testing and scenario analysis of credit risk*

Jimmy Skoglund and Wei Chen

21

*A model combination approach to developing robust models for credit risk stress testing: an application to a stressed economy*

Georgios Papadopoulos

49

*Point-in-time probability of default term structure models for multiperiod scenario loss projection*

Bill Huajian Yang

73

**ADDENDUM**

*Addendum to “A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration”*

Torsten Pyttlik, Mark Rubtsov and Alexander Petrov

95

---

Editor-in-Chief: Steve Satchell  
Subscription Sales Manager: Aaraa Javed  
Publisher: Nick Carver  
Global Key Account Sales Director: Michelle Godwin  
Journals Manager: Dawn Hunter  
Composition and copyediting: T&T Productions Ltd  
Printed in UK by Printondemand-Worldwide  

©Copyright Infopro Digital Limited, 2017. All rights reserved. No parts of this publication may be reproduced, stored in or introduced into any retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the copyright owners.

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
LETTER FROM THE EDITOR-IN-CHIEF

Steve Satchell
Trinity College, University of Cambridge

This issue contains the usual four papers along with an addendum to a paper that was previously published in *The Journal of Risk Model Validation*. This addendum has been submitted by the previous paper’s two authors with the help of a third. As an editor, I very much like receiving addenda and general correspondence about what we have published in the past. Apart from reassuring me that people actually read our papers, it is a sign of good scholarship and creates a community around the journal.

The issue’s first paper, “Asset correlations and procyclical impact” by Kung-Cheng Ho, Jiun-Lin Chen and Shih-Cheng Lee, is an examination of the behavior of asset correlations for companies in Taiwan under Basel’s asymptotic single-risk-factor (ASRF) approach. Using Merton’s model to estimate firm default probability from 1990 to 2013, the authors find that assets are positively correlated with firm size and negatively correlated with default probability. They also confirm the result that asset correlations are asymmetric and have a procyclical impact on the real economy after these effects are controlled for. This suggests that constant correlation modeling can be improved on. The challenge is to find a robust specification that allows for easy estimation and captures this phenomenon.

Jimmy Skoglund and Wei Chen are the authors of our second paper: “Rating momentum in the macroeconomic stress testing and scenario analysis of credit risk”. This looks at the addition of rating momentum to the popular factor model of credit risk. The conceptual feature that causes this augmented structure to differ from the usual setup is its non-Markovian nature. The authors find that models that take the stylized fact of rating momentum into account can accelerate the loss timing significantly when compared with a non-rating-momentum case. The exact effect depends on the scenario time horizon, the severity and the portfolio quality. In general, it takes more time for differences to be realized in good-quality portfolios; the effect on lower-quality portfolios, however, can be almost immediate, with significant loss underestimation. This new structure will also have a different explained versus idiosyncratic decomposition, with changes in exposure to factor risk as a result. This is an issue that regulators need to take into account.

“A model combination approach to developing robust models for credit risk stress testing: an application to a stressed economy” is the third paper in this issue of *The Journal of Risk Model Validation*. Here, Georgios Papadopoulos investigates the impact of macroeconomic shocks on bank-specific risk factors. The author’s concern is model frailty, and his proposed solution is to use a model combination approach to develop robust macrofinancial models for credit risk stress testing. The empirical part of his research utilizes data from the Greek economy, which experienced a sharp...
change from normal to distressed conditions; this makes it a particularly challenging case to forecast. The paper constitutes an interesting approach to this topic, and the highly visible problems of the Greek banking sector give it contemporary relevance.

Our fourth paper, “Point-in-time probability of default term structure models for multiperiod scenario loss projection” by Bill Huajian Yang, involves an important technical problem, which will be easier for me to quote than paraphrase:

Rating transition models have been widely used for multiperiod scenario loss projection for Comprehensive Capital Analysis and Review (CCAR) stress testing and International Financial Reporting Standard 9 (IFRS 9) expected credit loss estimation. Although the cumulative probability of default (PD) for a rating can be derived by repeatedly applying the migration matrix at each forward scenario sequentially, divergence between the predicted and realized cumulative default rates can be significant, particularly when the predicting horizon is extended.

The author offers a solution to this problem by directly modeling the forward PDs, validating them using a corporate portfolio. He reports greatly improved results relative to more conventional methods.

The issue concludes with the aforementioned addendum by Torsten Pyttlik, Mark Rubtsov and Alexander Petrov, which refers back to a 2016 paper titled “A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration” by Mark Rubtsov and Alexander Petrov (The Journal of Risk Model Validation 10(2), 83–112). In the addendum the authors offer an analytical solution to a problem that two of the authors had previously solved numerically.
Research Paper

Asset correlations and procyclical impact

Kung-Cheng Ho,¹ Jiun-Lin Chen² and Shih-Cheng Lee³

¹School of Finance, Zhongnan University of Economics and Law, 182# Nanhu Avenue, East Lake High-tech Development Zone, Wuhan 430073, People’s Republic of China; email: kcho731101@163.com
²School of Accounting and Finance, University of Adelaide, 10 Pulteney Street, Adelaide 5000, Australia; email: alex.chen@adelaide.edu.au
³College of Management, Innovation Center for Big Data and Digital Convergence, Yuan Ze University, No. 135, Yuandong Road, Zhongli District, Taoyuan City 320, Taiwan; email: sclee@saturn.yzu.edu.tw

(Received April 25, 2016; revised July 5, 2016; accepted August 18, 2016)

ABSTRACT

We examine the behavior of asset correlations for companies in Taiwan under the Basel Accord’s asymptotic single-risk-factor approach. Using Merton’s model to estimate a firm’s probability of default (PD) from 1990 to 2013, we find that asset correlations are positively correlated with firm size and negatively correlated with PD. We also find an industry effect, where, on average, firms in basic materials and technology industries have higher asset correlations. Most importantly, we confirm that asset correlations are asymmetric and have a procyclical impact on the real economy after these effects are controlled. They tend to decline during economic upturns and rise during economic downturns. In addition, the average decrease is smaller than the increase. Our findings provide important policy implications for future regulatory frameworks, as asset correlations may be underestimated during economic downturns.

Keywords: systemic risk; Basel Accord; asymptotic single risk factor; asset correlations; procyclical impact.
1 INTRODUCTION

After a series of revisions, the Basel Committee on Banking Supervision (BCBS) issued the Basel II Accord in 2006 to determine a bank’s capital requirement based on its asset quality (Basel Committee on Banking Supervision 2006), with the aim of effectively supervising the banking system and improving the risk management of banks. Under Basel II, a bank can choose the internal ratings-based (IRB) approach, under the asymptotic single-risk-factor (ASRF) model, to measure its minimum regulatory capital. Asset correlation, which measures the comovement between a firm’s asset returns and a common risk factor that summarizes general economic conditions, is a key parameter in the ASRF model. A higher asset correlation indicates a higher systematic risk and requires a bank to hold more capital (Basel Committee on Banking Supervision 2006). Hence, this approach provides a more risk-sensitive measure than the previous Basel I standard from 1988.

Banks’ greatest concern under Basel II is how much capital is required, as a higher capital requirement means less available funding for corporate loans. Meanwhile, a firm’s probability of default (PD) may also increase if it cannot get enough funding. The initial asset correlation prescribed by BCBS (2001a) was a constant value of 0.2 for all obligors. This was later modified to be between 0.1 and 0.2 depending on a firm’s PD (Basel Committee on Banking Supervision 2001b). The range was subsequently further adjusted to 0.08–0.24, depending on firm type and firm size (Basel Committee on Banking Supervision 2006). These changes indicate that the BCBS was aware of some potential problems in asset correlation and made adjustments correspondingly.

This paper explores the trade-off between risk sensitivity, asymmetric behavior and the impact of procyclicality on capital requirements in Taiwan under the Basel II framework. Previous Basel-related research in Taiwan mainly focuses on how the three pillars of Basel II (minimum capital requirements, supervisory review and market discipline) affect a bank’s risk taking, or discusses whether a particular variable will affect the relations between the three pillars and risk. Rarely has research studied the procyclical impact of asset correlations under the ASRF framework. Due to the moral hazard problems associated with credit risk, most studies suggest that the financial industry is cyclical. In general, small companies are usually affected by the nonsystematic risk, while large companies are more influenced by the business cycle. Although increasing the sensitivity of the capital requirement to credit risk mitigates moral hazard problems by aligning regulatory capital more closely with banks’ asset risks, it inadvertently increases the procyclicality of banks’ capital requirement. Procyclicality arises because banks tend to reduce their lending activities during economic downturns, when credit quality deteriorates. Meanwhile, a higher regulatory capital requirement due to higher credit risks may squeeze credits further, and the reverse applies during upturns. Therefore, a risk-sensitive capital requirement may
amplify each phase of the business cycle as banks respond by easing or tightening credits (Marcucci and Quagliariello 2009).

There is some evidence for the impact of Basel II on the procyclical behaviors of regulatory capital. For example, Lowe and Segoviano (2002) find large swings in capital requirement for Mexican private banks after the introduction of the IRB approach. Similarly, Kashyap and Stein (2004) estimate the degree of cyclicity in capital charges and find that capital charges may increase by 70–90% and 30–45% based on KMV and Standard & Poor’s ratings, respectively. Moreover, Heid’s (2007) model demonstrates that the way the minimum capital requirement is changed could cause a bubble, inflation and crisis. Altman et al. (2005) and Bruche and González-Aguado (2010) also find evidence of cyclicality under the capital requirement, which exacerbates the fluctuation of business cycles. In addition, some literature on the default rate (Crouhy et al. 2001), capital buffer (Ayuso et al. 2004), credit risk (Marcucci and Quagliariello 2009) and equity correlation (Longin and Solnik 1995, 2001) also documents the asymmetric and procyclical phenomena.

The economy of Taiwan is greatly affected by the US economy. Lee et al. (2011) study the asymmetric behavior and procyclical impact of asset correlations in US assets for 1988–2007 and conclude the following. First, even controlling for PD and firm size, the asset correlation still inversely correlates with real gross domestic product (GDP) growth rate. When real GDP growth rate reduces by 1%, asset correlations will increase by 12%. However, BCBS (2004) only imposes a range of 0.12–0.24 for asset correlations, without illustrating the detail of calculating the asset correlation. When we consider the asymmetric and procyclical effects, asset correlations will depart from this range. Second, asset correlations are negatively associated with PD but positively correlated with firm size. This finding is consistent with Lopez (2004) and the specification of Basel II but contrary to Dietsch and Petey (2004), who find an industry effect even after controlling for a company’s size effect, where “material” and “information industries” have higher asset correlations, while “health care” and “oil industries” have lower values. Also, industries with higher asset correlations have greater growth opportunities and are more subject to the overall economic impact.

However, Taiwan and the United States still have different industry structures, and there are fewer industries and companies in Taiwan. Our paper contributes to the literature by empirically examining the effect of PD, firm size and industry on asset correlations for Taiwanese firms. We also identify whether the asymmetric behavior and procyclical impact of asset correlations occur in Taiwan. Based on a sample over the period 1990–2009, we show that asset correlations are negatively associated with PD and positively associated with firm size. Also, the industry effect plays a distinct role: firms in “basic materials” and “technology” industries exhibit higher asset correlations, while firms in “health care” and “oil and gas” industries have lower asset correlations. Of all the industries, firms in the “financial” sector have the highest
Most importantly, our analysis confirms asymmetric and cyclical behavior of asset correlations in Taiwan that the current banking supervision authority may not have considered. After controlling for firm size, PD, industry and yearly fixed effects, we observe that asset correlations appear to increase in recessions and decrease in expansions. Also, the increase in asset correlations during economic downturns is more pronounced than the decrease during upturns, which means regulatory capital may be underestimated during economic downturns. Our study, therefore, has important policy implications in terms of increasing the average asset correlations through the business cycle to incorporate the variations in asset correlations.

The remainder of this paper is organized as follows. Section 2 provides an overview of Basel II and gives key parameters for asset correlations. The methodology of estimating asset correlations is illustrated in Section 3. Sections 4 and 5 present the descriptive statistics of our data and empirical results. Section 6 concludes the paper.

2 THE BACKGROUND TO THE BASEL ACCORD AND THE DETERMINANTS OF ASSET CORRELATIONS

2.1 Asset correlations under Basel II

Since the introduction of the Basel Accords, financial institutions in Taiwan have paid increasing attention to credit risk management and capital buffers. In the past, banks mainly assessed the credit risk of individual borrowers and considered a portfolio’s overall credit risk to be the weighted average of any individual borrower’s risk. However, individual borrowers in a credit portfolio actually interact with each other. When a huge crisis occurs, the overall credit risk of the portfolio should be greater than the sum of the individual borrowers’ risks, suggesting the overall risk is potentially underestimated (Hung et al 2006). In order to avoid this underestimation, the capital requirement formula in Basel II has introduced new methods to help assess a portfolio’s credit risk. In addition to the standardized approach, the BCBS introduced the foundation internal ratings-based (FIRB) approach and the advanced internal ratings-based (AIRB) approach. The overall IRB approach is based on the ASRF model developed by Gordy (2003). Asset correlation, which measures the correlation between an obligor’s asset returns and a common risk factor that summarizes
Asset correlations and procyclical impact

general economic conditions, becomes a key parameter in determining banks’ capital requirements.

To be more specific, capital requirements should properly reflect material systematic risk, \(X\), in credit losses over time and guarantee sufficient capital to cover losses, especially in adverse circumstances \(X = \chi_{99.9}\). Combining bank-reported risk parameters with a simple supervisory model determines the IRB risk weights that approximate the conditional expected loss (CEL) calculation for each exposure in a bank’s portfolio. The main method of estimating conditional default probability (CDP) in Basel II is the ASRF model developed by Gordy (2003). Accordingly, correlations between realized losses across exposures are driven by a single systematic risk factor that captures the effects of unexpected changes in economic conditions. Hence, the loss rate for a well-diversified credit portfolio depends only on the systematic risk factor rather than the idiosyncratic risk factors associated with individual exposures.

Although asset correlation is essentially the correlation between an asset return and the market return, calculating each individual asset correlation is not easy. Subsequently, Basel II provides various weights for industries to calculate asset correlations. For example, for corporate, sovereign and bank exposures (Basel Committee on Banking Supervision 2006), asset correlation can be estimated using the following equation:

\[
\rho = 0.12 \left( \frac{1 - e^{-50PD}}{1 - e^{-50}} \right) + 0.24 \left( 1 - \frac{1 - e^{-50PD}}{1 - e^{-50}} \right) = 0.24 - 0.12 \left( \frac{1 - e^{-50PD}}{1 - e^{-50}} \right),
\]

where \(PD\) is the probability of default. Equation (2.1) implies asset correlations are negatively related to \(PD\), which is consistent with the finding of BCBS (2004). When the credit risk of a firm increases, firm-specific risk factors usually become more important than systematic risk. Hence, the correlation of \(PD\) with the systematic risk factor will be negative.

\[2.2\] The effect of PD, firm size and industry on asset correlations

According to (2.1), asset correlations are negatively correlated with \(PD\). Lopez (2004) provides evidence that average asset correlations are a decreasing function of \(PD\). In general, firms with higher \(PD\) are likely to exhibit more idiosyncratic risk than systematic risk. Since asset correlations measure how asset values are subject to the influence of systematic risk (macroeconomic factors), asset correlations are therefore lower for firms with higher \(PD\). Meyer (2009), Lee et al (2011) and Yang (2013, 2014) report similar findings on the effect of \(PD\) on asset correlations.

However, De Servigny and Renault (2002) find that non-investment-grade firms have higher asset correlations than investment grade firms. Nickell et al (2000) report that the volatility of low-graded obligors increases sharply during a downturn in the
business cycle, implying a higher correlation with the systematic risk factor. Studying small and medium-sized enterprises (SMEs) in Germany and France, Dietsch and Petey (2004) find that the PD of larger firms due to nonsystematic risk factors is quite low. Meanwhile, asset correlations of SMEs are positively related to the PD. Therefore, they suggest that SMEs have a different risk-weight function, as their estimated correlations for SMEs are significantly lower than those specified by the Basel Committee.

Regarding the size effect, Dietsch and Petey (2004) find that large firms are more diversified and more subject to common risk factors than to firm-specific factors. Lopez (2004, 2009), Grunert and Weber (2009) and Lee et al (2011) find a similar positive association between asset correlations and firm size. Meanwhile, Chan et al (2007) suggest that the size and industry effects are different factors than asset correlations. They find SMEs in general have a higher PD than larger companies. However, in certain industries, such as manufacturing, automotive and construction, the asset correlations of SMEs are more vulnerable to systemic risk factors.

### 2.3 The asymmetry and cyclicality of credit risk factors

The PD of a company is affected by changes in macroeconomic conditions. Fama (1986), Das et al (2006), Duffie et al (2007) and Koopman et al (2009) report that the PD of companies increases during recession. Using bond information, Altman et al (2005) find that default rates were higher than 10% in the recession of 1990–1, compared with less than 5% in the boom of 1993–8, and increased again during the economic downturn from 2000 to 2001. Similarly, loss given default (LGD) swings substantially over business cycles and tends to rise during economic downturns and fall during upturns. Altman et al (2005) and Caselli et al (2008), after controlling for sensitivities to macroeconomic factors, find the LGD rate for SME loans still increases in a poor economy. In addition, Crouhy et al (2000, 2001) show that the increases in PD are larger in recessionary periods than the decreases in expansionary periods, suggesting the cyclicality of expected loss rates could be asymmetric. These results indicate that using banks’ internal rating methods to estimate LGD parameters during economic downturns may underestimate the overall risk faced by the banks (Basel Committee on Banking Supervision 2004).

In addition, the economy’s cyclical fluctuations can also affect stock returns (Longin and Solnik 1995, 2001; Ang and Chen 2002; Huang et al 2009). Many studies show that the financial industry is highly dependent on economic cyclicality. In economic downturns, not only will banks reduce lending activities when credit quality deteriorates, but a higher capital requirement due to higher credit risk may squeeze credits even further. Since a firm’s PD increases due to the poor economy, squeezed
credit will further lead to a higher default risk, making the economy deteriorate even more.

3 DEFAULT PROBABILITY AND ASSET CORRELATIONS

3.1 Probability of default

Traditionally, most researchers have used historical data to estimate when the next default may occur and estimate the PD. In this paper, we use Merton’s (1974) method to estimate the PD, as this approach does not require a large amount of historical data. Instead, it just needs some accounting information from the financial statements, and can reflect the company’s current condition better. Our method is the same as in Lee et al (2011), based on the structural model of Merton (1974) that treats the value of equity as a European call option on a firm’s underlying assets and the value of debt as the option’s premium. When a firm’s asset is less than its debt, its shareholders will give up the option and the firm will default. In this framework, the value of a firm’s total assets is assumed to follow a geometric Brownian motion:

\[ \frac{dV}{V} = \mu_A \, dt + \sigma_A \, dZ, \]  

(3.1)

where \( V \) is the value of a firm’s assets, \( \mu_A \) is the instantaneous expected return on the firm’s assets, \( \sigma_A \) is the standard deviation of a firm’s asset returns and \( Z \) is the Brownian motion.

The market value of a firm’s initial equity \( E \) can be solved as follows:

\[ E = VN(d_1) - e^{-rT} DN(d_2), \]  

(3.2)

where \( D \) is the level of debt at maturity \( T \),

\[ d_1 = \left( \frac{1}{\sigma_A \sqrt{T}} \right) \left( \ln \left( \frac{V}{D} \right) + \left( r + \frac{\sigma_A^2}{2} \right) T \right), \quad d_2 = d_1 - \sigma_A \sqrt{T}, \]

\[ r \] is the risk-free rate and \( N(\cdot) \) is the cumulative standard normal distribution. Applying Itô’s lemma,\(^1\) we get

\[ \sigma_E = \frac{V}{E} N(d_1) \sigma_A. \]  

(3.3)

Using (3.2) and (3.3) together, we can solve for the market value of total assets \( V \) and its standard deviation \( \sigma_A \). In (3.2), \( N(-d_2) \) is the PD of a firm under the risk-neutral assumption. To relax this assumption, we can replace the risk-free rate with the instantaneous expected return on firm’s assets (\( \mu_A \)) before calculating the default probabilities under an objective probability measure.

\(^1\) See Lee et al (2011, Appendix A) for a more detailed discussion of the model.
3.2 Asset correlations

Early studies apply Moody’s KMV model (Lopez 2004, 2009) or CreditMetrics (Finger 1999; Kim 1999) to estimate asset correlations. We use a simpler method to estimate individual asset correlation by first estimating asset beta as a function of equity beta. Based on Merton’s (1973) continuous-time version of the capital asset pricing model and the Black and Scholes (1973) option pricing model, Copeland et al (2005) and Lee et al (2011) show that the asset beta can be expressed as

\[
\rho = \left( \frac{\beta_E}{N(d_1)} \frac{E}{V} \frac{\sigma_M}{\sigma_A} \right)^2.
\]  

Equation (3.4) provides a convenient way of estimating asset correlations for the ASRF framework, as the equity beta, \(N(d_1)\), equity-to-value ratio and volatilities of market returns and asset returns can easily be obtained.

To compute the implied asset correlation for each firm, we require the following variables: the market value of the firm’s equity, the market value of the firm’s assets, the equity beta, the asset beta, the standard deviation of the market returns, the standard deviation of a firm’s asset returns, the payment due one period at a time and the monthly yield of government bonds as a proxy for the risk-free rate. Consistent with the literature, including Lee et al (2011) and Lee and Lin (2012), to estimate asset correlations, we measure the beta of a firm’s equity (\(\beta_E\)) by the covariance between asset returns and market returns divided by the standard deviation of the market returns in the sample with at least 100 daily data points per year. The annual measures of asset volatility (\(\sigma_A\)) and market volatility (\(\sigma_M\)) are then constructed as the daily variance measure \(\times \sqrt{\text{number of business days}}\). To measure the asset’s market value and asset volatility, we first compute the initial market value of a firm’s asset as the market value of equity plus the total debt for each firm, and use the Merton model to solve the simultaneous equations (3.2) and (3.3) with SAS software. In (3.2), it is important to note that \(N(-d_2)\) measures a firm’s PD. Finally, we calculate the asset correlations in (3.4). Consequently, we can use both yearly PD and asset correlation (AC) for the main regression analysis.

4 DATA

4.1 Sample selection and variable settings

Our main data source is the Datastream database, although some information, including the business cycle indicator, is from the website of Taiwan’s National Development Council (www.ndc.gov.tw) and the Taiwan Economic Journal’s database (www.finasia.biz). The sample firms include most listed companies in Taiwan from 1990 to 2013. We separate financial firms, as they are highly regulated and their
financial assets and liabilities are different from those of nonfinancial firms. Under Basel III, asset correlations for financial institutions are 25% higher than for non-financial firms. To illustrate the difference between the financial industry and other industries, we calculate the asset correlations for financial firms separately. To have sufficient observations to measure the standard deviation of equity returns and equity beta, firms with fewer than 100 observations of daily returns per year are excluded. In order to improve the accuracy of the calculation of the standard deviation of asset returns, we also remove firms without five consecutive years of asset values. In addition, companies without liabilities do not qualify for our sample, as we want to estimate the PD. Our final sample includes 10,768 firm–year observations over twenty years. We follow the Global Industry Classification Code (GICS) and the industry classification codes in Datastream to split our sample firms into ten sector categories: basic materials; consumer goods; consumer services; financial; health care; industrials; oil and gas; technology; telecommunications; and utilities. We use the monthly yield of government bonds as a proxy for the risk-free rate.

4.2 Summary statistics

The summary statistics of our data are listed in the first part of Table 1. The PD reached its peak during 2000–2, with values greater than 3%, while the average risk-neutral PD is about 0.87% for the whole sample period. The average value of asset correlations is about 32%, with a downward trend from 60% in 1990 to 12% in 2013. The second part of Table 1 presents the summary statistics of all the sample firms from 1990 to 2013; the calibrated asset correlations vary widely, with a mean value of 0.27 and a standard deviation of 0.19. Although the median value is a little lower than the mean value, both are within the specified range of regulatory asset correlations. Other variables also show a similar relationship between their mean and median values.

The summary statistics suggest PD could be cyclical. Specifically, PD tends to increase (decrease) during economic downturns (upturns) and shows a negative association with asset correlations. While the range of asset correlations is between 0.12 and 0.24 under Basel II, our estimates using the method of Lee et al (2011) for Taiwanese firms give between 0.12 and 0.60 (see Figure 1). The structural difference between Taiwan and the US industries may cause this disparity.

5 EMPIRICAL ANALYSIS

5.1 Default probability and firm size

For the period 1990–2013, we use the Merton method to estimate a firm’s PD and follow Lee et al (2011) to calculate the asset correlations. We then run the following
<table>
<thead>
<tr>
<th>Year</th>
<th>Volatility of equity returns</th>
<th>Volatility of asset returns</th>
<th>Risk-free rate</th>
<th>Equity beta</th>
<th>Risk-neutral PD (%)</th>
<th>Physical PD (%)</th>
<th>AC</th>
<th>GGDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>0.75</td>
<td>0.69</td>
<td>0.15</td>
<td>0.93</td>
<td>0.13</td>
<td>0.15</td>
<td>0.60</td>
<td>10.70</td>
</tr>
<tr>
<td>1991</td>
<td>0.48</td>
<td>0.35</td>
<td>0.08</td>
<td>0.91</td>
<td>0.18</td>
<td>0.07</td>
<td>0.59</td>
<td>11.90</td>
</tr>
<tr>
<td>1992</td>
<td>0.32</td>
<td>0.24</td>
<td>0.07</td>
<td>0.94</td>
<td>0.01</td>
<td>0.00</td>
<td>0.47</td>
<td>11.60</td>
</tr>
<tr>
<td>1993</td>
<td>0.39</td>
<td>0.26</td>
<td>0.08</td>
<td>0.75</td>
<td>0.16</td>
<td>0.05</td>
<td>0.37</td>
<td>10.40</td>
</tr>
<tr>
<td>1994</td>
<td>0.36</td>
<td>0.27</td>
<td>0.07</td>
<td>0.65</td>
<td>0.03</td>
<td>0.07</td>
<td>0.32</td>
<td>9.40</td>
</tr>
<tr>
<td>1995</td>
<td>0.39</td>
<td>0.29</td>
<td>0.07</td>
<td>1.19</td>
<td>0.05</td>
<td>0.04</td>
<td>0.50</td>
<td>8.90</td>
</tr>
<tr>
<td>1996</td>
<td>0.28</td>
<td>0.19</td>
<td>0.06</td>
<td>0.40</td>
<td>0.02</td>
<td>0.04</td>
<td>0.17</td>
<td>8.60</td>
</tr>
<tr>
<td>1997</td>
<td>0.46</td>
<td>0.31</td>
<td>0.06</td>
<td>0.78</td>
<td>0.34</td>
<td>0.28</td>
<td>0.24</td>
<td>8.50</td>
</tr>
<tr>
<td>1998</td>
<td>0.49</td>
<td>0.34</td>
<td>0.08</td>
<td>0.98</td>
<td>0.33</td>
<td>1.73</td>
<td>0.31</td>
<td>7.30</td>
</tr>
<tr>
<td>1999</td>
<td>0.49</td>
<td>0.30</td>
<td>0.05</td>
<td>0.70</td>
<td>0.84</td>
<td>1.82</td>
<td>0.20</td>
<td>4.80</td>
</tr>
<tr>
<td>2000</td>
<td>0.65</td>
<td>0.39</td>
<td>0.05</td>
<td>0.92</td>
<td>3.07</td>
<td>3.52</td>
<td>0.30</td>
<td>5.60</td>
</tr>
<tr>
<td>2001</td>
<td>0.58</td>
<td>0.29</td>
<td>0.05</td>
<td>0.78</td>
<td>3.26</td>
<td>5.39</td>
<td>0.20</td>
<td>-2.50</td>
</tr>
<tr>
<td>2002</td>
<td>0.55</td>
<td>0.28</td>
<td>0.03</td>
<td>0.91</td>
<td>2.54</td>
<td>4.01</td>
<td>0.27</td>
<td>4.80</td>
</tr>
<tr>
<td>2003</td>
<td>0.44</td>
<td>0.21</td>
<td>0.01</td>
<td>0.89</td>
<td>1.39</td>
<td>2.43</td>
<td>0.24</td>
<td>2.70</td>
</tr>
<tr>
<td>2004</td>
<td>0.44</td>
<td>0.23</td>
<td>0.01</td>
<td>0.96</td>
<td>1.24</td>
<td>2.11</td>
<td>0.33</td>
<td>6.30</td>
</tr>
<tr>
<td>2005</td>
<td>0.36</td>
<td>0.18</td>
<td>0.01</td>
<td>0.97</td>
<td>0.61</td>
<td>1.17</td>
<td>0.19</td>
<td>3.30</td>
</tr>
<tr>
<td>2006</td>
<td>0.41</td>
<td>0.20</td>
<td>0.02</td>
<td>1.07</td>
<td>1.02</td>
<td>2.38</td>
<td>0.26</td>
<td>4.30</td>
</tr>
<tr>
<td>2007</td>
<td>0.46</td>
<td>0.26</td>
<td>0.02</td>
<td>0.85</td>
<td>0.91</td>
<td>1.10</td>
<td>0.21</td>
<td>5.40</td>
</tr>
<tr>
<td>2008</td>
<td>0.58</td>
<td>0.34</td>
<td>0.03</td>
<td>1.17</td>
<td>2.15</td>
<td>3.35</td>
<td>0.48</td>
<td>-2.20</td>
</tr>
<tr>
<td>2009</td>
<td>0.46</td>
<td>0.21</td>
<td>0.01</td>
<td>0.87</td>
<td>1.48</td>
<td>1.93</td>
<td>0.28</td>
<td>-1.10</td>
</tr>
<tr>
<td>2010</td>
<td>0.38</td>
<td>0.22</td>
<td>0.01</td>
<td>1.12</td>
<td>0.42</td>
<td>0.43</td>
<td>0.29</td>
<td>8.60</td>
</tr>
<tr>
<td>2011</td>
<td>0.38</td>
<td>0.23</td>
<td>0.01</td>
<td>1.03</td>
<td>0.27</td>
<td>0.55</td>
<td>0.39</td>
<td>1.20</td>
</tr>
<tr>
<td>2012</td>
<td>0.32</td>
<td>0.17</td>
<td>0.01</td>
<td>0.94</td>
<td>0.28</td>
<td>0.97</td>
<td>0.27</td>
<td>2.70</td>
</tr>
<tr>
<td>2013</td>
<td>0.27</td>
<td>0.15</td>
<td>0.01</td>
<td>0.49</td>
<td>0.17</td>
<td>0.34</td>
<td>0.12</td>
<td>3.50</td>
</tr>
<tr>
<td>Average</td>
<td>0.4463</td>
<td>0.2765</td>
<td>0.04</td>
<td>0.8830</td>
<td>0.87</td>
<td>1.41</td>
<td>0.32</td>
<td>5.6125</td>
</tr>
</tbody>
</table>
### TABLE 1  Continued.

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volatility of equity returns</td>
<td>0.44</td>
<td>0.18</td>
<td>0.31</td>
<td>0.42</td>
<td>0.55</td>
</tr>
<tr>
<td>Volatility of asset returns</td>
<td>0.24</td>
<td>0.14</td>
<td>0.14</td>
<td>0.22</td>
<td>0.32</td>
</tr>
<tr>
<td>Risk-free rate</td>
<td>0.02</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Equity beta</td>
<td>0.91</td>
<td>0.48</td>
<td>0.58</td>
<td>0.90</td>
<td>1.23</td>
</tr>
<tr>
<td>Risk-neutral PD</td>
<td>0.01</td>
<td>0.03</td>
<td>6E−08</td>
<td>1E−04</td>
<td>0.01</td>
</tr>
<tr>
<td>Physical PD</td>
<td>0.02</td>
<td>0.07</td>
<td>2E−08</td>
<td>8E−05</td>
<td>0.01</td>
</tr>
<tr>
<td>AC</td>
<td>0.27</td>
<td>0.19</td>
<td>0.12</td>
<td>0.25</td>
<td>0.40</td>
</tr>
<tr>
<td>GGDP (%)</td>
<td>3.60</td>
<td>3.37</td>
<td>2.70</td>
<td>3.50</td>
<td>5.60</td>
</tr>
</tbody>
</table>

The average volatility of equity returns, volatility of asset returns, risk-free rate, equity beta, PD and asset correlation (AC). GGDP denotes the growth rate of real GDP. SD denotes standard deviation. N is the number of observations of sample firms for each year from 1990 to 2013.

#### FIGURE 1  Trend of asset correlations.

---

Regression analysis to examine the effects of firm size and PD on the asset correlations:

\[
AC_{i,t} = \alpha + b_1 PD_{i,t} + b_2 \ln sales_{i,t} + \sum_{j=1}^{10} \gamma_j DI_j + \sum_{y=1}^{24} \delta_y DY_y + \varepsilon_{i,t}. \tag{5.1}
\]

Here \( AC_{i,t} \) is the asset correlation for firm \( i \) at time \( t \), \( PD_{i,t} \) is the probability of default, \( \ln sales_{i,t} \) is the natural logarithm of total annual sales proxy for firm size,\(^2\)

\(^2\) Sales are used to proxy for firm size as we follow the new Basel Capital Accord document issued by the Bank for International Settlements (BIS), which classifies firm size in terms of annual total sales. For robustness checks, we also use equity value as a proxy for firm size. The results from equity values are similar to those reported in our paper.
DY$_y$ is the dummy variable for year $y$, DI$_j$ is the dummy variable for industry $j$ and $\epsilon_{i,t}$ is the error term. The dummy variable is incorporated to control for both year and industry fixed effects. Figure 2 reports the average asset correlations of each industry group over the sample period. Apparently, firms in “basic materials” and “technology” industries on average have higher asset correlations than those in “utilities” and “oil and gas”. Meanwhile, financial sector firms have the highest asset correlations of all the industries, which is in line with Basel III’s specification that asset correlations for financial institutions are 25% higher than for nonfinancial firms.

For a robustness check, we also replace PD with objective PD (without the risk-neutral assumption in Merton’s model) in (5.1). As asset correlations are bounded between 0 and 1, we use the fraction logit regression to measure the dependent variable taking values between 0 and 1 (Bellotti and Crook 2012; Pericoli et al 2013). All of these adjustments make our regression analysis more robust.

As shown in Table 2, asset correlation is inversely proportional to PD, regardless of whether the dependent variable is asset correlation in ordinary least squares (OLS) regression (models 1–4) or is fractional logit regression (models 5–8). When substituting risk-neutral PD with physical PD in the regression, we still find a significantly negative relationship, although the effect seems smaller. Adding size (ln sales) does not change the effect of PD on the asset correlation even when we use different measures of asset correlation and PD. Our results are therefore consistent with Lopez (2004) but contrary to Nickell et al (2000), De Servigny and Renault (2002) and
TABLE 2 Regression results of PD, firm size, industry and yearly fixed effect.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>OLS regression</th>
<th>Fractional logit regression APE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td>Constant</td>
<td>0.017</td>
<td>0.021</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(1.27)</td>
</tr>
<tr>
<td>Risk-neutral PD</td>
<td>−1.137***</td>
<td>−1.107***</td>
</tr>
<tr>
<td></td>
<td>(−21.77)</td>
<td>(−21.48)</td>
</tr>
<tr>
<td>Physical PD</td>
<td>−0.435***</td>
<td>−0.425***</td>
</tr>
<tr>
<td></td>
<td>(−19.34)</td>
<td>(−19.15)</td>
</tr>
<tr>
<td>Size</td>
<td>8.424***</td>
<td>8.522***</td>
</tr>
<tr>
<td></td>
<td>(17.29)</td>
<td>(17.42)</td>
</tr>
<tr>
<td>DY</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17.29)</td>
</tr>
<tr>
<td>DI</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(17.29)</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td>$N$</td>
<td>10 768</td>
<td>10 768</td>
</tr>
</tbody>
</table>

The regression results of asset correlation on PD, firm size, industry groups and yearly dummy variables. AC denotes asset correlation. Size is measured by ln sales. DY and DI are dummy variables for year and industry, respectively. $t$-statistics are reported in parentheses for a robustness check for the fractional logit regression average partial effect (APE). *** denotes statistical significance at the 1% level.
Dietsch and Petey (2004), who claim that the negative relationship is driven by the size effect.

We also confirm larger (smaller) firms are related to higher (lower) asset correlations in Table 2. Since asset correlations are estimated based on (3.4), which requires equity beta as an input, the size effect found in our regression suggests that the systematic risk for larger firms is higher (i.e., higher beta) than that for smaller firms (lower beta). Such findings show that the size effect is distinct from the effect of PD, and support the BCBS specifications on the effect of PD and firm size.

5.2 The procyclical influence of asset correlations

Lee et al (2011) show asset correlations of US firms are procyclical and tend to increase (decrease) in economic downturn (upturn). Our final analysis is to examine whether a similar effect exists in Taiwan after controlling for the above effects. We first include the growth rate of real GDP (GGDP) as the measure for macroeconomic condition to capture the potential procyclical behavior of asset correlation. Then, we use an interaction term of GGDP and a dummy variable for recessions (determined by the business cycle indicator) to explore the asymmetric effect. Table 3 reports the detail of business cycle indicator. The red light (45–38 points) indicates an economy boom, while the blue light (16–9 points) means a recession. As large firms are more diversified, they tend to be more sensitive to common risk factors than to firm-specific factors. When the economy is in recession, large firms are more subject to systematic risk and exhibit a higher PD than small firms.

Figure 3 shows the trend of business cycle indicator during our sample period. It suggests the economy reached a peak in 1994 and 2004 but fell to the bottom in 2001. We identify 1990, 2001, 2008–9 and 2011 as recession periods \((D_r = 1)\) and incorporate the interaction term into the following equation:

\[
AC_{i,t} = \alpha + b_1 \text{GGDP}_t + b_2 D_r \text{GGDP}_t + b_3 \text{PD}_{i,t} + b_4 \ln \text{sales}_{i,t} + \sum_{j=1}^{10} y_j \text{DI}_j + \sum_{y=1}^{24} \delta_y \text{DY}_y + \varepsilon, \tag{5.2}
\]

where \(\text{GGDP}_t\) is the GDP growth rate at time \(t\), and \(D_r \text{GGDP}_t\) is the interaction term in which the dummy variable \(D_r\) is equal to 1 in a recession period and 0 otherwise. A negative coefficient \((b_2)\) indicates that the increase of asset correlations during economic downturns is greater than the decrease during economic upturns.

Table 4 shows a robust result that both GGDP and \(D_r \text{GGDP}_t\) help explain asset correlations in different model specifications. After controlling for PD, size, industry and yearly fixed effects, lower economic growth still contributes to higher asset correlations. Furthermore, asset correlation’s procyclicality is exacerbated by its asymmetric
TABLE 3  Business cycle indicator.

<table>
<thead>
<tr>
<th>Component</th>
<th>Red (Booming)</th>
<th>Yellow-red</th>
<th>Green</th>
<th>Yellow-blue</th>
<th>Blue (Sluggish)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total score</td>
<td>45–38</td>
<td>37–32</td>
<td>31–23</td>
<td>22–17</td>
<td>16–9</td>
</tr>
<tr>
<td>Scores of component indicators</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Monetary aggregates M1B</td>
<td>15</td>
<td>12</td>
<td>6</td>
<td>2.5</td>
<td>➞</td>
</tr>
<tr>
<td>Direct and indirect finance</td>
<td>10</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>➞</td>
</tr>
<tr>
<td>Stock price index</td>
<td>24</td>
<td>11</td>
<td>-4</td>
<td>-22</td>
<td>➞</td>
</tr>
<tr>
<td>Industrial production index</td>
<td>9</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>➞</td>
</tr>
<tr>
<td>Nonagricultural employment</td>
<td>2.6</td>
<td>2.2</td>
<td>1.2</td>
<td>0.6</td>
<td>➞</td>
</tr>
<tr>
<td>Customs-cleared exports</td>
<td>15</td>
<td>11</td>
<td>5</td>
<td>1</td>
<td>➞</td>
</tr>
<tr>
<td>Imports of machinery and electrical equipment</td>
<td>25</td>
<td>16</td>
<td>7</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>Manufacturing sales</td>
<td>11</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>➞</td>
</tr>
<tr>
<td>Wholesale, retail and food services sales</td>
<td>8</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>➞</td>
</tr>
</tbody>
</table>

The components of the business cycle indicator in Taiwan. Individual components and checkpoints are in given terms of percentage changes over a one-year span. All components, except the stock price index, have been seasonally adjusted. Source: National Development Council (Taiwan).
behavior that an increase in asset correlations during recessions is more severe than a decline in asset correlations during expansions. For instance, a 1% decline in GDP growth is on average accompanied by an increase of 0.186% in asset correlations (column 1). Furthermore, an extra 0.133% increase should be considered during economic downturns because of the asymmetric effect of $D_rGGDP$ and the whole effect will be 0.319%.

Our empirical results show that asset correlations tend to rise during recessions with an asymmetric behavior. As banks tend to reduce their lending activities during downturns, firms may not get sufficient funding, and their PD rises accordingly. Meanwhile, a higher capital requirement due to higher credit risk may force banks to further squeeze credits, further deteriorating the economic condition. However, the current ASRF approach does not consider these effects and may amplify the fluctuation of business cycles when banks ease or tighten credits. More importantly, our findings suggest that banks may not hold sufficient capital during economic downturns.

6 CONCLUSIONS

The Basel II Accord of the BCBS explicitly ties the credit risk of a bank’s obligors to the bank’s regulatory capital requirement. While asset correlations are introduced to measure the comovement between a firm’s asset returns and the common risk factor returns under the ASRF framework, Basel II just released its specifications without clearly illustrating how to calculate asset correlations. In this study, we use the method of Lee et al (2011) to estimate asset correlations for Taiwanese firms and explore the effects of firm size, PD and industry on asset correlations.
### TABLE 4  Regression results of cyclical asset correlations.

<table>
<thead>
<tr>
<th>Explanatory variables</th>
<th>OLS regression</th>
<th>Fractional logit regression APE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model 1</td>
<td>Model 2</td>
</tr>
<tr>
<td></td>
<td>Model 3</td>
<td>Model 4</td>
</tr>
<tr>
<td>Constant</td>
<td>0.664***</td>
<td>0.674***</td>
</tr>
<tr>
<td></td>
<td>(19.00)</td>
<td>(19.19)</td>
</tr>
<tr>
<td>GGDP</td>
<td>-0.186***</td>
<td>-0.187***</td>
</tr>
<tr>
<td></td>
<td>(-18.45)</td>
<td>(-18.55)</td>
</tr>
<tr>
<td></td>
<td>-0.241***</td>
<td>-0.243***</td>
</tr>
<tr>
<td></td>
<td>(-18.62)</td>
<td>(-18.71)</td>
</tr>
<tr>
<td>$D_r\cdot$GGDP</td>
<td>-0.133***</td>
<td>-0.137***</td>
</tr>
<tr>
<td></td>
<td>(-7.86)</td>
<td>(-8.03)</td>
</tr>
<tr>
<td></td>
<td>-0.21***</td>
<td>-0.22***</td>
</tr>
<tr>
<td></td>
<td>(-10.79)</td>
<td>(-10.93)</td>
</tr>
<tr>
<td>Risk-neutral PD</td>
<td>-1.107***</td>
<td>-1.252***</td>
</tr>
<tr>
<td></td>
<td>(-21.48)</td>
<td>(-18.86)</td>
</tr>
<tr>
<td>Physical PD</td>
<td>-0.425***</td>
<td>-0.646***</td>
</tr>
<tr>
<td></td>
<td>(-19.15)</td>
<td>(-12.97)</td>
</tr>
<tr>
<td>Size</td>
<td>8.424***</td>
<td>8.522***</td>
</tr>
<tr>
<td></td>
<td>(17.29)</td>
<td>(17.42)</td>
</tr>
<tr>
<td></td>
<td>10.946***</td>
<td>10.872***</td>
</tr>
<tr>
<td></td>
<td>(3.15)</td>
<td>(3.14)</td>
</tr>
<tr>
<td>DY</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DI</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.33</td>
<td>0.32</td>
</tr>
<tr>
<td>$N$</td>
<td>10 768</td>
<td>10 768</td>
</tr>
</tbody>
</table>

The regression results of asset correlations on PD, firm size, industry groups, yearly dummy variables and real GDP growth rate. AC denotes asset correlation. Size is measured by ln sales. GGDP is the real GDP growth rate. $D_r\cdot$GGDP is the interaction term in which the dummy variable $D_r$ is equal to 1 in a recession year and 0 otherwise. DY and DI are dummy variables for year and industry, respectively. $t$-statistics are reported in parentheses for the OLS regression. Z-statistics for the robustness check with fractional logit regression average partial effect (APE) are reported in parentheses. *** denotes statistical significance at the 1% level.

Similarly to the results of Lopez (2004), and Lee and coworkers (Lee et al 2011, 2013; Lee and Lin 2012), we find that asset correlations are negatively associated with PD and positively associated with firm size. We also find an industry effect, where firms in “basic materials” and “technology” sectors on average have higher asset correlations. This finding is different from the study of Lee et al (2011) that US firms in the media, information technology and telecommunication sectors tend to have higher asset correlations. Taiwanese firms in the utilities and oil and gas sectors exhibit lower asset correlations, while US firms in retailing and consumer staples on average have lower values. Our findings suggest that industry effects are quite different in these two markets.

Our analysis also confirms an asymmetric and cyclical behavior of asset correlations that the current Basel Accord may not have fully considered. After controlling for firm size, PD, industry and yearly fixed effects, we still observe that asset correlations
appear to increase in recessions and decrease in expansions. More importantly, the increase in asset correlations during economic downturns is more pronounced than the decrease during upturns. Our findings suggest that regulatory capital may be underestimated during economic downturns and that that may have important policy implications in terms of incorporating the variations in asset correlations through the business cycle.

DECLARATION OF INTEREST

This research is funded by the Open Co-operation Project of the Institute of Guangdong, Taiwan, and Shantou University. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES


Research Paper

Rating momentum in the macroeconomic stress testing and scenario analysis of credit risk

Jimmy Skoglund and Wei Chen

SAS Institute Inc., 100 SAS Campus Drive, Cary, NC 27513-2414, USA; emails: jimmy.skoglund@sas.com; wei.chen@sas.com

(Received June 28, 2016; revised September 14, 2016; accepted September 16, 2016)

ABSTRACT

With a focus on multi-horizon macroeconomic credit loss projection models in stress testing and impairments, it is of interest to understand how, under stressed and best-estimate economic projections, different model assumptions can affect such projections. We focus on the popular factor model of credit risk with an added rating-momentum feature, which violates the Markov property of the model. While in retail credit models it is obvious that past delinquency is an important predictor of state path, commercial models are often implemented as Markovian, conditional on the macroeconomic paths. We find that models that do take into account the stylized fact of rating momentum can accelerate the timing of the loss significantly, compared with the non-rating-momentum case. The exact effect depends on the scenario’s time horizon, severity and portfolio quality. In general it takes longer for differences to materialize for good quality portfolios, while the effect on lower rating quality portfolios can be almost immediate, with significant loss underestimation. The factor model sensitivity to the explained part of the macroeconomic factors versus idiosyncratic effects is well known but must be recognized in practice by regulators, as models

Corresponding author: J. Skoglund

Print ISSN 1753-9579 | Online ISSN 1753-9587
Copyright © 2017 Infopro Digital Limited

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
with a small portion explained by the macroeconomic factors can protect the portfolio loss significantly.

Keywords: credit risk; macroeconomics; credit loss projection; stress testing; commercial models; rating momentum.

1 INTRODUCTION

Credit risk is arguably the most significant risk to every classical banking institution. The Basel Committee’s traditional regulatory credit risk approach to risk capital (risk weighted assets) takes a through-the-cycle (TTC) approach or even a downturn-adjusted approach for the risk parameters of probability of default, loss given default and exposure at default. This can be seen as an attempt over the years to stabilize banks’ regulatory capital requirements. The regulatory stress testing exercises, starting in the wake of the 2007 financial crisis, introduced a new credit risk modeling paradigm where point-in-time (PIT) or best-estimate macroeconomic credit risk projection approaches were introduced. Macroeconomic stress testing is being officially adopted as a risk management standard by many jurisdictions in the world. As a result, multi-horizon credit modeling, based on macroeconomic scenarios, is becoming increasingly important. The new international accounting standard for the recognition of the credit loss (impairments) known as International Reporting Standard 9 (IFRS 9, developed and approved by the International Accounting Standards Board (IASB); see www.ifrs.org) has taken the same approach. The US financial standard (Financial Accounting Standards Board (FASB); see www.fasb.org) is following suit as well.1

The key to the new macroeconomics-driven credit modeling requirements is to first call for a macroeconomics-sensitive credit model analysis focusing on survival through a series of projected economic horizons rather than a credit score or rating at a certain point in time. Macroeconomics-driven credit modeling can in principle also capture the underlying factors of systemic credit contagion. Another important requirement for credit risk modeling for the new stress testing and accounting standards is path dependency for a multi-horizon analysis. Most banking book credits are held to maturity, and there is rarely a jump-to-default. Defaults often happen as a consequence of migrations to increasing delinquency or rating grades worsening period by period. In the same time frame, accounts can recover and be considered healthy again. Careful credit risk analysis must recognize credit deterioration long before the actual default. Such recognition has also arisen through, for example, credit rating

1 See Skoglund (2017) for an overview of macroeconomics-driven credit models for lifetime impairments forecasting.
changes for commercial exposures and delinquency status tracking for retail loans. Such past behavior not only affects the forecasting of the default but can also model the cashflows in the future. For example, a forecasted delinquency might cause a shortage of cash inflow, which would affect the net income and liquidity in a future period. On the other hand, the cure of a delinquency might bring in more cashflows and accrued interest plus fees to the bank. In addition, in a multi-horizon analysis a prepayment or modification of a loan can also change cashflow patterns at the subsequent horizons, which again can significantly affect loss forecasting and cashflows. For commercial loans, the deterioration in a credit rating over time will certainly change the expected loss profile and fair value, but not necessarily the cashflows received.

All the major regulatory stress testing initiatives define multi-horizon regulatory macroeconomic scenarios in which the full balance-sheet stress testing must be projected. For example, US Comprehensive Capital Analysis and Review (CCAR) and Dodd–Frank Act stress tests (DFAST; see http://bit.ly/1wJvuhE) are based on nine-quarter scenarios, and, in Europe, European Banking Authority stress testing is based on twelve-quarter (three-year) scenarios. The credit models themselves can of course have a different frequency, eg, monthly, even if, say, quarterly projections are used. This is natural for retail portfolios, for example, where delinquency status is typically measured monthly due to the payment frequency. For the best estimate of expected credit loss in the new accounting standards (such as IFRS 9), when an exposure’s credit quality significantly deteriorates or is impaired, the exposure is subject to a full lifetime credit loss calculation based on a business-as-usual economic scenario. In other cases a twelve-month best estimate of expected credit loss is used. In practice, a bank might also use multiple economic scenarios as the basis for a best estimate of expected credit loss with probability weighting. The new stress testing and accounting standards also call for as much granularity as possible. In most cases the best practice is to project credit loss at the asset or loan level. For example, the Board of Governors of the Federal Reserve System requires the credit modeling for the US CCAR to be carried out at sufficient granularity to capture exposure that reacts differently to risk drivers under stress scenarios. IFRS 9 expected credit loss is similarly assessed at a sufficiently granular level.

In this paper we focus on the corporate stress testing models for credit risk. The model we consider has its foundation in corporate credit risk but is frequently used for retail portfolios as well. This is because, with observable default rates on a pool of retail credits, the model can be used for direct calibration of the macroeconomic sensitivity to observed default rates (see Skoglund and Chen 2015, p. 217).

---

2 The finalized US financial standard proposes to use lifetime credit loss as the impairment for all assets, regardless of whether they are credit impaired or not, and hence requires banks to set aside the full amount of lifetime loss at origination.

3 The model we consider has its foundation in corporate credit risk but is frequently used for retail portfolios as well. This is because, with observable default rates on a pool of retail credits, the model can be used for direct calibration of the macroeconomic sensitivity to observed default rates (see Skoglund and Chen 2015, p. 217).
assessed by external rating agencies such as Standard & Poor’s (S&P) or Moody’s. The rating assessment and the empirical transition probabilities in the transition matrix are frequently based on historical data covering many business cycles. Such long-term transition matrixes are the basis for TTC probability of default term structures. The macroeconomic impact and conversion from long-term to PIT rating transition probabilities are often based on a factor-model approach derived from the Merton (1974) model of defaultable corporate bond pricing. The single-factor version of the model (Vasicek 1991; Gordy 2003) is also the foundation of the Basel Committee’s advanced-internal-model-based capital requirement. Since the factor model allows corporate credit risk to be connected to macroeconomic factors it is also frequently used in banks’ internal economic capital models and is often referred to as the CreditMetrics (1997) approach (see also Finger 1999). With the introduction of macroeconomics-based stress testing and expected credit loss analysis requirements, traditional credit-portfolio economic-capital models are not obsolete, but in fact become more important. However, there is a requirement to deploy these models in path-dependent projections as well as capture more exactly the dynamics along the path. Such path dynamics can include different delinquency behavior as well as so-called rating momentum. On the retail side, it is obvious that the delinquency status and delinquency path are important predictors of future credit quality. On the commercial side, empirical evidence has been reported in the literature with respect to the impact of previous corporate downgrades or upgrades on future transitions (rating momentum).

In the macroeconomic projection of corporate expected credit loss, the standard factor model is conditionally Markovian on the macroeconomic factors with an analytic solution. The Markovian feature of the model has, however, been criticized. For example, Cantor and Hamilton (2004) find that rating history is important for predicting corporate default, violating the Markov feature for Moody’s transition matrixes. Bangia et al (2000) estimate different momentum transition matrixes on S&P data conditional on the last rating change being upward, downward or on there being no change. The rating change momentum hypothesis is supported by the data, with significant variation in transition matrixes conditional on recent rating changes. In particular, most downgrade probabilities for the downward-momentum matrix are larger than the corresponding values in the unconditional matrix. The exact opposite is true for the upward-momentum matrix, which exhibits smaller downgrade probabilities than the unconditional matrix. Guttler and Raupach (2008) analyze the one-year credit portfolio value-at-risk (VaR) differences using momentum-insensitive and momentum-sensitive transition matrixes. They find significant differences, with the momentum-insensitive VaR being underestimated (see also Djiambou (2008) on rating momentum and credit VaR models). However, while there is ample evidence for non-Markovian behavior in corporate ratings-based models, such features are rarely
accounted for in practice. This is in contrast to retail state-transition models used in expected loss stress testing, which often include both macroeconomics and loan path information, such as recent delinquency status, and spell and severity of the last delinquency. Since the model tracks past state information, often in the form of summary measures, such as “months since last delinquency”, it violates the Markov property. We refer the reader to Skoglund and Chen (2016) for multi-horizon expected-loss analysis using retail-dynamic state-transition models. In principle, non-Markovian models can always use state-transition simulation and subsequent averaging of losses to obtain an estimate of the expected credit loss path. However, due to the remote likelihood of default events it is preferable to avoid simulation if feasible, and use an expanded Markov iteration. Such expanded Markov iteration schemes are derived in Skoglund and Chen (2016).

Our analysis here focuses on the effect of the rating momentum on the projection of multi-horizon expected credit loss. While for extreme tail scenarios this is not so different from measures such as VaR, special considerations come into play when analyzing the model behavior over multi-horizon path-dependent scenarios as opposed to the behavior over a single horizon. Our aim is therefore to consider an empirical analysis of the impact of the rating momentum on corporate credit risk with special focus on the issues that arise with multi-horizon macroeconomic stress testing and best-estimate macroeconomic forecasts.

The paper is organized as follows. In Section 2 we introduce the corporate factor model and consider empirical analysis of the model’s behavior for different parameters, initial rating classes and stylized macroeconomic projections. In Section 3 we add the rating momentum to the model and analyze the difference between rating momentum and no rating momentum on the model for the stylized macroeconomic projections. Finally, in Section 4 we summarize our results and offer our conclusions.

2 CORPORATE FACTOR MODELS

The multivariate version of the Merton (1974) model of defaultable corporate bond pricing is often used to model bond default and rating migration. In the frequently used factor model (CreditMetrics 1997) inspired by Merton, the asset returns for issuer $i$, $Y_i$, are given by

$$Y_i = \sum_{h=1}^{H} \beta_{ih} Z_h + \lambda_i \varepsilon_i.$$  \hspace{1cm} (2.1)

In practical applications of the model, the unobserved asset value process, $Y_i$, is often approximated by equity and is referred to as the credit index of the firm. Here the $Z_h$ are interpreted as returns of the common credit factors $h = 1, \ldots, H$, $Z_h \sim N(0, \sigma_h)$, with $\beta_{ih}$ being the sensitivity of issuer $i$ to the $h$th index factor. The $\varepsilon_i$ are...
the idiosyncratic factors that represent the specific risk to issuer $i$. The $\varepsilon_i$ are here assumed to be independent and identically distributed standard normal variables that are independent of the $Z_h$.

Since the $Z_h$ are not necessarily standardized, we can obtain the standardized $Y_i$ as

$$
\hat{Y}_i = \phi_i \left( \sum_{h=1}^{H} \beta_{ih} Z_h \right) + \lambda_i \varepsilon_i, \quad \phi_i = \frac{1 - \lambda_i^2}{\beta_i \Sigma \beta_i'}, \tag{2.2}
$$

where $\Sigma$ is the covariance matrix of $(Z_1, \ldots, Z_H)$ and $\beta_i = (\beta_{i1}, \ldots, \beta_{iH})$. Note that $\hat{Y}_i \sim N(0, 1)$.

Denoting by $I$ the transformed (to normal) default threshold for issuer $i$, we have that the conditional (on $Z$) default probability is

$$
P(\hat{Y}_i < I) = N \left( \frac{1 - \phi_i \left( \sum_{h=1}^{H} \beta_{ih} Z_h \right)}{\sqrt{1 - R_i^2}} \right), \tag{2.3}
$$

where $R_i^2 = 1 - \lambda_i^2$ corresponds to the classical linear regression $R^2$. In the case of a multi-horizon projection, applicable in stress testing and lifetime expected credit loss modeling, credit rating migration should, however, be accounted for, rather than just default.

To extend the above to the macroeconomic conditional transition, we need to consider multiple thresholds, $\{I_{lr}\}_{r=1}^{K-1}$ for rating states $l = 1, \ldots, K - 1$ and $r = 1, \ldots, K$. Such thresholds are frequently obtained from empirical long-run transition matrices. Consider the S&P transition matrix in Table 1 estimated by Bangia et al (2000), which is an unconditional annual migration matrix. The rating thresholds, $\{I_{lr}\}_{r=1}^{K-1}$, are the implied one-year (standard normal distribution) rating thresholds that can be obtained from the transition matrix transition probabilities.

To derive the model’s conditional transition probabilities, $\{q_{lr}(t)\}_{r=1}^{K-1}$ for time $t = 1, \ldots, T$. Here we shall write $\{d_{ik}(t)\}_{t=1}^{T}$ for the $k = 1, \ldots, K$ state-transition classes. For firm $i$, assuming that the threshold for it to migrate to above state $r = 2$ (hence to state $r = 1$) is $I_{I2} < \infty$, we can now write

$$
P(d_{i1}(t) = 1) = P(\hat{Y}_i > I_{I2}) = 1 - P \left( \varepsilon_i \leq \frac{I_{I2} - \phi_i \left( \sum_{h=1}^{H} \beta_{ih} Z_h \right)}{\sqrt{1 - R_i^2}} \right).
$$

---

Note, however, that for the original regression (2.1) the regression $R^2$ is given by

$$
1 - \frac{\lambda_i^2}{\lambda_i^2 + \beta \Sigma \beta'} = \frac{\beta_i \Sigma \beta'}{\lambda_i^2 + \beta_i \Sigma \beta'}.
$$
TABLE 1  Standard & Poor's transition matrix.

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>91.63</td>
<td>7.46</td>
<td>0.48</td>
<td>0.08</td>
<td>0.04</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0.64</td>
<td>91.81</td>
<td>6.75</td>
<td>0.60</td>
<td>0.12</td>
<td>0.03</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0.07</td>
<td>2.27</td>
<td>91.69</td>
<td>5.11</td>
<td>0.25</td>
<td>0.01</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>BBB</td>
<td>0.04</td>
<td>0.27</td>
<td>5.56</td>
<td>87.88</td>
<td>4.83</td>
<td>1.02</td>
<td>0.17</td>
<td>0.24</td>
</tr>
<tr>
<td>BB</td>
<td>0.04</td>
<td>0.10</td>
<td>0.61</td>
<td>7.75</td>
<td>81.48</td>
<td>7.89</td>
<td>1.11</td>
<td>1.01</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.10</td>
<td>0.28</td>
<td>0.46</td>
<td>6.95</td>
<td>82.80</td>
<td>3.96</td>
<td>5.45</td>
</tr>
<tr>
<td>CCC</td>
<td>0.19</td>
<td>0</td>
<td>0.37</td>
<td>0.75</td>
<td>2.43</td>
<td>12.13</td>
<td>60.45</td>
<td>23.69</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

where the probability is conditional on \( \{Z\} \). Similarly, for firm \( i \) to migrate to any of the \( r = 2, \ldots, K-1 \) rating classes, we can write

\[
P(d_{ir}(t) = 1) = P(I_{r+1} \leq \tilde{Y}_i \leq I_r)
= P\left(\varepsilon_i \leq \frac{I_r - \phi_i(\sum_{h=1}^H \beta_h Z_h)}{\sqrt{1 - R^2}}\right) - P\left(\varepsilon_i \leq \frac{I_{r+1} - \phi_i(\sum_{h=1}^H \beta_h Z_h)}{\sqrt{1 - R^2}}\right).
\]

(2.4)

Finally, for migration to state \( K \), we obtain

\[
P(d_{iK}(t) = 1) = P(\tilde{Y}_i < I_{1K}) = P\left(\varepsilon_i \leq \frac{I_{1K} - \phi_i(\sum_{h=1}^H \beta_h Z_h)}{\sqrt{1 - R^2}}\right).
\]

We note that this corresponds to an ordered probit model (Amemiya 1985, Chapter 7) and that each transition probability is a marginal transition probability.

2.1 Macroeconomic conditional transition probabilities

We can now apply a conditional law of large numbers such that the expected transition fraction of a firm \( i \) that belongs to state \( r = 1, \ldots, K \) at time \( t = 1 \) is

\[
x_{ir}(t) = E[(d_{ir}(t) | Z)]
= N\left(\frac{I_r - \phi_i(\sum_{h=1}^H \beta_h Z_h)}{\sqrt{1 - R^2}}\right) - N\left(\frac{I_{r+1} - \phi_i(\sum_{h=1}^H \beta_h Z_h)}{\sqrt{1 - R^2}}\right) = \theta_{ir}^1(t),
\]

(2.5)

conditional on a realization of \( \{Z\} \), where \( N \) is the cumulative normal distribution and \( I_{11} = +\infty, I_{1K+1} = -\infty \). Obtaining the conditional transition probabilities
using the law of large numbers is similar to Lucas et al (1999), although here we do not attempt an approximation for the credit loss distribution and instead focus on the credit migration frequencies.

2.2 Conditional Markov iteration solution to macroeconomic scenarios

In a multiple-horizon scenario, analytic Markov iteration of the conditional transition matrix is used to generate the future state-transition distribution. We shall also show later that the iteration can be extended to incorporate that rating-momentum effect while retaining the same analytical tractability. Essentially, for each realization of \{Z\} we can obtain a \(K\)-dimensional vector \(x_i(t)\) from (2.5) with the expected state fractions \(0, 1\) for each of the possible \(r = 1, \ldots, K\) states. Clearly, \(\sum_{r=1}^{K} x_{ir}(t) = 1\). Now, at time \(t = 2\), our starting point is the vector \(x_i(1)\). This will in general have \(K\) nonzero elements, as the transition to state \(r = 1, \ldots, K\) is often possible from initial state \(l = 1, \ldots, K - 1\) (even if it is remotely possible with a very small probability). Hence, in order to compute expected state migrations at \(t = 2\) we need to calculate all the transition probabilities from the current state fractions (at \(t = 1\)), \(k = 1, \ldots, K - 1\), migrating to any of the states \(k = 1, \ldots, K\) (at \(t = 2\)). Conditional on \{Z\} we therefore obtain a \((K - 1 \times K)\)-dimensional transition matrix. Let \(A(t)\) be the \(K \times K\) matrix that is defined by adding a \(K\)th row of zeros to columns \(K - 1\) and unity for the \(K\)th column. That is,

\[
A^i(t) = \begin{bmatrix}
\theta_{i11}(t) & \cdots & \theta_{i1K}(t) \\
\vdots & \ddots & \vdots \\
\theta_{iK-11}(t) & \cdots & \theta_{iK-1K}(t) \\
0 & \cdots & 1
\end{bmatrix},
\]

where \(\theta_{ir}(t)\) is defined in (2.5). With this transition matrix we can now compute the expected state fractions at \(t = 2\) using a first-order Markov chain. Specifically,

\[x_i(2) = [A^i(2)]^t x_i(1)\]

such that, conditional on \{Z\}, we obtain the expected state fractions at \(t = 2\). It is now clear that, in general, for any \(t = 1, \ldots, T\) we obtain a simple expected state occupancy iteration relation, ie,

\[x_i(t) = [A^i(t)]^t x_i(t - 1). \tag{2.6}\]

This Markov iteration algorithm computes the expected transition probabilities (conditional on a realization of \{Z\}) for each possible state \(k = 1, \ldots, K\) for each time \(t = 1, \ldots, T\). Through the iteration equation we obtain the pathwise expected
rating transitions of the issuer, loan or bond for a given scenario of the economic factors, \( \{Z(t)\}_{t=1}^{T} \). Hence, we can obtain the expected loss in an economic stress scenario without sampling. This is because we can weight the state-dependent cash-flows with the state occupancy probabilities in \( x_i(t) \). Finally, we note that the reason we obtain this relatively simple representation of the Markov iteration in this case is that the transition probabilities depend only on time, through the realization of \( \{Z\} \) as in (2.5), and not on state-transition history. This type of model is referred to as a time-inhomogeneous (nonstationary) Markov state-transition model.

### 2.3 Computing expected losses

Our aim is to use the state-transition vector \( \{x_i(t)\}_{t=1}^{T} \) to compute the expected loss in a certain macroeconomic projection for \( \{Z(t)\}_{t=1}^{T} \). For simplicity, we will assume recovery is zero, and a constant notional over time with reduction due only to defaulted portions. Let the flow of defaulted assets be \( F(t) \). Then we can write

\[
F(t) = PD(t) \times E(t - 1),
\]

with \( E(0) \) the current exposure and \( PD(t) \) the conditional default probability at \( t \). Due to our constant notional assumption,

\[
E(t) = (1 - PD(t)) E(t - 1).
\]

For \( E(0) = 1 \), this now gives

\[
F(t) = x_K(t) - x_K(t - 1)
\]

and

\[
\sum_{s=1}^{t} F(s) = x_K(t), \tag{2.7}
\]

where \( x_K(t) \) is the \( K \)th element of \( x(t) \) having the defaulted fractions of the asset until \( t \), and \( x_K(0) \) is null as we start out with no defaulted fractions of the loan.

In our empirical analysis we assume a unit notional and, rather than comparing cumulative default rates as in (2.7), we model the full transition proportion over time and compare the unit notional remaining, ie, we will simply compare the term

\[
1 - x_K(t). \tag{2.8}
\]

See Skoglund and Chen (2016) on \( PD(t) \) being the conditional default probability of the model and on adjusting the notional with credit state specific growth or decay (amortization) rates.
2.4 An alternative simple version of the model

As we have seen above, the factor model produces a conditional default probability (or a PIT default probability) using the conditional transition rates and the Markov iteration in (2.6). Hence, it models an essential component through risk migration, which provides valuable information in a multi-horizon analysis, which is typical of macroeconomic expected-loss analysis and stress testing. As an aside, for completeness we also discuss a simpler (non-rating migration) version of the model formulation that considers only the conditional default probability in (2.3) in order to generate future PIT default probabilities from the implied long-term default probabilities in the transition matrix. This simpler approach is still often used in practice. In general, the long-term or TTC default term structure can still be generated by Markov iteration of the static unconditional long-term transition matrix. This yields a TTC default term structure $\delta(t)$ per grade that is used to generate the PIT default probability, $P$, as follows:

$$P(t) = N\left(\frac{N^{-1}(\delta(t)) - \phi(t) \beta_h Z}{\sqrt{1 - R^2}}\right).$$

(2.9)

More specifically, in order to give a direct connection to the Basel single-factor version of the model, (2.9) can be written as

$$P(t) = N\left(\frac{N^{-1}(\delta(t)) - \sqrt{R^2} Z}{\sqrt{1 - R^2}}\right),$$

which Carlehed and Petrov (2012) use to generate PIT-TTC adjustments to the Basel one-year probability of default models. In this setting the unobserved single factor, $Z$, can be backed out from observed default statistics such that

$$Z = \frac{N^{-1}(\delta(t)) - N^{-1}(P(t)) \sqrt{1 - R^2}}{\sqrt{R^2}}.$$

If we let $\mu$ and $\sigma$ denote the mean and volatility of $N^{-1}(P(t))$, we have that

$$Z = \frac{\mu - N^{-1}(P(t))}{\sigma},$$

which gives a direct connection to the Basel single-factor $Z$ model for risk weighted assets. It therefore allows the direct model connection to (and explanation of) the Basel Committee’s PIT-versus-TTC adjusted probability of default. The approach to extracting the unobserved single factor, $Z$, from observed default rates creates a syntectic $Z$-factor that may be undesirable for understanding the impact of the direct

---

5 It can of course also be estimated from historical default data, which is more appropriate for large nonrated pools of credits.
scenario analysis from observed macroeconomic factors on expected loss (e.g., unemployment rate and interest rates), although we can try to decompose the explained portion of the synthetic $Z$-factor observed using regression analysis. However, the connection is unnecessary and we can of course use a factor model with observed macroeconomics directly as in (2.9). The calibration procedure is the same as for the more general rating-migration-factor model, for example, using equity approximations for large corporates or actual observed default rates on a pool of credits with direct calibration of macroeconomic sensitivity to observed default rates. The model correlation coefficient ($R^2$) is often interpreted as the portfolio default rate correlation to the macroeconomics. The relationship is, however, nonlinear, and it is therefore not exactly a linear regression beta. Gordy (2000) uses a method of moments estimator, while Gordy and Heitfield (2002) consider the maximum likelihood method, using the binomial distribution and conditional independence (on the macroeconomic factors), in order to derive the joint likelihood for the observations. They also evaluate maximum likelihood estimators and the moment estimator in Monte Carlo simulations and express concern regarding the potential downward bias of unrestricted maximum likelihood estimators and the method of moments estimator. See also Yang and Du (2016), who calibrate the model to macroeconomics using the loglikelihood of observing the default or downgrade frequency.

In the remainder of the paper we shall focus on the full rating migration version of the model, which is more appropriate for the long-term multiple-horizon scenario projection. It is also required for incorporating and analyzing the rating-momentum impact.

### 2.5 Model scenario analysis and stress testing examples

To illustrate the behavior of the corporate factor model in macroeconomic stress testing and scenario analysis we consider a single-factor-model version of (2.1) for simplicity. With a single factor we can write

$$\tilde{Y}_i = \sqrt{1 - \lambda_i^2} Z_i + \lambda_i e_i,$$

where, as above, $R_i^2 = 1 - \lambda_i^2$ corresponds to the classical linear regression $R^2$. Hence,

$$\tilde{Y}_i = \sqrt{R_i^2} Z_i + \sqrt{1 - R_i^2} e_i,$$

such that, for a given transition matrix (or TTC default term structure), the model’s projected default rates depend on

- $R_i^2$,
- the specific $Z$ scenario,
- corporate credit quality, i.e., current rating.
Naturally, a high market correlation (high $R^2$) offers protection in upward and stable macro scenarios since there is little impact from idiosyncratic risk, $\varepsilon_i$, which can cause downward migration. On the other hand, a high market correlation gives more sensitivity to stressed outcomes of $Z$. For a severe $Z$ scenario and good initial rating grade, the impact of different $R^2$ can be substantial. With a high $R^2$ the severe scenario can quickly move to extreme values in $\tilde{Y}_i$ with less protection from the idiosyncratic risk component ($\varepsilon_i$). Similarly, a low $R^2$ offers protection against the severe scenario by mitigating the impact from systematic risk downturns. For poor credit quality rating grades there is less protection from low $R^2$, since less extreme asset shocks ($\tilde{Y}_i$) are needed to move the exposure into worse states or even default. In principle, we can also consider different $R^2$, depending on the scenario, ie, different loadings on the factors in different scenarios. In a normal macroeconomic scenario the idiosyncratic effect might be more significant, while in extreme stress scenarios exposures tend to have higher contagion and to be more affected by the macroeconomic environment. This is consistent with the prudent practice of using stressed correlations in risk models when estimating tail risk.

We now illustrate some of these model effects using stylized economic scenarios for $Z$ over a maximum of twenty-four horizons and the transition matrix in Table 1. Here, $Z$ is interpreted as a standard normal variable, with negative values representing downturns and positive values representing upturns. The stylized economic scenarios we consider are the following:

- a significant downturn scenario that has values $-1$, $-2$, $-3$ for the first three horizons and then $-3$ for all the remaining horizons;

- a severe scenario, with increasingly negative $Z$ over the twenty-four horizons, changing to a more severe $Z$ every three horizons (specifically, the severe scenario has $Z = -1$ for the first three horizons and is then (negatively) incremented for every three subsequent horizons), ie, the “same” three horizon scenario values are incremented as

$$Z = \{-1, -1.25, -1.75, -2, -2.25, -2.5, -2.75, -3\},$$

which is consistent with a severe very-long-term stress;

- a stylized stable scenario with $Z$ centered at $0$ for all twenty-four horizons;

- a significant upturn scenario, which is the positive scenario reciprocal of the severe scenario.

See Figure 1 for the different scenario projections.

Admittedly, a downturn or upturn does not last for twenty-four horizons (with a yearly horizon). However, we can think of the actual regulatory macroeconomic
FIGURE 1 Stylized scenarios evolution.

![Stylized scenarios evolution graph](image)

scenarios for stress testing being short-term versions of our severe and significantly severe scenarios, and we can think of the best-estimate credit impairment scenario for lifetime expected loss being represented approximately by our stable scenario. As mentioned above, our comparisons will focus on the unit corporate exposure remaining (see (2.8)).

Figure 2 shows the results for a significant severe stress test for AAA, BBB and CCC rated corporate exposures and \( R^2 \) equal to 0.2, 0.4, 0.6 or 0.8. Due to the significant severity of the scenario we only show the first six horizon results. For the AAA rated exposure we note that it is only after the first severe \( Z \) at horizon 3 that we can see migration and default effects. This is because a very large \( Z \) shock is needed to downgrade the AAA rated exposure (see the transition matrix in Table 1). After the initial downgrade has happened, defaulted portions of the exposure quickly occur due to the subsequent large negative shocks of \( Z = -3 \). We also note that there is

---

6 In practice, the best estimates for lifetime impairment in, for example, IFRS 9 use the current short-term business-cycle market-condition projection but revert to a stable (steady-state) outcome after a few years. It has also become standard practice to consider multiple scenarios with a weighted average of different scenario outcomes. This means that model expected loss impact (for a given corporate credit quality and \( R^2 \)) depend on multiple scenarios and their specific weighting.
FIGURE 2  Unit notional exposure remaining for initial rating AAA, BBB and CCC for different values of model $R^2$ and significant severe scenarios.
FIGURE 3  Unit notional exposure remaining for initial rating AAA, BBB and CCC for different model $R^2$ and severe scenario.
significant protection from low $R^2$ for the AAA rated exposure. In Figure 2 the BBB rated exposure does not need such a large shock initially to downgrade or default, and we hence see a quicker defaulted portion of the unit exposure, particularly with high $R^2$. The CCC rated exposure moves quickly to large defaulted portions since only a small $Z$ shock is needed for downward migration, and consequently there is less protection from low $R^2$.

Figure 3 displays the impact of the severe $Z$ downturn scenario for twenty-four projection horizons for initial rating grade AAA, BBB and CCC and $R^2$ equal to 0.2, 0.4, 0.6 or 0.8. It illustrates clearly that for good rating grades there can be substantial protection from low $R^2$, while for lower rating grades there is not so much protection. However, for good credits it takes much longer periods before the impact is seen with the severe scenario. In contrast, for poor quality ratings there is an immediate effect in the short term. This is an important insight, as most stress scenarios are short term. We can therefore conclude that, for this severe stress scenario and good quality portfolios, even if there is substantial difference in the long run, there may be very little difference between different $R^2$ in the actual stress scenarios after horizons of just a few years. This is because the initial negative shocks to $Z$ are not extreme enough to cause downward migration and, as we saw with the significantly severe scenario, there may be multiple very extreme sequential shocks needed to see ultimate defaults from good quality portfolios.

In Figure 4 we display results for the stable scenario case ($Z = 0$ for all time horizons) for rating grades BBB and CCC. For better rating grades than BBB there is almost no default impact due to the stable scenario protection from defaults. As mentioned before, the stable scenario is (in the long term) more reminiscent of the best-estimate lifetime projection scenarios used, for example, in IFRS 9 impairments. We note from Figure 4 that the BBB rated exposures with high $R^2$ have significant protection from default and migration to worse states, as the impact of idiosyncratic risk is low. Figure 5, for reference, also shows the impact of a significant upward scenario on the rating grades B and CCC (there is no default impact on rating grades higher than B regardless of the value of $R^2$). Given the strong up-trending scenario, it offers significant default projection for all the good rating grades and high $R^2$.

Clearly, the above results are simple consequences of the factor model specification in (2.1) and the structure of the transition matrix in Table 1. However, it is important to consider them. For example, a poor choice of macroeconomic credit factors, and consequently a low $R^2$, can offer protection from future projected default rates. In our severe scenario case this is true only for “good” rated corporate exposures in the very long term, while it is true in the short-term for lower quality credits. For the significantly severe scenario, low $R^2$ offered significant default protection even in the short term for the “good” rated corporate exposures. For long-term stable scenario projections, the impact of $R^2$ on the worst quality rating grades can be significant,
FIGURE 4  Unit notional exposure remaining for initial rating BBB and CCC for different model $R^2$ and stable scenario.

with a high $R^2$ offering more protection from default, and hence significantly affecting lifetime loss estimations.

3 ADDING RATING MOMENTUM

As we discussed in Section 1, quite some empirical evidence has been reported in the literature with respect to the effect of previous corporate downgrades or upgrades on future transitions (see, for example, Cantor and Hamilton 2004; Bangia et al 2000). The evidence of rating momentum in corporate state-transition matrixes necessitates the introduction of a state indicator that tracks the recent state movement(s). While the
FIGURE 5 Unit notional exposure remaining for initial rating B and CCC for different model $R^2$ and up-trending scenario.

impact of rating momentum on future state-transition rates has been documented, there is also a large complementary literature that focuses on the evidence of rating momentum based on stock and bond price jumps after rating or watchlist announcements. The information content of ratings can be examined by focusing on the impacts of rating changes on stock and bond prices. The effect on market prices from rating changes can, for example, come from new information (unanticipated rating change) or from increased funding costs. There can also be an anticipation of future downgrades. The results in the literature with respect to rating announcements actually influencing market prices have been mixed. For example, Kliger and Sarig (2000) show that
the reaction of market prices to ratings is only focused on the more precise and new pieces of rating information. Hull et al (2004) and Di Cesare (2006) find no significant price effects in the credit default swap markets after downgrades or upgrades. Grothe (2013) finds evidence of significant effects on market prices from rating actions in periods of crisis, while in favorable times the effects are not pronounced. Therefore, we should not neglect to take rating momentum into account in multiple-horizon credit-transition-based modeling, especially in stress testing scenarios, although it might not be as important for the expected loss calculation.

When we consider specifying a model that incorporates rating momentum we can of course, like Cantor and Hamilton (2004) and Bangia et al (2000), focus on transition matrices that are estimated conditional on rating history. However, we can also consider extending the corporate factor model in (2.1) to include jump effects conditional on rating changes. In the latter case we can interpret (2.1) as the approximate value of equity returns that would react to rating announcements. Of course, both these model approaches are similar. Using switching transition matrixes (conditional on past rating history) induces a jump in default and migration thresholds for a given factor model, and hence a jump in the TTC default term structure. Our analysis of the rating momentum will focus on the state-switching transition matrix (TTC default term structure) approach, because it is easier to estimate from collective migration behavior due to the effect of rating momentum on the migration probability. As in Bangia et al (2000) we consider only the case when the last state behavior is used as an indicator. This means that our analysis of the rating momentum will focus on the case when the state indicator only tracks the last period’s change in rating direction. That is, we will capture whether the most recent state change was upward, downward or experienced no change. This represents only a mild deviation from the Markov property, and we can condition on the recent state change as in Skoglund and Chen (2016) and use an expanded conditional Markov iteration to “solve” the model that is only marginally more complex than the iteration in (2.6). Of course, in practice the implementation may consider a moving window of, say, the last three periods and track any associated rating change history. This only requires a little more history carrying and can still use the expanded conditional Markov iteration methods in Skoglund and Chen (2016).

To capture the importance of past dependencies in the transition matrix with increasing probability of a further ratings downgrade for corporates that were previously downgraded, we split the ratings history into three states, \{S, U, D\}, which corresponds to

---

7 Zhou (1997) consider a jump–diffusion process for the asset value in the Merton model. Their main motivation is, however, to replicate market credit spreads and, in particular, explain that a perfectly healthy asset, which has a credit spread of almost zero in the original Merton model, can have a systematically positive credit spread.
TABLE 2 Stable transition matrix.

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>91.84</td>
<td>7.58</td>
<td>0.48</td>
<td>0.05</td>
<td>0.05</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0.58</td>
<td>91.60</td>
<td>7.02</td>
<td>0.60</td>
<td>0.07</td>
<td>0.12</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0.06</td>
<td>2.29</td>
<td>91.56</td>
<td>5.20</td>
<td>0.60</td>
<td>0.26</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BBB</td>
<td>0.06</td>
<td>0.33</td>
<td>5.96</td>
<td>87.57</td>
<td>4.53</td>
<td>1.12</td>
<td>0.23</td>
<td>0.19</td>
</tr>
<tr>
<td>BB</td>
<td>0.06</td>
<td>0.13</td>
<td>0.50</td>
<td>9.11</td>
<td>80.43</td>
<td>7.60</td>
<td>1.07</td>
<td>1.10</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0.16</td>
<td>0.29</td>
<td>0.42</td>
<td>8.38</td>
<td>80.82</td>
<td>3.76</td>
<td>6.17</td>
</tr>
<tr>
<td>CCC</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.77</td>
<td>1.54</td>
<td>12.31</td>
<td>70.00</td>
<td>15.38</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

to ratings that have been “stable”, reached through a recent “upgrade” in the last horizon and reached through a recent “downgrade” in the last horizon.⁸ The normal S&P transition matrix in Table 1 is thus extended to include states that are conditional on whether the last state movement was upward, downward or stable. As a consequence of the model, no corporate can directly reach an AAA rating through a downgrade because it is the highest possible rating. Also, no corporate can directly reach a CCC rating through an upgrade because it is the lowest rating. The last rating, D, which is an absorbing state, can only be reached through a downgrade.

To add rating momentum to the corporate factor model we need transition matrixes estimated conditionally on rating momentum. Bangia et al. (2000) estimate momentum transition matrixes on S&P data and find the conditional transition matrixes on past rating movements as in Tables 2–4 (corresponding to stable, downward and upward momentum transition matrixes, respectively). As Bangia et al. (2000) report the momentum hypothesis is supported by the data, as most downgrade probabilities for the downward momentum matrix are larger than the corresponding values in the unconditional matrix. The exact opposite is true for the upward momentum matrix, which exhibits smaller downgrade probabilities than the unconditional matrix. The upgrade probabilities of the upward momentum matrix for below-investment-grade classes are higher than for the unconditional matrix, while they are lower for investment grade categories. Overall, reduced upgrading and downgrade probabilities for the upward momentum matrix lead to an increased probability mass on the diagonal, while the downward momentum matrix displays the exact opposite. Only the maintain-momentum matrix does not display a systematic trend or differ to a great extent from the unconditional matrix in Table 1.

⁸As mentioned above, this approach can be expanded to tracking more events, such as immediate rating change or the number of rating changes in the last three years. In this approach we simply expand the event vector {S, U, D} accordingly.
### TABLE 3  Upward momentum transition matrix.

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>92.45</td>
<td>5.66</td>
<td>1.89</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0.39</td>
<td>95.33</td>
<td>4.28</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>A</td>
<td>0.26</td>
<td>1.29</td>
<td>93.54</td>
<td>4.91</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>3.78</td>
<td>90.12</td>
<td>4.94</td>
<td>0.87</td>
<td>0</td>
<td>0.29</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>1.38</td>
<td>8.30</td>
<td>85.81</td>
<td>3.46</td>
<td>0.69</td>
<td>0.35</td>
</tr>
<tr>
<td>B</td>
<td>0</td>
<td>0</td>
<td>1.96</td>
<td>5.88</td>
<td>5.88</td>
<td>82.35</td>
<td>1.96</td>
<td>1.96</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>—</td>
</tr>
</tbody>
</table>

### TABLE 4  Downward momentum transition matrix.

<table>
<thead>
<tr>
<th></th>
<th>AAA</th>
<th>AA</th>
<th>A</th>
<th>BBB</th>
<th>BB</th>
<th>B</th>
<th>CCC</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.60</td>
<td>90.48</td>
<td>7.74</td>
<td>0.60</td>
<td>0</td>
<td>0.60</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>AA</td>
<td>0</td>
<td>0.93</td>
<td>93.02</td>
<td>5.12</td>
<td>0.47</td>
<td>0.23</td>
<td>0</td>
<td>0.23</td>
</tr>
<tr>
<td>A</td>
<td>0</td>
<td>0.18</td>
<td>5.42</td>
<td>85.90</td>
<td>6.69</td>
<td>1.27</td>
<td>0</td>
<td>0.54</td>
</tr>
<tr>
<td>BBB</td>
<td>0</td>
<td>0</td>
<td>1.12</td>
<td>7.30</td>
<td>79.21</td>
<td>8.71</td>
<td>1.69</td>
<td>1.97</td>
</tr>
<tr>
<td>BB</td>
<td>0</td>
<td>0</td>
<td>0.24</td>
<td>0.73</td>
<td>6.05</td>
<td>75.79</td>
<td>7.99</td>
<td>9.20</td>
</tr>
<tr>
<td>B</td>
<td>0.52</td>
<td>0</td>
<td>1.05</td>
<td>0.52</td>
<td>4.19</td>
<td>9.95</td>
<td>42.93</td>
<td>40.84</td>
</tr>
<tr>
<td>CCC</td>
<td>0</td>
<td>—</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

### 3.1 Empirical analysis of the effect of rating momentum

In our empirical analysis of the effect of rating momentum we will again use a single-factor-model representation of the factor models as in (2.10). We aim to compare the effect of rating momentum against non-rating momentum using the stylized scenarios (significantly severe, severe, stable and upturn) we used in the empirical analysis of the model without rating momentum. As before, we will focus on the difference for different $R^2$, the credit quality of the issuer and scenario severity. The current status is a stable rating history, as we analyze behavior over longer time periods. To compare the rating-momentum model with the non-rating-momentum case we analyze the difference in the remaining unit notional between the non-rating-momentum case and the rating-momentum case. This means positive values reflect overestimates of remaining notional from the non-rating-momentum model, assuming the rating-momentum model is correct.

Figure 6 displays the difference in unit notional for initial rating grade AAA, A and BBB, respectively, and for $R^2$ equal to 0.2, 0.4, 0.6 or 0.8 for the severe scenario. For rating grade AAA we note that a longer survival is much more likely. Because of the severity of the scenario, eventually we will see almost all notional default...
FIGURE 6 Difference in unit notional exposure remaining for initial rating AAA, A and BBB for different model $R^2$ and severe scenario.
in both models, so they converge, except for the $R^2 = 0.2$ case, which has more severe scenario protection (see Figure 3). In Figure 6 only the AAA rating grade displays significant model differences for longer scenario horizons, although this is also true to some extent for the A rated exposure. Hence, while rating momentum affects the timing of the default, such that it occurs earlier for the AAA and A rated exposures (hence a larger defaulted portion of the exposure), the differences take time to materialize in the severe scenario case. This is an interesting aspect, as most stress testing is relatively short term, eg, between three and five years. However, considering a lower quality initial rating grade, eg, B, we observe that, while a fully defaulted portion of the unit notional occurs in this severe scenario, the difference in the effects of including and not including rating momentum can still be substantial. With rating momentum we consume the notional much more quickly, and, due to the lower quality initial credit rating and higher likelihood of further downgrade with rating momentum, we see the impact of rating momentum much more quickly than for the better quality rating grades. That is, for the B rated exposure the model difference is immediate.

Figure 7 shows the AAA rating grade unit notional difference for the first six horizons of the significant severe stress scenario. Due to the large shift in $Z = -3$ at horizon 3, we thereafter see immediate significant differences for the AAA rated exposure due to its downgrade. Naturally, with the significant severe scenario the model differences occur much earlier due to the early extreme shock values, especially at horizon 4, where the second extreme $Z = -3$ value occurs.

So far we have assumed the credit exposure is initially stable. That is, it has not recently been up- or downgraded. Figure 8 displays the severe scenario results for
rating case B, assuming it has recently been downgraded. Compared with the initial stable case for the B rated exposure in Figure 6, we have significantly more model differences, such that the combination of the B rated exposure being initially downgraded and using the corporate momentum model has significant effects. This is, of course, a consequence of the initial shift to the downgrade conditional transition matrix, which has more mass on further downgrade than the initially stable transition matrix.

While we have focused on individual corporate results, we can of course extrapolate these to portfolio effects. For portfolios dominated by stable AAA rated exposures with moderate to severe scenarios, the rating momentum will have little effect in the short term but potentially substantial effects in the long term. Similarly, for lower quality rating grade portfolios (that have a high probability of downgrade), we expect to see large, immediate model effects of the rating momentum, with significant underestimation of short-term loss projection. On the other hand, for significantly severe scenarios, with multiple sequential negative large shocks, the difference for good quality portfolios can be substantial, even in the short term. We do not show here any model difference results for the stylized stable and upturn scenario, but we did not find any significant difference with stable scenarios for good quality rating grades. For lower quality rating grades with low $R^2$ there is also only a minor difference. This is because with high $R^2$ there is significant protection from the defaulted portion in “good” and “stable” scenarios. Hence, we need significant idiosyncratic effects to cause large default portions, even with rating momentum. We can therefore conclude...
that rating momentum mainly has a significant effect in (very severe) stress scenarios. This is in contrast to best-estimate projections without significant decline. Of course, these results are consistent with the fact that rating momentum matters most for tail scenarios and, realistically, more for lower quality credit portfolios, given the time horizon or sequence of extreme \( Z \) shifts needed for differences to materialize for good quality portfolios. This is because rating-momentum transition matrixes mostly shift mass for poor quality rating grades. Hence, a large initial negative shock is needed in order for good quality portfolios to downgrade to poor quality ones, where rating momentum matters most.

4 SUMMARY AND CONCLUSIONS

The macroeconomic scenario analysis of credit models is gaining in importance, with new stress testing regimes and impairment rules focusing on loss projections under different macroeconomic environments. In this paper, we have focused on the factor model that is widely used in practice for (corporate) stress testing and scenario analysis. While the model’s sensitivity to \( R^2 \) is well known, it must be officially recognized, especially as there is concern over poor estimation of \( R^2 \) in stressed macroeconomic scenarios, which can protect the portfolio loss significantly.\(^9\) While the factor model has traditionally been implemented against Markovian, rating-history-independent, average long-term transition matrixes (or long-term default term structures), we have seen that taking into account the stylized fact of rating momentum can shorten the time to loss significantly compared with the non-rating-momentum case, causing loss underestimation in the model if the rating momentum is not captured. The exact effect depends, of course, on the scenario time horizon, severity and portfolio quality. It may take more time for differences to materialize for good quality portfolios if the scenario is not significantly severe with multiple sequential extreme shocks. However, the effect on lower rating quality portfolios can be almost immediate, with significant loss underestimation.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper. The opinions expressed in the paper are those of the authors and do not necessarily reflect the view of the SAS Institute.

---

REFERENCES


A model combination approach to developing robust models for credit risk stress testing: an application to a stressed economy

Georgios Papadopoulos

Department of Economics, Democritus University of Thrace, University Campus, 69100 Komotini, Greece; email: gpapad.gr@gmail.com

(Received June 28, 2016; revised August 17, 2016; accepted August 18, 2016)

ABSTRACT

An integral part of advanced stress-testing frameworks is the macrofinancial model, which maps the impact of macroeconomic shocks on bank-specific risk factors. The standard practice for the development of such models is the use of a single model. However, such an approach exposes the whole framework to model risk, since any individual model may be misspecified and/or affected by structural changes in ways that can be very difficult to predict in advance. To address this issue, we use a model combination approach to develop robust macrofinancial models for credit risk stress testing. The empirical part utilizes data from the Greek economy, which experienced a sharp change from normal to distressed conditions; this makes it a particularly challenging case to forecast. The results expose the inadequacy of any individual model to capture the evolution of credit risk and indicate that relying on a single model can result in significant misestimation of risk. This holds whether the estimation sample includes only precrisis information or data from both tranquil and stressed periods. However, model combination gives a better forecasting performance than
standard benchmark models. Overall, the proposed approach can lead to a more robust assessment of risk.

**Keywords:** model combination; stress testing; nonperforming loans (NPLs); forecast performance; structural change.

1 INTRODUCTION

The outbreak of the financial crisis in 2008 has spurred research to understand its root causes as well as develop new, and enhance existing, tools to prevent future crises. One such tool (or rather “toolbox”, according to Borio et al (2014)) is stress testing. A fundamental part of any modern stress-testing framework is the macrofinancial model, which maps the impact of macroeconomic shocks into bank-specific indicators related to various types of risk. Any such model should be able to capture the underlying relationship between the macroeconomic environment and the risk under study. It should also produce reliable forecasts of the potential risks to financial stability. In this paper, a model combination approach will be applied to develop robust macrofinancial models of credit risk, and their forecasting performances under adverse conditions will be evaluated and compared with those of standard individual models.

The primary objective of stress tests is to assess a system’s stability, with the more specific goal of “identify[ing] vulnerabilities in tranquil times and provid[ing] the basis for addressing them” (see Borio et al 2014). Although the usefulness of stress-testing exercises is established, there have been worries regarding their ability to both accomplish the aforementioned goal and identify the weaknesses of the financial system prior to the crisis (Haldane 2009; Galati and Moessner 2013; Borio et al 2014). This can be attributed to several factors, one of which is the ability of the employed econometric models to adequately capture the interactions among the variables under stress conditions, not least because the reduced-form models used rely on historical data, and such crisis events are rare, while data quality is generally low (Borio et al 2014). Therefore, the need for robust macrofinancial models that can produce reliable forecasts outside “normal” conditions is evident. In fact, Gadea Rivas and Perez-Quiros (2015) show that research provided clear evidence of the link between financial and real crises only after the financial crisis started and the profession had failed to foresee it. Put another way, models (structural or reduced-form ones) can successfully describe the past but are not able to infer the future.

The existing literature on models linking credit risk to macroeconomic conditions displays a significant degree of diversity in terms of the variables modelled and methodology used. The most frequently used proxies of credit risk are loan loss provisions and nonperforming loans (NPLs), along with their ratios to total loans.
This is mainly due to data-availability issues (Beck et al 2013). Two comprehensive surveys of the approaches followed by policymakers are those of Foglia (2009), which focuses on major central banks and supervisory authorities, and of Melecky and Podpiera (2012), which reviews the respective methods used by Central and South Eastern European Central Banks (CSEECB). The findings of both papers suggest that, at the time these surveys were conducted, the econometric methods used for the estimation of the macrofinancial models varied from simple ordinary least squares to time-series and panel data techniques. However, the main macroeconomic determinants of credit risk formed a rather small set, including gross domestic product (GDP) growth rate, unemployment rate, inflation and both short- and long-term interest rates. Also, despite minor country-specific differences, a common feature was the use of a single model to link credit risk with the macroeconomic environment.

A nonexhaustive list of the recent literature on this topic includes the following examples. Nkusu (2011), using panel data techniques on a data set from twenty-six advanced economies from 1998 to 2009, finds that NPLs are affected by a deterioration in the macroeconomic environment. In particular, there is a statistically significant relationship between a rise in NPLs and a decline in GDP growth, house and equity prices as well as a rise in unemployment. Louzis et al (2012) find that GDP growth, unemployment rate and lending rates have a significant impact on NPL growth in Greece. Vazquez et al (2012) identify the lagged values of NPLs and GDP growth rate as important determinants of credit risk in the Brazilian banking sector. In a multi-country study, Buncic and Melecky (2013) use a panel of fifty-four high- and middle-income countries and find that GDP growth rate, inflation and lending rate are the macroeconomic variables affecting NPLs. In the work of Beck et al (2013), the authors use static and dynamic panel data techniques to study the macroeconomic determinants of NPLs on a sample of seventy-five countries during the period 2000–10. Their findings suggest that real GDP growth, share prices, the exchange rate and the lending interest rate have a significant effect on NPLs. For an overview of the earlier literature, the works of Quagliariello (2009) and Sorge (2004) provide excellent references. Despite the heterogeneity in the methods used and variables examined, all of these studies have one feature in common: they make use of a single model to identify the factors affecting credit risk and forecast it.

Apart from in academia and policymaking institutions’ research, this single-model approach is also expected to be followed by individual financial intermediaries to forecast credit risk in their loan books. Although detailed information is inaccessible, “the documentation provided by the participating banks very clearly confirm[s] that virtually all institutions operate, indeed, with single equation approaches” (Gross and Población 2015). A notable exception to the single model approach is from the European Central Bank. Gross and Población (2015) use a Bayesian model averaging approach to model credit risk, proxied by corporate distance to default for eighteen
European Union (EU) countries. Their illustrative scenario-conditional forecasts show that the use of individual models might lead to the underestimation of risk, although these models are sound from an economic and econometric point of view.

Despite the existence of a general consensus regarding some key macroeconomic variables and their impact on credit risk, the model diversity used in the relevant literature indicates that no individual model predominates the field or is all-encompassing. A reasonable extension of the current approach for the development of a macrofinancial model for stress-testing purposes would be using an apparatus in order to combine these models and/or their output into a single final product.

Since the seminal paper of Bates and Granger (1969), the advantages of forecast combination have been explored and well documented in both empirical (McAdam and Hughes Hallett 1999; Stock and Watson 2001, 2004; Marcellino 2004; McAdam and McNelis 2005) and simulation (Winkler and Makridakis 1983; Gupta and Wilton 1987; Hendry and Clements 2004) studies. The works of Clemen (1989), Timmermann (2006) and Wallis (2011) provide an excellent overview of this. The most prominent finding is that, on average, combining forecasts results in improved performance when compared with the use of the best single model.

Timmermann (2006) summarizes several arguments from the respective literature advocating the use of forecast combinations, three of which make this methodology particularly appealing for adoption in a stress-testing framework. The first is that different models may be affected by the presence of structural breaks in various ways. This touches the very essence of stress testing, which is not to predict the average expected outcome, but rather to predict the effect of a severe, yet plausible, event (Drehmann 2008). Even with a long-enough time series (which is hardly the case in credit risk stress testing) capturing one or more full business cycles, it would be quite unrealistic to assume that past statistical relationships will hold under future stress events. Indeed, Alfaro and Drehmann (2009) find that most of the reduced-form models used in their study experience a structural break around the beginning of a crisis. Moreover, Mestre and McAdam (2010) find that changes in the underlying data-generation processes affect small time series and large-scale, institutional models alike. Therefore, combining forecasts from models that are affected differently by a break (such as the adverse scenario of a stress test) can result in improved forecasts, as shown by Hendry and Clements (2004), Pesaran and Timmermann (2007) and Kapetanios et al (2008), among others.

Another reason for combining forecasts, which is of special relevance to stress testing, is that this methodology addresses in an objective way the "forecaster's bias" issue: that is, the difference in the loss function on which each user may base their forecasts. On the one hand, it is plausible to assume that financial institutions have an incentive to use models that yield more optimistic results, thus underpredicting the evolution of risk under adverse conditions (as mentioned by Gross and Población
2015). On the other hand, supervisory authorities might prefer to overpredict the levels of risk for reasons of prudence. Thus, combining forecasts provides a balanced and objective framework that can merge these two opposing approaches, bridging the distance between policymakers and supervised institutions. In that context, Gross and Población (2015) suggest that the combined model forecast be used as a threshold by the supervisory authorities, which the models of the financial institutions must pass in a stress-testing exercise. This ensures that the risks will be more adequately captured, and a more-level playing field across banks with similar risk characteristics will be created.

The final argument (Timmermann 2006) is based on the recognition that individual models, being local simplifications of a complex data-generating process, might be affected by some form of misspecification bias. Consequently, it is unlikely that a specific model will always outperform every other one in the future, or that its parameters will remain constant over time. Hardy and Schmieder (2013), studying a very large sample of banks covering almost 170 countries in the period from 1996 to 2011, provide evidence that the relationship between GDP and credit losses is twice as strong under severe stress situations as it is under moderate stress situations. That leads them to the conclusion that “it is misleading to take a parameter from a moderate GDP shock and use it to project the effects of a severe shock”. Thus, combining forecasts can be regarded as a hedge against such misspecification biases, assisting in obtaining more robust results.

By using historical data covering both normal and severely deteriorated conditions, this paper contributes to the current literature on macrofinancial model development for credit risk stress testing in three main ways. First, it examines the potential misestimation of risk by individual models that are estimated in normal times. The models are estimated using standard econometric methodology, and the conditions of the stressed part of the period under study are used as an input to these models in order to forecast credit risk. To the best of our knowledge, no other study provides an assessment of this using real data. Second, it employs methods from the forecast combination literature and evaluates the forecast performance of linear combinations of the aforementioned models under the same adverse conditions. The objective is to construct a framework for the development of robust models that can provide more accurate estimates of risk under macroeconomic shocks, even if their constituent parts are estimated in generally tranquil times. Finally, our paper studies the potential and performance of the model combination approach in a data-rich case and compares this to that of individual models. The rest of the paper is organized as follows. In Section 2, we present the data used and discuss its evolution in time. In Section 3, we describe the methodology used, elaborating on the steps of the proposed approach from model space definition to model estimation and the formation of combinations. In Section 4, the results of the empirical application are reported and discussed. Section 5 concludes.
2 DATA

The analysis is based on a data set constructed from quarterly, publicly available data for both the bank-specific and the macroeconomic variables of the Greek economy. It spans from 2006 Q4 until 2015 Q4, thus covering both normal and turbulent periods. The analysis will be performed in two parts to examine the potential of the proposed model development approach. The first part covers the period from the beginning of the sample until the onset of the sovereign debt crisis in Greece, the beginning of which can be placed at 2010 Q2, during which the Greek government requested financial assistance from the euro area member states and the International Monetary Fund (European Commission 2010). The second part will use all available data apart from the last six quarters to construct the various models. In this way, the robustness of the results will be examined when information is available for tranquil as well as distressed times.

The source of the macroeconomic variables is Eurostat. The indicators include those regularly used by the authorities for stress-testing exercises (Foglia 2009; Hardy and Schmieder 2013; Jobst et al. 2013), such as GDP, Harmonised Index of Consumer Prices (HICP), inflation (INF), unemployment rate (URT) and government debt to GDP ratio (GDEBTR), as well as household consumption expenditure (HHCE), compensation of employees (CE), net disposable income (NDI) and long-term unemployment rate (ULT). Altogether, they cover a significant part of the macroeconomic conditions.

The frequently studied ratio of NPLs to total loans (NPL) is used as a proxy for credit risk. A large part of the related institutional practice (Foglia 2009; Melecky and Podpiera 2012) and academic literature (Quagliariello 2009; Sorge 2004) utilizes NPLs for modelling credit risk. This is mainly due to data-availability issues: relevant variables such as probabilities of default (PDs) are often scarce. Besides offering the advantage of longer time series, NPLs can also be used to approximate PDs through the use of simple rules of thumb (Hardy and Schmieder 2013). The data for NPLs is collected from the financial statements for each of the four banks that qualify as significant according to the criteria of Single Supervisory Mechanism (SSM; 2015). Special care was taken to ensure that the definition of NPLs remains constant and homogeneous across banks for the whole period under investigation. This led to an unbalanced data set, since, in their reports, banks adopted the definition of NPLs as loans in arrears above ninety days at different starting dates. This also resulted in the removal of one bank from the sample because its respective reports begin significantly later than 2006 Q4, which is the earliest date around which the remaining three banks make their data available.

All variables are transformed in year-on-year (YoY) growth rates in percent. For variables already expressed as ratios, such as NPLs, both unemployment indicators,
A model combination approach to developing robust models for credit risk stress testing

### TABLE 1 Descriptive statistics of the data in the two subsamples.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Subsample</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>NPL</td>
<td>1</td>
<td>0.529</td>
<td>1.091</td>
<td>0.686</td>
<td>-1.400</td>
<td>2.300</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5.057</td>
<td>4.034</td>
<td>3.700</td>
<td>0.300</td>
<td>17.200</td>
</tr>
<tr>
<td>GDP</td>
<td>1</td>
<td>2.679</td>
<td>3.860</td>
<td>4.051</td>
<td>-3.557</td>
<td>9.344</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-5.104</td>
<td>3.214</td>
<td>-5.844</td>
<td>-9.160</td>
<td>0.465</td>
</tr>
<tr>
<td>HHCE</td>
<td>1</td>
<td>4.875</td>
<td>4.615</td>
<td>6.692</td>
<td>-4.331</td>
<td>10.330</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-4.887</td>
<td>3.127</td>
<td>-4.726</td>
<td>-10.190</td>
<td>0.332</td>
</tr>
<tr>
<td>CE</td>
<td>1</td>
<td>5.221</td>
<td>3.738</td>
<td>5.798</td>
<td>-1.308</td>
<td>10.900</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-6.374</td>
<td>5.157</td>
<td>-7.086</td>
<td>-12.570</td>
<td>2.732</td>
</tr>
<tr>
<td>NDI</td>
<td>1</td>
<td>2.253</td>
<td>4.567</td>
<td>4.597</td>
<td>-8.140</td>
<td>6.959</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>-8.228</td>
<td>2.558</td>
<td>-7.827</td>
<td>-13.280</td>
<td>-4.308</td>
</tr>
<tr>
<td>INF</td>
<td>1</td>
<td>3.075</td>
<td>1.361</td>
<td>3.223</td>
<td>0.897</td>
<td>5.117</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.924</td>
<td>2.839</td>
<td>0.070</td>
<td>-2.710</td>
<td>6.273</td>
</tr>
<tr>
<td>URT</td>
<td>1</td>
<td>0.416</td>
<td>1.362</td>
<td>-0.500</td>
<td>-1.000</td>
<td>2.400</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.557</td>
<td>3.169</td>
<td>3.100</td>
<td>-2.000</td>
<td>7.300</td>
</tr>
<tr>
<td>ULT</td>
<td>1</td>
<td>-0.166</td>
<td>0.616</td>
<td>-0.438</td>
<td>-0.831</td>
<td>1.295</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>2.439</td>
<td>2.369</td>
<td>2.367</td>
<td>-1.808</td>
<td>5.981</td>
</tr>
<tr>
<td>GDEBTR</td>
<td>1</td>
<td>6.981</td>
<td>8.048</td>
<td>4.000</td>
<td>-3.800</td>
<td>18.800</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>8.009</td>
<td>14.110</td>
<td>12.100</td>
<td>-16.700</td>
<td>25.900</td>
</tr>
</tbody>
</table>

SD denotes standard deviation.

the HICP and government debt to GDP, their lag-4 seasonal differences are used: $y_t = x_t - x_{t-4}$. For the rest of the macroeconomic variables, their YoY growth rate is calculated as $y_t = 100(x_t - x_{t-4})/x_{t-4}$.

The economic rationale behind the expected form of interaction of various macroeconomic indicators with credit risk is discussed by Kalirai and Scheicher (2002), who provide a detailed analysis. Therefore, the increase of variables that are related to a thriving (weakening) economy is expected to be negatively (positively) associated with NPLs.

Table 1 presents the descriptive statistics of the macroeconomic variables and NPLs, split in two subsamples. The first spans from 2006 Q4 to 2010 Q1, while the second spans from 2010 Q2 to 2015 Q4. This division illustrates the abrupt change the economy experienced in ways that are very difficult, if not impossible, to predict in advance.

Table 1 reveals the weak macroeconomic fundamentals that prevailed in the economy from 2010 Q2 onward. All indicators related to a prosperous economy such as GDP, HHCE, CE and NDI are severely deteriorated, with their average and median YoY growth rates turning negative. The fact that their maximums are close to zero or negative highlights the degree of distress. On the other hand, the average values of both unemployment indicators increase significantly as the median reverses from
negative to positive. NPLs, which are fairly stable before 2010 Q2 and at low levels, as shown by this variable’s standard deviation and mean, respectively, undergo a sharp change, with its average YoY growth increasing almost tenfold. In Figure 1 are plotted some representative indicators illustrating the conditions pictured in Table 1.

These specific economic conditions can provide an appropriate basis for the subsequent analysis in two ways. First, the adverse macroeconomic situation of the second period can effectively serve the purpose of an adverse scenario, since it is – by definition – severe, plausible and self-consistent. Second, the state change from a generally normal economic environment to an economy under stress observed in the two subsamples makes this data set particularly suitable for studying the forecasting performance of macrofinancial models under macroeconomic shocks. It should be noted that the first subsample consists of thirty-two observations, which are distributed among the three individual banks. In particular, the respective numbers of observations for each bank are $N_1^1 = 12$, $N_1^2 = 10$ and $N_1^3 = 10$. In order to strike a balance between the need to include as many predictors as possible to study their impact on NPLs and avoid the risk of overfitting, the data is treated as a panel, and the models are estimated using panel data techniques.
3 METHODOLOGY

The proposed framework was inspired by forecast combination literature. The procedure involves several distinct steps: model space formation, model estimation, scenario-conditional forecast generation for each model, forecast combination and, finally, performance assessment and comparison with specific benchmark models.

3.1 Model space formation

The formation of the model space is bounded by the maximum size \( K \) of the model. This is guided by the rule of thumb of one-in-ten, regarding the predictor-to-observations ratio, in order to avoid the risk of overfitting (Harrell 2015). Therefore, for the period from 2006 Q4 to 2010 Q1, the maximum model size is set to \( K = 3 \), while for 2006 Q4–2014 Q1, \( K = 8 \).

The number of all \( k \)-combinations from the set of \( n = 9 \) (eight macroeconomic and the dynamic term \( \text{NPL}_{t-1} \)) variables forming the potential model space is

\[
M = \sum_{k=1}^{K} \binom{n}{k} = \sum_{k=1}^{K} \left( \frac{n!}{k!(n-k)!} \right).
\]  

The general form of the models is

\[
\text{NPL}_{i,t} = f(\text{NPL}_{i,t-1}, X_{i,t}, u_i, e_{i,t}),
\]

where \( X_{i,t} \) is the matrix of macroeconomic variables, \( u_i \) denotes the individual effects and \( e_{i,t} \) is the error term.

3.2 Model estimation

Each of these \( M \) models is estimated using the appropriate panel data estimator. For the static models, the choice between the fixed and random effects estimator is guided by Hausman’s specification test (Hausman 1978). The bias-corrected least squares dummy variables (LSDVC) estimator (Kiviet 1995; Bun and Kiviet 2003; Bruno 2005a) is used for the estimation of the dynamic panel data models. This estimator addresses the problem of bias in the estimation of the dynamic term’s coefficient (Nickell 1981); it is found to outperform other dynamic panel data estimators in short panels (Judson and Owen 1999; Bruno 2005b), as is the case for the application at hand. From the pool of \( M \) estimated models are discarded those that do not meet the expected sign criteria (as in Gross and Población 2015). The purpose of this constraint is to ensure that the scenario-conditional forecast of each model will be in line with the purposes of a stress-testing exercise (eg, an adverse scenario will lead to an increase in NPLs). The last step in order to approve a model’s inclusion in the final model...
space is to check the stationarity of its residuals. This test is performed using the Fisher-type tests proposed by Maddala and Wu (1999) and Choi (2001), which allow for unbalanced panels. Monte Carlo simulations by Maddala and Wu (1999) suggest that when the time dimension exceeds the number of individuals in a panel data set, a Fisher-type test has greater power than its counterparts.

The resulting set of $M^* \leq M$ models, which are economically (reflecting the expected relationships among variables) and econometrically (with well-behaved residuals) sound, is utilized to produce forecasts for NPLs for typical forecast horizons (six to eight quarters ahead, depending on the period under study) in stress-testing exercises (Schuermann 2014).

### 3.3 Forecast/model combination

For the combination of forecasts, three combination schemes are employed: the equal weights (average), the median and the Bates–Granger weights. Several studies (Palm and Zellner 1992; Stock and Watson 2001; Hendry and Clements 2004; Timmermann 2006) have shown that simple combination schemes, such as the average or median forecast, oftentimes outperform more complex approaches.

The equal weights (EW) combination scheme is defined as the sum of a group of distinct forecasts divided by their number:

$$f^{EW}_{t+h} = \frac{1}{M^*} \sum_{i=1}^{M^*} f_{i,t+h},$$

(3.3)

where $i = 1, \ldots, M^*$ is the number of forecasts and $f_{i,t+h}$ denotes the individual forecasts.

Apart from its very good empirical track record, the EW is shown to be optimal when the variance and pairwise correlation ratios between any individual forecast errors are close to unity (Timmermann 2006).

Another equally simple combination scheme that is found to perform satisfactorily (Armstrong 1989; Stock and Watson 2001; Hendry and Clements 2004) is the median forecast (Md). This is simply defined as

$$f^{Md}_{t+h} = \text{median}\{f_{i,t+h}\}_{i=1}^{M^*}.$$  

(3.4)

Both of these methods have the advantage that they do not require any estimation of weights, and thus avoid introducing parameter estimation errors.

In a more general setting, where the forecast error variances may be unequal, Bates and Granger (1969) propose the use of estimated weights based on pseudo out-of-sample forecast variances. The Bates–Granger weights are estimated as follows:

$$\omega^{BG}_{i} = \frac{\text{RMSE}_{i}^{-2}}{\sum_{i=1}^{M^*} \text{RMSE}_{i}^{-2}}.$$  

(3.5)
where \( \text{RMSE}_i \) is the pseudo out-of-sample root mean squared error of forecast \( i \), and \( M^* \) is the number of forecasts. Thus, the combined forecast under the Bates–Granger scheme is

\[
 f_{t+h}^{BG} = \sum_{i=1}^{M^*} \omega_i^{BG} f_{i,t+h}.
\]

Despite the fact that these weights are proven to be optimal when the errors of the combined forecasts are uncorrelated, it is found that this combination method performs adequately in practice, even when this condition is not met (Bjørnland et al 2011).

Additionally, the previous combination schemes are applied in subsets of the final model space, which include a number of the best-performing models according to their pseudo out-of-sample RMSEs. Many studies have found evidence in support of trimming the model space. Aiolfi and Favero (2005), for example, find that combining the top 20% of the models in their study gives the best results. Also, Granger and Jeon (2004) suggest trimming the model space by discarding a much smaller amount than Aiolfi and Favero (2005), ranging from 5% to 10% of the worst performers. Hence, the combinations of the top 20% and 80% of the models, ranked by their pseudo out-of-sample RMSEs, are formed, and their forecasting performance is evaluated.

When the underlying data is provided, forecast combination is a form of model combination. Another method for model combination, the potential use for stress testing of which has recently started being explored (Gross and Población 2015), is Bayesian model averaging. Despite being a promising approach and considering a similar view regarding the nonexistence of a single “true” model, it has two disadvantages. The first is that it is data demanding. In a data-constrained environment such as that in the empirical application, this approach is hardly feasible; this is because of the need for maximum likelihood estimation in order to form the combination weights. On the one hand, estimating separate time series models would be very restrictive in terms of the number of regressors included without the risk of overfitting. On the other hand, while the use of static panel data estimators is possible, the appropriate maximum likelihood dynamic panel data estimators demand greater numbers of individuals than those present in the studied data set. Therefore, the combination methods examined in the present study have the advantage of being feasible and applicable, even in the case of small banking systems (or systems dominated by very few players) and/or where the time series are short. The second drawback recognized in the literature (Levine et al 2012; Geweke and Amisano 2011, 2012) is more fundamental: Bayesian selection criteria can result in degenerate odds. That is, most of the weight is assigned to a single model, rendering the contribution of the rest negligible. The combination schemes used in this study exhibit more balanced behavior (with the EW
scheme being perfectly balanced), thus exploiting the information from an adequately large pool of models.

### 3.4 Performance assessment

The forecasting performance of the various combination schemes is compared with the root mean squared forecast error (RMSFE) of two benchmark models. The first is selected according to its forecasting performance, while the second is a model that includes standard indicators typically used in stress-testing exercises. The selection of the benchmark models reflects the standard procedure a stress tester would follow in order to choose a suitable individual model. The first benchmark model is that with the lowest pseudo out-of-sample RMSE. The second benchmark model is the largest model that includes GDP as a regressor, since this variable is the most commonly used for stress testing, and that for which the supervisory authorities always provide scenarios (Jobst et al 2013; Hardy and Schmieder 2013). Before estimating the respective RMSEs and RMSFEs, all forecasts are back-transformed from YoY growth rates to the level of NPLs.

### 4 EMPIRICAL RESULTS

In the empirical application, we study the forecasting performance, conditional on a real adverse scenario, of individual models estimated on an incomplete part of the business cycle. This reflects the reality for many macrofinancial models developed for credit risk stress testing (Segoviano Basurto and Padilla 2006; Ong et al 2010; Oura and Schumacher 2012; Gross and Población 2015). We examine whether combining models can be beneficial, resulting in improved forecasting performance compared with certain benchmark models. The analysis is conducted in two periods. In the first period of study (period 1), the models are estimated during tranquil times, and the subsequent macroeconomic shocks are used as input to make scenario-conditional forecasts. In the second period (period 2), the estimation sample exploits as much information as possible, including both tranquil and a significant part of deteriorated macroeconomic conditions. In this way, the performance of the individual and combined models is assessed in two settings that essentially cover the potential cases a stress tester might encounter.

---

1 The goodness-of-fit measure of mean absolute forecast error (MAFE) was also used to assess the performance, yielding similar results.
4.1 The case of estimation ahead of deteriorated economic conditions

The estimation sample starts as early as possible before the onset of the financial crisis in 2008, and stops on the brink of the Greek sovereign debt crisis in 2010 Q2. In order to exploit the available information to the maximum, the pseudo out-of-sample period used for model RMSE estimation is just 2010 Q2. Hence, following (3.1), each one of the $M_1 = 129$ models is estimated in the period from 2006 Q4 to 2010 Q1. The forecast horizon is set to $h = 8$ quarters, starting from 2010 Q3; this is a time interval typical for stress-testing purposes, since it is short enough to justify the static bank balance sheet assumption while, at the same time, being long enough to allow for the impact of the applied scenario to fully materialize (Schuermann 2014).

The values of the macroeconomic variables as they actually evolved from 2010 Q3 to 2012 Q2 are used for the construction of the adverse scenario. These are reported in Table A.1 in the online appendix. This choice serves several purposes. First, the scenario is, by construction, severe, plausible and self-consistent. As regards its severity, it can be seen in Table A.1 that all macroeconomic fundamentals (as we indicated earlier) are persistently weak, while variables related to unemployment and government debt are continuously rising. It should be noted that the sharp decline of GDEBTR is due to the restructuring of the Greek sovereign debt in 2012 Q1 (Eurogroup 2012). This is expected to affect the forecasts produced by the models that include this variable as a regressor. However, the overall macroeconomic conditions remain significantly deteriorated, a fact that can cause the purely model-based forecasts to diverge from the real evolution of NPLs after 2012 Q1. Second, using the conditions as they actually happened allows us to focus on how individual models would have reacted had the hypothesized adverse scenario coincided with the actual evolution of the macroeconomic conditions.

The models that have the expected signs and are econometrically sound are $M_{16}^*$ and form the final model space, as reported in Table A.5 in the online appendix. From Table A.5, it follows that the first benchmark model, with the lowest pseudo out-of-sample RMSE, is $M_3$. The second benchmark model is $M_7$, which is the largest model with GDP as a regressor. These two specific models are mentioned as $BM_1$ and $BM_2$, respectively. The weights for each combination scheme are reported in Table A.2 in the online appendix.

The historical values of Table A.1 are used as an input for the $M_{16}^*$ models to create the scenario-conditional forecasts. Their combinations are formed by applying the weights of Table A.2 to the respective models. The RMSFEs are estimated for each model and combination scheme at each time step. The performance of the combined models is reported in Table 2 in terms of RMSFE, relative to the respective benchmark model. A ratio above 1 indicates that the benchmark model performs better. The
opposite holds when the ratio is below 1. \( \text{EW}_1, \text{Md}_1 \) and \( \text{BG}_1 \) denote the equal, median and Bates–Granger weights, respectively. The number following the name signifies the percentage of the top models used in that specific combination scheme. The combination of the full model space is denoted without any numbers.

There are several interesting findings in Table 2. A general observation is that, on average, \( \text{EW}_1 \cdot 20, \text{BG}_1, \text{BG}_1 \cdot 80 \) and \( \text{Md}_1 \cdot 80 \) outperform both benchmark models. Closer examination of the results shows that all combination schemes perform better than \( \mathcal{B} \cdot \mathcal{M}_1 \) after 2011 Q3, with the simple ones (\( \text{EW}_1, \text{EW}_1 \cdot 80, \text{Md}_1 \) and \( \text{Md}_1 \cdot 80 \)) achieving approximately 70% and 80% lower RMSFEs for the last two forecast steps, respectively. This pattern reverses for \( \mathcal{B} \cdot \mathcal{M}_2 \). The performance of the majority of combination schemes is better than \( \mathcal{B} \cdot \mathcal{M}_2 \) until 2011 Q3, but it deteriorates increasingly after this. Both patterns are attributable to the restructuring of Greek sovereign debt. All models (including \( \mathcal{B} \cdot \mathcal{M}_1 \)) with GDEBTR as a regressor are affected by this restructuring, while the rest (such as \( \mathcal{B} \cdot \mathcal{M}_2 \)) better capture the evolution of NPLs due to the elevated uncertainty reflected by the other macroeconomic indicators.

Trimming the model space generally improves performance, but the results are mixed. For the Bates-Granger scheme, \( \text{BG}_1 \) exhibits better average performance than \( \text{BG}_1 \cdot 20 \) and \( \text{BG}_1 \cdot 80 \) for any benchmark. For the other combination schemes, the top 20% model combination has a better average performance than the top 80% and full model space combinations for the case of \( \mathcal{B} \cdot \mathcal{M}_1 \), while the opposite holds for \( \mathcal{B} \cdot \mathcal{M}_2 \). An important observation is that the \( \text{EW}_1 \cdot 80 \) and \( \text{Md}_1 \cdot 80 \) schemes outperform \( \mathcal{B} \cdot \mathcal{M}_2 \) in every step of the forecast horizon and, on average, both \( \mathcal{B} \cdot \mathcal{M}_1 \) and \( \mathcal{B} \cdot \mathcal{M}_2 \). An overall picture is presented in Figure 2.

It is evident from Figure 2 that \( \mathcal{B} \cdot \mathcal{M}_1 \) always underpredicts (in an ever-increasing way) the true evolution of NPLs under stress, yielding substantially milder results. On the contrary, \( \mathcal{B} \cdot \mathcal{M}_2 \) overpredicts NPLs for the first two banks (B1, B2) over the whole forecast horizon as well as for almost half of the forecast steps for the third (B3) bank. The gray-shaded area in Figure 2 marks the range of the scenario-conditional forecasts from the individual models. As one moves away from the forecast origin, the range (and, thus, the uncertainty) of the predictions from the individual models increases. However, the representative \( \text{Md}_1 \cdot 80 \) exhibits a balanced behavior, and its forecasts remain reasonably close to the materialized values of NPLs.

Figure 2 shows that the use of a single model can potentially lead to significant misestimation of risk. While the consequences of underestimating it are evident, overestimation can jeopardize financial stability as well. In particular, from a policymaker’s point of view, it is crucial that the support measures needed to repair the weaknesses in banks’ balance sheets are available (Angeloni 2014). Thus, disclosing excessively pessimistic results, and needing a significant amount of resources to fix them, might undermine market confidence and adversely affect financial stability (Oura and Schumacher 2012).
### TABLE 2
The forecasting performance of the combined models in period 1.

<table>
<thead>
<tr>
<th>$h_1$</th>
<th>Benchmark (BM)</th>
<th>$EW_1$</th>
<th>$EW_{120}$</th>
<th>$EW_{180}$</th>
<th>$BG_1$</th>
<th>$BG_{120}$</th>
<th>$BG_{180}$</th>
<th>$Md_1$</th>
<th>$Md_{120}$</th>
<th>$Md_{180}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010 Q3</td>
<td>$BM_1$</td>
<td>2.241</td>
<td>0.738</td>
<td>1.757</td>
<td>0.732</td>
<td>0.834</td>
<td>0.692</td>
<td>2.565</td>
<td>0.827</td>
<td>1.876</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.964</td>
<td>0.317</td>
<td>0.756</td>
<td>0.315</td>
<td>0.358</td>
<td>0.298</td>
<td>1.103</td>
<td>0.356</td>
<td>0.807</td>
</tr>
<tr>
<td>2010 Q4</td>
<td>$BM_1$</td>
<td>1.033</td>
<td>0.773</td>
<td>0.704</td>
<td>0.503</td>
<td>0.874</td>
<td>0.563</td>
<td>1.199</td>
<td>0.892</td>
<td>0.806</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>1.009</td>
<td>0.755</td>
<td>0.688</td>
<td>0.492</td>
<td>0.854</td>
<td>0.550</td>
<td>1.172</td>
<td>0.871</td>
<td>0.788</td>
</tr>
<tr>
<td>2011 Q1</td>
<td>$BM_1$</td>
<td>1.393</td>
<td>0.779</td>
<td>0.946</td>
<td>0.464</td>
<td>0.878</td>
<td>0.535</td>
<td>1.522</td>
<td>0.904</td>
<td>0.986</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.948</td>
<td>0.530</td>
<td>0.644</td>
<td>0.316</td>
<td>0.598</td>
<td>0.364</td>
<td>1.037</td>
<td>0.615</td>
<td>0.671</td>
</tr>
<tr>
<td>2011 Q2</td>
<td>$BM_1$</td>
<td>2.108</td>
<td>0.599</td>
<td>1.520</td>
<td>0.259</td>
<td>0.757</td>
<td>0.267</td>
<td>2.343</td>
<td>0.697</td>
<td>1.520</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.956</td>
<td>0.272</td>
<td>0.690</td>
<td>0.118</td>
<td>0.344</td>
<td>0.121</td>
<td>1.063</td>
<td>0.316</td>
<td>0.689</td>
</tr>
<tr>
<td>2011 Q3</td>
<td>$BM_1$</td>
<td>0.797</td>
<td>0.819</td>
<td>0.583</td>
<td>0.583</td>
<td>0.888</td>
<td>0.637</td>
<td>0.871</td>
<td>0.849</td>
<td>0.573</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.937</td>
<td>0.963</td>
<td>0.686</td>
<td>0.685</td>
<td>1.044</td>
<td>0.749</td>
<td>1.024</td>
<td>0.998</td>
<td>0.674</td>
</tr>
<tr>
<td>2011 Q4</td>
<td>$BM_1$</td>
<td>0.720</td>
<td>0.840</td>
<td>0.544</td>
<td>0.618</td>
<td>0.903</td>
<td>0.670</td>
<td>0.786</td>
<td>0.875</td>
<td>0.600</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.948</td>
<td>1.106</td>
<td>0.717</td>
<td>0.814</td>
<td>1.189</td>
<td>0.883</td>
<td>1.036</td>
<td>1.153</td>
<td>0.790</td>
</tr>
<tr>
<td>2012 Q1</td>
<td>$BM_1$</td>
<td>0.340</td>
<td>0.843</td>
<td>0.328</td>
<td>0.644</td>
<td>0.900</td>
<td>0.687</td>
<td>0.367</td>
<td>0.848</td>
<td>0.332</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.867</td>
<td>2.149</td>
<td>0.836</td>
<td>1.642</td>
<td>2.295</td>
<td>1.753</td>
<td>0.937</td>
<td>2.163</td>
<td>0.847</td>
</tr>
<tr>
<td>2012 Q2</td>
<td>$BM_1$</td>
<td>0.204</td>
<td>0.843</td>
<td>0.249</td>
<td>0.643</td>
<td>0.899</td>
<td>0.685</td>
<td>0.216</td>
<td>0.840</td>
<td>0.217</td>
</tr>
<tr>
<td></td>
<td>$BM_2$</td>
<td>0.814</td>
<td>3.360</td>
<td>0.939</td>
<td>2.561</td>
<td>3.583</td>
<td>2.730</td>
<td>0.859</td>
<td>3.346</td>
<td>0.863</td>
</tr>
</tbody>
</table>
FIGURE 2  The historical values of NPLs with the scenario-conditional forecasts of individual models and their Md180 and EW180 combinations for each bank.
4.2 The case of estimation including deteriorated economic conditions

To study the robustness of the results, the same procedure is applied using the maximum amount of data, covering tranquil and turbulent periods almost evenly. The estimation sample ranges from 2006 Q4 to 2014 Q1, and 2014 Q2 is used for pseudo out-of-sample RMSE estimation. The predictor NDI is not considered due to missing data toward the end of the forecast horizon. Following \(\text{(3.1)}\), the model space consists of \(M_2 = 255\) models.

The forecast horizon is \(h_2 = 6\) from 2014 Q3 until 2015 Q4, and the actual values of the macroeconomic indicators during that period, reported in Table A.3 (available online), are used as a scenario. The final model space consists of \(M^*_2 = 20\) models presented in Table A.6 in the online appendix. \(M^*_2\) exhibits the lowest pseudo out-of-sample RMSE and, thus, will be the first benchmark (\(BM_2\)). The largest model including GDP as a predictor is \(M_3^{13}\), which makes it the second benchmark (\(BM_2^{13}\)), as previously mentioned. The combinations are again formed by applying the respective weights of Table A.4 in the online appendix to the respective models. The results are presented in Table 3, in the same way as before.

The pattern in Table 3 is similar to that of the previous case. \(BM_2\) exhibits lower RMSFE than model combinations initially, but this changes quickly. \(BG_2\), \(BG_3^{80}\) and \(Md_2^{80}\) outperform it when \(h_2 \geq 2015\) Q1, while all top-20% schemes perform better when \(h_2 \geq 2014\) Q4. Compared with \(BM_2\), all model combinations display better average performances. Specifically, \(EW_2\) and \(EW_2^{80}\) outperform \(BM_2\) for the whole \(h_2\), while the rest of the model combinations perform remarkably well for the majority of the forecast steps. Despite the fact that more information is used in this case, it is a very challenging environment to forecast. Economic and political tensions and bank idiosyncratic reasons have an impact on NPLs, while the purely model-based forecasts are affected by events, such as the reduction of GDEBTR due to scheduled debt repayments, or a minor decrease in unemployment indicators. However, overall, model combinations exhibit a very good performance, and they provide more accurate forecasts than their individual counterparts. Figure 3 provides a visual representation of the above findings.

5 CONCLUSIONS

We implemented a model combination approach for the development of macrofinancial models in the context of a credit risk stress-testing framework. The intuition is that, since any individual model is, at best, a local approximation of the true data-generating process, a pool of models will carry useful extra information and, therefore, outperform a single model. While model combination methods have been applied successfully in many scientific fields in general (Rosenfeld 2000), and within economics...
<table>
<thead>
<tr>
<th></th>
<th>Benchmark</th>
<th>$h_2$</th>
<th>EW</th>
<th>EW$_{20}$</th>
<th>EW$_{80}$</th>
<th>BG</th>
<th>BG$_{20}$</th>
<th>BG$_{80}$</th>
<th>Md</th>
<th>Md$_{20}$</th>
<th>Md$_{80}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014 Q3</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>7.139</td>
<td>2.334</td>
<td>5.895</td>
<td>3.885</td>
<td>1.631</td>
<td>3.405</td>
<td>5.902</td>
<td>2.468</td>
<td>4.833</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.856</td>
<td>0.280</td>
<td>0.707</td>
<td>0.466</td>
<td>0.196</td>
<td>0.408</td>
<td>0.707</td>
<td>0.296</td>
<td>0.579</td>
</tr>
<tr>
<td>2014 Q4</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>2.073</td>
<td>0.954</td>
<td>1.645</td>
<td>1.188</td>
<td>0.949</td>
<td>1.085</td>
<td>1.583</td>
<td>0.962</td>
<td>1.178</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.757</td>
<td>0.348</td>
<td>0.601</td>
<td>0.434</td>
<td>0.347</td>
<td>0.396</td>
<td>0.578</td>
<td>0.351</td>
<td>0.430</td>
</tr>
<tr>
<td>2015 Q1</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>1.552</td>
<td>0.886</td>
<td>1.210</td>
<td>0.945</td>
<td>0.917</td>
<td>0.896</td>
<td>1.060</td>
<td>0.876</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.710</td>
<td>0.405</td>
<td>0.553</td>
<td>0.432</td>
<td>0.419</td>
<td>0.410</td>
<td>0.485</td>
<td>0.401</td>
<td>0.400</td>
</tr>
<tr>
<td>2015 Q2</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>1.121</td>
<td>0.772</td>
<td>0.822</td>
<td>0.633</td>
<td>0.848</td>
<td>0.639</td>
<td>0.759</td>
<td>0.744</td>
<td>0.610</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.653</td>
<td>0.450</td>
<td>0.479</td>
<td>0.369</td>
<td>0.494</td>
<td>0.373</td>
<td>0.442</td>
<td>0.433</td>
<td>0.355</td>
</tr>
<tr>
<td>2015 Q3</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.558</td>
<td>0.900</td>
<td>0.588</td>
<td>0.712</td>
<td>0.927</td>
<td>0.764</td>
<td>0.597</td>
<td>0.882</td>
<td>0.788</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.769</td>
<td>1.241</td>
<td>0.810</td>
<td>0.981</td>
<td>1.278</td>
<td>1.054</td>
<td>1.082</td>
<td>1.215</td>
<td>1.087</td>
</tr>
<tr>
<td>2015 Q4</td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.651</td>
<td>0.920</td>
<td>0.693</td>
<td>0.788</td>
<td>0.944</td>
<td>0.826</td>
<td>0.708</td>
<td>0.906</td>
<td>0.832</td>
</tr>
<tr>
<td></td>
<td>BM</td>
<td>B.M$_{1/2}$</td>
<td>0.933</td>
<td>1.318</td>
<td>0.992</td>
<td>1.129</td>
<td>1.351</td>
<td>1.182</td>
<td>1.014</td>
<td>1.298</td>
<td>1.191</td>
</tr>
</tbody>
</table>
FIGURE 3  The historical values of NPLs with the scenario-conditional forecasts of
individual models and their Md280 and EW280 combinations for each bank.
in particular (Moral-Benito 2013), their potential use for stress-testing purposes has only recently started to be explored (Gross and Población 2015).

Apart from the expected improvement in forecasting performance, there are arguments in support of the use of model combination that are particularly relevant for the macrofinancial models used in stress testing. First, individual models might experience the impact of a structural break in various degrees. Second, models might suffer from misspecification bias because their estimated parameters may not be invariable over time and across the various states of an economy. Finally, model combinations can address the “forecaster’s bias” issue by providing an objective way to merge forecasts from users who would prefer to err on the positive side with those who would prefer to err on the negative side.

In the empirical part of the study, the recent experience of the financial and sovereign debt crises in the Greek economy was used, as a kind of “laboratory”, to test the forecasting performance of individual models and their combinations under real, extreme conditions. This is a particularly challenging application, since the economy experienced an abrupt state change from generally normal conditions to severely deteriorated ones. Individual models linking NPLs with macroeconomic indicators were estimated from the earliest possible moment until the wake of the financial crisis; these were used to produce forecasts using the adverse economic conditions that followed as a stress scenario. The analysis fully confirmed the previous arguments regarding the risks of using a single model. Model combinations outperformed all benchmark models when subjected to the adverse scenario. In addition, the forecasts from model combinations exhibited balanced behaviors between too-mild and too-conservative estimations of risk. The application showed that the dominance of model combinations over individual models also holds in the case where a sufficient amount of information is used to estimate each model. In line with the existing literature, our analysis also demonstrates that simple combination schemes, such as the EW and median schemes, perform remarkably well. Moreover, trimming the model space results in an improved performance. Out of the two levels that were examined, the best results were obtained from the less-aggressive level of trimming.

An apparent disadvantage of the proposed approach is the potential combinatorial explosion of the model space, which can make its application very intensive in computational resources as the number of predictors increases. The maximum number of models estimated in this study, conditional on limits to avoid overfitting, was 255. Although model spaces of this size can be easily managed by modern computers, the $2^n$ combinations of $n$ possible predictors can make this approach challenging. One way to narrow down the set of possible predictors would be to use the relevant literature as a guide as to which variables are the key determinants and focus on these. Another alternative is to use stochastic search techniques (Furnival and Wilson 2000) to reduce the number of models that need to be estimated.
This study focused on the development of robust models for credit risk for use in a stress-testing framework. However, the proposed model combination approach is not limited to credit risk. Future research can examine its applicability and performance for other types of risk, in which the current standard practice is dominated by the use of a single model. Overall, this study showed that model combination can provide an adequate estimate of credit risk under extreme conditions, despite the fact that the individual models are estimated in generally tranquil times. Therefore, in an ever-changing environment, where a single, all-encompassing model is very difficult, or maybe impossible, to find, model combination provides a flexible framework that may pass Keynes’s test: it enables us to be roughly right, rather than precisely wrong.

DECLARATION OF INTEREST

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

This research has been cofinanced by the European Union (European Social Fund (ESF)) and Greek national funds through the Operational Program “Education and Lifelong Learning” of the National Strategic Reference Framework (NSRF): “Research Funding Program: THALES. Investing in knowledge society through the European Social Fund”.

REFERENCES


A model combination approach to developing robust models for credit risk stress testing


Rosenfeld, J. (2000). The butterfly that roared: chaos bedevils meteorological computer models. That's why even the best ones can't reliably predict more than 14 days ahead. Scientific American 11(1), 22–27.


Single Supervisory Mechanism (2015). The list of significant supervised entities and the list of less significant institutions. Report, SSM.


Research Paper

Point-in-time probability of default term structure models for multiperiod scenario loss projection

Bill Huajian Yang

Enterprise Stress Testing, Royal Bank of Canada, 155 Wellington Street West, Toronto, M5V 3H6, Canada; email: bill.yang@rbc.com

(Received May 24, 2016; revised July 29, 2016; accepted July 29, 2016)

ABSTRACT

Rating transition models have been widely used for multiperiod scenario loss projection for Comprehensive Capital Analysis and Review (CCAR) stress testing and International Financial Reporting Standard 9 (IFRS 9) expected credit loss estimation. Although the cumulative probability of default (PD) for a rating can be derived by repeatedly applying the migration matrix at each forward scenario sequentially, divergence between the predicted and realized cumulative default rates can be significant, particularly when the predicting horizon is extended. We propose approaches to model the forward PDs directly. The proposed models are structured via a credit index, representing the systematic risk for the portfolio explained by a series of macroeconomic variables, together with the risk sensitivity with respect to the credit index, for each rating and each forward term. We also propose a parameter estimation algorithm based on the maximum likelihood of observing the default frequency for each nondefault rating and each forward term. Our proposed models and approaches are validated on a corporate portfolio, to which a forward PD model and a point-in-time rating transition model are fitted. Both models accurately predict the portfolio quarterly default rate (ie, for a one-term horizon), but the term model in general outperforms the transition model as the predicting horizon increases (eg, for a two-year horizon), due to the fact
that the term model is calibrated over a longer horizon. We believe that the proposed models will provide practitioners with a new, robust tool for directly modeling the PD term structure for multiperiod scenario loss projection, for CCAR stress testing and for IFRS 9 expected credit loss estimation.

**Keywords:** Comprehensive Capital Analysis and Review (CCAR) stress testing; International Financial Reporting Standard 9 (IFRS 9) expected credit loss; probability of default (PD) term structure; forward PD; credit index; risk sensitivity.

1 INTRODUCTION

Let $p_k(t_k)$ denote the forward probability of default (PD) for a loan in the $k$th period $(t_{k-1}, t_k]$ after the initial observation time $t_0$, i.e., the conditional probability of default for the loan in this period given that the loan does not default before the period. Then the marginal PD for the loan in the $k$th period is given by

$$(1 - c_{k-1}(t_{k-1})) p_k(t_k),$$

where $c_{k-1}(t_{k-1})$ denotes the cumulative PD for the period $(t_0, t_{k-1}]$, and $(1 - c_{k-1}(t_{k-1}))$ is the survival probability for the loan in the period $(t_0, t_{k-1}]$.

Let $l_i(t_i)$ and $e_i(t_i)$ respectively denote the point-in-time (PIT) loss given default (LGD) and exposure at default (EAD) factors for the $i$th period after $t_0$. Let $f_i(t_i) = [l_i(t_i)][e_i(t_i)]$. Given the PIT PD term structure, the expected credit loss for a loan from $t_0$ up to the $k$th period can be estimated, assuming the PIT EAD and LGD term structures, by

$$\text{loss} = p_1(t_1) f_1(t_1) + (1-c_1(t_1)) p_2(t_2) f_2(t_2) + \cdots + (1-c_{k-1}(t_{k-1})) p_k(t_k) f_k(t_k).$$

(1.1)

Rating transition models (Bluhm and Overbeck 2007; Miu and Ozdemir 2009; Yang and Du 2015, 2016) have been widely used for multiperiod scenario loss projection for Comprehensive Capital Analysis and Review (CCAR) (Board of Governors of the Federal Reserve System 2016) stress testing and International Financial Reporting Standard 9 (IFRS 9) (Ankarath et al 2010; Basel Committee on Banking Supervision 2015a,b) expected credit loss estimation. Though cumulative PDs for a rating can be derived by repeatedly applying the migration matrix at each forward scenario sequentially, divergence between the predicted and realized cumulative default rates can be significant when the number of iterations increases (Bluhm and Overbeck 2007). The forward-looking PIT PD term structure comes into play as an option.

A credit index, as introduced in Yang and Du (2016) and summarized in the next section, is a linear combination of given macroeconomic variables that best predict
the default risk of the portfolio under some appropriate assumptions. The linear combination is normalized to have zero mean and one standard deviation. As shown in Theorem 2.2, forward PDs for a nondefault risk rating $R_i$ and a forward term can be structured via the credit index by using the following three types of parameters:

(a) the coefficients of macroeconomic variables for the credit index, which are common for all nondefault ratings and forward terms, at the portfolio level;

(b) the risk sensitivity with respect to the credit index for each rating and each forward term;

(c) the threshold value for each rating and each forward term.

The threshold values in (c) can be estimated separately (Lemma 2.1(b)). For the parameters in (a) and (b), we propose estimation approaches based on the maximum likelihood of observing the default frequency for each rating and each forward term.

The advantages of our proposed forward PD model for PD term structure include the following.

(1) Analytical formulations for forward PDs can be derived under the Merton model framework.

(2) The model is structured via a credit index, representing the part of systematic risk for the portfolio explained by a series of given macroeconomic variables, together with the risk sensitivity with respect to the credit index, for each rating and each forward term. This means, given the credit index, the model for a rating and a forward term is determined by the sensitivity and the threshold value (for the intercept).

(3) Parameter estimation is based on the maximum likelihood of observing historical forward term default frequency, which can be implemented by using, for example, the SAS procedure PROC NLMIXED (Wolfinger 2008).

The paper is organized as follows. In Section 2, we define the credit index for a portfolio, and derive the forward PD model under the Merton model framework. In Section 3, we show how a PD term structure can be derived based on forward PDs and how loss can be evaluated over a multiperiod scenario using the PD term structure. In Section 4, we determine the loglikelihood function for observing the term default frequency. In Section 5, we propose an algorithm for fitting the forward PD model. The proposed model and parameter estimation approaches are validated in Section 6, where we fit a forward PD model and a PIT rating transition model for a corporate portfolio. Backtest and out-of-sample test results are provided. Conclusions are given in Section 7.
2 PROPOSED MODELS FOR FORWARD PROBABILITY OF DEFAULT

Given a borrower with a nondefault risk rating \( R_i \) at \( t_0 \), assume the borrower did not default in the period \([t_0, t_{k-1}]\). We assume that the default risk for the borrower in the period \((t_{k-1}, t_k)\) is governed by a latent random variable \( z_{ik}(t) \), called the firm’s normalized asset value, which splits into two parts under the Merton model framework (Gordy 2003; Merton 1974; Miu and Ozdemir 2009; Vasicek 2002; Yang and Du 2015, 2016) as

\[
z_{ik}(t) = s(t) \sqrt{\rho_{ik}} + \varepsilon_{ik}(t) \sqrt{1 - \rho_{ik}},
\]

\[
0 < \rho_{ik} < 1, \ s(t) \sim N(0, 1), \ \varepsilon_{ik}(t) \sim N(0, 1), \quad (2.1)
\]

where \( s(t) \) denotes the systematic risk (common to all nondefault ratings and all terms) at time \( t \), and \( \varepsilon_{ik}(t) \) is the idiosyncratic risk independent of \( s(t) \). The quantity \( \rho_{ik} \) is called the asset correlation given the initial risk rating \( R_i \) and forward term number \( k \). It is assumed that there exist threshold values \( \{b_{ik}\} \) such that the borrower will default in the \( k \)th period \((t_{k-1}, t_k)\) if the normalized asset value \( z_{ik}(t) \) falls below the threshold value \( b_{ik} \). We call \( b_{ik} \) the default point for the \( k \)th forward term for a borrower whose initial risk rating is \( R_i \) at time \( t_0 \).

For simplicity, where there is likely to be no confusion we suppress the time notation \( t \) in \( z_{ik}(t) \), \( s(t) \) and \( \varepsilon_{ik}(t) \), and write these as \( z_{ik}, s \) and \( \varepsilon_{ik} \), respectively.

2.1 Forward PD

For a borrower with a nondefault initial risk rating \( R_i \), the \( k \)th forward PD is the conditional probability that the borrower defaults in the \( k \)th period \((t_{k-1}, t_k)\) given that the borrower does not default in the period \([t_0, t_{k-1}]\). For a given sample, the forward PD can be estimated by

\[
d_{ik}(t_k) / n_{ik}(t_k), \quad (2.2)
\]

where \( n_{ik}(t_k) \) denotes the number of borrowers who survived the period \([t_0, t_{k-1}]\) with an initial risk rating \( R_i \) at \( t_0 \), and \( d_{ik}(t_k) \) is the number of such borrowers who defaulted in the period \((t_{k-1}, t_k)\).

Let \( p_{ik}(s) \) denote the \( k \)th forward PD given the systematic risk \( s \) in the \( k \)th period. Under model (2.1), we have

\[
p_{ik}(s) = P(z_{ik} < b_{ik} \mid s) = \Phi \left[ \frac{b_{ik} - s \sqrt{\rho_{ik}}}{\sqrt{1 - \rho_{ik}}} \right], \quad (2.3)
\]
where $\phi$ denotes the standard normal cumulative distribution. Let

$$ r_{ik} = \frac{\sqrt{\rho_{ik}}}{\sqrt{1 - \rho_{ik}}}. \quad (2.4a) $$

which implies

$$ \rho_{ik} = \frac{r_{ik}^2}{1 + r_{ik}^2}, \quad \frac{1}{\sqrt{1 - \rho_{ik}}} = \sqrt{1 + r_{ik}^2}. \quad (2.4b) $$

By (2.3) and (2.4), we have

$$ p_{ik}(s) = \phi(b_{ik} \sqrt{1 + r_{ik}^2} - r_{ik}s). \quad (2.5) $$

We can interpret the quantity $r_{ik}$ as the risk sensitivity for the $k$th forward PD, namely, $p_{ik}(s)$, with respect to the systematic risk factor $s$.

Given a nondefault rating $R_i$ at $t_0$ and a forward term $k$, the risk sensitivity $r_{ik}$ can be estimated by maximizing the likelihood, given by (2.5), of observing the default frequency for the rating and the forward term, using, for example, the SAS procedure PROC NLMIXED (Wolfinger 2008; Yang and Du 2015, 2016).

### 2.2 The proposed forward PD models

Let $E_s(\cdot)$ denote the expectation with respect to $s$. The threshold value $b_{ik}$ can be derived from the through-the-cycle (TTC) average of the $k$th forward PDs, as shown in Lemma 2.1(b).

**Lemma 2.1**

(a) We have

$$ E_s[\phi(a_0 + a_1s)] = \phi\left(\frac{a_0}{\sqrt{1 + a_1^2}}\right), \text{ where } s \sim N(0, 1). $$

(b) $\phi(b_{ik}) = E_s[p_{ik}(s)]$.

**Proof of Lemma 2.1** Part (a) is given in Rosen and Saunders (2009). Part (b) follows from (2.5) by applying part (a). $\square$

Given a series of macroeconomic variables $x_1, x_2, \ldots, x_m$ with means $u_1, u_2, \ldots, u_m$, let $w(x)$ be a linear combination:

$$ w(x) = a_1 x_1 + a_2 x_2 + \cdots + a_m x_m. \quad (2.6) $$

Let $\tilde{x}_i = (x_i - u_i)$. Normalize $w(x)$ by setting the credit index for the portfolio to be

$$ \text{ci}(x) = \frac{w(x) - u}{v} = \frac{(a_1 \tilde{x}_1 + a_2 \tilde{x}_2 + \cdots + a_m \tilde{x}_m)}{v}. \quad (2.7) $$
where \( u \) and \( v \) denote the mean and standard deviation of \( w(x) \), respectively. We assume that, given the series of macroeconomic variables, the systematic risk factor \( s \) splits into two parts as follows:

\[
s = -\lambda \, \frac{\mu}{\sigma} - e \sqrt{1 - \lambda^2}
= -\left[ \lambda \left( \bar{a}_1 \bar{x}_1 + \bar{a}_2 \bar{x}_2 + \cdots + \bar{a}_m \bar{x}_m \right) + \sigma e \right], \quad e \sim N(0, 1), \quad 0 < \lambda < 1, \quad (2.8)
\]

where \( \bar{a}_i = a_i / v \) and \( \sigma = \sqrt{1 - \lambda^2} \).

By (2.5) and (2.8), we have

\[
p_{ik}(s) = \Phi \left[ b_{ik} \sqrt{1 + r_{ik}^2} + r_{ik} (\lambda \, \frac{\mu}{\sigma} + \sigma e) \right]
= \Phi \left[ b_{ik} \sqrt{1 + r_{ik}^2} + r_{ik} (\bar{a}_1 \bar{x}_1 + \bar{a}_2 \bar{x}_2 + \cdots + \bar{a}_m \bar{x}_m) + r_{ik} \sigma e \right]. \quad (2.9a)
\]

Let \( p_{ik}(x) = E[p_{ik}(s) \mid x] \) be the expected value of \( p_{ik}(s) \) given macroeconomic variables \( x = (x_1, x_2, \ldots, x_m) \). We call \( p_{ik}(x) \) the forward PD given scenario \( x \).

Applying Lemma 2.1(a) to (2.9a) and (2.9b), we obtain the following theorem for forward PDs.

**Theorem 2.2** Given a series of macroeconomic variables \( x_1, x_2, \ldots, x_m \), assume that the residual \( e \) in (2.8) is independent of \( x_1, x_2, \ldots, x_m \). Under (2.1), we have

\[
p_{ik}(x) = \Phi \left[ b_{ik} \sqrt{1 + \tilde{r}_{ik}^2} + \tilde{r}_{ik} \mu \mu \right]
= \Phi \left[ b_{ik} \sqrt{1 + \tilde{r}_{ik}^2} + \tilde{r}_{ik} (\bar{a}_1 \bar{x}_1 + \bar{a}_2 \bar{x}_2 + \cdots + \bar{a}_m \bar{x}_m) \right]. \quad (2.10a)
\]

where

\[
\tilde{r}_{ik} = \frac{r_{ik} \lambda}{\sqrt{1 + r_{ik}^2 \sigma^2}} = \frac{\lambda}{\sqrt{1 + r_{ik}^2 (1 - \lambda^2)}}. \quad (2.11)
\]

**Proof** By (2.11), the definition of \( \tilde{r}_{ik} \), we have

\[
\sqrt{1 + \tilde{r}_{ik}^2} = \frac{\sqrt{1 + r_{ik}^2}}{\sqrt{1 + r_{ik}^2 (1 - \lambda^2)}} \implies b_{ik} \frac{\sqrt{1 + r_{ik}^2}}{\sqrt{1 + r_{ik}^2 (1 - \lambda^2)}} = b_{ik} \sqrt{1 + \tilde{r}_{ik}^2}.
\]

We need only to show (2.10a). Applying Lemma 2.1(a) to (2.9a), we have

\[
p_{ik}(x) = E[p_{ik}(s) \mid x]
= \Phi \left[ b_{ik} \sqrt{1 + \tilde{r}_{ik}^2} + \tilde{r}_{ik} \mu \mu \right] + \mu \mu \tilde{r}_{ik} \lambda \frac{\lambda}{\sqrt{1 + \tilde{r}_{ik}^2 (1 - \lambda^2)}} \right]
= \Phi \left[ b_{ik} \sqrt{1 + \tilde{r}_{ik}^2} + \tilde{r}_{ik} \mu \mu \right].
\]
There are many choices for (2.7). Given the asset correlations \( \{ \rho_{jk} \} \) in (2.1) (thus \( \{ r_{ik} \} \)), we define the credit index for a portfolio to be the ci(x) in (2.7) satisfying the following conditions.

(a) The residual \( e \) in (2.8) is independent of \( x_1, x_2, \ldots, x_m \).

(b) \( ci(x) \) is normalized from a linear combination \( a_1 \tilde{x}_1 + a_2 \tilde{x}_2 + \cdots + a_m \tilde{x}_m \) with which the model \( \{ \tilde{p}_i(x) \} \) best predicts (via the maximum likelihood, as stated more precisely in Section 5) the default probability of the portfolio, where

\[
\tilde{p}_i(x) = \Phi[c_i + r_{i1}(a_1 \tilde{x}_1 + a_2 \tilde{x}_2 + \cdots + a_m \tilde{x}_m)]
\]

(2.12)

is a model predicting the default probability for the initial rating \( R_i \) in the one-term horizon, and the corresponding risk sensitivity \( \tilde{r}_{i1} \) is driven by (2.11). No constraint is imposed for \( \{ a_1, a_2, \ldots, a_m \} \) or the intercept parameters \( \{ c_{ik} \} \).

**Remark 2.3** The forward PDs in models (2.10a), (2.10b) are given after the portfolio credit index is determined. The fact that no constraint is imposed for intercepts \( \{ c_{ik} \} \) ensures a full optimization is possible for parameters \( \{ a_1, a_2, \ldots, a_m \} \).

**Remark 2.4** The portfolio credit index is fitted by targeting the portfolio default risk for only a single term horizon. It can be extended to cover a longer horizon when data sparsity is not an issue and the risk pattern is persistent for the extended horizon.

Similarly to the quantities \( r_{ik} \) and \( \rho_{ik} \), which are defined with respect to the systematic risk factor \( s \) by (2.1), the quantity \( \tilde{r}_{ik} \) can be interpreted as the risk sensitivity for the \( k \)th forward PD with respect to the credit index \( ci(x) \), and a quantity \( \tilde{p}_{ik} \) can be defined by

\[
\tilde{p}_{ik} = \rho_{ik} \lambda^2.
\]

**Proposition 2.5** The following three equations hold:

\[
\tilde{r}_{ik} = \frac{\sqrt{\rho_{ik}}}{\sqrt{1 - \rho_{ik}}}, \quad \tilde{p}_{ik} = \frac{\tilde{r}_{ik}^2}{1 + \tilde{r}_{ik}^2}, \quad \frac{1}{\sqrt{1 - \rho_{ik}}} = \sqrt{1 + \tilde{r}_{ik}^2}.
\]

**Proof of Proposition 2.5** We show only the first relation. Note that \( \sigma^2 = 1 - \lambda^2 \).

By (2.11), we have

\[
\tilde{r}_{ik} = \frac{r_{ik} \lambda}{\sqrt{1 + r_{ik}^2 \sigma^2}} \implies 1 + \tilde{r}_{ik}^2 = 1 + \frac{r_{ik}^2 \lambda^2}{1 + r_{ik}^2 \sigma^2} = \frac{1 + r_{ik}^2}{1 + r_{ik}^2 \sigma^2}
\]

\[
\implies \frac{\tilde{r}_{ik}^2}{1 + \tilde{r}_{ik}^2} = \frac{r_{ik}^2 \lambda^2}{1 + r_{ik}^2 \sigma^2}.
\]

(2.13)
By (2.4a), we have
\[ r_{ik} = \frac{\sqrt{\rho_{ik}}}{\sqrt{1 - \rho_{ik}}} \Rightarrow \rho_{ik} = \frac{r_{ik}^2}{1 + r_{ik}^2}. \] (2.14)

By (2.13) and (2.14), we have
\[ \frac{r_{ik}^2}{1 + r_{ik}^2} = \rho_{ik}\lambda^2 = \tilde{\rho}_{ik} \Rightarrow \tilde{r}_{ik} = \frac{\sqrt{\rho_{ik}}}{\sqrt{1 - \tilde{\rho}_{ik}}}. \]

Consequently, by (2.10a) and (2.10b), for the determination of the forward PDs \( \{ p_{ik}(x) \} \), the following parameters are required:

(a) parameters \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) for macroeconomic variables in credit index \( ci(x) \), common to all nondefault ratings and all forward terms;

(b) risk sensitivities \( \{ \tilde{r}_{ik} \} \), with one sensitivity for each nondefault risk rating and each forward term;

(c) threshold values \( \{ b_{ik} \} \), with one value for each nondefault risk rating and each forward term.

The threshold values \( \{ b_{ik} \} \) can be estimated separately by using Lemma 2.1(b). Therefore, the key to the probabilities \( \{ p_{ik}(x) \} \) is the determination of parameters \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) and \( \{ \tilde{r}_{ik} \} \).

**Remark 2.6** When the number of ratings is large and data sparsity is an issue, fitting the rating level sensitivities \( \{ \tilde{r}_{ik} \} \) could be a problem. In practice, we can regroup the risk ratings into fewer classes, e.g., into the grades “investment”, “subinvestment” and “problematic”, while the forward term numbers can be regrouped based on the risk patterns observed from the historical term structure. For example, for quarterly terms, forward term numbers can be regrouped into \( \{ 1 \} \), \( \{ 2 \} \), \( \{ 3, 4 \} \) and one group for every four consecutive terms after time \( t_4 \).

### 2.3 A review of the benchmark PIT rating transition probability models

The first PIT rating transition probability model was proposed by Miu and Ozdemir (2009). This was extended by Yang and Du (2015, 2016) to facilitate rating level asset correlation.

Let \( t_{ij}(x) \) denote the expected value of transition probability from an initial rating \( R_i \) at \( t_0 \) to rating \( R_j \) at the end of the horizon, given macroeconomic variables.
\( x = (x_1, x_2, \ldots, x_m) \). Under the Merton model framework (with \( k \) in (2.1) being set to 1), it can be shown (Yang and Du 2016), similarly to (2.10a), (2.10b), that

\[
\begin{align*}
t_{ij}(x) &= \Phi(\tilde{q}_{i(k-j+1)} + \tilde{r}_i \text{ci}(x)) - \Phi(\tilde{q}_{i(k-j)} + \tilde{r}_i \text{ci}(x)) \\
&= \Phi[\tilde{q}_{i(k-j+1)} + \tilde{r}_i (\tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 + \cdots + \tilde{a}_m \tilde{x}_m)] \\
&\quad - \Phi[\tilde{q}_{i(k-j)} + \tilde{r}_i (\tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 + \cdots + \tilde{a}_m \tilde{x}_m)].
\end{align*}
\]

where \( \tilde{q}_{ih} = q_{ih} \sqrt{1 + \tilde{r}_i^2} \), and \( \text{ci}(x) \) is the portfolio credit index defined similarly to (2.12). The quantities \( \{q_{ij}\} \) are the threshold values with \( q_{ij} = \Phi^{-1}(\tilde{p}_{ij}) \), where \( \tilde{p}_{ij} \) is the TTC transition probability from rating \( R_i \) to rating \( R_j \), which can be estimated from the historical sample. The key parameters for this rating transition probability model are \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) and \( \{\tilde{r}_i\} \), which can be estimated (Yang and Du 2016) by an approach similar to the algorithm described in Section 5.

### 3 THE DERIVED PROBABILITY OF DEFAULT TERM STRUCTURE AND MULTIPERIOD LOSS PROJECTION

In this section, we describe how a PIT PD structure can be derived from the forward PDs, and how loss can be projected over a multiperiod scenario given the PD term structure or given a PIT rating migration model.

#### 3.1 The PIT PD term structure derived from forward PDs

Let \( x(t_k) \) denote the vector of the values of macroeconomic variables \( x_1, x_2, \ldots, x_m \) at time \( t_k \). Let \( p_{ik}[x(t_k)] \) be the forward PD for the \( k \)th forward term given the scenario \( x(t_k) \). For a borrower with a nondefault initial risk rating \( R_i \) at \( t_0 \), the cumulative probability of default \( c_{ik}(t_k) \) over the period \( (t_0, t_k) \) can be derived from the forward PDs as follows:

\[
\begin{align*}
c_{i1}(t_1) &= p_{i1}[x(t_1)], \\
c_{i2}(t_2) &= c_{i1}(t_1) + [1 - c_{i1}(t_1)]p_{i2}[x(t_2)], \\
&\vdots \\
c_{ik}(t_k) &= c_{i(k-1)}(t_{k-1}) + [1 - c_{i(k-1)}(t_{k-1})]p_{ik}[x(t_k)].
\end{align*}
\]

Note that the quantity \( (1 - c_{ik}(t_k)) \) is the survival probability for the period \( [t_0, t_k] \). The following proposition demonstrates the relationship between the forward PD and survival probability.

**Proposition 3.1** The following factorization holds for the survival probability:

\[
1 - c_{ik}(t_k) = (1 - p_{i1}[x(t_1)])(1 - p_{i2}[x(t_2)]) \cdots (1 - p_{ik}[x(t_k)]). \tag{3.1}
\]
Factorization (3.1) follows from the equation below by induction:

\[ 1 - c_{ik}(t_k) = [1 - c_{ik-1}(t_{k-1})](1 - p_i[x(t_k)]). \]

\[ \square \]

### 3.2 Multiperiod scenario loss projection

Given the PIT PD term structure, the expected credit loss for the period \([t_0, t_k]\) for a loan of a borrower with initial rating \(R_i\) at \(t_0\) can be evaluated as follows (using the notation of (1.1)):

\[
\text{loss}_i(t_k) = p_{i1}[x(t_1)]f_1(t_1) + [1 - c_{i1}(t_1)]p_{i2}[x(t_2)]f_2(t_2) + \cdots \\
+ [1 - c_{ik-1}(t_k)]p_{ik}[x(t_k)]f_k(t_k).
\] (3.2)

The marginal PD for the period \([t_{k-1}, t_k]\) is given by \([1 - c_{ik-1}(t_{k-1})]p_{ik}[x(t_k)]\).

Given the PIT rating transition probability and a scenario \(x(t_k)\), let \(T[x(t_k)] = \{t_{ij}(t_k)\}\) denote the rating migration matrix, and \(t_{ij}(t_k)\) the probability that a rating \(R_i\) will migrate to \(R_j\) in a single term horizon. Assume that a higher index rating carries a higher default risk and that there are, for example, twenty-one ratings, with \(R_{21}\) the default rating. Then the last column of the matrix contains the PIT PDs for all risk ratings, and the last row of the matrix is set as

\[
v_{21j}(t_k) = 0 \quad \text{if } 1 \leq j \leq 20, \\
v_{2121}(t_k) = 1.
\]

With this notation, the cumulative PD for the period \([t_0, t_k]\) for a loan of a borrower whose initial risk rating is \(R_i\) can be derived by the following matrix multiplication:

\[
u_iT[x(t_1)]T[x(t_2)] \cdots T[x(t_k)],
\] (3.3)

where \(u_i\) is a row vector with all components equal to zero except for the \(i\)th component, which is 1. Consequently, marginal PDs can be derived and multiperiod scenario loss can be evaluated using a methodology similar to (3.2).

### 4 Loglikelihood Functions for Observing Term Default Frequency

In this section, we introduce a concept called forward loglikelihood, corresponding to the forward PD for a forward term. We show how, by observing the multistage term default frequency, the loglikelihood can be formulated using the forward loglikelihoods. The loglikelihood function expressions (4.1) and (4.3) will be used for parameter fitting in Section 5.
Recall the following notation from Section 2.1:

- \( n_{ik}(t_k) \) is the number of borrowers who survived the period \([t_0, t_{k-1}]\) with an initial risk rating \( R_i \) at \( t_0 \);
- \( d_{ik}(t_k) \) is the number of borrowers who defaulted in \((t_{k-1}, t_k]\).

Given the historical data for a risk-rated portfolio, a time series of the form \( \{n_{ik}(t_k), d_{ik}(t_k)\} \) can be derived. For each pair \((i, k)\), the \( k \)th forward term and the initial rating \( R_i \) at time \( t_0 \), using \( p_{ik}[x(t_k)] \), i.e., the forward PD for the term \((t_{k-1}, t_k]\), the forward loglikelihood is defined as

\[
FL_{ik} = \sum_{t_k} \{[n_{ik}(t_k) - d_{ik}(t_k)] \log(1 - p_{ik}[x(t_k)]) + d_{ik}(t_k) \log(p_{ik}[x(t_k)])\},
\]

with \( t_k \) varying smoothly over the sample time window. Here, we assume that the term default count follows a binomial distribution. The binomial coefficient, which is independent of the parameters for \( p_{ik}[x(t_k)] \) (as given by (2.10a) or (2.12)), has been dropped. Expression (4.1) is the actual loglikelihood over the conditional probability space given that borrowers have survived the period \([t_0, t_{k-1}]\).

In general, we are interested in the loglikelihood for a forward period \([t_h, t_{h+k}]\) with \( k \) terms. We assume that there is no withdrawal in the sample, and a borrower either defaults or survives at the end of a period.

For a borrower with initial rating \( R_i \) at \( t_0 \), let \( L_i(h, h+k) \) denote the loglikelihood over the period \([t_h, t_{h+k}]\) given that the borrower survived the period \([t_0, t_{h-1}]\), where the time window \([t_h, t_{h+k}]\) varies smoothly over the sample time window in (4.1). Similarly, let \( L(h, h+k) \) be the loglikelihood over the period \([t_h, t_{h+k}]\) for all borrowers of the portfolio with a nondefault initial risk rating at \( t_0 \) given that the borrowers survived the period \([t_0, t_{h-1}]\), where the time window \([t_h, t_{h+k}]\) varies smoothly over the sample time window.

**Proposition 4.1** Under the assumption of no withdrawal, the following equations hold (up to a constant independent of the parameters for \( \{p_{ij}[x(t_k)]\} \) as given by (2.10a) or (2.12)):

\[
L_i(h, h+k) = FL_{i[h+1]} + FL_{i[h+2]} + \cdots + FL_{i[h+k]},
\]

\[
L(h, h+k) = \sum_i L_i(h, h+k).
\]

Expression (4.2) demonstrates an additive property of the loglikelihood function: the loglikelihood for a forward period of consecutive forward terms is the sum of the individual forward loglikelihoods for the forward terms. This is expected because of the multiplicative property of the conditional probability for a multistage event.
Proof of Proposition 4.1  Equation (4.3) follows directly from (4.2). We show only (4.2) and the case when \( h = 0 \). For simplicity, we replace \( n_{ij}(t_j), d_{ij}(t_j), p_{ij}[x(t_j)] \) and \( c_{ij}(t_j) \) by \( n_j, d_j, p_j \) and \( c_j \), respectively.

Note that the marginal probability that a borrower with an initial rating \( R_i \) defaults in the period \((t_j-1, t_j]\) is \((1 - c_{j-1}) p_j \). Thus, the likelihood of observing \( d_j \) defaults in period \((t_j-1, t_j]\) is \((1 - c_{j-1})^d_j p_j^{d_j} \) (up to a factor given by the binomial coefficient of choosing \( d_j \) defaulters from \( n_j \) borrowers). Consequently, the likelihood of observing a sequence \( \{d_{ij}\}_{j=1,2,\ldots,k} \) of defaults in the period \([t_0, t_k]\), with \( d_j \) defaults in each period \((t_j-1, t_j]\), is

\[
\Delta(t_k) = p_1^{d_1} p_2^{d_2} \cdots p_k^{d_k} (1 - c_1)^{d_2} (1 - c_2)^{d_3} \cdots (1 - c_{k-1})^{d_k} (1 - c_k)^{n_1 - (d_1 + d_2 + \cdots + d_k)}
\]

(4.4)

(up to a constant factor given by binomial coefficients), where the last factor, \((1 - c_k)^{n_1 - (d_1 + d_2 + \cdots + d_k)}\), is the likelihood of surviving the entire period \([t_0, t_k]\).

Because of the no-withdrawal assumption, the following holds:

\[
n_i = n_1 - (d_1 + d_2 + \cdots + d_{i-1}).
\]

(4.5)

By (3.1), we have

\[
(1 - c_1)^{d_2} (1 - c_2)^{d_3} \cdots (1 - c_{k-1})^{d_k} (1 - c_k)^{n_1 - (d_1 + d_2 + \cdots + d_k)}
= (1 - p_1)^{d_2} [(1 - p_1)(1 - p_2)]^{d_3} \cdots [(1 - p_1)(1 - p_2) \cdots (1 - p_{k-1})]^{d_k}
\times [(1 - p_1)(1 - p_2) \cdots (1 - p_k)]^{n_1 - (d_1 + d_2 + \cdots + d_k)}
= (1 - p_1)^{d_2 + d_3 + \cdots + d_k} (1 - p_2)^{d_3 + d_4 + \cdots + d_k} \cdots (1 - p_{k-1})^{d_k}
\times [(1 - p_1)(1 - p_2) \cdots (1 - p_k)]^{n_1 - (d_1 + d_2 + \cdots + d_k)}
= (1 - p_1)^{n_1 - d_1} (1 - p_2)^{n_1 - (d_1 + d_2)} \cdots
\]

\[
(1 - p_{k-1})^{n_1 - (d_1 + d_2 + \cdots + d_{k-1})} (1 - p_k)^{n_1 - (d_1 + d_2 + \cdots + d_k)}
= (1 - p_1)^{n_1 - d_1} (1 - p_2)^{n_2 - d_2} \cdots (1 - p_{k-1})^{n_{k-1} - d_{k-1}} (1 - p_k)^{n_k - d_k}.
\]

(4.6)

The last equality in (4.6) follows from (4.5). By (4.4), we have the following loglikelihood for the period \([t_0, t_k]\):

\[
\log(\Delta(t_k)) = [d_1 \log(p_1) + (n_1 - d_1) \log(1 - p_1)]
+ [d_2 \log(p_2) + (n_2 - d_2) \log(1 - p_2)] + \cdots
+ [d_k \log(p_k) + (n_k - d_k) \log(1 - p_k)].
\]

Letting the period \([t_0, t_k]\) vary smoothly over the sample time window, we obtain the following loglikelihood:

\[
L_i(0, k) = \sum_{t_k} \log(\Delta(t_k)) = FL_{i1} + FL_{i2} + \cdots + FL_{ik}.
\]
A function is log concave if its logarithm is concave. If a function is concave, a local maximum is actually a global maximum, and the function is unimodal. This property is important for estimating the maximum likelihood.

**Proposition 4.2** The loglikelihood function (4.1), where $p_{ik}[x(t_k)]$ is given by (2.12), is concave as a function of $c_i, a_1, a_2, \ldots, a_m$, and it is concave as a function of $\bar{r}_{ik}$, where $p_{ik}[x(t_k)]$ is given by (2.10a). This concavity of (4.1) holds when the cumulative standard normal distribution $\Phi$ is replaced by any cumulative probability distribution that is log concave (eg, the cumulative distribution for logistic distribution).

**Proof** It is well known that the cumulative standard normal distribution is log concave, and the sum of concave functions is again concave. It is also known that if $f(x)$ is log concave, then so is $f(Az + b)$, where $Az + b : \mathbb{R}^m \to \mathbb{R}^1$ is any affine transformation from the $m$-dimensional Euclidean space to the one-dimensional Euclidean space. This means both the cumulative distribution $\Phi(x)$ and $F(x) = \Phi(-x)$ are log concave, and (4.1) is concave as a function of $c_i, a_1, a_2, \ldots, a_m$, where $p_{ik}[x(t_k)]$ is given by (2.12).

For the concavity of (4.1) as a function of $\bar{r}_{ik}$, where $p_{ik}[x(t_k)]$ is given by (2.10a), it suffices to show that the second derivative of the function

$$ L(r) = \log[\Phi(b\sqrt{1 + r^2 + ra})] \tag{4.7} $$

is nonpositive for any constants $a$ and $b$. The second derivative, $d^2[L(r)]/dr^2$, is given by

$$ \left( \frac{br}{\sqrt{1 + r^2}} + a \right)^2 \left\{ -\frac{\varphi(b\sqrt{1 + r^2 + ra})^2}{\Phi(b\sqrt{1 + r^2 + ra})^2} + \frac{\varphi'(b\sqrt{1 + r^2 + ra})}{\Phi(b\sqrt{1 + r^2 + ra})} \right\} 
+ \frac{\varphi(b\sqrt{1 + r^2 + ra})(b + r^2)(1 + r^2)^{-3/2}}{\Phi(b\sqrt{1 + r^2 + ra})} = I + II \tag{4.8} $$

where $\varphi$ and $\varphi'$ denote the first and second derivatives of $\Phi$, respectively. Because the factor in the first summand of (4.8),

$$ \left\{ -\frac{\varphi(b\sqrt{1 + r^2 + ra})^2}{\Phi(b\sqrt{1 + r^2 + ra})^2} + \frac{\varphi'(b\sqrt{1 + r^2 + ra})}{\Phi(b\sqrt{1 + r^2 + ra})} \right\}, $$

corresponds to the second derivative of $\log \Phi(x)$, it is nonpositive. Thus, the first summand in (4.8) is nonpositive. The second summand in (4.8) is nonpositive if $b \leq 0$. For the $b > 0$ case, we can change $b$ back to the negative case using the function $F(x) = \Phi(-x)$ and repeat the same discussion to obtain the nonpositivity of the second derivative in (4.7).
5 PARAMETER ESTIMATION BY MAXIMUM LIKELIHOOD APPROACHES

In this section, we assume that the threshold values \( \{b_{ik}\} \) are known and so are \( \{r_{ik}\} \), where \( r_{ik} \) is the risk sensitivity given by (2.1) and (2.4a) for the initial rating \( R_i \) and term \( k \) with respect to the latent systematic risk factor \( s \). Note that both \( \{b_{ij}\} \) and \( \{r_{ik}\} \) are defined before observing any macroeconomic condition \( x = (x_1, x_2, \ldots, x_m) \) (see Section 2.1 for the estimation of \( \{r_{ik}\} \), and Lemma 2.1(b) for \( \{b_{ik}\} \)).

As indicated at the end of Section 2.2, the key to the forward PDs \( \{p_{i,k}(x)\} \) is the determination of the coefficients \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) for the credit index, and of the rating level risk sensitivities \( \{\tilde{r}_{ik}\} \). The credit index enters the model via (2.8) and is determined by the parameters \( \lambda, \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \). Recall that by (2.11) the following relation is satisfied:

\[
\tilde{r}_{ik} = \frac{r_{ik}\lambda}{\sqrt{1 + r_{ik}^2(1 - \lambda^2)}}. \tag{5.1}
\]

Given \( \{b_{ik}\} \) and \( \{r_{ik}\} \), recall that the coefficients \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) for the credit index are derived from a normalization of a linear combination \( a_1\tilde{x}_1 + a_2\tilde{x}_2 + \cdots + a_m\tilde{x}_m \), with which the model \( \{\tilde{p}_i(x)\} \) best predicts the default probability of the portfolio for initial ratings for the one-term horizon, where \( \tilde{p}_i(x) \) is given by (2.12) as

\[
\tilde{p}_i(x) = \Phi[c_i + \tilde{r}_{i1}(a_1\tilde{x}_1 + a_2\tilde{x}_2 + \cdots + a_m\tilde{x}_m)]. \tag{5.2}
\]

This can be implemented by using the loglikelihood function (4.1), with \( p_{i,k}(x) \) being replaced by the \( \tilde{p}_i(x) \) above. Maximize the corresponding total loglikelihood for parameters \( \lambda, \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \).

When \( \lambda, \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) are known, \( \{\tilde{r}_{ik}\} \) can be determined by a calibration for each \( \tilde{r}_{ik} \) at the rating level by maximizing the total loglikelihood using (4.1) for the initial rating \( R_i \) and term \( k \), with \( p_{i,k}(x) \) being given by (2.10a), ie, the final term structure model is given by (2.10a).

We thus propose the following two-step approach.

**Step 1** (Estimate \( \tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m \) for the credit index) Get the first estimates for \( \lambda, a_1, a_2, \ldots, a_m \) by maximizing the total loglikelihood using (4.1) and (5.2) for all initial ratings for one forward term as a function of \( \lambda, a_1, a_2, \ldots, a_m \). To ensure these first estimates are the global maximum likelihood estimates, the following series of additional searches are performed. Let \( \lambda \in (0, 1) \) vary through the set of values \( \{i/N \mid 1 \leq i < N\} \) for large integer \( N \) (eg, \( N \geq 10 \)). For each value of \( \lambda \), calculate \( \{\tilde{r}_{ik}\} \) using (5.1). Find the maximum likelihood estimates for \( a_1, a_2, \ldots, a_m \) using the total loglikelihood from (4.1) and (5.2) for all initial ratings and one term. By the concavity of (4.1), as a function of \( a_1, a_2, \ldots, a_m \), any of the local maximum likelihood estimates \( a_1, a_2, \ldots, a_m \) are the global maximum likelihood estimates for a given \( \lambda \).
Use these estimates as the initial values for $\lambda, a_1, a_2, \ldots, a_m$, and remaximize the total loglikelihood from (4.1) and (5.2) as a function of $\lambda, a_1, a_2, \ldots, a_m$. Repeat this process to obtain the global maximum likelihood estimate for $\lambda, a_1, a_2, \ldots, a_m$.

Normalize the linear combination $a_1 \tilde{x}_1 + a_2 \tilde{x}_2 + \cdots + a_m \tilde{x}_m$ to obtain the estimate for $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_m$.

**Step 2** (Estimate $\tilde{r}_{ik}$ for each initial rating $R_i$ and term $k$ separately) Calculate the credit index $c_i(x)$ as

$$
c_i(x) = \frac{(\tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 + \cdots + \tilde{a}_m \tilde{x}_m)}{v},
$$

where $v$ is the standard deviation of $\tilde{a}_1 \tilde{x}_1 + \tilde{a}_2 \tilde{x}_2 + \cdots + \tilde{a}_m \tilde{x}_m$. We then calibrate and estimate $\tilde{r}_{ik}$ by maximizing the total loglikelihood using (4.1) for the initial rating $R_i$ and term $k$, with $p_{ik}[x(t_k)]$ being given by (2.10a) as

$$
p_{ik}(x) = \Phi[b_{ik} \sqrt{1 + \tilde{r}_{ik}^2} + \tilde{r}_{ik} c_i(x)].
$$

We implemented the above two-step optimization process by using the SAS PROC NLMIXED procedure.

**6 AN EMPIRICAL EXAMPLE: THE PROBABILITY OF DEFAULT TERM STRUCTURE FOR A CORPORATE PORTFOLIO**

The sample is created synthetically from a historical data set of a corporate portfolio containing a quarterly rating level default frequency (the sample default rate does not represent the original portfolio default rate). There are twenty-one ratings for the portfolio, with rating $R_1$ being the best quality rating, and $R_{21}$ the default rating. The higher the index rating, the higher the default risk.

Figure 1 depicts the trend of forward default rates, averaged over the sample time window, for twenty forward terms (ie, twenty forward quarters):

- at portfolio level;
- for investment rating;
- for subinvestment rating.

It is observed that the simple average forward default rates tend to converge after about twenty terms. For this reason, we can focus on terms covering a period of four years (sixteen quarters). For terms beyond sixteen quarters, a constant forward rate is assumed for all ratings. This constant rate can be estimated, for example, by the portfolio level average forward default rate for the fifth year.
FIGURE 1  Simple average forward default rate.

Macroeconomic data is sourced from the US Federal Reserve. It is merged with the term default frequency sample by matching the final quarter of a term to the calendar quarter of the macroeconomic data. Inclusion of macroeconomic variables is subject to a governance review process. All variables should pass the unit root tests. We consider four lag versions for each macroeconomic variable: lag 0 (current), lag 1 (lag one quarter), lag 2 (lag two quarters), lag 3 (lag three quarters). Each lag-version variable is named by prefixing with the label “L” followed by the lag number.

Table 1 shows the nine macroeconomic variables we use. The Pearson correlation to the quarterly portfolio level default rate is reported in the last four columns for the four lag versions of each variable.

In the remainder of this section, we focus on model fitting.

**STEP 1** (Variable selection for term models) Let \( m \) denote the number of variables in a model. Due to the limited number of data points in the time series sample, we consider only models with \( m = 2, 3 \). A preliminary model selection process is performed via SAS logistic regression with the model selection option being set to “Score”, targeting the portfolio level default frequency over the sample. The best 5000 models for each value of \( m \) are selected for further processing.

This SAS selection option calculates the score statistics for each possible variable combination without performing a full regression analysis, then selects the specified number of models based on the higher score statistics.

**STEP 2** (Forward PD model fitting) For each series of macroeconomic variables \( x_1, x_2, \ldots, x_m \) from step 1, follow the steps proposed in Section 5 to fit for coefficients \( \hat{a}_1, \hat{a}_2, \ldots, \hat{a}_m \) and sensitivities \( \{\hat{r}_{i,k}\} \).
<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Description</th>
<th>L0</th>
<th>L1</th>
<th>L2</th>
<th>L3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GDP_GQOQ_COM</td>
<td>Growth rate of US GDP</td>
<td>-0.38</td>
<td>0.39</td>
<td>0.39</td>
<td>0.53</td>
</tr>
<tr>
<td>2</td>
<td>LURC_DQOQ</td>
<td>Increase in US civilian unemployment rate**</td>
<td>-0.32</td>
<td>0.49</td>
<td>0.16</td>
<td>0.38</td>
</tr>
<tr>
<td>3</td>
<td>PCREPI_GQOQ_COM</td>
<td>Growth rate of US commercial real estate price*</td>
<td>-0.06</td>
<td>0.01</td>
<td>0.16</td>
<td>0.38</td>
</tr>
<tr>
<td>4</td>
<td>PPSDJT_GQOQ_COM</td>
<td>Growth rate of Dow Jones total stock market index*</td>
<td>0.13</td>
<td>0.15</td>
<td>0.58</td>
<td>0.51</td>
</tr>
<tr>
<td>5</td>
<td>RCBBB_DQQQ</td>
<td>Increase of US BBB 10Y corporate yield</td>
<td>0.16</td>
<td>0.06</td>
<td>0.31</td>
<td>0.43</td>
</tr>
<tr>
<td>6</td>
<td>RT10Y_DQOQ</td>
<td>Increase in US constant maturity treasury yield, 10Y**</td>
<td>0.14</td>
<td>0.02</td>
<td>0.03</td>
<td>0.31</td>
</tr>
<tr>
<td>7</td>
<td>RTB_DQOQ</td>
<td>Increase of US 3M treasury bill: secondary market rate**</td>
<td>-0.16</td>
<td>0.06</td>
<td>0.30</td>
<td>0.43</td>
</tr>
<tr>
<td>8</td>
<td>VIX_FED</td>
<td>US implied volatility**</td>
<td>0.52</td>
<td>0.36</td>
<td>0.43</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Quarter over quarter annualized by compounding. **Quarter over quarter annualized. ***Maximum of daily values per quarter.
<table>
<thead>
<tr>
<th>#</th>
<th>V1</th>
<th>V2</th>
<th>V3</th>
<th>Q1</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>L0_PPSDJT_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.66</td>
<td>0.63</td>
<td>0.59</td>
<td>0.71</td>
<td>0.32</td>
</tr>
<tr>
<td>2</td>
<td>L3_PCREPI_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.61</td>
<td>0.70</td>
<td>0.76</td>
<td>0.79</td>
<td>0.36</td>
</tr>
<tr>
<td>3</td>
<td>L0_VIX_FED</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.60</td>
<td>0.61</td>
<td>0.61</td>
<td>0.73</td>
<td>0.35</td>
</tr>
<tr>
<td>4</td>
<td>L1_RTB_DQOQ</td>
<td>L1_RT10Y_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.60</td>
<td>0.69</td>
<td>0.75</td>
<td>0.78</td>
<td>0.34</td>
</tr>
<tr>
<td>5</td>
<td>L0_GDP_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.59</td>
<td>0.61</td>
<td>0.61</td>
<td>0.72</td>
<td>0.24</td>
</tr>
<tr>
<td>6</td>
<td>L3_GDP_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.58</td>
<td>0.69</td>
<td>0.79</td>
<td>0.80</td>
<td>0.40</td>
</tr>
<tr>
<td>7</td>
<td>L2_PCREPI_GQOQ_Com</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.58</td>
<td>0.69</td>
<td>0.77</td>
<td>0.78</td>
<td>0.36</td>
</tr>
<tr>
<td>8</td>
<td>L3_RCBBB_RT10Y</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.57</td>
<td>0.68</td>
<td>0.78</td>
<td>0.78</td>
<td>0.38</td>
</tr>
<tr>
<td>9</td>
<td>L1_PCREPI_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>L0_LURC_DQOQ</td>
<td>0.56</td>
<td>0.68</td>
<td>0.77</td>
<td>0.73</td>
<td>0.40</td>
</tr>
<tr>
<td>10</td>
<td>L0_RCBBB_RT10Y</td>
<td>L1_PCREPI_GQOQ_COM</td>
<td>L1_RTB_DQOQ</td>
<td>0.56</td>
<td>0.68</td>
<td>0.77</td>
<td>0.83</td>
<td>0.42</td>
</tr>
</tbody>
</table>
FIGURE 2  Predicted versus realized cumulative portfolio default rate.

Table 2 shows the top ten forward PD models ranked by their $R^2$-squared (RSQ) portfolio level quarterly default rate (eg, first term), and the average RSQ for portfolio level forward default rates for forward terms in the first, second, third and fourth years (labeled Y1–Y4, respectively).

**STEP 3  (Benchmarking and backtests for the selected forward PD model)**

The top term model selected from Table 2 is scored over the development sample. Two rating migration models are used for benchmarking:

1. the TTC rating transition model;
2. a PIT rating transition model, using the same series of macroeconomic variables as the selected top forward term model.

The TTC transition matrix is calculated from the rating migration frequency across time (Lemma 2.1(b); see Yang and Du (2015) for a detailed calculation). The PIT transition model is developed following the approaches reviewed in Section 2.3.

Figure 2 plots the quarterly portfolio level default rate (eg, one term horizon), and the portfolio level cumulative default rate for one-, three- and four-year horizons.
Table 3 shows the model RSQ for backtesting the selected PIT term structure model and PIT transition model, to predict the portfolio cumulative default rates over the entire sample.

For the out-of-sample test for the model methodologies, we face a limitation on the availability of the number of data points (number of quarters) and the number of downturn periods in the time series sample. We split the time series sample into two parts: training and validation. The term and transition models are refitted over the training sample consisting of data points up to 2010, based on the final quarter of the forward term, using the same series of macroeconomic variables for both models as the top selected term model. The sample after 2010 is used as the validation sample. Model RSQ over the training sample is reported in Table 4. Since the training sample contains the stress period (2008 Q1–2009 Q4) when the default rate was very high, while the validation sample covers the five-year period after 2010 when the default rate was very low, for a fair comparison between training and validation we report the model RSQ only over the combined sample.

The following can be observed from our empirical example.

- The PIT term model slightly underperforms the PIT transition model when predicting the portfolio default rate for one-quarter horizon. Both models demonstrate strengths in predicting the portfolio quarterly default rate, in particular for the downturn period 2008 Q1–2010 Q1.
- The PIT term model generally outperforms the PIT transition model when the prediction horizon is extended, due to the fact that the term model is calibrated over a longer horizon.
- The TTC transition model is poor at picking up the trends during economic recession or expansion.

7 CONCLUSIONS

Models that directly fit the forward term default rate were proposed in this paper. These models are structured via a credit index, representing the part of systematic portfolio risk explained by a series of macroeconomic variables, given together with the risk
TABLE 4  Out-of-sample test-model RSQ for predicting portfolio cumulative default rate.

<table>
<thead>
<tr>
<th>(a) Training</th>
<th>PIT model</th>
<th>Q1</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>0.62</td>
<td>0.91</td>
<td>0.91</td>
<td>0.90</td>
<td>0.54</td>
<td></td>
</tr>
<tr>
<td>Transition</td>
<td>0.74</td>
<td>0.83</td>
<td>0.69</td>
<td>0.76</td>
<td>0.21</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(b) Training and validation</th>
<th>PIT model</th>
<th>Q1</th>
<th>Y1</th>
<th>Y2</th>
<th>Y3</th>
<th>Y4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>0.63</td>
<td>0.86</td>
<td>0.87</td>
<td>0.82</td>
<td>0.83</td>
<td></td>
</tr>
<tr>
<td>Transition</td>
<td>0.70</td>
<td>0.76</td>
<td>0.58</td>
<td>0.33</td>
<td>0.11</td>
<td></td>
</tr>
</tbody>
</table>

sensitivity for each nondefault initial risk rating and each forward term. An algorithm for parameter fitting is proposed by using the maximum likelihood of observing the term default frequency. We believe our proposed model and approaches will provide practitioners with a new, robust tool for modeling the PD term structure for multiperiod scenario loss projection, for CCAR stress testing and for IFRS 9 expected credit loss estimation.

DECLARATION OF INTEREST

The views expressed in this paper are not necessarily those of Royal Bank of Canada or any of its affiliates. The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

The author thanks Carlos Lopez for his consistent support for this research. Special thanks go to Clovis Sukam for his critical reading of this manuscript, and Glenn Fei, Biao Wu and Zunwei Du for many valuable discussions, insights and comments.

REFERENCES


Addendum

Addendum to Rubtsov and Petrov (2016): “A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration”

Torsten Pyttlik,1 Mark Rubtsov2 and Alexander Petrov3

1zeb.rolfes.schierenbeck.associates gmbh, Hammer Straße 165, 48153 Münster, Germany; email: TPyttlik@zeb.de
2Group Model Risk Control, Nordea Bank Norge, NH5A, Essendrops gate 7, 0368 Oslo, Norway; email: mark.rubtsov@nordea.com
3Group Credit Risk, Nordea Bank AB, H450, Smålandsgatan 15, 10571 Stockholm, Sweden; email: alexander.petrov@nordea.com

(Received November 30, 2016; revised December 9, 2016; accepted December 15, 2016)

In June 2016, The Journal of Risk Model Validation published a paper by Rubtsov and Petrov (2016) called “A point-in-time–through-the-cycle approach to rating assignment and probability of default calibration”. On p. 102 of the paper, the authors solved a system of equations (5.7)–(5.9) numerically; these equations are reproduced below as (1)–(3):

\[
\mathbb{E}[\Phi^{-1}(d_r)] = \frac{\mu_r - \sqrt{\rho_r} \mathbb{E}(\hat{Z})}{\sqrt{1-\rho_r}}, \\
\mathbb{E}[(\Phi^{-1}(d_r))^2] = \frac{1}{1-\rho_r} \left[ \mathbb{E}(B_r^2) - 2\sqrt{\rho_r} \mathbb{E}(B_r \hat{Z}) + \rho_r \mathbb{E}(\hat{Z}^2) \right] \\
= \frac{1}{1-\rho_r} \left[ (\gamma_r + \mu_r^2) - 2\sqrt{\rho_r} \mu_r \mathbb{E}(\hat{Z}) + \rho_r \mathbb{E}(\hat{Z}^2) \right].
\]

(Rceived November 30, 2016; revised December 9, 2016; accepted December 15, 2016)
Torsten Pyttlik has recently proposed an analytical solution to this system, and we present the details of that solution below. We believe it adds substantial extra value to the original material.

Let $Y_r := \Phi^{-1}(d_r)$ for brevity. The original equations (5.7)–(5.9) then become

\[
\mathbb{E}[Y_r] = \frac{\mu_r - \sqrt{\rho_r} \mathbb{E}[\hat{Z}]}{\sqrt{1 - \rho_r}}, \tag{4}
\]

\[
\mathbb{E}[Y_r^2] = \frac{1}{1 - \rho_r} [\gamma_r + \mu_r^2 - 2\sqrt{\rho_r} \mu_r \mathbb{E}[\hat{Z}] + \rho_r \mathbb{E}[\hat{Z}^2]], \tag{5}
\]

\[
\mathbb{E}[Y_r^3] = (1 - \rho_r)^{-3/2} [\mu_r \gamma_r - \sqrt{\rho_r} (\mu_r^2 + \gamma_r) \mathbb{E}[\hat{Z}]] + \rho_r \mu_r \mathbb{E}[\hat{Z}^2] - \rho_r^{3/2} \mathbb{E}[\hat{Z}^3]. \tag{6}
\]

Rearranging (4) gives

\[
\mu_r = \sqrt{1 - \rho_r} \mathbb{E}[Y_r] + \sqrt{\rho_r} \mathbb{E}[\hat{Z}]. \tag{7}
\]

Taking the square of (4) and subtracting that from (5) and then rearranging gives us

\[
\mathbb{E}[Y_r^2] - \mathbb{E}[Y_r]^2 = \frac{1}{1 - \rho_r} [\gamma_r + \rho_r (\mathbb{E}[\hat{Z}^2] - \mathbb{E}[\hat{Z}])]. \tag{8}
\]

\[
\gamma_r = (1 - \rho_r) \mathbb{V}[Y_r] - \rho_r \mathbb{V}[\hat{Z}]. \tag{9}
\]

Here, we have introduced the variance, defined as

\[
\mathbb{V}[X] := \mathbb{E}[X^2] - \mathbb{E}[X]^2.
\]

Note that (9) might result in $\gamma_r < 0$ if $\rho_r > 0$, which is undesirable since $\gamma_r$ was defined as a variance when the original system of equations was set up. Negative values of $\rho_r$ could therefore be considered, which would require an extensive modification of (1)–(3), using $-\sqrt{-\rho_r}$ and changing signs in several places.

Taking the third power of (4) and subtracting this from (6) gives

\[
\mathbb{E}[Y_r^3] - \mathbb{E}[Y_r]^3 = (1 - \rho_r)^{-3/2} [3 \mu_r \gamma_r - \sqrt{\rho_r} \gamma_r \mathbb{E}[\hat{Z}] + \rho_r \mu_r \mathbb{V}[\hat{Z}]]
 - \rho_r^{3/2} (\mathbb{E}[\hat{Z}^3] - \mathbb{E}[\hat{Z}^3]).
\]
Inserting (7) and (9) into the inner square brackets on the right-hand side yields, after rearranging, an expression that is solvable for $\rho_r$ alone:

$$E[Y_r^3] - 3E[Y_r]V[Y_r] - E[Y_r]^3 = -\left(\frac{\rho_r}{1 - \rho_r}\right)^{3/2} \left[E[\hat{Z}]^3 - 3E[\hat{Z}]V[\hat{Z}] - E[\hat{Z}]^3\right].$$

$$\rho_r = \frac{1}{1 + \left(S[Y_r]/S[\hat{Z}]\right)^{-2/3}}. \quad (10)$$

Here, we have defined


which is the nonnormalized skewness (to obtain normalized skewness, multiply $S[X]$ by $V[X]^{-3/2}$).

After evaluating $\rho_r$ from (10), use (9) and (7) to obtain values for $\gamma_r$ and $\mu_r$, respectively.

Note that if the distribution of $\hat{Z}$ is symmetrical, ie, $S[\hat{Z}] = 0$, then (10) has no solution if $S[Y_r] \neq 0$. There is no unique solution if both $S[\hat{Z}] = 0$ and $S[Y_r] = 0$. For the limiting case $\rho_r = 1$, the whole system of equations (4)–(6) would be invalid.

REFERENCES

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net