- A new bootstrap test for multiple assets joint risk testing
  David Ardia, Lukasz Gatarek and Lennart F. Hoogerheide

- Default risk charge: modeling framework for the Basel risk measure
  Sascha Wilkens and Mirela Predescu

- A review of the fundamentals of the Fundamental Review of the Trading Book: standard foreign exchange rules are highly asymmetric with respect to reporting currencies
  Hany M. Farag

- How risk managers should fix tracking error volatility and value-at-risk constraints in asset management
  Luca Riccetti
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.

To subscribe to a Risk Journal visit Risk.net/subscribe or email info@risk.net
The Journal of Risk
EDITORIAL BOARD

Editor-in-Chief
FARID AITSAHLIA University of Florida

Editor Emeritus and Chairman of the Editorial Board
STAN URYASEV University of Florida

Associate Editors
CARLO ACERBI MSCI Group
ALPER ATAMTURK University of California, Berkeley
TURAN G. BÅLĢ Georgetown University
JEREMY BERKOWITZ University of Houston
PETER CARR Morgan Stanley
New York University
PATRICHERIDITO Princeton University
JON DANIELSSON London School of Economics
ROBERT F. ENGLE New York University
KAY GIESECKE Stanford University
JOHN HULL University of Toronto

SUSSI KEPO National University of Singapore
PAUL H. KUPIEC Federal Deposit Insurance Corporation
FRANÇOIS LONGIN ESSEC
ELISA LUCIANO University of Torino & Fondazione Collegio Carlo Alberto
NEIL D. PEARSON University of Illinois at Urbana–Champaign
RICCARDO REBONATO PIMCO
TIL SCHUERMANN Oliver Wyman
RENÉ M. STULZ Ohio State University
CASPER G. DE VRIES Erasmus University
Rotterdam & Tinbergen Institute

SUBSCRIPTIONS

The Journal of Risk (Print ISSN 1465-1211 | Online ISSN 1755-2842) is published six times per year by Infopro Digital, 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 Email: info@risk.net

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

GENERAL SUBMISSION GUIDELINES

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, Infopro Digital,
Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK
Tel: +44 207 316 9300; Fax: +44 (0)20 7488 9758

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

This international refereed journal publishes a broad range of original research papers that aim to further develop understanding of financial risk management. As the only publication devoted exclusively to theoretical and empirical studies in financial risk management, *The Journal of Risk* promotes far-reaching research on the latest innovations in this field, with particular focus on the measurement, management and analysis of financial risk.

*The Journal of Risk* is particularly interested in papers in the following areas.

- Risk management regulations and their implications.
- Risk capital allocation and risk budgeting.
- Efficient evaluation of risk measures under increasingly complex and realistic model assumptions.
- The impact of risk measurement on portfolio allocation.
- Theoretical development of alternative risk measures.
- Hedging (linear and nonlinear) under alternative risk measures.
- Financial market model risk.
- Estimation of volatility and unanticipated jumps.
- Capital allocation.

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

Letter from the Editor-in-Chief vii

RESEARCH PAPERS

A new bootstrap test for multiple assets joint risk testing 1
David Ardia, Łukasz Gatarek and Lennart F. Hoogerheide

Default risk charge: modeling framework for the Basel risk measure 23
Sascha Wilkens and Mirela Predescu

A review of the fundamentals of the Fundamental Review of the Trading Book: standard foreign exchange rules are highly asymmetric with respect to reporting currencies 51
Hany M. Farag

How risk managers should fix tracking error volatility and value-at-risk constraints in asset management 79
Luca Riccetti
To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
LETTER FROM THE EDITOR-IN-CHIEF

Farid AitSahlia  
Warrington College of Business,  
University of Florida

This issue of *The Journal of Risk* addresses risk-constrained portfolio management, the practical implementation of the 2016 Basel capital requirements in the face of market and default risk, and the joint testing of marginal specifications for risk models.

In our first paper, “A new bootstrap test for multiple assets joint risk testing”, David Ardia, Łukasz Gatarek and Lennart F. Hoogerheide present a simulation-based technique for the joint statistical testing of marginal models that is particularly well suited to dependent time series and copula models. In particular, the authors provide an illustration on generalized autoregressive conditional heteroscedasticity (GARCH) models with a view toward value-at-risk (VaR) estimation in a variety of settings, including portfolios contrasted along size and sector dimensions.

In “Default risk charge: modeling framework for the ‘Basel’ risk measure”, the second paper in this issue, Sascha Wilkens and Mirela Predescu contribute a study that helps us better understand the challenges in implementing the proposed 2016 Basel regulations on default credit charges in trading portfolios. These new rules are a direct consequence of the undercapitalization of trading books observed during the 2008 crisis. In their paper, the authors focus on calibration issues to estimate default correlations in order to capture systemic, industry and firm-specific effects. They also highlight the computational issues that arise in default risk charge calculations.

The issue’s third paper, “A review of the fundamentals of the Fundamental Review of the Trading Book: standard foreign exchange rules are highly asymmetric with respect to reporting currencies”, sees Hany Farag highlighting further specific issues arising from the implementation of the 2016 Basel regulations with regard to market risk for trading portfolios. In particular, Farag discusses the significant effects of currency reporting, which are further exacerbated by the aggregation mechanism used for nonlinear portfolios.

In our fourth and final paper, “How risk managers should fix tracking error volatility and value-at-risk constraints in asset management”, Luca Riccetti puts the spotlight on the trade-off between risk management, as measured by tracking error volatility (TEV) relative to a benchmark, and portfolio optimization. Riccetti develops a strategy that helps to identify the most efficient portfolios subject to a TEV constraint. This study also supports the adoption of a maximum VaR constraint, instead of a maximum variance constraint, for active portfolio management.
CALL FOR PAPERS
RESEARCH CONFERENCE
Safeguarding retirement assets and income
11–12 September 2017

Overview
The University of Birmingham and VU University Amsterdam, supported by EIOPA and Netspar, are jointly organizing a research conference on safeguarding retirement assets and income from defined contribution (DC) pension plans.

Mission
The conference will bring together academics, researchers, think tanks, regulators, investors, corporations and consultants to discuss key issues around safeguarding DC-based retirement assets and income. The conference will deepen the evidence base as well as support and encourage dialogue and concepts across organizational and national boundaries.

Structure and time frame
The one-and-a-half-day conference will take place at EIOPA, Frankfurt, Germany, on September 11–12, 2017. It is linked to a special issue of The Journal of Risk on the topic.

Concept
How we save and access our pensions is fundamentally changing. Rising longevity, demographic change and low interest rates have sharpened the challenges that defined benefit (DB) and collective defined contribution (CDC) plans face, accelerating their windup or presaging a scaling back in the benefits they offer.

At the same time, more people than ever are members of DC pension plans, where individual investment accounts are offered with fund options chosen by the individual alongside a preset option for nonchoosers. In several countries, the number of people contributing to DC pension plans now far exceeds the number contributing to pay-as-you-go or funded DB pension plans. For example, in the United Kingdom, 90% of people contributing to private sector pension funds now do so via DC plans.

The demographic, financial markets and policy changes we are seeing is prompting a dramatic shift from income-based later life to asset-based later life. Relatively stable and regular later life income from annuities, DB and CDC pensions is giving way
to flexibly accessed lump sums from DC pension pots, property and assets held in tax-incentivized savings accounts. While for most individuals the shift has not been of their choosing, for some individuals, at least, control of assets seems to be enjoyed and preferred.

Many of today’s pension savers are not well equipped to make decisions based on an asset-based retirement and the risks this brings. The influence that market, inflation and longevity risk can have on the ability to sustain consumption and income in retirement is disliked, and most people find navigating the decisions they are being asked to make challenging and, at times, emotionally driven.

The growing proportion of asset-based retirement wealth raises significant policy and practical questions as to how far DC retirement assets and income can and should be safeguarded, as well as how risks can be dispersed, pooled and shared for the possibly substantial proportion of people wanting this (such as smoothing of benefits or investment returns, inter- or intra-generational risk sharing or augmenting pension contributions with elements of insurance). It is highly likely that there will be trade-offs with any one proposition, eg, flexibility versus guarantees; support and advice versus safety nets; and riskier, higher consumption and income versus low risk, lower consumption and income. A mix-and-match approach may be more desirable than one idealized solution. The availability of choice may be important even if there is low expectation of mass usage. One suggestion for mass usage is the design of standardized, lifecycled/lifestyle glidepaths for accessing assets and income.

The conference will explore how assets and income in retirement can and should be safeguarded for people who have growing proportions of asset-based retirement wealth from DC plans, and how the risks can be dispersed, pooled and shared for those wanting or needing this. The focus is on the workplace (Pillar 2) and personal (Pillar 3) DC markets within both single-country and EU cross-border contexts.

Manuscripts and articles for presentation at the conference and for the special issue of The Journal of Risk are welcomed from

- academics and researchers,
- investment professionals, pension funds and pension delivery organizations,
- companies and consultants,
- law makers, regulators, standards boards and committees, and
- think tanks and nonprofits making significant contributions to these topics.

*Journal of Risk* **19**(4)

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
Suggested areas of research interest include, but are not limited to, the following.

- With the trend toward asset-based retirement and “freedom and choice”, to what extent can and should assets and income be safeguarded in retirement?

- Innovative ways of pooling asset class, investment and individual as well as macro-longevity risks, and how the benefits of risk pooling and sharing may compare with the “benefits” of scale delivered by DC plans with no risk sharing or pooling.

- The reliability of promises or ambitions from risk pooling and sharing in DC plans. In the case of voluntary DC plan participation, do such promises change the likelihood that people will contribute, or the amount they contribute, to the pension plan? How can different levels of contributions and benefits be incorporated?

- Challenges in persuading countries, trade unions and employers to set up and permit risk pooling and sharing, as well as their design, governance, rules and strategies; how overheads can be shared; communication to members; alignment of members’ expectations; and the potential risks to members around changes in expected values, benefits and funding position.

- The identification of individual property rights at fair market prices to provide members with the freedom to stop contributions, make withdrawals when permitted or exit a plan altogether.

- Is pooling and sharing risk with DC plan assets more effective during the accumulation or spending phase? Is this for the whole or just a part of the phase?

- Evidence of preferences for accessing DC pension wealth flexibly as a series of cash lump sums with or without stable income.

- The amount of DC plan wealth needed to achieve stable and sustainable consumption in different EU countries.

- Applicability of standardized, lifecycled/lifestyled glidepaths, defaults and choice, as well as regular income versus flexible access to capital, both at and in retirement.

- Merits and drawbacks of augmenting DC pension contributions to provide improvements in financial resilience of members and their households, for example, by an amount of the overall contribution paying for death-in-service benefit, critical illness, health insurance or long-term care.
**Publication**

We encourage papers or presentations to be submitted by standard setters, regulators, practitioners, academics, researchers, think tanks and the third sector. All papers submitted are eligible to be considered for publication in a special issue of *The Journal of Risk* devoted to the research conference. If you wish your paper to be considered for 2018 publication in *The Journal of Risk*, please indicate so in your cover letter. If this is the case, your paper should not be under review with any other journal. Papers intended for publication will be subject to peer review. Subject to satisfactory revision, a selection of these papers will be published in a special issue of *The Journal of Risk*. Note that the acceptance of a paper to the research conference is not a guarantee of publication.

A contribution to travel and accommodation may be provided by the conference organizers to authors of the best selected papers.

**Electronic submission**

Authors are invited to submit a paper. The first page of the paper should contain the title, name of each author, postal addresses, telephone numbers and email addresses. Please indicate in your cover letter whether you would be willing to serve as a session chair and/or discussant. All submitted papers must be accompanied by an abstract explaining the contribution of the paper. You are asked to follow *The Journal of Risk*’s Author and Submission Guidelines, available at www.risk.net/static/risk-journals-submissionguidelines.

**Deadlines**

The deadline for submission of manuscripts is June 1, 2017. Authors will be notified by July 1, 2017.

**Further information**

A conference website will soon be available at the Netspar events page: www.netspar.nl/en/events/.

**Contact**

Please contact us with any questions.

Paul Cox, Senior Lecturer of Finance, University of Birmingham;
email: p.cox@bham.ac.uk

Fieke van der Lecq, Professor of Pension Markets, Faculty of Economics and Business, VU University Amsterdam; email: s.g.vander.lecq@vu.nl

*Journal of Risk* 19(4)
A new bootstrap test for multiple assets joint risk testing

David Ardia,1,2 Lukasz Gatarek3,4 and Lennart F. Hoogerheide4,5

1Institute of Financial Analysis, University of Neuchâtel, Rue A.-L. Breguet 2, CH-2000 Neuchâtel, Switzerland; email: david.ardia@unine.ch
2Department of Finance, Insurance and Real Estate, Laval University, Pavillon Palasis-Prince, 2325 Rue de la Terrasse, Ville de Québec, QC G1V 0A6, Canada
3Institute of Econometrics and Statistics, University of Lodz, Ulica Rewolucji 1905, 90-214 Lodz, Poland; email: gatarek@tlen.pl
4Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands
5Department of Econometrics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands; email: f.f.hoogerheide@vu.nl

(Received January 22, 2016; revised July 18, 2016; accepted July 28, 2016)

ABSTRACT

In this paper, a novel simulation-based methodology is proposed to test the validity of a set of marginal time series models, where the dependence structure between the time series is taken directly from the observed data. The procedure is useful when one wants to summarize the test results for several time series in one joint test statistic and p value. The proposed test method can have higher statistical power than a test for a univariate time series, especially for short time series. Therefore, our test for multiple time series is particularly useful if one wants to assess value-at-risk (or expected shortfall) predictions within a small time frame (eg, a crisis period). We apply our
method to test generalized autoregressive conditional heteroscedasticity (GARCH) model specifications for a large panel data set of stock returns.

**Keywords:** bootstrap test; generalized autoregressive conditional heteroscedasticity (GARCH); marginal models; multiple time series; value-at-risk (VaR).

1 INTRODUCTION

In several situations, one may desire to evaluate the validity of a set of models for dependent univariate time series. Therefore, a test of the joint validity is needed. In this paper, a novel simulation-based methodology is proposed to test the validity of a set of marginal time series models, where the dependence structure between the time series is taken directly from the observed data. Since the dependence structure is not tested, the procedure is useful in the following situations. The first case is where one has specified the dependence structure but specifically desires to test only the marginal models, eg, as an additional tool for assessing the univariate specifications in a copula model, without testing the validity of the particular copula choice. That is, a separate goodness-of-fit test can be applied for the copula specification. The second case is where no dependence structure is specified, eg, when one wants to summarize the test results of a proposed model or method for several time series in one joint test statistic and \( p \) value.

We apply our method to test generalized autoregressive conditional heteroscedasticity (GARCH) model specifications for a large panel data set of stock returns; from a practical viewpoint, it is desirable to have a single model for many time series. The proposed test method can have higher statistical power than a test for a univariate time series, especially for short time series. The reference test by Christoffersen (1998) for testing the value-at-risk (VaR) performance may require a large number of observations in order to have a reasonable power. Therefore, our test for multiple time series is particularly useful if one wants to assess VaR (or expected shortfall) predictions over small time frames, such as crisis periods. In particular, we test a range of GARCH models for left-tail density as well as VaR predictions for the universe of Standard & Poor’s (S&P) 500 equities. We use the new methodology to discuss the application of the models over the whole universe and for subsets of the universe (eg, sizes or sectors). We also make use of the cross-sectional power of the approach to test the models over sub-windows, particularly during a crisis period.

A well-known alternative is the Bonferroni correction method, where in a situation of multiple testing the significance level per test is computed as the total significance level (or family-wise error rate (FWER)), divided by the number of tests. An extension is the Holm–Bonferroni method (Holm 1979). However, for the null hypothesis that all marginal models are correct, this Holm-Bonferroni method reduces to the original
Bonferroni correction. These Bonferroni correction methods may lead to very low power, especially in cases of large numbers of tests. Furthermore, the Bonferroni correction only leads to rejection or nonrejection in a conservative setting, without explicitly giving an exactly appropriate \( p \) value. This also implies that, in certain situations, where one wants to retain, not reject, the null hypothesis, the Bonferroni correction is actually nonconservative. Šidák (1967) proposed an alternative way to improve the power of the Bonferroni correction method, but this method requires the independence of the multiple tests.

Another alternative method to test for the validity of the marginal models is the false discovery rate (FDR) methodology of Storey (2002), in combination with the confidence interval for the percentage of correct marginal models (Barras et al 2010). One could attempt to test the hypothesis that all marginal models are correct by investigating whether the confidence interval for the percentage of correct marginal models contains the value 100\%. However, the FDR methodology assumes independence between the different time series (or a limited type of dependence, such as independence between a large enough number of subsets), an assumption that may be substantially violated by, for example, time series of asset returns. Storey and Tibshirani (2003) discuss the FDR under different types of dependence. Under their most general assumptions on the dependence, their estimator is (conservatively) biased. Furthermore, they consider DNA data, which arguably has a substantially different dependence structure than time series in economics and finance.

This paper continues as follows. In Section 2, we introduce the bootstrap test and discuss its applications. In Section 3, we perform a Monte Carlo study to test the performance of the new methodology, in which we consider the size and power for simulated data sets. In Section 4, we apply the new method to GARCH models on several equity universes. In particular, we test the forecasting performance of the models over subperiods such as crisis periods. In Section 5, we discuss some possible extensions and applications. Section 6 concludes.

2 BOOTSTRAP TEST PROCEDURE AND ITS APPLICATIONS

We consider variables \( y_{ti} (t = 1, \ldots, T; i = 1, \ldots, N) \), where we have \( N \) time series of \( T \) periods. We present the test in the context of financial returns for convenience with regard to the latter applications, where \( y_{ti} \) is the ex post return (or profit/loss) of asset \( i \) at time \( t \). However, it can be applied in any situation where the validity of a set of marginal models is tested.

\( H_0 \): all marginal models are correct.

\( H_1 \): at least one marginal model is incorrect.
Define the probability integral transform (PIT) or $p$ scores (Diebold et al 1998) $x_{t,i}$ as

$$x_{t,i} = \Phi^{-1}(y_{t,i}) = \int_{-\infty}^{y_{t,i}} \hat{f}_{t|t-1,i}(u) \, du,$$

with $\hat{f}_{t|t-1,i}$ the ex ante forecasted return density, and $\hat{F}_{t|t-1,i}$ the ex ante forecasted return cumulative density function. Only under a correct model specification, that is, if both the model and the parameter values are correctly specified, do the PITs have uniform distributions that are independent over time. The PITs have this property asymptotically if consistent parameter estimators are used (Rosenblatt 1952). If an invalid model specification is used, then the PITs will have a nonuniform distribution and/or the PITs will be dependent over time. Therefore, it is natural to investigate the PITs of a model in order to assess the validity of a marginal model.

Define the matrix $X$ as the $(T \times N)$ matrix of $x_{t,i}$ ($t = 1, \ldots, T; i = 1, \ldots, N$). First, compute the test statistic for each time series (i.e., for each column of $X$), e.g., the likelihood ratio (LR) statistic in a particular test for the validity of the marginal model. Examples of such LR tests include the test for correct unconditional coverage, independence or correct conditional coverage (Christoffersen 1998); the tests for correctly distributed or independent durations between violations of the VaR (Christoffersen and Pelletier 2004); and the test of Berkowitz (2001), with the null hypothesis that $z_{t,i} = \Phi^{-1}(x_{t,i})$ has a standard normal (mean = 0, variance = 1) distribution without autocorrelation.

The joint test statistic for all columns (assets) is now a function (yielding a scalar) of the test statistics over the $N$ columns (assets). In this paper, we consider the sum of the $N$ test statistics; however, a different function such as the maximum could also be used. This choice obviously affects the power against different violations of $H_0$. For example, if each marginal model suffers from a minor misspecification, then the sum of the $N$ test statistics seems a reasonable choice. If only one marginal model has a substantial misspecification, whereas the other models are correctly specified, then the maximum of the $N$ test statistics seems a reasonable choice. If the $N$ test statistics are LR statistics, then this sum of $N$ test statistics would be the valid joint LR test statistic (asymptotically having a $\chi^2$ distribution under $H_0$) if the columns (assets) were independent. However, we do not make this (unrealistic) assumption of independence. In order to take the dependence between the columns/assets into account, we compute the $p$ value by simulating the distribution of $X$ under $H_0$.

**Step 1**

Compute the $(T \times N)$ matrix $R$, where each column contains the ranks of the corresponding column of $X$ (where rank 1 corresponds to the lowest value and rank $T$ corresponds to the highest value). For example, if the second column of $X$ is $(0.4, 0.8, 0.5, 0.2)'$, then the second column of $R$ is $(2, 3, 4, 1)'$.
Step 2  Simulate $X$ under $H_0$: simulate independent and identically distributed (iid) $\mathcal{U}(0, 1)$ elements per column, with dependence between elements per row. That is, we generate numbers that when considered within the same column are iid $\mathcal{U}(0, 1)$ but when considered within the same row are allowed to be dependent. For this purpose, we perform the following steps.

1. Simulate a $(T \times 1)$ vector $v$ containing $T$ iid draws from the discrete uniform distribution on $\{1, 2, \ldots, T\}$. 

2. Define the $(1 \times N)$ vector $s_t$ ($t = 1, 2, \ldots, T$) as the $v_t$th row of $R$, where $v_t$ is element $t$ of vector $v$.

3. Simulate $u_{ti}$ ($t = 1, \ldots, T; i = 1, \ldots, N$) from a Beta distribution $\text{Beta}(s_{ti}, T + 1 - s_{ti})$.

Note that we make use of the result that the $s_{ti}$th order statistic of $T$ iid $\mathcal{U}(0, 1)$ variables has a $\mathcal{B}(s_{ti}, T + 1 - s_{ti})$ distribution. Also, note that for $N = 1$ this is a clumsy way of simulating a $(T \times 1)$ column vector of iid $\mathcal{U}(0, 1)$ draws (by first simulating the order statistic from $\{1, 2, \ldots, T\}$ and second simulating it from the relevant Beta distribution, conditional on the simulated order statistic). Now we have a $(T \times N)$ matrix in which each column has iid $\mathcal{U}(0, 1)$ draws, but with dependence between the elements within any given row. The reason is that we simulate rows independently, and the ranking numbers per row are dependent (if the columns of the original data are dependent). The dependence between the columns is preserved. For example, if all columns are highly correlated (with correlations close to 1), then many rows of $R$ contain ranks that are close to each other. Therefore, each row of the simulated $X$ contains draws from similar Beta distributions (with approximately the same means and small variances if $T$ is moderate or large). Therefore, most rows of $X$ will contain only high values (close to 1), only low values (close to 0) or only moderate values. In this way, the dependence structure of high correlation is preserved in the simulated $X$.

Step 3  Repeat the simulation of $X$ under $H_0$ and compute the $p$ value as the fraction of simulated data sets for which $H_0$ is not rejected. Typically, the test statistic can be computed from the PITs; otherwise, a time series of the original variable can be computed from the PITs and the model.

A scheme of the procedure is presented in Figure 1. The computer code of the above steps is straightforward; see the online appendix for implementation in the R statistical language (Ardia 2017; R Core Team 2016).
We can generate many matrixes $X$ from the distribution under $H_0$, compute the test statistic (i.e., the sum of $N$ test statistics) for each simulated data set, and compute the $p$ value by comparing the test statistic for the empirical data set with the test statistics for the simulated data sets under $H_0$.

As mentioned before, for the LR tests, the sum of the LR statistics for the univariate time series provides a natural joint test statistic for the set of time series, since this amounts to the LR statistics if the series were independent. The bootstrap method serves as a way to correct the distribution under the null hypothesis, replacing the no longer valid (asymptotic) $\chi^2$ distribution. However, other test statistics for univariate time series can also be summed to produce a joint test statistic, e.g., F-statistics or Cramér–von Mises test statistics.

We stress that we do not make the assumption that the series are independent. Our bootstrap method for the computation of the distribution of the test statistic under the null hypothesis assumes dependence of the series, which implies that the simulated distribution of the test statistic under the null hypothesis is valid under dependence of the series, so that the resulting critical value and/or $p$ value is valid under dependence of the series. The proposed method would also work for test statistics other than the sum of LR statistics, since the bootstrap method would also yield the correct critical
value and/or $p$ value under dependence of the series for any other test statistic. Choosing a test statistic other than the sum of LR statistics will only lead to a different power of the test. However, both the Monte Carlo experiments and empirical application in Sections 3 and 4 show that the specific choice of the sum of LR statistics leads to a rather powerful test.

We preserve the dependence between the time series. That is, we take the dependence structure between the time series from the data. For this purpose, we do not require that the dependence structure is constant over time. In the case of a time-varying dependence structure of a stationary process, for large enough samples the distribution of the dependence structure tends to its unconditional distribution. Our experiments in Section 3 illustrate that our approach also works well, in the sense of a $U(0,1)$-distributed $p$ value under the null hypothesis for the continuously distributed test statistic of Berkowitz (2001), for two data-generating processes (DGPs) involving time-varying dependence structures: one DGP where the time-varying dependence structure is exogenous (a two-regime Markov switching model), and one DGP where the time-varying dependence structure is endogenous (a dynamic conditional correlation model).

In this paper, we will consider the following tests.

Berkowitz (2001). We test whether $z_{ti}/NUL_{1,i} = \Phi^{-1}(x_{ti})$ is standard normal and independent over time. That is, in the model

$$z_{ti} - \mu_{i} = \rho_{i}(z_{t-1,i} - \mu_{i}) + \varepsilon_{ti}, \quad \varepsilon_{ti} \sim \text{iid } \mathcal{N}(0, \sigma_{i}^{2}), \quad (2.2)$$

we perform the LR test with a null hypothesis that $\mu_{i} = 0$, $\rho_{i} = 1$ and $\sigma_{i}^{2} = 1$. Under the null hypothesis, the LR statistic asymptotically follows a $\chi^{2}$ distribution with three degrees of freedom; however, we do not make use of this asymptotic distribution, as we only use our simulation-based approach to determine the $p$ value. We also consider an extension of this LR test, which was suggested by Christoffersen and Pelletier (2004), focusing specifically on the left tail. In that extended version of the test, we only select $x_{ti}$ that are smaller than $\alpha = 0.05$ or $\alpha = 0.01$, after which we test whether $z_{ti} = \Phi^{-1}(x_{ti}/\alpha)$ is standard normal and independent over time.

Christoffersen (1998). We perform the LR test for correct conditional coverage of the 95% and 99% VaR, that is, the fifth and first percentile of the predicted return’s distribution. That is, for $\alpha = 0.05$ and $\alpha = 0.01$, we perform the LR test with a null hypothesis that the binary variables

$$I_{ti} = \begin{cases} 1 & \text{if } x_{it} < \alpha, \\ 0 & \text{if } x_{it} \geq \alpha, \end{cases} \quad (2.3)$$
are independent and Bernoulli distributed, with $P\{I_{ti} = 1\} = \alpha$, against the alternative that the $I_{ti}$ follow a first-order Markov chain. This LR test combines the LR tests for correct unconditional coverage, $P\{I_{ti} = 1\} = \alpha$, and independence between $I_{ti}$ and $I_{t-1,i}$.

3 MONTE CARLO STUDY

3.1 Illustration of correct size

In this section, we consider the size and power of our proposed approach. We analyze the finite sample performance of the method via some (necessarily limited) simulations. First, we illustrate that our test yields a $p$ value that is $\mathcal{U}(0,1)$ distributed under the null hypothesis. For this purpose, we use the test of Berkowitz (2001), for which the test statistic has a continuous distribution. Note that in the tests of Christoffersen (1998) for correct unconditional coverage, independence and correct conditional coverage, the test statistic has a discrete distribution, which implies that even under the null hypothesis the distribution of the $p$ value is not $\mathcal{U}(0,1)$.

We consider different DGPs with different dependence structures between the time series. For the marginal distributions, we consider the convenient $\mathcal{N}(0, 1)$ distribution, since the choice of the marginal distribution does not matter in these experiments. Anyway, the marginal is correct under the null hypothesis, so the PIT will thus be $\mathcal{U}(0, 1)$ distributed for the marginal models. We consider the following three DGPs.

**Constant correlation model.** The DGP for the $(N \times 1)$ vector $y_t$ is given by

$$y_t \sim \text{iid } \mathcal{N}(0, \Sigma),$$

$$\Sigma = \rho J + (1 - \rho) I,$$

with $J$, an $(N \times N)$ matrix of ones, and $I$, the $(N \times N)$ identity matrix. We set $\rho = 0.9$.

**Two-regime Markov switching correlation model.**

$$y_t \sim \mathcal{N}(0, \Sigma_{s_t}),$$

$$\Sigma_{s_t} = \rho_{s_t} J + (1 - \rho_{s_t}) I,$$

with $\{s_t\}$ a sequence that is assumed to be a stationary, irreducible Markov process with discrete state space $\{1, 2\}$, and transition matrix $P = [P_{ij}]$, where $P_{ij} = P\{s_{t+1} = j \mid s_t = i\}$. We consider $\rho_1 = 0.9$, $\rho_2 = 0$, $P_{11} = 0.9$ and $P_{22} = 0.9$. Note that the time-varying dependence structure is exogenous here, in the sense that the dependence structure at time $t$ only depends on the dependence structure at time $t - 1$, and not on $y_{t-1}$ (given the dependence structure at time $t - 1$).
A new bootstrap test for multiple assets joint risk testing

Dynamic conditional correlation model. The DGP for the \((N \times 1)\) vector \(y_t\) is given by

\[
y_t \sim \mathcal{N}(0, \Sigma_t),
\]

\[
\Sigma_t \doteq (I - A - B) \Sigma + Ay_{t-1}y_{t-1}' + B \Sigma_{t-1},
\]

with \(\Sigma \doteq \rho J + (1 - \rho)I\), \(A\) is an \((N \times N)\) diagonal matrix with 0.02 in the diagonal, and \(B\) is an \((N \times N)\) matrix with 0.97 in the diagonal. That is,

\[
y_t \sim \mathcal{N}(0, \Sigma_t),
\]

\[
\Sigma_t \doteq (1 - a - b) \Sigma + ay_{t-1}y_{t-1}' + b \Sigma_{t-1},
\]

with \(a = 0.02\) and \(b = 0.97\). Again, we set \(\rho = 0.9\). Note that the time-varying dependence structure is endogenous here, in the sense that the dependence structure at time \(t\) not only depends on the dependence structure at time \(t\) but also on \(y_{t-1}\) (given the dependence structure at time \(t - 1\)).

We compare rejection frequencies under the null hypothesis, based on 500 simulated data sets per scenario. For each data set, the bootstrapped \(p\) value is computed using 500 replications. Our results are reported in Figure 2. For various \(N\) and \(T\), we display the average and 95\% confidence bands of the frequency with which the test rejects the null hypothesis (under \(H_0\)) at the 1\%, 5\% and 10\% significance levels. Figure 2 indicates that the size seems correct for the three DGPs, including the two DGPs involving different time-varying dependence structures.

3.2 Power of the test as the number of series \(N\) (all with wrongly specified marginals) increases

In order to measure the power of our test procedure as well as the possible gain in power obtained with a larger number of time series \(N\), we consider the constant correlation (CC) model, where we generate data sets with shifted mean \(\mu > 0\) of the marginal distributions ranging from 0 to 0.5; we wrongly assume \(\mu = 0\) in the PITs (i.e., in \(\hat{F}_{t|t-1,i}(y_{ti})\) and \(\hat{f}_{t|t-1,i}(u)\) in (2.1)). We simulate 500 data sets, and each time we use 500 simulated data sets in our bootstrap test procedure. We consider \(T = 100\) observations for \(N = 1\), \(N = 10\) and \(N = 100\). The results for a cross-correlation of 0, 0.5, 0.7 and 0.9, respectively, are reported in Figure 3. The plots report the 95\% confidence bands of the frequency of rejection of the null hypothesis (at the 5\% significance level).

As expected, we observe the following two results. First, the power obviously increases as the actual mean \(\mu\) in the marginal distributions of the DGP is further away from 0. Second, the gain in power from using a larger number of time series \(N\) is smaller for higher values of the cross-correlation. This makes sense, since if
FIGURE 2  Size results.

Size results for the (a) constant correlation (CC) model, (b) two-regime Markov switching correlation (2RC) model and (c) dynamic conditional correlation (DCC) model. Various sample sizes $T$ and numbers of time series $N$ are considered for 500 Monte Carlo replications. The plots display the average and the 95% confidence bands of the frequency with which the null hypothesis is rejected, at the (a) 1%, (b) 5% and (c) 10% significance levels. For each simulated data set, the number of bootstrap replications in the test is set to 500.
the cross-correlation tends to 1, then the addition of extra time series (which are then all (scaled) versions of the same time series) does not increase the power of the test. However, the median cross-correlation between daily returns on equities within a stock index (and between different worldwide equity indexes) is typically not as high as 0.7 or 0.9, but rather in the neighborhood of 0.5, as can be seen in the empirical section.

If all cross-correlations are equal to $\rho > 0$, and if each time series has the same mean $\mu \neq 0$ (whereas the marginal models assume the misspecified mean 0), then the power does not increase to 1 if $N \to \infty$. Intuitively, each additional $i$th time series can be better explained than the previous $i-1$ time series, where the extra information in the additional $i$th time series decreases to 0. This explains why the difference in power (for $\rho = 0.5, 0.7$ or 0.9) between $N = 1$ and $N = 10$ is substantially larger than the difference in power between $N = 10$ and $N = 100$. However, in practice, not all marginal models of the time series have the same amount of misspecification (here incorporated by the value of $\mu \neq 0$ in the DGP).

3.3 Power of the test as the number of series with wrongly specified marginals (out of ten series) increases

Figure 4 displays the power of our methodology for $N = 10$ times series, where the marginal distribution is wrongly specified for $M (M = 1, 5, 10)$ of the $N = 10$ series. We generate $M$ time series with shifted mean $\mu > 0$ of the marginal distributions ranging from 0 to 0.5, whereas we simulate $N - M$ time series with mean $\mu = 0$.

The increase in power if $M$ increases from 1 to 10 in Figure 4 is larger than the increase in power if $N$ increases from 1 to 10 in Figure 3. The reason is that the number of nuisance series $N - M$ (for which the null hypothesis is correct) decreases from 9 to 0 in Figure 4.

Obviously, even for large numbers of time series, the gain in power from adding even more time series can still be substantial if the models for the added time series suffer from substantial misspecification that was not present in the models for the other time series. Therefore, if one desires to test for the validity of a set of marginal models, it is obviously recommended to test for the validity of all these marginal models, which may give substantially higher power than testing for the validity of a subset of the models, even if all the time series are highly correlated.

4 EMPIRICAL APPLICATION: BACKTESTING GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY MODELS ON VARIOUS EQUITY UNIVERSES

The investigation of volatility dynamics has attracted many academics and practitioners, as this is of substantial importance for risk management, derivatives pricing
FIGURE 3 Power results.

Power results for the constant correlation (CC) model with shifted mean $\mu > 0$ of the marginal distributions ranging from 0 to 0.5, with a correlation value of (a) $\rho = 0$, (b) $\rho = 0.5$, (c) $\rho = 0.7$ and (d) $\rho = 0.9$, for $T = 100$ observations and various numbers of time series $N$. The plot displays the 95% confidence bands of the frequency of rejecting the null hypothesis as a function of the shifted mean $\mu$. The number of Monte Carlo replications is set to 500, and each time we use 500 bootstrap replications in our test procedure.

and portfolio optimization. Since the seminal paper by Bollerslev (1986), GARCH-type models have been widely used in financial econometrics for the forecasting of volatility. Nowadays, these are standard models in risk management; They are easy to understand and interpret, and they are available in many statistical packages. For a review on GARCH models, we refer the reader to Bollerslev et al. (1992) and Bollerslev et al. (1994).

In our application, we consider the AR(1)–GJR(1,1)\(^1\) specification for the log returns $\{r_t\}$:

\[
\begin{align*}
  r_t &= \mu + \rho r_{t-1} + \eta_t \quad (t = 1, \ldots, T), \\
  \eta_t &= \sigma_t \varepsilon_t \quad \varepsilon_t \sim \text{iid } f_{\varepsilon}, \\
  \sigma_t^2 &= \omega + (\alpha + \gamma I\{u_{t-1} \leq 0\})u_{t-1}^2 + \beta \sigma_{t-1}^2,
\end{align*}
\]  

\(4.1\)

\(^1\)AR: autoregressive; GJR: Glosten–Jagannathan–Runkle.
FIGURE 4  Power results (number of series with wrongfully specified marginals out of ten series).

Power results for the constant correlation (CC) model with shifted mean $\mu > 0$ of the marginal distributions ranging from 0 to 0.5 for $M = 1,5,10$ out of $N = 10$ series, and $\mu = 0$ for the remaining $N - M$ series. The correlation values of (a) $\rho = 0$, (b) $\rho = 0.5$, (c) $\rho = 0.7$ and (d) $\rho = 0.9$ are considered for $T = 100$ observations and various numbers of time series $N$. The plot displays the 95% confidence bands of the frequency of rejecting the null hypothesis as a function of the shifted mean $\mu$. The number of Monte Carlo replications is set to 500, and each time we use 500 bootstrap replications in our test procedure.

where $\omega > 0$ and $\alpha, \gamma, \beta \geq 0$ to ensure a positive conditional variance. $I \{ \cdot \}$ denotes the indicator function, which has a value of one if the constraint holds, and zero otherwise. No constraints have been imposed to ensure covariance stationarity; however, as a sensitivity analysis, we repeated the whole study with additional covariance stationarity constraints, which yielded approximately the same results with qualitatively equal conclusions. Another sensitivity analysis that included variance targeting (where the unconditional variance in the GARCH model is set equal to the sample variance, reducing the number of parameters to be estimated by one) led to similar results. If we impose $\gamma = 0$, this results in the symmetric GARCH model. For the distribution $f_x$, we consider the simple Gaussian and Student $t$ distributions, together with a nonparametric Gaussian kernel estimator. The Student $t$ distribution is probably
the most commonly used alternative to the Gaussian for modeling stock returns, and it allows us to model fatter tails than the Gaussian. The kernel approach gives a non-parametric alternative, which can deal with skewness and fat tails in a convenient manner.

Models are fitted by quasi-maximum likelihood. For the nonparametric model, the bandwidth is selected by the rule-of-thumb of Silverman (1986) on the residuals of the quasi-maximum likelihood fit; alternative bandwidth choices lead to similar results. We rely on the rolling window approach, where 1000 log returns (approximately four trading years) are used to estimate the models. Similar results were obtained for windows of 750 and 1500 observations. The next log return is used as a forecasting window. The model parameters’ estimates are updated every day.

We test the performance of the models on several universes:

(i) a set of nine international stock market indexes: the S&P 500 (US), Financial Times Stock Exchange (FTSE) 100 (UK), Cotation Assistée en Continu (CAC) 40 (France), Deutscher Aktienindex (DAX) 30 (Germany), FTSE Milano Italia Borsa (MIB; Italy), Toronto SE 300 (Canada), All Ordinaries (AORD) index (Australia), Taiwan Stock Exchange Corporation (TSEC) weighted index (Taiwan) and Hang Seng index (Hong Kong);

(ii) equity universe of the DAX 30 (as of June 2013);

(iii) equity universe of the EURO STOXX 50 (as of June 2013); and

(iv) equity universe of the S&P 100 (as of June 2013).

For all data sets, we take the daily log return series defined by \( y_t = \log(P_t) - \log(P_{t-1}) \), where \( P_t \) is the adjusted closing price (index) on day \( t \). The period ranges from January 1990 to June 2013. Our data is downloaded from Datastream. The data is filtered for liquidity following Lesmond et al (1999). In particular, we remove the time series with less than 1500 data points of history, with more than 10% of zero returns and with more than two trading weeks of constant price. This filtering approach reduces the databases of all universes. More precisely, this leads to

(i) nine equity indexes,

(ii) twenty-four equities for the DAX 30,

(iii) thirty-nine equities for the EURO STOXX 50 and

(iv) eighty-five equities for the S&P 100.

The numbers of observations also vary among the universes. We have 3824 observations for the nine indexes, 3367 for the DAX 30, 3549 for the EURO STOXX 50 and 3348 for the S&P 100.
FIGURE 5 One-year rolling window cross-correlation of daily log returns.

(a) Universe of nine stock indexes. (b) DAX 30 equities. (c) EURO STOXX 50 equities. (d) S&P 100 equities. The black line represents the median value of all the cross-correlations, while the red lines report the 5th and 95th percentiles of the distribution of cross-correlations. The date in the horizontal axis indicates the end of the one-year rolling window.

In Figure 5, we display the median (and 5th and 95th percentiles) of the one-year rolling window correlation of the daily log returns for the four considered universes. Although the cross-correlations vary over time, the median cross-correlation remains closer to (and often below) 0.5 than to 0.7 or 0.9. Hence, we expect that our bootstrap test procedure will have a reasonable-to-good ability to detect invalid specifications of the marginal models.

We perform our bootstrap test procedure for the extension of the test of Berkowitz (2001), as suggested by Christoffersen and Pelletier (2004), with a specific focus on the left tail (with $\alpha = 0.05$ and $\alpha = 0.01$). We also test for correct conditional coverage of the 95% and 99% VaR, using the test of Christoffersen (1998). The results for the four universes are reported in Table 1.
The GJR model with Student t innovations seems to be the best model. For the \( N = 24 \) equities in the DAX 30 and the \( N = 39 \) equities in the EURO STOXX 50 indexes, it is not rejected at a significance level of 5% and 1%, respectively. The other models have at least two \( p \) values of 0.0000 (for the two Berkowitz tests) for each of the four universes. The \( p \) values of 0.000 for the correct conditional coverage test of the 99% VaR, which is known to have low power for a univariate time series, reflect the gain in power of our bootstrap test procedure for multiple time series.

4.1 Precrisis, crisis and postcrisis periods

The forecast sample period covers the financial crisis of 2008. The performance of the models may vary between the financial crisis period and the precrisis and postcrisis periods. It is important to know if the rejection of most of the models’ forecasts can be ascribed to the crisis period only, or not. We thus present a comparison of the models’ forecasting performance between the crisis (September 2007–February 2009), precrisis (January 2005–August 2007) and postcrisis (March 2009–June 2011) periods (see Ardia and Boudt 2013). The results are reported in Table 2, which shows that the rejection of the density forecasts from the GARCH(1,1) and GJR(1,1) models with Gaussian innovations, the GJR(1,1) model with a Gaussian kernel density estimate for the innovations and the exponentially weighted moving average (EWMA) for the variance estimate is not only due to the crisis period. For the precrisis and postcrisis subperiods, the validity of these forecasts is also rejected. However, the GJR model with Student t distributed innovations is rejected for the crisis period; this is also the case for the DAX 30 and EURO STOXX 50 universes, in which the GJR model was not rejected for the whole period (due to its good performance in the precrisis period). Therefore, none of the marginal models seems appropriate in the crisis subperiod. This stresses the usefulness of our bootstrap test procedure, since it enables us to perform a relatively powerful test, while focusing on a relatively short (crisis) subperiod of eighteen months, in the context of tests for correct conditional coverage that are known to have relatively low power for univariate time series.

5 POSSIBLE EXTENSIONS OF THE PROCEDURE AND THE APPLICATIONS

In this section, we make several remarks on further possibilities regarding the novel test procedure. First, the procedure can easily cope with situations in which the time series are not observed in exactly the same periods. If for some of the time series the PIT is not observed in the first part of the period, then one simply includes missing values or not-a-number (NaN or NA) values for the corresponding elements in these rows of the matrixes \( X \) and \( R \). For the simulated data sets under the null hypothesis,
TABLE 1  Performance results for the whole considered period.

(a) Worldwide equity indexes: $N = 9$, $T = 3824$

<table>
<thead>
<tr>
<th></th>
<th>Berk/5% CC95</th>
<th>Berk/1% CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.022</td>
<td>0.050</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.824</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

(b) DAX 30 equity universe: $N = 24$, $T = 3367$

<table>
<thead>
<tr>
<th></th>
<th>Berk/5% CC95</th>
<th>Berk/1% CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.184</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.168</td>
<td>0.052</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.672</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

(c) EURO STOXX 50 equity universe: $N = 39$, $T = 3549$

<table>
<thead>
<tr>
<th></th>
<th>Berk/5% CC95</th>
<th>Berk/1% CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.026</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.222</td>
<td>0.136</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.196</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

(d) S&P 100 equity universe: $N = 85$, $T = 3348$

<table>
<thead>
<tr>
<th></th>
<th>Berk/5% CC95</th>
<th>Berk/1% CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.002</td>
<td>0.144</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.944</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Bootstrapped $p$ values for the 5% lower tail (Berk/5%) and 1% lower tail (Berk/1%) test of Christoffersen and Pelletier (2004), which extends Berkowitz (2001), and for the conditional coverage test of Christoffersen (1998) for the VaR95 (CC95) and VaR99 (CC99). All models contain an AR(1) part. GARCH-N: symmetric GARCH(1,1) with Gaussian innovations; GJR-N: asymmetric GJR(1,1) with Gaussian innovations; GJR-S: asymmetric GJR(1,1) with Student $t$ innovations; GJR-K: GJR(1,1) with a Gaussian kernel density estimate for the innovations. EWMA denotes the exponentially weighted moving average for the variance estimate, where a Gaussian distribution is assumed for the innovations. The bootstrap test is computed with 500 replications.

www.risk.net/journal  
To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
### TABLE 2 Performance results for the precrisis, crisis and postcrisis subperiods. [Table continues on next two pages.]

(a) Precrisis

<table>
<thead>
<tr>
<th></th>
<th>Berk/5%</th>
<th>CC95</th>
<th>Berk/1%</th>
<th>CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worldwide equity indexes: N = 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.444</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.952</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.794</td>
<td>0.513</td>
<td>0.000</td>
<td>0.370</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>1.000</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>DAX 30 equity universe: N = 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.010</td>
<td>0.000</td>
<td>0.004</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.206</td>
<td>0.150</td>
<td>0.316</td>
<td>0.364</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.034</td>
<td>0.000</td>
<td>0.608</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.012</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>EURO STOXX 50 equity universe: N = 39</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.032</td>
<td>0.000</td>
<td>0.002</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.420</td>
<td>0.160</td>
<td>0.342</td>
<td>0.134</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.030</td>
<td>0.000</td>
<td>0.426</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.040</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>S&amp;P 100 equity universe: N = 85</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.008</td>
<td>0.000</td>
<td>0.000</td>
<td>0.024</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.436</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

The NaN values are simply discarded when computing the test statistic for a column, so the test statistic is computed for a consecutive time series.

Second, the replications in our bootstrap test procedure (i.e., the bootstrapped data sets) are independent, and the test statistics for each time series are computed separately. These two properties facilitate parallel implementation on multiple central processing units (CPUs) or graphics processing units (GPUs). Such an implementation may enormously reduce the required computing time, so that one can consider larger universes of stocks within a reasonable period. Or, alternatively, this enables the use of rather time-consuming tests for each time series, such as a bootstrap test.
TABLE 2  Continued.

(b) Crisis

<table>
<thead>
<tr>
<th></th>
<th>Berk/5%</th>
<th>CC95</th>
<th>Berk/1%</th>
<th>CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td>Worldwide equity indexes: $N = 9$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.192</td>
<td>0.000</td>
<td>0.486</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.016</td>
<td>0.000</td>
<td>0.232</td>
<td>0.022</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>DAX 30 equity universe: $N = 24$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.048</td>
<td>0.000</td>
<td>0.156</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>EURO STOXX 50 equity universe: $N = 39$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.084</td>
<td>0.000</td>
<td>0.396</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.004</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>S&amp;P 100 equity universe: $N = 85$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.114</td>
<td>0.000</td>
<td>0.328</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.020</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

procedure for the validity of the expected shortfall or Cramér–von Mises tests. This would then lead to a bootstrap within a bootstrap test.

Third, it is also possible to add different related test statistics for the same time series, in order to produce one joint test statistic (and a corresponding $p$ value). For example, in our empirical application, we could add the four test statistics (of the two Berkowitz tests and the two conditional coverage tests) for each time series to construct one $p$ value that summarizes the four $p$ values in Table 1 or Table 2. If in each data set simulated under the null hypothesis these four test statistics are computed from the same column of simulated PIT values, then the simulated distribution under the null hypothesis automatically takes into account the dependence between the four test statistics.
TABLE 2 Continued.

<table>
<thead>
<tr>
<th></th>
<th>Berk/5%</th>
<th>CC95</th>
<th>Berk/1%</th>
<th>CC99</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Worldwide equity indexes: N = 9</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.216</td>
<td>0.004</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.158</td>
<td>0.018</td>
<td>0.000</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.492</td>
<td>0.054</td>
<td>0.680</td>
<td>0.184</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.184</td>
<td>0.270</td>
<td>0.002</td>
<td>0.604</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>DAX 30 equity universe: N = 24</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.010</td>
<td>0.000</td>
<td>0.382</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.016</td>
<td>0.000</td>
<td>0.310</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.426</td>
<td>0.328</td>
<td>0.146</td>
<td>0.800</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.040</td>
<td>0.000</td>
<td>0.488</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.234</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>EURO STOXX 50 equity universe: N = 39</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.322</td>
<td>0.000</td>
<td>0.004</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.234</td>
<td>0.000</td>
<td>0.004</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.474</td>
<td>0.188</td>
<td>0.000</td>
<td>0.240</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.282</td>
<td>0.000</td>
<td>0.694</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.024</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td><strong>S&amp;P 100 equity universe: N = 85</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GARCH-N</td>
<td>0.000</td>
<td>0.004</td>
<td>0.000</td>
<td>0.050</td>
</tr>
<tr>
<td>GJR-N</td>
<td>0.000</td>
<td>0.012</td>
<td>0.000</td>
<td>0.178</td>
</tr>
<tr>
<td>GJR-S</td>
<td>0.168</td>
<td>0.064</td>
<td>0.080</td>
<td>0.748</td>
</tr>
<tr>
<td>GJR-K</td>
<td>0.000</td>
<td>0.014</td>
<td>0.000</td>
<td>0.718</td>
</tr>
<tr>
<td>EWMA</td>
<td>0.000</td>
<td>0.296</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Bootstrapped p values for the 5% lower tail (Berk/5%) and 1% lower tail (Berk/1%) test of Christoffersen and Pelletier (2004), which extends Berkowitz (2001), and for the conditional coverage test of Christoffersen (1998) for the VaR95 (CC95) and VaR99 (CC99). All models contain an AR(1) part. GARCH-N: symmetric GARCH(1,1) with Gaussian innovations; GJR-N: asymmetric GJR(1,1) with Gaussian innovations; GJR-S: asymmetric GJR(1,1) with Student t innovations; GJR-K: asymmetric GJR(1,1) with a Gaussian kernel density estimate for the innovations. EWMA is the exponentially weighted moving average for the variance estimate, where a Gaussian distribution is assumed for the innovations. The bootstrap test is computed with 500 replications. The window sizes are 2305, 1848, 2030 and 1829 for the precrisis period for the four universes, 370 for the crisis period and 1149 for the postcrisis period.

Fourth, our proposed method is also applicable if $T < N$ (for example, if one wants to test the validity of a large number of marginal models over a very short period, such as a few days around a crash). In such cases, it would even be difficult to compute a useable covariance matrix.

Fifth, our method can be used to summarize a set of p values for the validity of models for multiple data sets (with which existing papers in the literature may...
conclude in order to prove the validity of the proposed model) in one summarizing $p$ value.

Sixth, the method can be used to investigate the effect of the particular dependence of a given data set on a procedure that implicitly assumes independence (or a limited type of dependence, such as independence between a large enough number of subsets); this is an assumption that may be substantially violated by, for example, time series of asset returns. One important example is the FDR methodology of Storey (2002). For example, we can investigate how the estimated percentage of correct models (and the confidence interval for this percentage) is affected, under the assumption that all models are correct, or under various assumptions on the percentage and nature of misspecified models, where a block bootstrap extension of our method may be used for the latter. This may facilitate the generation of more robust results for the FDR methodology in case of highly dependent time series. Note that in the case of highly dependent time series, we do not have independently uniformly distributed $p$ values under the null hypothesis, where the latter basically forms the underlying principle of the FDR methodology. Storey and Tibshirani (2003) discuss the FDR under different types of dependence. Under their most general assumptions on the dependence, their estimator is (conservatively) biased. Further, they consider DNA data, which arguably has a substantially different dependence structure than, for example, time series in economics and finance.

6 CONCLUSION

We have introduced a novel simulation-based methodology to test the validity of a set of marginal time series models, where the dependence structure between the time series is taken directly from the observed data. We have illustrated its correct size and its power for typically realistic levels of cross-correlation, and have shown its potential usefulness in empirical applications involving GARCH-type models for daily returns on stocks or stock indexes. We have also outlined several possible extensions of the procedure and its applications, which we will consider in future research.

DECLARATION OF INTEREST

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

ACKNOWLEDGEMENTS

We are grateful to Kris Boudt, William Doehler and a referee for numerous useful comments. David Ardia gratefully acknowledges FQRSC, IFM2 and Industrielle Alliance for financial support. Lukasz Gatarek is grateful to National Science Center,
Poland, for funding with grant Preludium No. 2013/09/N/HS4/03751. Any remaining errors or shortcomings are the authors’ responsibility.

REFERENCES

Default risk charge: modeling framework for the “Basel” risk measure

Sascha Wilkens and Mirela Predescu

Risk Analytics & Modelling, BNP Paribas, 10 Harewood Avenue, London NW1 6AA, UK; emails: sascha.wilkens@uk.bnpparibas.com, mirela.predescu@uk.bnpparibas.com

(Received January 6, 2016; revised August 12, 2016; accepted August 15, 2016)

ABSTRACT

As a result of the Basel Committee on Banking Supervision’s Fundamental Review of the Trading Book, revised standards for capital requirements for market risk in banks’ trading books have been issued. Under the new standards, default risk needs to be measured and capitalized through a dedicated default risk charge (DRC). Although quantitative impact studies are ongoing and banks are preparing for these regulatory changes, this paper is the first to present a modeling framework for the DRC measure that projects losses over a one-year capital horizon at a 99.9% confidence level. We discuss selected risk factor models, which we use to derive simulation-based loss distributions and associated default risk figures. The model’s properties, aspects of its implementation and a comparison with the standardized approach for default risk are explored through the use of example portfolios.

Keywords: banking regulation; risk modeling; market risk; Fundamental Review of the Trading Book (FRTB); default risk charge (DRC).

1 INTRODUCTION AND REGULATORY BACKGROUND

Since the mid-1990s, banks have been allowed to use internal models to calculate market risk capital requirements for activities in their trading books. These models
are subject to approval by supervisory authorities. They are supposed to reflect large, adverse market moves and are usually based on a 99% ten-day value-at-risk (VaR). Owing to the fact that VaR models usually do not account for the potential illiquidity of trading positions, and since market losses can be driven by large cumulative price moves, in the aftermath of the worldwide 2007–8 financial crisis, new capital charges were proposed to complement the existing market risk capital requirements (Basel Committee on Banking Supervision 2009, 2011; European Banking Authority 2012). The concept of stressed VaR and its associated capital charges were introduced in January 2013 (in Europe); these were supplemented by the incremental risk charge (IRC), which was intended to cover market risks from credit rating migrations and defaults for flow instruments, such as bonds and credit default swaps (CDSs). More complex instruments such as collateralized debt obligations (CDOs) have been made subject to a separate, extended market risk model: the comprehensive risk measure (CRM). These changes, widely referred to as “Basel 2.5”, can be seen as rather short-term fixes, as they do not provide a consistent overall risk capital framework.

With the aim of creating a more consistent market risk capital framework, a Fundamental Review of the Trading Book (FRTB) was recently conducted by the Basel Committee on Banking Supervision (BCBS). The result of the FRTB has been the revision of the minimum capital requirements for market risk (see Basel Committee on Banking Supervision 2016). The main risk figure determining capital requirements will be changed to an expected shortfall (ES) at a 97.5% confidence level. In addition, this figure shall account for different liquidity horizons of trading positions and their associated risk factors, and it will be calibrated to a period of market stress. It is to be accompanied by a default risk charge (DRC) for nonsecuritization products, which will measure the trading portfolio’s default risk based on the 99.9% loss over a one-year capital horizon, or, at an institution’s discretion, a minimum of sixty days for equity subportfolios. In contrast to the IRC, the DRC measure does not consider rating migration risk and rules out any portfolio rebalancing assumptions. The coverage is extended to equity positions, and projected losses need to reflect stressed market conditions under the new rule set. The DRC model is to be applied in conjunction with the ES-based market risk capital charge. Should the latter fail, given quality criteria such as backtesting and profit and loss (P&L) attribution, either at the bank-wide or trading-desk level, a standardized approach (SA) will need to be applied for both ES and DRC. Otherwise, the DRC can be calculated using an institution’s own internal model, which requires validation by the relevant home supervisor and has to be computed at least on a weekly basis. For securitization products (correlation trading and noncorrelation trading), the DRC measure needs to be calculated using a dedicated SA; internal models are not allowed.

While the industry is participating in ongoing quantitative impact studies (QISs) and preparing for the implementation of the revised market risk capital standards, no
literature on a suitable model for DRC has been published so far. The only notable exception is Laurent et al (2016), in which the regulation-imposed constraints on correlation matrixes and the factor structures that underpin the dependence between defaults are analyzed. In the case of IRC and CRM, Wilkens et al (2013) spearheaded the modeling discussion. By presenting a comprehensive DRC model, this paper closes a substantial gap in the research and can serve as a basis for an industry-wide discussion. Our model development is accompanied by example calculations and a discussion of implementation aspects. As in the case of IRC, high confidence levels and long projection horizons, in conjunction with limited backtesting feasibility, leave a substantial model risk.

The paper is organized as follows. Section 2 discusses the fundamental elements of a modeling framework for DRC. Section 3 focusses on the modeling of marginal, joint default and recovery rate risk. Section 4 is dedicated to the generation of P&L distributions, which are illustrated using a range of example portfolios. Aspects such as convergence and sensitivity analyses are also addressed. A quantitative comparison between the DRC based on the developed model and the SA is provided as well. Section 5 concludes. Some details on model data and calibration as well as a discussion of some model aspects and results have been relegated to an online appendix.

2 FUNDAMENTAL ELEMENTS OF A MODELING FRAMEWORK

The DRC has to capture default risk in a bank’s trading book (Basel Committee on Banking Supervision 2016, p. 60).1 Extending the coverage of the former IRC, the affected instruments are those which are not subject to standardized charges, and whose valuations do not depend solely on commodity prices or foreign exchange rates. Hence, bonds (including defaulted debt positions) and vanilla credit derivatives such as CDSs on single names and indexes as well as equity positions are in scope. While this is already the case for IRC, the regulation emphasizes the inclusion of sovereign exposures (including those denominated in the sovereign’s domestic currency). The risk is to be measured by means of a VaR-type measure at a 99.9% confidence level for a one-year capital horizon (ie, assuming a constant portfolio), or, at an institution’s discretion, a sixty-day horizon for designated equity subportfolios. This setup propagates the use of instantaneous shocks, ie, any time-value changes as well as cashflows are to be ignored.

With regard to the granularity of the simulation, the regulation foresees the obligor and its default risk as the primary level, while accounting for different losses from

---

1 The previously contemplated incremental default risk (IDR) was supposed to allow for the removal of any double counting between ES and default charge. This option has been removed with the modification leading to the DRC.
different instruments backed by the same obligor. As an example, if a parent company is the guarantor for the bond issues of two subsidiaries, the default risk for long positions amounts to the sum of the present values minus the bond-specific (simulated) recovery rates. The DRC model needs to reflect the probability of default of the obligor itself in conjunction with both (conditional) recovery rates.

3 CAPTURING DEFAULT AND RECOVERY RATE RISK

3.1 Marginal default risk

As a first step, default probabilities of corporates and sovereigns need to be determined. These can be implied from market prices of bonds and credit derivatives (known as risk-neutral default probabilities) or calculated from historical default observations (known as historical or objective default probabilities). The difference between risk-neutral and historical default probabilities is discussed in Hull et al (2005), among many others. A shortcoming of market-implied probabilities is that they embed market risk premiums, which tend to bias the prediction of the actual default frequency. Regulation emphasizes that a correction of the market-implied default probabilities would be mandatory to arrive at objective probabilities of default (Basel Committee on Banking Supervision 2016, p. 61). Given the difficulties in estimating market risk premiums and the large uncertainty around the estimates, the use of historical default probabilities is usually preferable.

In an attempt to increase the accuracy of future predictions, historical default rates, by rating, can be differentiated by attributes such as type of obligor (e.g., corporate, sovereign), region and industry. A high granularity of default probabilities needs to be balanced against the data availability of historical default observations and whether differences are statistically significant (see, for example, Moody’s 2011). In addition, certain highly rated types of obligors that did not experience defaults in the past (e.g., AAA-rated sovereigns) will lead to a default probability of zero unless additional assumptions are imposed.

In general, default rates tend to vary over business cycles, with more defaults observed during recessions (see Altman and Kalotay (2014), among many others). Changes in default risk can be traced further, for example, to account for global and industry effects (see Aretz and Pope 2013). One can generally distinguish between through-the-cycle and point-in-time default probabilities. While point-in-time estimates tend to be more risk sensitive, as they better reflect current economic conditions, they can imply instability in the forecast as well as potentially procyclical risk measures. As such, they can lead to procyclical capital requirements. For the application in the context of DRC, through-the-cycle probabilities therefore seem preferable. Based on these considerations, in the following, through-the-cycle default
Default risk charge

probabilities differentiated by corporates and sovereigns as provided by Standard & Poors (S&P) are used.

In practice, regulation requires that banks with an internal ratings-based approach (IRBA) use probabilities of default from their own internal framework (see Basel Committee on Banking Supervision 2016, p. 62). The spirit of the regulation is therefore to encourage banks not to rely on external ratings and corresponding default probabilities for the determination of capital, but rather to have their own assessment of the credit quality of different obligors or types of securities.

Table 1 provides the average one-year default probabilities between 1981 and 2012 for corporates and between 1975 and 2012 for sovereigns. For corporates with ratings better than AA−, and for sovereigns with ratings better than BB+, default probabilities are floored at 3 basis points (bps), which is the minimum value prescribed by regulation. The sensitivity of the DRC measure to the floor is likely material, given that banks’ portfolios tend to comprise significant positions in securities of well-rated sovereigns and corporates.

The default probabilities in Table 1 determine the asset return thresholds that trigger defaults within the simulation model.

3.2 Correlation of defaults across obligors

3.2.1 Overview

The DRC model needs to reproduce the dependence between defaults of different obligors. Regulation prescribes the use of a default correlation model with two types of systematic factors, based on listed equity prices or CDS spreads (Basel Committee on Banking Supervision 2016, p. 60). The correlations must be based on “a calibration period of at least ten years that includes a period of stress” as well as “measured over a one-year liquidity horizon” (Basel Committee on Banking Supervision 2016, p. 61).

These regulatory requirements provide only loose guidance on the correlation estimates to be used in the default model. Different proxies have been considered in the literature for the objective of measuring default correlations. Moody’s (2008) makes the case for using asset correlations (estimated via a structural Merton-type model (Merton 1974)) as predictors for subsequently realized default correlations.

---

2 This requirement is similar in nature to the rules in the United States, where the Dodd–Frank Wall Street Reform and Consumer Protection Act (2010) requires federal agencies to remove any references to external credit ratings from federal regulations, and to replace such ratings with different standards of credit quality; see Section 939A, paragraph (b).

3 See the critical assessment by Chourdakis and Jena (2013) on the levels of default probability that can be inferred for events with few or even no occurrences in history, such as sovereign defaults. In general, the regulatory requirement of a floor can be incorporated into more advanced methods to calibrate default probability curves (see, for example, Tasche 2013).
TABLE 1 Default probabilities by rating.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td>AA</td>
<td>0.00*</td>
<td>0.00*</td>
</tr>
<tr>
<td>AA–</td>
<td>0.02*</td>
<td>0.00*</td>
</tr>
<tr>
<td>AA−</td>
<td>0.03*</td>
<td>0.00*</td>
</tr>
<tr>
<td>A+</td>
<td>0.06</td>
<td>0.00*</td>
</tr>
<tr>
<td>A</td>
<td>0.07</td>
<td>0.00*</td>
</tr>
<tr>
<td>A−</td>
<td>0.07</td>
<td>0.00*</td>
</tr>
<tr>
<td>BBB+</td>
<td>0.14</td>
<td>0.00*</td>
</tr>
<tr>
<td>BBB</td>
<td>0.20</td>
<td>0.00*</td>
</tr>
<tr>
<td>BBB–</td>
<td>0.35</td>
<td>0.00*</td>
</tr>
<tr>
<td>BB+</td>
<td>0.47</td>
<td>0.10</td>
</tr>
<tr>
<td>BB</td>
<td>0.71</td>
<td>0.41&lt;sup&gt;&amp;&lt;/sup&gt;</td>
</tr>
<tr>
<td>BB–</td>
<td>1.21</td>
<td>1.70</td>
</tr>
<tr>
<td>B+</td>
<td>2.40</td>
<td>2.06&lt;sup&gt;&amp;&lt;/sup&gt;</td>
</tr>
<tr>
<td>B</td>
<td>5.10</td>
<td>2.50</td>
</tr>
<tr>
<td>B−</td>
<td>8.17</td>
<td>6.30</td>
</tr>
<tr>
<td>CCC/C</td>
<td>26.85</td>
<td>34.00</td>
</tr>
</tbody>
</table>

Source: Standard & Poor’s (2013a, 2013b). The average historical one-year default rates (in %) by rating for corporate and sovereign issuers, to serve as probabilities for future defaults. The rates are calculated as weighted averages of the one-year default rates for each year in the sample, with weights based on the number of issuers rated at the beginning of each year. Regulation prescribes the use of a 3bps floor; therefore, the default rates marked with “*” are to be floored accordingly. In order to ensure the monotonicity of default rates by rating, the historical values for BB and B+ sovereigns (0.00% and 0.60%, respectively) have been fitted from the adjacent rates, indicated by “&”.

Alternatively, equity correlations have been shown to considerably overestimate asset correlations (see, for example, Düllmann et al 2008) and to be, at best, noisy indicators of default correlations (see, for example, De Servigny and Renault 2002; Qi et al 2015). Vassalou and Xing (2004) analyze the intimate relationship between default risk and equity returns, also providing a risk-based interpretation of certain stylized effects (eg, company size). In general, equity (and CDS) prices embed much more codependence than the simple default correlation. With default being an absorbing state, market prices of nondefaulted names by nature cannot provide a straightforward measure of default correlation. Some recent approaches, such as Liu et al (2012), allow the direct use of equity correlations in the prediction of joint defaults; however, they rely on the correlations implied by actual defaults as an additional calibration source.

While one should take into account all the previous findings on default correlation estimates, the explicit reference in the regulation to listed equity prices (or CDS spreads) points toward the use of a direct measure of correlation between them. This requirement rules out alternatives, such as structural-type models or other approaches,
which use additional information (e.g., debt information, asset values, actual default correlations). Using directly observed equity (or CDS spread) correlations has a significant advantage for portfolio risk measurement, management, and analysis, as it allows us to clearly explain the drivers and the magnitude of the co-movements between the names in the portfolio or subportfolios, and perform stress tests and sensitivity analyses with respect to the market level of correlations.

3.2.2 Data

In this paper, equity prices are used in the estimation of the correlation model. The application to CDS spreads would be analogous. In general, equity data offers a significantly wider coverage than CDS spreads for corporate obligors; hence, it is preferable, given the wide scope of the DRC, which encompasses both equity and credit products.

From a very fundamental perspective, “correlation” can be defined and measured in many different ways. This poses a series of challenges and pitfalls, especially in the context of tail events (see, for example, Embrechts et al. 1999). Regulation does not provide any guidance on the intended interpretation; given the simple implementation and measurement, we can assume that the framework refers to a classical linear (Pearson) correlation, in spite of its known shortcomings. For the actual parameter determination, return windows and the measurement interval need to be chosen. While the targeted forecasting horizon refers to a one-year period, it is not viable to measure correlations based on (nonoverlapping) one-year equity returns, since the great uncertainty in the measurement prevents this. Along the same lines, the measurement interval needs to have a minimum length in order to allow for a meaningful estimation of return correlations. As an illustration, in the case of a measured (Pearson) correlation between two equities of +10%, the 95% confidence interval around a correlation estimate amounts to approximately [-86%, +90%] for a sample of five observations, and shrinks to about [-36%, +52%] for twenty observations and [-24%, +42%] for thirty-five observations. Hence, the combination of length of return window and measurement interval needs to allow for a certain minimum number of observations in order to render the exercise reasonable.

One possible interpretation of the prescribed calibration period is to measure correlations over rolling time windows covering at least ten years, thereby identifying a period of stress. The latter is not defined further; it is likely supposed to point to a period for which the DRC model provides the comparatively highest loss estimate for a bank. This leaves two routes for investigation. As a first approach, the period can be identified as that with the highest (pairwise) correlation. This is a suitable route in case the DRC is ceteris paribus higher the higher the correlation. For a directional portfolio (for example, holding bonds), this assumption is valid for reasonable marginal
distributions. In order to put this approach into practice, the running pairwise correlation over a period of at least ten years can serve for the stress identification. As a second approach, the DRC can be run multiple times for a series of correlation measurement periods, and the stress period can be identified as that leading to the highest loss. While this approach does not require an ex ante assumption on the DRC as a function of correlation levels, it is much more involved from a computational perspective, and likely less stable, since it is portfolio dependent. The latter, in particular, is not a desirable feature for a risk measure with a long projection horizon. The example portfolios in Section 4.2 readdress this topic and study the influence of correlation assumptions on the DRC by means of sensitivity analyses.

Another possible interpretation of the prescribed calibration period is the direct use of at least ten years of data to measure correlations, justifying that it covers a period of stress. Under this interpretation, the correlation estimates would exhibit more of a through-the-cycle character. Nevertheless, one would likely need to revert to the tools described for the first interpretation.

In the following, the first approach is adopted, targeting the identification and use of a period of stress for the correlation input data used to calibrate the model. The practical interpretation chosen here refers to monthly (nonoverlapping) returns over three-year periods, which leads to thirty-five returns for each measured pairwise correlation. This can be seen as a reasonable compromise between the forecasting horizon of the model and the parameter estimation from historical data; robustness checks with regard to the choice are generally recommended. The embedded assumption is that correlations measured over monthly and annual intervals are identical (and a good predictor for future one-year correlations). This hypothesis can be challenged (see studies such as Erb et al (1994) and Turley (2012), among many others), but, given the uncertainty of the correlation measurement itself (confer the width of the confidence intervals), it is hard to reject from a statistical perspective.

For the subsequent illustrations, historical equity prices of the constituents of a broad set of worldwide equity indexes are used. Using the index compositions at end-December 2013, the time series for 1990–2013 of members from EURO STOXX 50, Swiss Market Index (SMI), Financial Times Stock Exchange (FTSE) 100, S&P 500,
Australian Securities Exchange (ASX) 200, Hang Seng index (HSI) and Nikkei 225 are selected. Section A of the online appendix provides further details on the data.

Pairwise equity correlations are calculated over a rolling window spanning three years and using monthly returns. Figure 1 shows different percentiles of the distribution of correlations as well as the average width of the 95% confidence level around the individual correlation estimates. The peak of the correlation is found around the 2008–9 time period, as expected, with a median correlation of about 45%; the corresponding surrounding three-year period is marked in the figure and chosen as September 2007 through to September 2010 in the following. While the estimation uncertainty is high, as indicated by the confidence intervals, the correlation pattern over time clearly points toward this crisis period. Notably, this also holds when conducting the analysis with different return windows (e.g., quarterly) and measurement periods (e.g., two years; not shown here).

3.2.3 Model description

One way of addressing the requirement that factor correlation models have two types of systematic factors, proposed by Laurent et al (2016), consists of calibrating empirically observed correlations to a “nearest” correlation matrix with a given number of factors. While this approach is a best fit in mathematical terms, it will usually be difficult to attribute an economic meaning to the resulting factors. Furthermore, the stability of such fitted matrixes over time might not be sufficient.

An alternative approach pursued here is to define the economic factors driving equity returns. The first systematic factor is a global factor that is common to all obligors, reflecting the overall state of the economy and, in particular, potential periods of stress. For the second systematic factor, we consider both a country- and an industry-specific factor. While this in essence reflects a full set of country and industry factors,

---

6 Since the DRC covers both equity- and credit-sensitive trading book positions, one might enlarge the equity data set to explicitly cover the constituents from the main credit indexes, such as iTraxx and the credit default swap index (CDX). Many of the names, however, are already covered by the equity index selection used here, for example, iTraxx Europe is well represented by EURO STOXX 50 and FTSE 100. The selection from the main equity indexes puts an inherent focus on larger companies. If a bank’s portfolio were mainly concentrated on smaller companies with listed equities, the selection could be amended as required in order to capture potentially different correlation structures among names.

7 Figure 1 shows correlations only between mid-1991 and mid-2012, since the three-year window requires one-and-a-half years of return data before and after the reference date.

8 As noted in footnote 6, other factors such as company size could be relevant (see, for example, Fama and French 1993). With regard to the prescribed regulation, it is unclear how the requirement to “reflect all significant basis risks in recognising these correlations, including, for example, maturity mismatches, internal or external ratings, vintage, etc” (Basel Committee on Banking Supervision 2016, p. 62) could be interpreted and accommodated.
FIGURE 1 The evolution of equity correlation over time, based on the index constituents of EURO STOXX 50, SMI, FTSE 100, S&P 500, ASX 200, HSI and Nikkei 225.

The equity index compositions reflect the status at end-December 2013. The return correlation is calculated from nonoverlapping monthly (log) returns and conducted on a rolling three-year window. The distribution of the pairwise equity correlations is illustrated by means of the tenth, twenty-fifth, fiftieth, seventy-fifth and ninetieth percentile. The shaded area reflects the three-year window (September 2007–September 2010) surrounding the peak in correlation (median > 40%). The bars on the bottom of the graph (related to the secondary axis) provide the width of the average 95% confidence interval from the correlation estimation.

it is in line with the rule set, which refers to two types of systematic factors. Nevertheless, with these constraints, only the country or the industry would be amenable as explanatory factors. Subsequently, and in accordance with industry practice and the established modeling technique adopted for the IRC charge, the most general form with a global factor and country and industry factors is explored, in addition to the simplified variants. For example, Ford Motor Company would be associated with the United States and/or consumer cyclicals, respectively. In this regard, Aretz and Pope (2013), for example, argue that changes in default risk depend most strongly on global and industry effects, with country effects usually being more dependent on the sample period. The last return model component is a corporate-specific, idiosyncratic factor. As for countries themselves (or, say, municipals and similar noncorporate obligors), one can express returns as a function of a global and a country-specific return factor.
The first practical step in model building and calibration consists of standardizing each of the individual time series of corporate returns \((r_{i,t})\) to a mean of zero and a standard deviation of one. At each time point, global \((r_{G,t})\), country \((r_{C,j,t})\) and industry returns \((r_{I,k,t})\) are derived from the relevant cross-section of the corporate returns.\(^9\) Let \(N_c\) and \(N_I\) denote the number of countries and industries, respectively. The resulting factor time series, then, all have a mean of zero. In order to capture the dependence of country and industry returns on global returns, the following linear regressions are run:

\[
\begin{align*}
    r_{C,j,t} &= \beta_{C,j} r_{G,t} + \varepsilon_{C,j,t}, \\
    r_{I,k,t} &= \beta_{I,k} r_{G,t} + \varepsilon_{I,k,t},
\end{align*}
\]

(3.1)

where \(\beta_{C,j}\) and \(\beta_{I,k}\) are weights given to the global factor, and \(\varepsilon_{C,j,t}\) and \(\varepsilon_{I,k,t}\) are the country- and industry-specific residuals. Let \(\sigma_G\) denote the standard deviation of \(r_{G,t}\), and let \(\sigma_{C,j}\) and \(\sigma_{I,k}\) denote that of the residuals \(\varepsilon_{C,j,t}\) and \(\varepsilon_{I,k,t}\), respectively.

The details of the regression analysis run on the basis of (3.1) for the identified stress period are provided in Section B of the online appendix. Most of the country and industry returns move in line with the global returns (\(\beta_{C,j}\) and \(\beta_{I,k}\) are not statistically different from one), and the explanatory content reflected by \(R^2\) is high (between 65% and 95%), indicating the dominant role of the global factor.

Moving on to the case of single corporates and their returns, one can postulate that these are, in the most general case, a function of the respective global, country and industry returns. Since the country and industry returns have already been expressed via the global factor in (3.1), only the residuals \(\varepsilon_{C,j}\) and \(\varepsilon_{I,k}\) are considered as additional explanatory factors for the corporate returns. It is worth noting that reusing the coefficients \(\beta_{C,j}\) and \(\beta_{I,k}\) from (3.1) as the weights for the global factor is not possible without introducing assumptions on the relationship of country and industry returns vis-à-vis corporates. Therefore, in the following, the sensitivity to the global factor itself is expressed via a separate coefficient. The sensitivities to the country-specific and industry-specific factors are \(\gamma_{C(1)}, \gamma_{C(2)}, \ldots, \gamma_{C(N_c-1)}\) and \(\gamma_{I(1)}, \gamma_{I(2)}, \ldots, \gamma_{I(N_I-1)}\), respectively:

\[
\begin{align*}
    r_{i,t} &= \gamma_G r_{G,t} + \sum_{j=1}^{N_c-1} \gamma_{C(j)} \varepsilon_{C(j),t} 1_{\{C(i)=C(j)\}} \\
    &\quad + \sum_{k=1}^{N_I-1} \gamma_{I(k)} \varepsilon_{I(k),t} 1_{\{I(i)=I(k)\}} + \varepsilon_{i,t},
\end{align*}
\]

(3.2)

\(^9\)At least five observations in a cross-sectional return set are required here for a valid systematic return.

www.risk.net/journal

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
where \( \mathbf{1} \) is the indicator function. In order to address the problem of multicollinearity in the model setup – each corporate belongs to one country and one industry – one country (here, United States) and one industry (here, utilities) is omitted from (3.2); this means that excess returns stemming from the country and industry factors are expressed relative to this base case.

The panel regression in (3.2) uses the full sample information in order to identify the systemic factor structure explaining the cross-sectional and time series covariation in corporate returns. The goodness-of-fit for (3.2) is assessed via an overall \( R^2 \) coefficient. The parameter estimates from (3.2), \( \hat{\gamma}_G, \hat{\gamma}_{C(i)}, \ldots, \hat{\gamma}_{C(N_C-1)}, \hat{\gamma}_{I(i)}, \hat{\gamma}_{I(N_I-1)}, \), are then used to calculate an aggregated systematic return for each obligor. All names in a specific country \( C(i) \) and industry \( I(i) \) have a common aggregated systematic return component, as given by \( \hat{\gamma}_G r_{G,i} + \hat{\gamma}_{C(i)} \epsilon_{C(i),i} + \hat{\gamma}_{I(i)} \epsilon_{I(i),i} \). The sensitivity of each obligor to this systematic return component as well as name-specific values of \( R^2 \) are obtained in a second estimation step by regressing the returns of each obligor against the corresponding systematic component:

\[
    r_{i,t} = \beta_i (\hat{\gamma}_G r_{G,i,t} + \hat{\gamma}_{C(i)} \epsilon_{C(i),i} + \hat{\gamma}_{I(i)} \epsilon_{I(i),i}) + \epsilon_{i,t},
\]

where \( \hat{\gamma}_G, \hat{\gamma}_{C(i)} \) and \( \hat{\gamma}_{I(i)} \) represent the parameter estimates from (3.2).

The model proposed here and described in (3.1)–(3.3) bears some similarity to the Moody’s KMV (GCorr) model, to the extent that it captures global, country and industry factors as the common drivers of asset returns (see Bohn and Stein (2009, Chapter 8) for a description of the Moody’s KMV model). However, there are some important differences between the two models. First, the correlation model in this paper is based on directly observable equity returns, while the GCorr model is based on (unobserved) asset returns estimated from equity prices and balance sheet information. As discussed above, the observability is a feature that the regulatory text emphasizes. Second, the GCorr model is a multifactor model with a number of systematic factors that exceeds the maximum according to the regulation.10

### 3.2.4 Estimation results

Table 2 summarizes the results of the multilinear panel regression analysis run on the basis of (3.2) for the stress period. Using only the global factor as an explanatory variable accounts for about 43% of the variation in equity returns. Unsurprisingly, the corresponding best-fitting return model equals the global return \( r_{i,t} \approx r_{G,i,t} \), ie, on average, the global return is the best predictor for an individual corporate return. Adding country factors increases the explanatory power to about 48%. The estimates for \( \gamma_{C(i)} \) are nearly all equal to one, ie, the best predictor for an individual corporate return.

10 For instance, there are fourteen common systematic factors in addition to industry- and country-specific factors.
return is, on average, the sum of global and country-specific excess returns. The joint use of all contemplated factors leads to an $R^2$ of about 51%; the factor structure as reflected by the set of coefficients $\gamma_{C_l(j)}$ and $\gamma_{I(k)}$ becomes a bit more diverse.

The descriptive statistics for the name-specific $R^2$, estimated on the basis of (3.3), are shown at the bottom of Table 2. The corresponding values range from 0% to about 90%, with an average that corresponds approximately to the overall $R^2$ for the four models. When studying the correlations (not shown here), it is noteworthy that there are a few names with negative values, i.e., an anticyclical equity performance.

More details on the empirical correlations that the factor model is supposed to reproduce and the resulting differences between model-implied and empirically measured correlations are provided in Section B of the online appendix. In conclusion, to a varying degree, all four model flavors for corporate names result, on average, in an adequate estimation of the correlations (with an average difference of about 5%). Adding country and/or industry factors to the global factor does not change the picture too much, at least on average, as is expected from the results in Table 2. The country/country correlation tends to be well estimated by the factor model as well (with an average difference of about 2%).

### 3.2.5 Model simulation

Section C of the online appendix outlines the way the calibrated model can be applied in a simulation context.

### 3.3 Integrating recovery rate risk

The DRC model has to reflect the “dependence of the recovery [rates] on the systemic risk factors” (Basel Committee on Banking Supervision 2016, p. 62). This renders recovery rates dependent on the economic cycle. So, during economic downturns, their simulated values will tend to be comparatively lower.

In the literature, different approaches have been considered to capture the dependence between recovery rates and the economic cycle. Altman et al (2005) estimate a linear inverse relation between historical observed recoveries and annual default rates, signaling the indirect dependence of both recovery rates and default rates on the economic cycle, which also drives defaults. Other papers take a different approach and calibrate a recovery rate distribution, assuming that the recovery rate is linked to one unobservable global risk factor and proxying for the economic cycle. A range of

---

11 Given that “United States” is treated as a base case here, the excess return of a country is expressed relative to the global return that encompasses the US-specific excess country return.
<table>
<thead>
<tr>
<th>Model</th>
<th>(1) Global factor only</th>
<th>(2) Global and country factors</th>
<th>(3) Global and industry factors</th>
<th>(4) Global, country and industry factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>36,124</td>
<td>36,124</td>
<td>36,124</td>
<td>36,124</td>
</tr>
<tr>
<td>Coefficients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global</td>
<td>1.0021</td>
<td>1.0025</td>
<td>1.0022</td>
<td>1.0031</td>
</tr>
<tr>
<td>Australia</td>
<td>1.0017</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>1.0000</td>
<td></td>
<td>0.9728</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>1.0000</td>
<td></td>
<td>0.9301</td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>1.0000</td>
<td></td>
<td>1.0442</td>
<td></td>
</tr>
<tr>
<td>Hong Kong</td>
<td>0.9998</td>
<td></td>
<td>0.8970</td>
<td></td>
</tr>
<tr>
<td>Ireland</td>
<td>0.9943</td>
<td></td>
<td>0.9537</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>1.0000</td>
<td></td>
<td>0.9091</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>0.9998</td>
<td></td>
<td>0.9696</td>
<td></td>
</tr>
<tr>
<td>Jersey</td>
<td>0.0000</td>
<td></td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Mexico</td>
<td>0.0000</td>
<td></td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Netherlands</td>
<td>1.0000</td>
<td></td>
<td>0.9038</td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td>0.2867</td>
<td></td>
<td>0.3374</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>0.0000</td>
<td></td>
<td>0.0000</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>1.0000</td>
<td></td>
<td>0.9822</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>0.9995</td>
<td></td>
<td>0.9380</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.0002</td>
<td></td>
<td>0.9546</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic materials</td>
<td></td>
<td>0.9877</td>
<td>0.8178</td>
<td></td>
</tr>
<tr>
<td>Communications</td>
<td></td>
<td>0.9901</td>
<td>0.9667</td>
<td></td>
</tr>
<tr>
<td>Consumer cyclical</td>
<td></td>
<td>0.9998</td>
<td>0.8617</td>
<td></td>
</tr>
<tr>
<td>Consumer noncyclical</td>
<td></td>
<td>0.9990</td>
<td>1.0437</td>
<td></td>
</tr>
<tr>
<td>Diversified</td>
<td></td>
<td>1.0000</td>
<td>0.5506</td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td></td>
<td>0.9996</td>
<td>1.0166</td>
<td></td>
</tr>
<tr>
<td>Financial</td>
<td></td>
<td>0.9987</td>
<td>0.9084</td>
<td></td>
</tr>
<tr>
<td>Industrial</td>
<td></td>
<td>1.0458</td>
<td>0.4605</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td>0.9987</td>
<td>0.9968</td>
<td></td>
</tr>
<tr>
<td>Utilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
recovery rate distribution functions is considered in the literature, such as the beta (see Moody’s 2005; RiskMetrics Group 2007), lognormal (see Bade et al 2011a, 2011b), logit (see Wilkens et al 2013) and Vasicek-type (see Frye 2014).

In addition to the economic cycle, Schürmann (2004) finds other factors that drive the differences between historical recovery rate distributions. These include senior-ity in the capital structure (senior versus subordinated debt), debt collateralization (secured versus unsecured debt) and industry conditions. Similar factors are confirmed by Altman and Kalotay (2014), who take into account the observed differences in recovery rate distributions by means of a mixture-distribution model for recovery rates.

Here, a lognormal recovery rate model, similar to that proposed in Bade et al (2011a, 2011b) and based on previous work by Pykhtin (2003), is adopted. This model captures the correlation between the default process and recovery rate given default via a common systematic factor. It allows for the simultaneous estimation of the default and recovery rate model parameters by taking into account the fact that realized recoveries are observed only when defaults occur. In a separate paper, Bade et al (2011b) conclude that their model has better predictive power for future recoveries than alternative models, in which (log) recovery rates are estimated separately from the default model as linear functions of observable variables.
Adopting the setup in Bade et al (2011a), the following linear factor model is assumed for the log recovery rate of obligor $i$: \[ Y_i = y' x_i + \sigma' \varepsilon_i (\sqrt{\rho^Y} Z_G + \sqrt{1 - \rho^Y} \varepsilon_i^Y). \] (3.4)

The drivers of the log recovery rate are the systematic return ($Z_G$) and an idiosyncratic recovery factor ($\varepsilon_i^Y$), both of which are assumed to follow independent standard normal distributions. The systematic factor is the same global return as in the asset return model for sovereigns and corporates (see Section C of the online appendix). The parameter $\rho^Y$ controls the extent to which the recovery rate is influenced by the global asset return. The deterministic vectors

\[
\gamma = (\gamma_1^{\text{Corp}}, \gamma_2^{\text{Corp}}, \ldots, \gamma_L^{\text{Corp}}, \gamma_1^{\text{Sov}}, \gamma_2^{\text{Sov}}, \ldots, \gamma_L^{\text{Sov}})
\]

and

\[
\sigma = (\sigma_1^{\text{Corp}}, \sigma_2^{\text{Corp}}, \ldots, \sigma_L^{\text{Corp}}, \sigma_1^{\text{Sov}}, \sigma_2^{\text{Sov}}, \ldots, \sigma_L^{\text{Sov}})
\]

reflect specific calibration parameters for corporates and sovereigns across the spectrum of ratings. $x_i = (\ldots, 0, 1, 0, \ldots)$ indicates the type (corporate or sovereign) and rating of obligor $i$. The recovery rate itself is given by $\text{RR}_i = \exp(Y_i)$.

Here, two main debt categories are considered for the estimation of the recovery rate model parameters: corporate senior and sovereign. This is in line with the categories for the default probabilities. More granular recovery rate model calibrations (e.g., by industry, debt seniority, etc) can be carried out if necessary, given sufficient data availability. If separate parameters were estimated for different debt seniorities, one might need to additionally enhance the recovery rate model to enforce the absolute priority rule (APR), which states that the junior creditors would only recover something if the more-senior creditors fully recovered their claims. In other words, the recovery rate for junior debt should be greater than zero only if the senior debt has a 100% recovery rate. \(^{13}\)

As for the calibration, Table 3 summarizes the main parameters. In Altman and Kalotay (2014), the marginal distribution of the recovery rates conditional on default is derived from a large set of defaulted bonds over the period 1988–2011. The data provides the empirical mean ($\mu^{\text{RR}}$ in Table 3) and standard deviation ($\sigma^{\text{RR}}$ in Table 3) that the model should reflect. The objective is to fit the empirical recovery rate distribution.

---

\(^{12}\) The setup in (3.4) differs from that in Bade et al (2011a) in that, here, no correlation between the idiosyncratic asset return and the idiosyncratic recovery rate components is assumed.

\(^{13}\) Notably, there is an established empirical literature, starting with Franks and Torous (1989), documenting violations of the strict APR. This is due to the fact that junior creditors have the ability to delay bankruptcy resolutions, while senior creditors may be willing to accept less and thus reduce additional costs from lengthy resolutions.
Importantly, the fitting needs to be carried out conditional on default, i.e., the empirical distribution needs to be matched given $V_i < 0$. When we differentiate by corporates and sovereigns as well as rating-implied default probabilities (as given in Table 1), separate tuples are estimated in an iterative numerical procedure designed to fit the empirical conditional recovery rate distribution. Specifically, the objective function that is minimized in order to obtain the parameters for the corporate recovery rate model reads as follows:

$$
\min_{\hat{\mu}_{RR}^C, \hat{\sigma}_{RR}^C} \mathbb{E}(RR \mid V < 0) - \hat{\mu}_{RR}^C)^2 + (\mathbb{E}(RR \mid V < 0) - \hat{\sigma}_{RR}^C)^2
$$

such that

$$
\hat{P}(RR > 1 \mid V < 0) \leq \phi.
$$

(3.5)

where $\hat{\mu}_{RR}$ and $\hat{\sigma}_{RR}$ are the empirical estimates for $\mu_{RR}$ and $\sigma_{RR}$ (shown in part (a) of Table 3). The probability mass attributed to recovery rates larger than one is limited to $1\%$ ($\phi = 0.01$) in order to avoid an ill-fitted model. The estimates $\hat{E}(RR \mid V < 0)$, $\hat{\sigma}(RR \mid V < 0)$ and $\hat{P}(RR > 1 \mid V < 0)$ are determined from the joint simulation of the default process (see Section C of the online appendix) and the recovery rate process (as per (3.4)). The sovereign parameters are estimated similarly to the corporate case using a fitting procedure, as in (3.5). The differences in the calibration procedure are the estimates $\hat{E}(RR \mid V < 0)$, $\hat{\sigma}(RR \mid V < 0)$ and $\hat{P}(RR > 1 \mid V < 0)$, which depend on the sovereign return correlations ($\rho_{V,Sov}$).

Note that our calibration of the marginal distribution of recovery rates conditional on default differs from the approach in Bade et al (2011a); there, all default and recovery rate process parameters are determined jointly via a maximum likelihood estimation. Due to the joint fitting of all parameters, the moments of the marginal distribution of the recovery rate conditional on default in Bade et al (2011a) may differ from the empirical estimates to a greater extent than those in this paper, as our approach is designed to fit the two moments directly. In addition, the fitting approach used here can be easily applied in practice to estimate recovery rate parameters for more granular debt categories.

Given limited data availability, and to avoid calibration instability, one parameter from the comprehensive study in Bade et al (2011a) is reused: the correlation between log recovery rates, which implies that $\rho^Y$ in (3.4) is set equal to $\rho^{ln(RR)}$. Bade et al (2011a) use a data set of approximately 188,000 annual observations for nonfinancial
TABLE 3 Recovery rate model: calibration. [Table continued on next page.]

(a) External parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (%)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_{V,\text{Corp}}^Y$</td>
<td>42.82</td>
<td>Equity correlation model, Section 3.2</td>
</tr>
<tr>
<td>$\rho_{V,\text{Sov}}^Y$</td>
<td>81.88</td>
<td>Equity correlation model, Section 3.2</td>
</tr>
<tr>
<td>$\mu^\text{RR}$</td>
<td>44.90</td>
<td>Altman and Kalotay (2014, Table 1) [used in fitting]</td>
</tr>
<tr>
<td>$\sigma^\text{RR}$</td>
<td>37.90</td>
<td>Altman and Kalotay (2014, Table 1) [used in fitting]</td>
</tr>
<tr>
<td>$\rho_{\ln(\text{RR})}$</td>
<td>4.11</td>
<td>Bade et al (2011a, Table 5)</td>
</tr>
</tbody>
</table>

(b) Resulting parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho^Y$</td>
<td>4.11</td>
</tr>
<tr>
<td>$\rho_{V,Y,\text{Corp}}$</td>
<td>13.27</td>
</tr>
<tr>
<td>$\rho_{V,Y,\text{Sov}}$</td>
<td>18.35</td>
</tr>
</tbody>
</table>

bonds, spanning the period 1982–2009, with a default rate of about 1%. Their derived value for $\rho_{\ln(\text{RR})}$ of about 4% suggests a low correlation between log recovery rates, implying a near-zero correlation between individual recoveries conditional on default. In other words, their estimates imply that, on average, recovery rates conditional on default are driven mostly by idiosyncratic factors. The correlation parameter ($\rho^Y$) in this paper, estimated from the corporate data in Bade et al (2011a), is assumed to have the same value for both corporates and sovereigns. A separate estimation of this parameter for sovereigns would likely be difficult or result in significant measurement uncertainty, given the very small number of historical sovereign default events. Note that the correlation between the asset return and recovery rate model is a function of the asset return correlation; therefore, the correlation parameters $\rho_{V,\text{Corp}}$ and $\rho_{V,\text{Sov}}$ from the equity correlation model in Section 3.2 influence the model behavior. The properties of the asset-recovery model are explored in further detail in Section D of the online appendix.

Equity positions are assumed to have a zero recovery rate. Defaulted debt positions need also be captured in the DRC measure (Basel Committee on Banking Supervision 2016, p. 61). The recovery rate model can be used to model changes in values of defaulted debt. Specifically, one can assume that the recovery rate distribution is equal to the fitted conditional one, and then sample from it to generate recovery rate scenarios.
TABLE 3  Continued.

(c) Calibrated parameters

<table>
<thead>
<tr>
<th>Rating</th>
<th>$\gamma_{\text{Corp}}$</th>
<th>$\sigma_{\text{Corp}}$</th>
<th>$\gamma_{\text{Sov}}$</th>
<th>$\sigma_{\text{Sov}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA</td>
<td>-0.7131</td>
<td>0.4301</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>AA+</td>
<td>-0.7131</td>
<td>0.4301</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>AA</td>
<td>-0.7131</td>
<td>0.4301</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>AA−</td>
<td>-0.7131</td>
<td>0.4301</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>A+</td>
<td>-0.7461</td>
<td>0.4824</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>A</td>
<td>-0.7476</td>
<td>0.4834</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>A−</td>
<td>-0.7476</td>
<td>0.4834</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>BBB+</td>
<td>-0.7517</td>
<td>0.4534</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>BBB</td>
<td>-0.7615</td>
<td>0.4361</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>BBB−</td>
<td>-0.7764</td>
<td>0.4108</td>
<td>-0.6275</td>
<td>0.4192</td>
</tr>
<tr>
<td>BB+</td>
<td>-0.7876</td>
<td>0.4117</td>
<td>-0.6819</td>
<td>0.4555</td>
</tr>
<tr>
<td>BB</td>
<td>-0.7964</td>
<td>0.4081</td>
<td>-0.7080</td>
<td>0.4271</td>
</tr>
<tr>
<td>BB−</td>
<td>-0.8114</td>
<td>0.4129</td>
<td>-0.7634</td>
<td>0.4173</td>
</tr>
<tr>
<td>B+</td>
<td>-0.8324</td>
<td>0.4182</td>
<td>-0.7690</td>
<td>0.4151</td>
</tr>
<tr>
<td>B</td>
<td>-0.8474</td>
<td>0.4158</td>
<td>-0.7778</td>
<td>0.4160</td>
</tr>
<tr>
<td>B−</td>
<td>-0.8577</td>
<td>0.4165</td>
<td>-0.8099</td>
<td>0.4229</td>
</tr>
<tr>
<td>CCC/C</td>
<td>-0.8927</td>
<td>0.4165</td>
<td>-0.8769</td>
<td>0.4188</td>
</tr>
</tbody>
</table>

The parameterization of the recovery rate model according to (3.4). The asset correlation parameters $\rho^V_{\text{Corp}}$ and $\rho^V_{\text{Sov}}$ stem from the equity correlation model in Section 3.2. In particular, they are set to the $R^2$ for corporates (Table 2, with model (1) as a proxy) and the average $R^2$ across countries for sovereigns (Table B.1 in the online appendix). The calibration aims to match the empirically observed mean ($\mu_{\text{RR}}$) and standard deviation ($\sigma_{\text{RR}}$) of recovery rates of a set of 2828 corporate bonds over the period 1988–2011, according to data in Altman and Kalotay (2014). These moments of the recovery rate distribution need to be matched conditional on default. With probabilities of default varying across the rating spectrum, separate tuples $(\gamma, \sigma)$ are calibrated, each in an iterative procedure. In order to avoid an ill-fitting, with significant probability mass attributed to recovery rates greater than one, $P(\text{RR} > 1) \leq 0.01$ is added as a constraint to the fitting procedure. Given limited data availability, and to avoid calibration instability, one parameter from the comprehensive study in Bade et al (2011a) is reused, namely the correlation between log recovery rates ($\rho^{\text{lnRR}}$). The log recovery rate correlation implies the weight of the systematic asset return in the recovery rate model: $\rho^V = \rho^{\text{lnRR}}$. The correlation between the asset return and recovery model $(\rho^V)$ is given by $\rho^V \gamma = \sqrt{\rho^V \rho^T}$ and, hence, is different for corporates and sovereigns, given the different asset (equity) correlations.

4 P&L GENERATION AND DISTRIBUTIONS

4.1 Overview

In order to generate P&L distributions and derive associated tail measures, joint realizations of the risk factors are applied as instantaneous shocks to the deals in the DRC coverage. The risk factors in scope are the obligors’ defaults and the associated recovery rates.

www.risk.net/journal  
Journal of Risk

To subscribe to a Risk Journal visit subscriptions.risk.net/journals or email info@risk.net
In practice, for large portfolios, the computational requirements for the DRC calculation can pose a concern. In order to ease the computational burden, one could potentially pre-generate the P&L per obligor and only read the corresponding values in each Monte Carlo scenario. Given that the DRC captures losses from defaults, only the P&L needs to be generated, and only for the default scenarios, since it is zero in all nondefault scenarios. For the default scenarios, a P&L grid can be pre-generated using a discretization of the recovery rate: a grid between 0% and 100% with steps of, say, 5%. Notably, this technique is not applicable in case the P&L is not separable by obligor, for example, for CDS index option positions, or certain multi-underlying equity derivatives, for which the P&L cannot be decomposed into that of the constituents. Another challenge arises for path-dependent derivatives (eg, variance swaps): without simulating an actual path to a potential default, the associated P&L is ill-defined.

In spite of the one-year projection horizon with constant positions, it may be necessary to simulate (joint) defaults at shorter horizons as well, for example, to reflect potential mismatches between the maturity of a position and its hedge. Section E of the online appendix discusses this aspect further.

4.2 Example portfolios

The properties of the DRC model are explored using example portfolios. For these portfolios, the DRC model results are also compared with those based on the SA for default risk. Note that the DRC is a capital measure and, as such, a (nonnegative) loss figure. For ease of illustration and discussion, the tables and figures in this paper show the actual (signed) P&L corresponding to the 0.1th (for DRC) or other percentiles. Table 4 illustrates the P&L distributions by means of selected percentiles (including the 0.1th for DRC) for a set of bond and equity portfolios. The results are based on the calibrations according to Section 3 and simulations with one million scenarios each. For a straightforward comparison, the liquidity horizon is assumed to be one year.

15 For multi-underlying equity derivatives only, subject to supervisory approval, the rule set allows “simplified modelling approaches (for example […] that rely solely on individual jump-to-default sensitivities to estimate losses)” (Basel Committee on Banking Supervision 2016, p. 62).
16 One possible approach, although it contradicts the paradigm of instantaneous shocks, consists of simulating an actual random default time within the one-year period and complementing this with a Brownian bridge to define a path to default.
17 The SA is expected to serve as “a credible fallback for, as well as a floor to, the IMA [internal model approach]” (Basel Committee on Banking Supervision 2016, p. 1). The SA for the default risk charge is described in Basel Committee on Banking Supervision (2016, pp. 43–46).
18 Recovery rates are capped at 100% in the simulation. Given the fitting procedure (see Section 3.3), this capping would affect not more than 1% of the default scenarios.
year for all portfolios. Relative estimation errors at a 95% confidence level for the P&L percentiles are shown in parentheses.

For the long investment grade bond portfolio (A), the DRC is around €663,000, representing about 7% of the total absolute notional (€10 million). If we change the setup to a long/short portfolio (B), the DRC reduces to about €196,000, or 2% of the total absolute notional. The equivalent high-yield portfolios (C and D) yield larger DRC figures of around €3.1 million and €750,000, respectively, reflecting higher default risk in these portfolios compared with the investment grade cases. Applying the model to long (portfolio E) and long/short positions (portfolio F) in sovereign bonds results in DRC figures of around €899,000 and €466,000, respectively. The DRCs for the selected sovereign portfolios lie between those for investment grade portfolios and high-yield corporate portfolios. This is to be expected, given that the sovereign portfolio has a mix of investment grade and subinvestment grade (high-yield) issuers, with examples such as Ukraine (CCC) and Vietnam (BB+) among the high-yield ones. In order to reflect a balanced long/short portfolio across investment grade and high-yield corporate as well as sovereign bonds, portfolio G represents the aggregation of portfolios B, D and F. One can observe that the DRC for the aggregate portfolio G is around €1 million, reflecting a diversification benefit of about 25% when compared with the sum of the DRC figures for portfolios B, D and F. It is worth noting that pure short portfolios in bonds and/or equities are not considered, since they only allow for nonnegative P&Ls and thus render the DRC figures equal to zero. For a long equities portfolio (H), the DRC measure amounts to €1 million (10% of the total absolute notional). This corresponds to a scenario in which five names default, given that the recovery rate is 0% for equities and that each name in portfolio H has an equal weight. The corresponding long/short equities portfolio (I) has a DRC of €400,000. Finally, reflecting a reasonably well-diversified long/short portfolio of bonds and equities with obligors of different credit quality, portfolio J represents the aggregation of portfolios B, D, F and I. The DRC for the aggregate portfolio J is around €1.1 million, reflecting a diversification benefit of around 40% when compared with the sum of DRC figures for portfolios B, D, F and I.

A comparison between the DRC results based on the model in this paper and the corresponding SA yields interesting results. The SA-based DRC is smaller than the model-based one for all example portfolios. For portfolio J, the SA-based DRC amounts to about 80% of the model-based one. For some portfolios, the ratio between the SA-based and the model-based DRC can be as low as 20% (portfolio I). These differences suggest that the SA-based DRC corresponds to a less extreme scenario

---

19 If one were to use a liquidity horizon of sixty days, the DRC figures for portfolios H and J would change to €400,000 and €200,000, a reduction of 60% and 50% compared with the DRC for a one-year horizon.
TABLE 4  P&L distributions and DRC figures for selected example portfolios.

<table>
<thead>
<tr>
<th>Percentile of P&amp;L distribution</th>
<th>10%</th>
<th>1%</th>
<th>0.1%</th>
<th>SA</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Investment-grade bonds:</td>
<td>0</td>
<td>-184,696</td>
<td>-662,808</td>
<td>-347,400</td>
</tr>
<tr>
<td>long position</td>
<td>(0.0%)</td>
<td>(1.6%)</td>
<td>(3.3%)</td>
<td></td>
</tr>
<tr>
<td>B. Investment-grade bonds:</td>
<td>0</td>
<td>-66,048</td>
<td>-196,464</td>
<td>-91,584</td>
</tr>
<tr>
<td>long/short position</td>
<td>(0.0%)</td>
<td>(1.0%)</td>
<td>(3.1%)</td>
<td></td>
</tr>
<tr>
<td>C. High-yield bonds:</td>
<td>-629,037</td>
<td>-1,768,448</td>
<td>-3,116,519</td>
<td>-1,669,500</td>
</tr>
<tr>
<td>long position</td>
<td>(0.4%)</td>
<td>(0.6%)</td>
<td>(1.3%)</td>
<td></td>
</tr>
<tr>
<td>D. High-yield bonds:</td>
<td>-137,433</td>
<td>-421,220</td>
<td>-749,059</td>
<td>-452,250</td>
</tr>
<tr>
<td>long/short position</td>
<td>(0.3%)</td>
<td>(0.6%)</td>
<td>(1.2%)</td>
<td></td>
</tr>
<tr>
<td>E. Sovereign bonds:</td>
<td>-208,180</td>
<td>-386,529</td>
<td>-899,212</td>
<td>-465,441</td>
</tr>
<tr>
<td>long position</td>
<td>(0.1%)</td>
<td>(0.5%)</td>
<td>(3.6%)</td>
<td></td>
</tr>
<tr>
<td>F. Sovereign bonds:</td>
<td>-207,003</td>
<td>-351,149</td>
<td>-466,263</td>
<td>-248,713</td>
</tr>
<tr>
<td>long/short position</td>
<td>(0.1%)</td>
<td>(0.6%)</td>
<td>(0.8%)</td>
<td></td>
</tr>
<tr>
<td>G. Portfolios B, D and F together</td>
<td>-264,601</td>
<td>-646,645</td>
<td>-1,064,022</td>
<td>-793,828</td>
</tr>
<tr>
<td>H. Equity index constituents:</td>
<td>0</td>
<td>-200,000</td>
<td>-1,000,000</td>
<td>-396,000</td>
</tr>
<tr>
<td>long position</td>
<td>(0.0%)</td>
<td>(0.0%)</td>
<td>(0.0%)</td>
<td></td>
</tr>
<tr>
<td>I. Equity index constituents:</td>
<td>0</td>
<td>-200,000</td>
<td>-400,000</td>
<td>-84,000</td>
</tr>
<tr>
<td>long/short position</td>
<td>(0.0%)</td>
<td>(0.0%)</td>
<td>(0.0%)</td>
<td></td>
</tr>
<tr>
<td>J. Portfolios B, D, F and I together</td>
<td>-263,129</td>
<td>-651,687</td>
<td>-1,105,615</td>
<td>-877,292</td>
</tr>
</tbody>
</table>

The DRC figures for a set of example portfolios, including the DRC calculated using the model developed in this paper and the DRC based on the SA. While the DRC is a capital measure and, as such, a (nonnegative) loss figure, this table shows the actual (signed) P&L for the model and the SA. Some additional P&L percentiles are shown for comparison. A one-year liquidity horizon is assumed throughout. Portfolio A consists of equally weighted investment-grade corporate bonds with issuers from the iTraxx (125 names, series 20) and assumes an overall notional of €10 million. Portfolio B represents a long/short corporate bond portfolio, where about half of the bond positions (sixty-three) are long, while the others (sixty-two) are short. Portfolios C and D represent the equivalent setup, but with high-yield bonds and issuers from the iTraxx Crossover (100 names, series 20). The long and long/short portfolios in sovereign bonds are reflected in portfolios E and F; the issuers are from the SovX Western Europe (series 8), SovX CEEMEA ex-EU (series 8) and SovX Asia Pacific (series 8), for a total of thirty-four names. Portfolio G is a combination of portfolios B, D and F. A long position in constituents of the EURO STOXX 50 is reflected in portfolio H, and portfolio I represents a long/short equity portfolio, where half of the positions are long and the others are short (twenty-five each). Finally, portfolio J aggregates portfolios B, D, F and I into a mixture of long/short positions in bond and equity positions across a wide spectrum of issuers. Note that short positions in either bonds or equities are not considered, since they only allow nonnegative P&Ls and thus render the DRC figures equal to €0. All positions are assumed to be equally weighted with respect to the notional; bonds are considered as trading at par, which implicitly puts a slightly higher weight on lesser-rated bonds whose P&Ls ceteris paribus would be lower than those of better-rated ones. Probabilities of default use the data from Table 1 and reflect the S&P rating of the issuers as of December 2014; Moody’s ratings serve as primary fallback, and the index average ratings serve as secondary fallback. Initial recovery rates (subject to simulation) reflect the estimate and $P_{40\%}^{\text{RR}}$ from Table 3; equity positions are assumed to have a 0% recovery rate. With not all issuers’ countries being part of the correlation model calibration, fallbacks based on geographical location and type of country (developed versus emerging) are used. Bloomberg is the data source for index members and ratings. The P&L figures (in euros) are derived by means of one million simulations each. The estimation error in parentheses is calculated as $\max(P_{97.5\%} - \hat{P}, 0) = P_{97.5\%}/\hat{P}$, with $\hat{P}$ as the percentile estimate and $P_{40\%}^{\text{RR}}$, as the upper and lower bounds representing a 95% confidence interval. The last column shows the P&L calculated using the SA.
Default risk charge

than the 0.1th percentile from the P&L distribution generated by the model. Indeed, the SA-based DRC reflects the 0.3th and third percentile of the P&L distribution for portfolios J and I, respectively. Section F of the online appendix sheds further light on the results for the SA.

Focusing on the model-based DRC, established tools for VaR-type risk measures such as marginal contributions can be applied as well. For portfolio J, for example, one finds Ukraine is the biggest contributor, with a marginal DRC of about €187 000. Another way of investigating the figures consists of analyzing the simulated scenarios close to the DRC one. In the case of portfolio J, one finds that the average P&L across the ten next-worst and ten next-best scenarios provides a very similar picture to that from the marginal DRC, again with Ukraine as the biggest single contributor. Laurent et al (2016) suggest a factor decomposition of the DRC. Applying this technique to portfolio J (in a simplified setup, with a one-factor model and constant recovery rates) yields a decomposition of the DRC into approximately 5% coming from the expected loss and 25% stemming from the systematic factor. Note that, in this second approach, a single scenario (corresponding to the 0.1th P&L percentile) is used to determine the marginal contributions to the DRC.

If we analyze the relationship between the different tail measures for the test portfolios, one can observe that the P&L distributions generally have fatter tails than a standard normal distribution. For example, for portfolios A and B, the ratio between the 0.1th and first percentile is around 3.6 and 3.0, respectively, compared with 1.33 for the standard normal distribution. Similar characteristics can be observed across the P&L distributions for all portfolios, with the exception of the sovereign bond portfolios E and F, where the tail is slightly thinner than the corresponding one for a standard normal distribution.

4.3 Model properties

4.3.1 Convergence

The convergence of a simulation model is an important property to investigate. It allows the user to gauge an acceptable relationship between computational burden and accuracy of the model estimate. For typical large-scale bank portfolios, with the DRC stretching over many business lines, the overall P&L distribution can be asymmetric, fat-tailed and nonsmooth, thus requiring a high number of simulations. Using the example of portfolio J from Section 4.2, Figure 2 shows the simulated P&L distribution (part (a)) and the DRC estimate as a function of the number of simulations, in conjunction with the corresponding 95% confidence interval (part (b)). The pattern reveals that the confidence interval around the DRC decreases with the number of simulations, as expected. Using 10 000 simulations results in an uncertainty, measured
as the relative width of the confidence interval, of around 27%. This decreases to around 2% when using one million simulations.

In order to improve convergence, standard variance reduction techniques such as importance sampling (see Glasserman and Li (2005), among many others) can be
used in the Monte Carlo setup. Alternatively, or in conjunction with this, extreme
value theory (EVT) is a technique that could be used to address the problem of
usually nonsmooth P&L distribution tails, which tend to coincide with large estimation
uncertainty and missing robustness of extreme quantiles (see McNeil et al (2005,
pp. 264–326) for an overview of EVT). In the concrete case of DRC, as evidenced by
the examples in Table 4 and Figure 2, an acceptable accuracy can already be achieved
with pure Monte Carlo simulation and a limited number of scenarios.

4.3.2 Parameter sensitivity

Beside the inherent model risk, the parameterization itself poses an important chal-
lenge. Section G of the online appendix provides a series of sensitivity studies for the
example portfolios and discusses the resulting effects. The analyses therein illustrate
a reasonable degree of parameterization risk for the suggested DRC model.

5 CONCLUSION AND OUTLOOK

The revised standards for capital requirements for market risk in the trading book
issued as a result of the FRTB aim to consolidate several building blocks from the
existing rules for market risk capital. Intensive discussions between the industry and
regulators as well as work-intense QISs over the past few years have helped to achieve
more coherent risk and capital measures. Although this comprehensive regulation
will only become binding in 2019 at the earliest (as the discussion now stands), the
banking industry needs to prepare for this fundamental step, not least due to the current
background of ongoing QIS exercises and potential revisions to the rules. This paper
is the first to present a comprehensive model framework for DRC that is compliant
with the revised regulatory framework. The DRC is supposed to complement short-
term (continuous) risk measures such as VaR and ES by incorporating event risk in
the form of defaults. It requires model components for marginal default and recovery
rate risk as well as a factor correlation model to link them together.

As a general conclusion, an extreme tail measure and a long projection horizon,
in conjunction with very limited backtesting feasibility, leave a substantial model
and parameterization risk, as in the case of the Basel 2.5 risk (capital) measures.
Following industry trends (see, for example, the approach of Glasserman and Xu
2014), analyzing the model risk further might be a point for future research. The
same applies to aspects such as variance reduction techniques and EVT. From a
practitioner’s point of view, the DRC, or a DRC-like quantity, might actually serve
as a useful risk measure instead of reflecting only a capital charge. In light of the
imposed model assumptions and restrictions, however, it is not obvious that this goal
can be achieved.
DECLARATION OF INTEREST

The views expressed in this paper are those of the authors and do not necessarily reflect the views and policies of BNP Paribas.

ACKNOWLEDGEMENTS

We are very thankful to Hamid Skoutti, who provided valuable ideas for the model design. We also appreciate helpful input and suggestions from Andrei Greenberg, Jean-Paul Laurent, Lee Moran and Bruno Thiery as well as an anonymous referee.

REFERENCES


European Banking Authority (2012). EBA guidelines on the incremental default and migration risk charge (IRC). Report, May, EBA.


Moody’s (2007). Moody’s ultimate recovery database. Special Comment, April, Moody’s Investors Service, Inc.


Research Paper

A review of the fundamentals of the Fundamental Review of the Trading Book: standard foreign exchange rules are highly asymmetric with respect to reporting currencies

Hany M. Farag

PO Box 1, Toronto, ON M5L 1A2, Canada; email: hany.farag@cibc.com

(Received May 9, 2016; revised September 14, 2016; accepted September 14, 2016)

ABSTRACT

We examine the mathematical formalism underlying the Basel Committee on Banking Supervision’s “Revised standardized approach of the Fundamental Review of the Trading Book”. One of its goals is to provide a simple, uniform methodology for market risk that applies to all banks independently of their internal models. The formalism exhibits subtle nonlinearities that can have deep implications and deserve careful analysis. We demonstrate that the capital charge for the linear part of the foreign exchange book can differ substantially, depending on the reporting currency. Indeed, we prove that the linear charge is the same for any reporting currency if and only if the correlation between risk-weighted positions in the currencies is 50%, and the intra-bucket correlation is 100%. We show that the capital charge for the nonlinear part of the book fails to be invariant regardless of the correlation values. We demonstrate (in the most recent version of the rules) up to 58% and 300% anomalies for the linear and nonlinear charges, respectively, depending on the book. We also show the rather surprising result that, for a given reporting currency, the capital charges are only functionally dependent on the original underlying rates of the portfolio and
are not necessarily dependent on the rates referencing the reporting currency. The anomaly is therefore due to the noninvariance of the aggregation mechanism under a change of reporting currency. We solve the problem by offering a proposal that carefully restores the invariance, using the same regulatory framework and without changing the underlying conservatism. Our analysis also has the practical application of facilitating the transformations of the computations in any reporting currency, possibly circumventing the need for system changes in some cases.

**Keywords:** Fundamental Review of the Trading Book (FRTB); standardized approach (SA); FX Delta charge; FX curvature charge; reporting currency.

1 INTRODUCTION

The global financial crisis of 2007–8 has had a major impact on the financial markets, the banking industry and society in general. In response to the crisis, the Basel Committee on Banking Supervision (BCBS) introduced some expedient revisions to the market risk framework in order to restore stability to the banks and to reduce the chances of another crisis in the foreseeable future. This was part of “Basel 2.5” (Basel Committee on Banking Supervision 2009). At the same time, however, the BCBS realized that these reforms were merely a temporary solution until a more comprehensive framework could be put in place. Accordingly, the BCBS began a “fundamental review” of the trading book regime. The first consultative paper on that review was published in May 2012 (Basel Committee on Banking Supervision 2012). A second consultative paper was published in October 2013 (Basel Committee on Banking Supervision 2013) and a final one was published in December 2014 (Basel Committee on Banking Supervision 2014). Additionally, several “Basel III monitoring” exercises and quantitative impact studies (QISs) have taken place. Recently, the final rules (Basel Committee on Banking Supervision 2016) have been published.

The implementation of this new market risk regime is expected to be fully in place by late 2019 or early 2020.

1.1 Two approaches

1.1.1 The internal models approach

The market risk paradigm of the Fundamental Review of the Trading Book (FRTB), which is our focus in this paper, has two versions. The first is the internal models approach (IMA), which essentially uses the bank’s internal models to assess market risk. The most popular flavor of that approach is full revaluation of derivatives to estimate the full nonlinear impact of shocks of the relevant risk factors, and it is typically combined with historical simulations on data from a stress period. The latter
A review of the fundamentals of the Fundamental Review of the Trading Book

can be from a moving window of history but mainly forces the models to estimate the impact during times of market stress by choosing the worst such period for the bank’s portfolio. Under the FRTB, the simulations must reflect so-called liquidity horizons, which are a reflection of how long it takes certain assets to be liquidated in times of stress. While it is generally sensible, the IMA suffers from the fact that each bank has a private view of risk, and many risks may or may not be reflected in the models. In turn, these models might unjustifiably reduce the capital charges intended to protect the institution and its stakeholders in times of severe market stress. In other words, they are still somewhat subjective even if they appear quantitative. Additionally, as part of the FRTB, some trading desks might, going forward, fail to meet certain requirements that allow them to use such internal models. In such cases, the regulators would like to enforce a standardized model of some form as a fallback. This leads us to the revised standardized approach. From this point forward, we will refer to this as simply the SA, now that the rules have been finalized.

1.1.2 The SA

As we already mentioned, the standardized approach is needed, at least, to be a fallback in case of the failure of the IMA for some trading desks. Similarly, small banks with less complex products are expected to be able to use the SA to manage their market risk entirely without the IMA. Additionally, and just as importantly, the SA is needed to provide some uniformity between banks. For example, the approach will be required to be reported and disclosed, even if it is not directly used for capital charges. Furthermore, it is being considered as a possible floor for capital charges as well. Whichever its final usage, it is likely to be material. Given that it is quite prescriptive, as we will discuss in more detail for our problem, it is not flexible. In many ways, the SA is essentially a set of mathematical rules that are moderately complex but essentially reflect some mathematical model that is not fully disclosed. For now, we concern ourselves only with the capitalization of the foreign exchange (FX) trading book in this approach.

1.2 FX capital charges in the SA

The rules for capitalization of FX risk in general, and an FX trading book in particular, require that the FX risk factors be defined as the exchange rates between foreign currencies and the reporting currency of the bank. In principle this sounds quite reasonable (and it is). After all, this is a symmetric rule applying equally to all banks. Furthermore, we will see that the prescribed correlation for aggregating risk-weighted positions takes a uniform value, regardless of the currencies involved or the reporting currency with respect to which they are measured. The value of the regulatory correlation at time of writing is 60% for the linear charge, and the square
of that value (0.36) for the nonlinear charge. In fact, the risk weights themselves were originally equal for all currencies, with a value of 15%. At time of writing they are 30%, with some permitted exceptions for certain pairs (this has further exacerbated the anomaly). Aside from the exceptions on risk weights, we can argue that the rules are symmetric with respect to the reporting currencies. Furthermore, the aggregation formulas prescribed for the capital charges, namely the Delta (or linear) charge or the curvature (or nonlinear part, in excess of linear) charge are the same for all reporting currencies. As yet there is no explicit bias, and we certainly would not suggest that there is any intended bias.

The problem, however, is that the FX book itself may not be symmetric under permutation of the currencies, as we will see. Why should it? After all, it is well known that trading volumes in EUR/USD are the highest among all pairs, followed by USD/JPY, and then the volumes decay relatively quickly from those high values. Now things get interesting, as we start to think that the same FX book should be charged the same capital under any reporting currency (after converting the capital charge to a unified currency of course). More generally, it is reasonable to expect that we should be charged an amount of capital that accounts for the FX risk of a portfolio and nothing else. Any asymmetry would be an unpleasant anomaly in the formalism. After all, we imagine the large banks are competing globally on the same kind of business. We would expect that their portfolios are somewhat similar to each other. However, once the capital charges favor one reporting currency over another, we introduce an unnecessary inefficiency in the market place. This issue will be the focus of the rest of the paper.

In Section 2, we explain the FX capitalization rules and give some analogies from physics to motivate our approach to the analysis and possible solutions. In Section 3, we completely characterize the invariance of the linear (also known as Delta) capital charge as a function of inter-currency correlation and the intra-bucket correlation. We show that for the current regulatory values of the inter-currency correlation, a nonexotic FX book can be charged up to 58% more in capital depending on the reporting currency. We also demonstrate that the Delta charge is in fact a function only of the underlying rates of the portfolio, and that the anomaly is due to the noninvariance of the aggregation coefficients under change of reporting currency. In Section 4, we demonstrate the nonlinear (also known as curvature) capital charge anomaly, explain the impossibility of a direct fix and show that it can create up to 300% discrepancy in the charge in different reporting currencies for some portfolios. Note that when we compare charges we always measure them in the same units (so these anomalies have nothing to do with currency conversion). With such strong anomalies, it is reasonably clear to the author that a US-dollar-reporting currency is likely to be the most beneficial for a typical FX book. We show that this capital charge is directly dependent only on the underlying rates, and also that the anomaly is due to
the noninvariance of the aggregation formula. In Sections 5 and 6, we discuss an entire family of solutions, including a simple, efficient and essentially optimal solution to the problem. Our proposed solutions do not require a complete rewrite of the rules or a reduction in flexibility or conservatism, yet they minimize the anomaly and restore a level playing field. In the process of our proofs, we produce formulas that should be helpful to the practitioner in computing sensitivities with respect to one currency in terms of those of another (often needed if the existing system will only recognize the explicit/underlying exchange rates given to it). In particular, our solution(s) give a capital charge calculation using the regulatory formula, but with reference only to the explicitly given exchange rates in the portfolio (also called the underlying rates). In other words, we no longer have to reference the reporting currency in the formula, except for a harmless conversion factor at the end.

Some lack of invariance in the charges depending on reporting currency was pointed out by other practitioners before the author, and long before the final release of the rules. There was a feeling that (for example) when we shock an option that depends on USD/JPY, by instead shocking GBP/USD and GBP/JPY (with the same percentage shock) for a GBP reporting bank, we must be “double counting” the shocks. Numerical examples were produced with differences in the capital charge. To the best of our knowledge, however, the true nature of the problem was not appropriately identified due to the complexity of the formalism. Indeed, we will see below that, for the linear charge, the anomaly is quite subtle, as there are cancellations allowed between the above two shocks, and for properly chosen values of the correlations (and uniform risk weights) the anomaly can disappear altogether in spite of the perceived double counting. For the nonlinear charge, the anomaly is stronger, but double counting is not really correct there either; the situation is more convoluted than that. In this paper, we analyze the problem in its full generality and essentially solve it, by offering a solution from within the very formalism of the rules (we do not even touch the prescribed correlations).

2 THE CAPITAL CHARGE FUNCTIONS FOR FOREIGN EXCHANGE RISK

2.1 Definitions

The SA requires that asset classes be defined by their sensitivity to various risk factors, which generally speaking are diverse enough for most purposes but not as granular as a bank’s internal model might be. These are then divided into general interest rate risk (GIRR), credit spread risk (CSR), equity, commodity and FX. There are other components to the SA, such as the jump to default, but we are mainly interested in the FX portion in this paper. It is expected that Deltas (also known as sensitivities) are
computed with respect to the risk factors, essentially by computing first derivatives, or the linear terms in a Taylor expansion. These sensitivities are then multiplied by various risk weights assigned by the regulators as parameters calibrated (typically) by historical simulations. In other words, for each risk factor labeled by index $k$, we have the weighted sensitivity $WS$ given by

$$WS_k = RW_k \times s_k,$$  \hspace{0.5cm} (2.1)

where

$$s_k = \frac{\partial V}{\partial (p_k)} p_k$$

is the sensitivity of the portfolio with respect to the risk factor $p_k$, and $RW_k$ is the risk weight assigned to the risk factor. In GIRR and CSR, the definition is slightly different, but we will not need that here. Once the weighted sensitivities have been calculated, we then have an aggregation formula to compute the capital charge. Aggregation in the SA was initially multi-leveled (also known as cascading) (Basel Committee on Banking Supervision 2015a), and that is the most general case, which we study for completion and in case there is a need for further change or reversion to the older version. Later, we specialize as desired. Essentially, instruments of a certain asset class are assigned to various buckets $b$, and aggregated within each bucket first. Then we have another aggregation across different buckets. There can be more than two levels. From Basel Committee on Banking Supervision (2015a), we find that the aggregation within a bucket $b$ is given by

$$K_b = \sqrt{\sum_{k} WS_k^2 + \sum_{i \neq j} \rho_{i,j} WS_i WS_j},$$  \hspace{0.5cm} (2.2)

where $\rho$ is the intra-bucket “correlation matrix” and the sum is over all risk factors within the bucket. For FX, the buckets are the different foreign currencies, and the sum in (2.2) is a sum over sensitivities for the same spot exchange rate but with different maturities distinguished for each deal (or position). In fact, $\rho_{i,j}$ takes values $\rho_{i,j}^+$ or $\rho_{i,j}^-$, $\rho_{i,j}^+ \geq \rho_{i,j}^- \geq 0$, if the signs of the sensitivities for the risk positions $i$ and $j$ are the same, or different, respectively. The reader need not worry about this asymmetry for now. We only use it once later on and we simplify it in that case. This is a way to capture the risk of deviations from expected movements at different maturities as functions of the spot rate. Once $K_b$ has been computed for each bucket, we then aggregate across buckets using

$$\text{linear risk capital charge} = \sqrt{\sum_{b} K_b^2 + \sum_{a \neq b} \gamma_{a,b} S_a S_b},$$  \hspace{0.5cm} (2.3)
where now the sum is over buckets (or exchange rates of foreign currencies against the reporting currency), $\gamma$ is the intra-bucket correlation matrix and

$$S_b = \sum_l WS_l^2,$$

(2.4)

where the sum is over the risk factors in bucket $b$ (in FX, this is a sum over certain maturity buckets assigned as less than one year, one to three years and more than three years). In other words, each foreign currency represents a bucket, and within that bucket we have three labels denoting in which “term bucket” (which are actually sub-buckets of the larger bucket $b$) the deal, or position, in that currency happens to be. We actually prefer the equivalent, but possibly more intuitive, representation

$$\sum_{(l,a)} WS^2_{(l,a)} + \sum_{(k,a)\neq(l,b)} \pi_{(k,a),(l,b)} WS_{(k,a)} WS_{(l,b)},$$

(2.5)

where the index $(l,a)$, indicates risk factor $l$ in bucket $a$, and the matrix $\pi$ is defined by

$$\pi_{(k,a),(l,b)} = \begin{cases} \rho_{k,l}, & a = b, \\ \gamma_{a,b}, & a \neq b. \end{cases}$$

(2.6)

In FX, $\gamma$ denotes the correlation between any two foreign currency risk factors of the form $p[X_1, Y]$ and $p[X_2, Y]$, and RW was originally the same for all exchange rates (at time of writing, however, the regulatory values are 0.6 and 0.3, respectively; the latter was changed from the value 0.15 in Basel Committee on Banking Supervision (2015a)). As discussed in Section 3, it is not possible to have meaningful values of $\gamma \neq 0.5$ as a true correlation. However, there is nothing to stop the regulatory risk measures from being constructed as norms on vector spaces (eg, Banach spaces) that do not behave like a Euclidean metric or a Hilbert metric. So, we will not consider the term correlation to be genuine here and we are willing to move beyond this objection.

The final Basel rules have now removed the need for maturity buckets or asymmetry in the inter-bucket correlations, but for generality and completion we allow their use here. Again, the different risk factors within a bucket are defined as the spot exchange rate of the currency of the bucket to reporting currency, but they are distinguished by the maturity of the deal. In other words, we distinguish between the sensitivity of a deal expiring in five years from that of one expiring in five months. For our purposes, all we need to know is that $\rho$ is not necessarily a uniform value of the correlations between the different risk factors in each bucket. For completeness, however, we give the reader the current values of $\rho_{i,j}$: for positions of the same sign (see Basel Committee on Banking Supervision 2015a),

$$\begin{pmatrix} 1 & 0.95 & 0.70 \\ 0.95 & 1 & 0.85 \\ 0.70 & 0.85 & 1 \end{pmatrix}.$$
and for positions of the opposite sign,

\[
\begin{pmatrix}
1 & 0.90 & 0.65 \\
0.90 & 1 & 0.80 \\
0.65 & 0.80 & 1
\end{pmatrix}.
\]

The linear risk capital charge is often referred to as the Delta charge by practitioners. We use the terms interchangeably. For the nonlinear or curvature capital charge for FX, we start by defining a curvature risk position for each risk factor, namely

\[
\text{CVR}_k = \min \left\{ \sum_{\alpha} V(x_{\alpha,k} + \text{RW}_k \times x_{\alpha,k}) - V(x_{\alpha,k}) - \text{RW}_k \times s_k, \right. \\
\left. \sum_{\alpha} V(x_{\alpha,k} - \text{RW}_k \times x_{\alpha,k}) - V(x_{\alpha,k}) + \text{RW}_k \times s_k \right\}, \tag{2.7}
\]

where \(x_{\alpha,k}\) is the original value of the risk factor (often called the base value by practitioners), \(V(\cdot)\) is the value of the instrument in a given position \(\alpha\) as a function of the risk factors, and the sum is over all instruments in the portfolio. From this point on we will drop the sum over the instruments and just think of \(V(\cdot)\) as the value of the entire portfolio (that is, the sum over positions is implicit). Observe that the curvature risk position in risk factor \(k\) is the risk of profit-and-loss change of the portfolio that is in excess of the linear (delta) part. It is thus intended to isolate the nonlinearities, and hence is mainly relevant for options rather than linear positions such as spot positions or forwards. Now the curvature capital charge has more nonlinearities than the Delta charge but follows a similar aggregation scheme,

\[
K^2_b = \max \left( 0, \sum_k (\max(0, \text{CVR}_k))^2 \right)
\]

\[
+ \sum_{i \neq k} \rho'_{i,k} \times \text{CVR}_k \text{CVR}_i \psi(\text{CVR}_k, \text{CVR}_i) \right) \right)^{1/2}, \tag{2.8}
\]

where the sum is over all risk factors in bucket \(b\), \(\rho'\) is an intra-bucket correlation matrix not necessarily the same as \(\rho\) and \(\psi(x, y)\) is a function that takes the value 0 if both its arguments are negative and 1 otherwise. Again, to aggregate, we start with

\[
S'_b = \sum_k \text{CVR}_k,
\]

where the sum is over the risk factors in bucket \(b\). Recall that the buckets and maturity term labels are the same for both delta and curvature. Then we have

\[
\text{curvature risk capital charge} = \sqrt{\max \left( 0, \sum_b K'^2_b + \sum_{i \neq k} \gamma'_{a,b} \times S_b S_a \psi(S_a, S_b) \right)}, \tag{2.9}
\]
where the sum is now over the buckets, and $\gamma'$ is an inter-bucket correlation matrix not necessarily the same as $\gamma$. At time of writing, the (regulatory) values of the entries of $\rho'$ and $\gamma'$ are prescribed to be the squares of the corresponding entries of $\rho$ and $\gamma$, respectively. In particular, $\gamma'_{a,b}$ between the different currencies is 0.36.

**Remark 2.1** Before we proceed further, and as alluded to earlier, we liberate ourselves from the concept of correlations as a geometric concept. In what follows, we do not worry about whether or not any of the matrices above are possible genuine correlation matrices. From our point of view, these can be looked at as aggregation coefficients or association parameters. We consider that naming them “correlations” is a bit unfortunate at times. The reader should think of the charge functions as just functions – and we may take a purely algebraic point of view without the usual geometric meaning. The reason we do this is that we anticipate the regulators will change their matrices in conservative ways regardless of the geometric meaning. For the problem at hand, all we want to study is the invariance (or lack thereof) of the capital charge functions.

Finally, we note that the risk weights have recently been changed in the final Basel rules (Basel Committee on Banking Supervision 2016). They are now 30% in general, but banks are allowed to divide this value by $\sqrt{2}$ at their discretion for the following currency pairs: USD/EUR, USD/JPY, USD/GBP, USD/AUD, USD/CAD, USD/CHF, USD/MXN, USD/CNY, USD/NZD, USD/RUB, USD/HKD, USD/SGD, USD/TRY, USD/KRW, USD/SEK, USD/ZAR, USD/INR, USD/NOK, USD/BRL, EUR/JPY, EUR/GBP, EUR/CHF and JPY/AUD. This newly introduced asymmetry in risk weights causes even bigger anomalies.

### 2.2 Transformation between reporting currencies

Now we come to the issue under consideration. Suppose we have a portfolio with value $V(p[x, y])$, where $p[x, y]$ is the exchange rate between currencies $x$ and $y$. Conventionally, we let this denote the value of one unit of currency $x$ in units of currency $y$. For example, the USD/CAD rate $p$ is the value of one US dollar in Canadian dollars. The problem begins with the following transformation. Suppose $V$ is the value in units of $y$ (we will remove this assumption in Section 3, but for now it is convenient). Recall the triangulation property for FX rates, namely that

$$p[x, y] \times p[y, z] = p[x, z]. \quad (2.10)$$

We now have many ways to express $V$. Indeed, we have

$$V(p[x, y]) \text{ in units of } y = p[y, z] \times V(p[x, z] \times p[z, y]) \text{ in units of } z, \quad (2.11)$$

for any currency $z$. So, right from the beginning, the representation is not unique. The regulatory requirement is to consider the risk factors to be the exchange rates referencing the reporting currency of the bank. So, in principle, the functional dependency
for each bank will be different depending on the reporting currency. The conversion to units of \( y \) or \( z \) is harmless, as we will see. It is the functional dependency that is problematic. We are now led to the following question: do we get the same Delta and curvature charges for the same portfolio, for two different banks, with different reporting currencies?

### 2.3 Definitions

At this stage, we are ready to give some definitions. Given currency \( Z \), and for a portfolio \( V(p) \) valued in currency \( W \), with \( p \) a vector of exchange rates \( p_i = p[X_i, Y_i] \), let \( U(q) \) be the function \( V \), with each \( p_i \) replaced by \( (p[X_i, Z]/p[Y_i, Z]) \). We then have the following.

**Definition 2.2** We call the function \( (p[W, Z]) \times U(q) \) the \( Z \)-currency transformation of \( V \), or simply \( V \) in \( Z \)-currency. We also call the mapping

\[
p[X_i, Y_i] \rightarrow \left( \frac{p[X_i, Z]}{p[Y_i, Z]} \right)
\]

the \( Z \)-currency transformation of \( p[X_i, Y_i] \).

Now, for reporting currency \( Z \), we recall that the regulations require us to express sensitivities \( s_k \) using the partial derivative of a portfolio \( V \), valued in \( Z \), with respect to \( p[X_k, Z] \). When we compute such sensitivities, we can think of them as a list of numbers with an aggregation rule according to (2.1). For our purposes, there is no harm or loss of generality in thinking of some objects \([X_i]_Z\) being associated with the foreign currencies \( X_i \). We may think of these objects as vectors in an inner product space if the aggregation matrix happens to be a genuine correlation matrix with all the standard properties, or, more generally, elements of a general vector space if we need to drop the latter requirement. For now, however, it is enough to think of objects \([X_i,\alpha]_Z\), where the index \( \alpha \) is the time-bucket label (\( \alpha \) is a discrete label denoting regulatory maturity buckets within the currency bucket and should not be confused with the currency bucket). We also added the \( Z \) label to remind us that these vectors represent sensitivities in \( Z \)-currency. We typically drop the label \( Z \) when the meaning is clear from the context. We can define some product of these objects such that

\[
[X_i,\beta] \cdot [X_k,\alpha] = \pi(\alpha, \beta).
\]

Again, this does not have to be a genuine scalar (dot) product; it may merely be an abstract product to help us with bookkeeping. What matters to us is the capital charge function, which we treat purely algebraically. Geometry is merely for bookkeeping here. With this in mind, we have the following definition.
For risk weights $RW_i$, and $V$ the value of a portfolio in units of currency $Z$, the $Z$-Delta representation in the $Z$-currency is defined as a function of all exchange rates between foreign currencies $X_i$ and currency $Z$ by

$$
\sum_{i,a} RW_i \times \left( \frac{\partial V_a}{\partial (p[X_i, Z])} \times p[X_i, Z] \right) [X_i, a]_Z,
$$

with $a$ being the index of the term bucket. We indicate by $V_a$ that we only take the partial derivative of the value of the portfolio that depends on maturity bucket $a$.

With this definition, we observe that the weighted sensitivity $WS_{i,a}$ is now the coefficient of $[X_i, a]_Z$. This representation will in particular be quite powerful in our analysis of the invariance of the Delta charge.

### 3 CHARACTERIZATION OF THE INVARIANCE OF THE LINEAR (DELTA) CAPITAL CHARGE

In this section, we describe the Delta charge’s invariance and its dependence on the choices of the matrixes $\gamma$ and $\rho$. The regulatory matrix has $\rho_{aa} = 1$. To start, suppose we take an FX book that depends only on the first time bucket, so that only time index $\alpha = 1$ is used and we only need to focus on $\gamma$. Recall also that $\gamma_{ij} = 1$ if $i = j$ and $\gamma_{ij} = \gamma$ if $i \neq j$.

Before we proceed, we must explain that it is not possible to expect invariance for a portfolio of nonzero mark-to-market (MtM). Indeed, suppose you have a portfolio that is only a function of $p[Y, X]$, so that $V = V(p[Y, X])$, valued in units of $X$. The base value of the MtM of this portfolio, for fixed $p[Y, X]$, has zero FX risk in $X$-currency, but can be nonzero in another currency (think of a simple cash position in $X$). More precisely, we are referring to the constant part of $V$, which does not depend on any other exchange rate when expressed in units of $X$ units. It is essentially a cash position in $X$. We handle this issue later on, but for now it is not meaningful to expect invariance of the Delta charge, unless the MtM in $X$ units is zero. Additionally, note that for a bank that reports in currency $X$, the constant value of $V$ (with $V$ measured in units of $X$) in the Taylor expansion ($V_{\text{base}}$) has no FX risk and no Delta charge. Therefore, there is no loss of generality in working with this condition for the problem under consideration. We therefore study the Delta charge of $V - V_{\text{base}}$, where $V_{\text{base}}$ is the base value of the portfolio (ie, the value if no risk factor moves). We also denote the base value of $p[Y, X]$ by $\tilde{p}[Y, X]$. We start by computing the Delta charge in $X$-currency. Let

$$
a = \frac{\partial V(p)}{\partial p} \bigg|_{\tilde{p}[Y, X]}.
$$
Now we compute the Delta charge in $X$-currency, and units of $X$:

$$\text{Delta charge in } X\text{-currency and units of } X = \text{RW} \times a \times \tilde{\rho}[Y, X]. \quad (3.1)$$

On the other hand, if $Z$ is a third currency, we find that in $Z$-currency and units of $Z$, the $Z$-currency transformation of $V$ is

$$\tilde{\rho}[X, Z]\left\{ V\left(\frac{p[Y, Z]}{p[X, Z]}\right) - V_{\text{base}} \right\}, \quad (3.2)$$

and, using the chain rule, the $Z$-Delta representation is

$$\text{RW} \times p[X, Z]a\left\{ \frac{\tilde{\rho}[Y, Z]}{\tilde{\rho}[X, Z]} [Y]_Z - \frac{\tilde{\rho}[Y, Z]}{\tilde{\rho}[X, Z]} [X]_Z \right\}$$

$$= \text{RW} \times a \times \tilde{\rho}[Y, Z]\{[Y]_Z - [X]_Z\}. \quad (3.3)$$

Now, the Delta charge, using $\pi$, or equivalently $\gamma$ (since $\rho$ does not play a role here), is given by

$$\text{RW} \times a \times \tilde{\rho}[Y, Z]\sqrt{2 - 2\gamma}. \quad (3.4)$$

Finally, the Delta charge in $Z$-currency but in units of $X$ is

$$\text{RW} \times a \times \tilde{\rho}[Y, X]\sqrt{2 - 2\gamma}. \quad (3.5)$$

It is now clear that invariance can occur only if $\gamma = 0.5$. To the best of the author’s knowledge, this value was first pointed out by Youngsuk Lee (personal communication, 2015) as a consequence of the triangular property of exchange rates, consistency of correlations and equal volatilities of the three rates in the triangle (with drifts set to zero). Here we are treating the capital charge more algebraically and without any requirement for the existence of true correlations between the currencies. In other words, we are treating them as aggregation parameters rather than true correlations (which seems to be the regulatory approach), but we still end up with $\gamma = 0.5$ as a requirement. Because we deal with the problem purely algebraically, however, we are able to pursue the solution without having to change the value of $\gamma$ from the regulatory 0.6. We will see this later on.

Another feature to observe is that in units of $X$ it does not matter which reporting currency we use, and the dependence of the charge on the exchange rates involves only the “underlying” rate. No other exchange rate is needed in the final expression, even though we needed to work in $Z$-currency initially. This is an important observation that we will expand on later. At this point, let us take two deals with the same $V$ as above, but this time they lie in different maturity buckets, $a = 1, 2$, with correlation
\( \rho_{1,2} = \beta \). In \( X \)-currency, we have the following \( X \)-Delta representation:

\[
\text{RW} \times a \times \bar{p}[Y, X]\{[Y_1]_X + [Y_2]_X\},
\]

(3.6)

and the Delta capital charge is given by

\[
\text{Delta charge in } X \text{-currency and units of } X = \text{RW} \times a \times \bar{p}[Y, X]\sqrt{2 + 2\beta}. \quad (3.7)
\]

In \( Z \)-currency, however, the \( Z \)-Delta representation is

\[
\text{RW} \times a \times \bar{p}[Y, Z]\{[Y_1]_Z - [X_1]_Z + [Y_2]_Z - [X_2]_Z\},
\]

(3.8)

and the Delta capital charge is given by

\[
\text{Delta charge in } Z \text{-currency and units of } X = \text{RW} \times a \times \bar{p}[Y, Z]\sqrt{4 - 8\gamma + 4\beta}. \quad (3.9)
\]

In units of \( X \), we have

\[
\text{Delta charge in } Z \text{-currency and units of } X = \text{RW} \times a \times \bar{p}[Y, X]\sqrt{4 - 8\gamma + 4\beta}. \quad (3.10)
\]

For \( \gamma = 0.5 \), we can see that we have invariance of the Delta charge only if \( \beta = 1 \) or \( \rho \) is the identity matrix. Now we return to the issue of asymmetric “correlations”, which we only address here briefly in order to give a full characterization theorem below. To handle this case it is sufficient to consider two maturities with labels 1 and 2, but with opposite signs for sensitivity. In this case, we have the same type of computation, except that now the \( X \)-Delta and \( Z \)-Delta representations are given by

\[
\text{RW} \times a \times \bar{p}[Y, X]\{[Y_1]_X - [Y_2]_X\}, \quad (3.11)
\]

\[
\text{RW} \times a \times \bar{p}[Y, Z]\{[Y_1]_Z - [X_1]_Z - [Y_2]_Z + [X_2]_Z\}, \quad (3.12)
\]

respectively. We therefore get

\[
\text{Delta charge in } X \text{-currency and units of } X = \text{RW} \times a \times \bar{p}[Y, X]\sqrt{2 - 2\beta} \quad (3.13)
\]

and

\[
\text{Delta charge in } Z \text{-currency and units of } X = \text{RW} \times a \times \bar{p}[Y, X]\sqrt{4 - 4\beta}. \quad (3.14)
\]

Once again, we have invariance only if \( \beta = 1 \). Similarly, the conclusion holds for any pair of indices \( i, j \), with \( i \neq j \). We conclude that the matrix \( \rho \) has to be the identity, with no asymmetry allowed, if we want invariance. In fact, this example shows that in \( Z \)-currency the square of the charge can double for some reporting currency. In fact, it is doubled in any \( Z \)-currency where \( Z \) is different from \( X \) and \( Y \). So the charge can be 40% higher for \( \beta \neq 0 \), and indeed Basel Committee on
Banking Supervision (2015a) had three time buckets in the regulatory framework with corresponding values for $\rho \neq 1$. This is certainly not an exotic condition. It simply means that if your strategy is based on different maturity buckets hedging each other and you hedge your book on a currency pair basis just like in our example, you would be in this situation. Fortunately, maturity buckets were removed in Basel Committee on Banking Supervision (2016), so this type of anomaly is not present anymore. However, let us examine the impact of the exceptions introduced for the special currency pairs in Basel Committee on Banking Supervision (2016). Suppose we have an option on EUR/CAD and our reporting currency is Canadian dollars. Suppose the same option is also held by a US-dollar-reporting bank. EUR/USD and CAD/USD are in the exception list. Then, according to the exception on risk weights and the regulatory value of $\gamma = 60\%$, the Delta charges for the same option under Canadian dollar reporting would be $\sqrt{2} \times (\sqrt{2 - 2 \times 0.6})^{-1} \approx 1.58$ times as big as the Delta charge under US dollar reporting.

Again, we will not attempt to change any matrixes in our solution later on, as we understand the regulatory conservatism behind them and we merely think of these as aggregation parameters. We are now ready to characterize invariance for a general FX book. If the book is valued in $X$-currency, in units of $X$, at the present time the constant part of the portfolio (the part with no dependency on rates) has no FX risk and is certainly not capitalized. This is because the Delta charge aggregates the sensitivities, which are zero for the constant part. It also needs to be subtracted from the portfolio before we transform it into any other reporting currency, and it is that excess beyond the constant term for which we seek invariance of Delta charge.

Armed with the techniques above, we can now show the following result.

Theorem 3.1

1. The square of the Delta capital charge function is the dot product of the Delta representation with itself (by construction) and is invariant under currency transformations (e.g., change in reporting currency) if and only if risk weights are equal without exceptions, $\gamma = 0.5$ and $\rho$ is the identity matrix (without asymmetry).

2. The current regulatory values for the risk weights (with the list of exceptions) and $\gamma = 0.6$ as in Basel Committee on Banking Supervision (2016) allow an at least 58% difference in capital charges for some FX portfolios. Other anomalies exist, even for uniform risk weights, if $\rho$ is not the identity (as in Basel Committee on Banking Supervision (2015a)), with differences of at least 40% for some portfolios.
(3) For fixed currency $X$, and as functions of the exchange rates, the Delta capital charge and the Delta representation coefficients depend only on the underlying rates, except possibly for a conversion factor depending on the currency in which they are measured. The lack of invariance is therefore a consequence of the aggregation mechanism and the structure of the Delta charge function, and not the functional dependence on the exchange rates against the reporting currency.

**Proof** We have already proved the “only if” part of (1). We also proved part (2), and we demonstrated (3) in the simple examples above. To prove the invariance for $\gamma = 0.5$ and $\rho = 1$, we represent the exchange rates in a consistent way as $p[X_i, X_j]$, with $i = j + 1, \ldots, n$ and $j = 0, \ldots, n$, where, without loss of generality, we order our currencies in any way we like. We start with a portfolio $V$, in units of $X_0$, and defined as a function of the vector $p$ of currency pairs by

$$V(p) = V(p_{1,0}, p_{2,0}, \ldots, p_{n,0}, p_{2,1}, \ldots, p_{n,1}, \ldots, p_{n,n-1}).$$  \hspace{1cm} (3.15)

We also denote by $t_{i,j}$ the partial derivatives of $V$ with respect to the $p_{i,j}$, evaluated at the base values $\bar{p}_{i,j}$: that is,

$$t_{i,j} = \frac{\partial V}{\partial p_{i,j}} \bigg|_{\bar{p}}.$$  \hspace{1cm} (3.16)

In $X_0$-currency, and in units of $X_0$, we have, without loss of generality,

$$V(p) - V_{\text{base}} = \left\{ \sum_{i=1}^{n} t_{i,0} \times (p[X_i, X_0] - \bar{p}[X_i, X_0]) \right. \hspace{0.5cm} + \sum_{i=2}^{n} t_{i,1} \times \left( \frac{p[X_i, X_0]}{p[X_1, X_0]} - \frac{\bar{p}[X_i, X_0]}{\bar{p}[X_1, X_0]} \right) \right. \hspace{0.5cm} + \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \left( \frac{p[X_i, X_0]}{p[X_j, X_0]} - \frac{\bar{p}[X_i, X_0]}{\bar{p}[X_j, X_0]} \right) \right\} + \text{h.o.t.},$$  \hspace{1cm} (3.17)

where “h.o.t.” denotes higher-order terms.

Using (2.10), the $X_0$-Delta representation is now

$$\text{RW} \left\{ \sum_{i=1}^{n} (t_{i,0} \times \bar{p}[X_i, X_0]) [X_i] + \sum_{i=2}^{n} t_{i,1} \times \bar{p}[X_i, X_1] ([X_i] - [X_1]) \right. \hspace{0.5cm} + \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \bar{p}[X_i, X_j] ([X_i] - [X_j]) \right\}.$$  \hspace{1cm} (3.18)
We note that the weighted sensitivity of the portfolio to the risk factor \( p[X_i, X_0] \) in \( X_0 \)-currency, and in units of \( X_0 \), is the coefficient of \( [X_i] \) in (3.18). We denote this by \( S^X_0 \). Thus,

\[
S^X_0 = \begin{cases} 
\text{RW}\left\{ t_{i,0} \times \tilde{p}[X_i, X_0] + t_{i,1} \times \tilde{p}[X_i, X_1] \right. \\
+ \sum_{j=2}^{i-1} t_{i,j} \times \tilde{p}[X_i, X_j] - \sum_{j=i+1}^{n} t_{j,i} \times \tilde{p}[X_j, X_i] \}, & i > 1, \\
\text{RW}\left\{ t_{1,0} \times \tilde{p}[X_1, X_0] - \sum_{i=2}^{n} t_{i,1} \times \tilde{p}[X_i, X_1] \right. \\ \left. \right. & i = 1.
\end{cases}
\tag{3.19}
\]

To avoid having to take lengthy derivatives unnecessarily, we observe the following.

**Lemma 3.2** In \( X_1 \)-currency, the higher-order terms in (3.17) do not contribute to the linear term in the Taylor expansion.

The lemma essentially says that on substituting

\[
\left( \frac{p[X_i, X_1]}{p[X_j, X_1]} \right)
\]

for \( p[X_j, X_j] \) in the quadratic or higher-order terms for \( V \) and considering this as a function of \( p[X_i, X_1] \), with \( p[X_j, X_1] \) fixed at base value, only the linear term presented in (3.17) gives linear terms in \( p[X_i, X_1] \). A similar result is obtained for \( p[X_j, X_1] \). The proof is trivial for \( p[X_i, X_1] \), and for \( p[X_j, X_1] \) we just need to observe that

\[
\left( \frac{1}{f + \delta f} - \frac{1}{f} \right)^n
\]

has no terms of order less than \( n \) in \( \delta f \).

The lemma allows us to ignore the higher-order terms in writing the \( X_1 \)-currency transformation of \( V(p) - V_{\text{base}} \), which is now

\[
p[X_0, X_1] \times \left\{ \sum_{i=1}^{n} t_{i,0} \times \left( \frac{p[X_i, X_1]}{p[X_0, X_1]} - \frac{\tilde{p}[X_i, X_1]}{\tilde{p}[X_0, X_1]} \right) \\
+ \sum_{i=2}^{n} t_{i,1} \times (p[X_i, X_1] - \tilde{p}[X_i, X_1]) \\
+ \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \left( \frac{p[X_j, X_1]}{p[X_j, X_j]} - \frac{\tilde{p}[X_j, X_1]}{\tilde{p}[X_j, X_1]} \right) \right\} + \text{h.o.t.} \tag{3.20}
\]
Finally, using (2.10), we have the $X_1$-Delta representation, in units of $X_1$:

\[
\text{RW} \times p[X_0, X_1] \\
\times \left\{ -(t_{1,0} \times \bar{p}[X_1, X_0])[X_0] + \sum_{i=2}^{n} (t_{i,0} \times \bar{p}[X_i, X_0])([X_i] - [X_0]) \\
+ \sum_{i=2}^{n} t_{i,1} \times \bar{p}[X_i, X_1][X_i] + \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \bar{p}[X_i, X_j]([X_i] - [X_j]) \right\}.
\]

(3.21)

In units of $X_0$, this becomes

\[
\text{RW} \times \left\{ -(t_{1,0} \times \bar{p}[X_1, X_0])[X_0] + \sum_{i=2}^{n} (t_{i,0} \times \bar{p}[X_i, X_0])([X_i] - [X_0]) \\
+ \sum_{i=2}^{n} t_{i,1} \times \bar{p}[X_i, X_1][X_i] + \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \bar{p}[X_i, X_j]([X_i] - [X_j]) \right\}.
\]

(3.22)

We also have the sensitivities in $X_1$-currency, in units of $X_1$, given by

\[
S_i^{X_1} = p[X_0, X_1] \\
\times \left\{ (t_{i,0} \times \bar{p}[X_i, X_0]) + t_{i,1} \times \bar{p}[X_i, X_1] \\
+ \sum_{j=2}^{i-1} t_{i,j} \times \bar{p}[X_i, X_j] - \sum_{j=i+1}^{n} t_{j,i} \times \bar{p}[X_j, X_i], \quad i > 1, \\
- \sum_{j=1}^{i} (t_{j,0} \times \bar{p}[X_j, X_0]), \quad i = 0, \right\}
\]

(3.23)

and in units of $X_0$, we get

\[
S_i^{X_0} = \left\{ (t_{i,0} \times \bar{p}[X_i, X_0]) + t_{i,1} \times \bar{p}[X_i, X_1] \\
+ \sum_{j=2}^{i-1} t_{i,j} \times \bar{p}[X_i, X_j] - \sum_{j=i+1}^{n} t_{j,i} \times \bar{p}[X_j, X_i], \quad i > 1, \\
- \sum_{j=1}^{i} (t_{j,0} \times \bar{p}[X_j, X_0]), \quad i = 0. \right\}
\]

(3.24)

The capital charge can then be computed using (2.13). As we can see in (3.18)–(3.24), the functional form of the Delta representation is only dependent on the original
underlying rates in $V$ when we convert to the same currency. Therefore, the same holds for the Delta charge. In other words, the anomaly is that the vector representation has changed. This proves part (3).

**Remark 3.3** The change in vector representation is important because we hoped to find a method of computing capital charges such that they are dependent only on the underlying rates of the portfolio and not on the rates referencing the reporting currency. Part (3) of Theorem 3.1 shows that this is already the case within the framework. The problem is not the functional form, but rather that the anomaly is embedded in the mechanism of aggregation (in the coefficients) and that the Delta representation in each reporting currency is a bit different.

To complete the proof, we observe that the square of capital charge is now a second degree multi-variable polynomial in $t_{i,j}$, and therefore it is enough to prove that the coefficients for each $t_{i,j} \times t_{k,l}$ are the same for the aggregations of (3.18) and (3.22). We start by observing that in both representations the coefficients of $t_{i,j}$ have exactly the same dependency on the exchange rate. The only part that needs to be checked is the numerical factor that arises from the dot product that we defined on the representation vectors. Both (3.18) and (3.22) have three main summands, which we call $I_1$, $I_2$ and $I_3$ for (3.18), and $I'_1$, $I'_2$ and $I'_3$ for the corresponding terms of (3.22):

$$I_1 = \sum_{i=1}^{n} (t_{i,0} \times \tilde{p}[X_i, X_0])[X_i],$$

$$I_2 = \sum_{i=2}^{n} t_{i,1} \times \tilde{p}[X_i, X_1][(X_i) - [X_1]),$$

$$I_3 = \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \tilde{p}[X_i, X_j][(X_i) - [X_j]).$$

$$I'_1 = -(t_{1,0} \times \tilde{p}[X_1, X_0])[X_0] + \sum_{i=2}^{n} (t_{i,0} \times \tilde{p}[X_i, X_0])([X_i] - [X_0]),$$

$$I'_2 = \sum_{i=2}^{n} t_{i,1} \times \tilde{p}[X_i, X_1][X_i],$$

$$I'_3 = \sum_{j=2}^{n-1} \sum_{i=j+1}^{n} t_{i,j} \times \tilde{p}[X_i, X_j][(X_i) - [X_j]).$$

The case when $i = 1$ represents an exceptional case, so we deal with it first. The numerical coefficient for $t_{1,0}^2$ is 1 in $I_1 \cdot I_1$, and the same is true for $I'_1 \cdot I'_1$. The coefficient of $t_{1,0}t_{j,0}$ arising from $I_1 \cdot I_1$ is $\gamma$, whereas for $I'_1 \cdot I'_1$ it is $(1 - \gamma)$, and these coincide for $\gamma = 0.5$. Similar terms arising from $I_1 \cdot I_2$ give $(\gamma - 1)$, whereas
for $I'_1 \cdot I'_2$ we get $\gamma$, which again coincides for $\gamma = 0.5$. Finally, the coefficients of $t_{1,0} f_{1,0}$, arising from $I_1 \cdot I_3$ and $I'_1 \cdot I'_3$, are proportional to $(\gamma - \gamma) = 0$.

Now, we consider the case when $i \neq 1$. We note that for $I_1 \cdot I_1$ the terms corresponding to $(t_i f_i)^2$ have a numerical factor 1 from the dot product, as they arise from $[X_i] \cdot [X_i]$. For $I'_1 \cdot I'_1$, the corresponding coefficient comes from $([X_i] - [X_0]) \cdot ([X_i] - [X_0]) = (2 - 2\gamma)$. Therefore, for $\gamma = 0.5$, we have equality. Now, the cross-terms for $I_1 I_1$ have a numerical factor arising from $[X_i] \cdot [X_i] = \gamma$, $i \neq 1$. For $I'_1 I'_1$, we have

$$( [X_j] - [X_0] ) \cdot ( [X_i] - [X_0] ) = ( 1 + \gamma - 2\gamma ) = ( 1 - \gamma ).$$

Again, we have equality for $\gamma = 0.5$. For $I_2, I'_2$, we have the same situation, except $[X_1]$ replaces $[X_0]$. The terms of $I_3, I'_3$ are already identical to begin with. Now, we consider the terms arising from $I_1 \cdot I_2$. These have numerical factors that come from $[X_i] \cdot ( [X_j] - [X_0] )$, which are 0 if $i$ and $j$ are distinct, and $(1 - \gamma)$ if $i$ and $j$ are the same. For $I'_1 \cdot I'_2$, again we have $[X_i] \cdot ( [X_j] - [X_1] )$, and the same consideration applies. For $I_1 \cdot I_3$, we have terms of the form $[X_k] \cdot ( [X_i] - [X_j] )$. If $k$ is distinct from $i$ and $j$, we get 0, and if $k = i$, we get $1 - \gamma$. For $I'_1 \cdot I'_3$, we have factors of the form $( [X_k] - [X_0] ) \cdot ( [X_i] - [X_j] )$. Here we note that the dot product involving $X_0$ vanishes because $k$ and $l$ are nonzero by construction. Therefore, we have a similar consideration as for the term for $I_1 \cdot I_3$. Finally, we consider the terms for $I_2 \cdot I_3$. Here we have $( [X_k] - [X_0] ) \cdot ( [X_i] - [X_j] )$. Again, by construction, $i$ and $j$, are different from 1, and therefore this reduces to $[X_k] \cdot ( [X_i] - [X_j] )$. For $I'_2 \cdot I'_3$, the terms arise from $[X_k] \cdot ( [X_i] - [X_j] )$, which is exactly the same form. We conclude equality for all coefficients under the constraint $\gamma = 0.5$. This concludes our proof of the theorem.

\[ \square \]

4 THE NONINVARIANCE OF THE NONLINEAR (CURVATURE) CAPITAL CHARGE

We start by showing the failure of invariance in an example of a portfolio that depends on one currency pair $X_1/X_0$, with rate $p[X_1, X_0]$. Consider a portfolio $V(p[X_1, X_0])$ in $X_0$-currency, valued in units of $X_0$. Suppose $S^{X_0} \Gamma^X$ is the sensitivity of $V$ with respect to $p[X_1, X_0]$, in $X_0$-currency, and in units of $X_0$. We also allow $R$ to be a signed variable that takes values $\pm$RW. With $\bar{p}$ denoting the base value of the rate, we can then compute the curvature representation in $X_0$-currency:

$$CVR = \min_{R \in \{-RW,RW\}} \{ V((1+R)\bar{p}[X_1, X_0]) - V(\bar{p}[X_1, X_0]) - S^{X_0} \Gamma^X \} [X_1]. \quad (4.1)$$

For a currency $Y$ to be distinct from $X_1$ and $X_0$, we transform it to $Y$-currency and units of $Y$. We have the following $Y$-currency transform of $V$, in units of $Y$:

$$p[X_0, Y] \times V\left( \frac{p[X_1/Y]}{p[X_0, Y]} \right). \quad (4.2)$$
and suppose also that we have the sensitivities $S^Y_1, S^Y_0$ in $Y$-currency and units of $Y$. Then we have the $Y$-curvature representation

$$-\tilde{p}[X_0, Y] \times \left\{ \min_{R \in \{-RW,RW\}} \{ [V((1 + R)\tilde{p}[X_1, X_0]) - V(\tilde{p}[X_1, X_0])] - S^Y_1 R \} [X_1] \\
+ \min_{R \in \{-RW,RW\}} \{ [(1 + R)V((1 + R)^{-1}\tilde{p}[X_1, X_0]) - V(\tilde{p}[X_1, X_0])] - S^Y_1 R \} [X_0] \right\}. \tag{4.3}$$

where we have observed that shocking $p[X_1, Y]$ only, to $(1 + R)p[X_1, Y]$, produces an equivalent shock to $p[X_1, X_0]$, whereas the same shock to $p[X_0, Y]$ produces a shock $(1 + R)^{-1}$ to $p[X_1, X_0]$. We will generalize this theme shortly. Let us first compute the curvature charge for this portfolio, assuming quadratic behavior. In other words, in $X_0$-currency, we assume

$$V(r) = V_0 + a(r - \tilde{p}) - b(r - \tilde{p})^2 + \text{h.o.t.} \tag{4.4}$$

and $b > 0$ (otherwise, curvature charge will be zero), and we are basically “short gamma”. We have already computed the sensitivities for the relevant reporting currencies in (3.1) and (3.3). The curvature charge in $X_0$-currency is trivial, namely

$$b \times (\tilde{p}[X_1, X_0])^2RW^2,$$

whereas an easy computation shows that the linear terms cancel the sensitivities so that, in $Y$-currency and units of $Y$, the curvature representation reduces to

$$\tilde{p}[X_0, Y] \times b \left\{ RW^2(\tilde{p}[X_1, X_0])^2[X_1] + \left( \frac{RW^2}{1 - RW} \right)(\tilde{p}[X_1, X_0])^2[X_0] \right\}. \tag{4.5}$$

We can then compute the curvature charge in units of $X_0$ using (2.9), the regulatory value $\gamma' = 0.36$ and $RW = 0.3$, to get

$$RW^2(\tilde{p}[X_1, X_0])^2b \times \sqrt{1 + \frac{1}{(1 - RW)^2} + \frac{2\gamma'}{(1 - RW)}} \approx 2 \times RW^2(\tilde{p}[X_1, X_0])^2. \tag{4.6}$$

We also note that there is no value for $\gamma'$ that can restore invariance. Our example here is not too stylistic. Indeed, a (short) butterfly option would have a region of values for the spot rate where the behavior would be quadratic. In addition, (4.1) and (4.3) are quite general, and the lack of invariance could be shown on many other more complicated portfolios, except that the computation would require full revaluation.
Regardless of the complexity of the option, we can see that the part of the curvature representation in the \([X_1]\) direction in both reporting currencies (but measured in the same units) is the same. Therefore, the anomaly persists, at least if, for example, we have positive values for both components. So, quadratic behavior is not required at all, but helps to demonstrate the computation. So far we have not used the exceptional set of currency pairs. Suppose the above example was on an option on EUR/GBP. The curvature charge in Canadian dollar reporting would then be \((\sqrt{2})^2 \times 2 = 4\) times the charge under GBP reporting, due to the same anomaly above and exacerbated by the fact that the risk weight for EUR/GBP is allowed to be divided by \(\sqrt{2}\) (which can be applied in GBP reporting), whereas this is not allowed for EUR/CAD or GBP/CAD (which would be the case in Canadian dollar reporting).

With the above example in mind, and \(V(\cdot)\) as in (3.15), we define the following.

**Definition 4.1** We define \(F_i^{X_0}(R)\) as the finite shock of \(V\) in \(X_0\)-currency, and units of \(X_0\), with respect to \(X_i\), where

\[
F_i^{X_0}(R) = V((1 + R)^\delta(1,0) \bar{p}_{1,0}, (1 + R)^\delta(2,0) \bar{p}_{2,0}, \ldots, (1 + R)^\delta(n,0) \bar{p}_{n,n-1}) - V_0, \quad (4.7)
\]

and

\[
\delta(j,k) := \begin{cases} 1, & j = i, \\ -1, & k = i, \\ 0, & \text{otherwise}. \end{cases} \quad (4.8)
\]

Note that what was special about \(X_0\) is that it was the reporting currency in which \(V\) was defined to have its initial functional form. In any other reporting currency, we must transform \(V\) as in (2.12).

In \(X_1\)-currency, and units of \(X_1\), we can similarly define \(F_i^{X_1}(R)\) for \(i \neq 1\):

\[
F_i^{X_1}(R) = (1 + R)^\varepsilon \times \bar{p}_{0,1} \{ V((1 + R)^\delta(1,0) \bar{p}_{1,0}, (1 + R)^\delta(2,0) \bar{p}_{2,0}, \ldots, (1 + R)^\delta(n,0) \bar{p}_{n,n-1}) - V_0 \}, \quad (4.9)
\]

where

\[
\varepsilon = \begin{cases} 1, & i = 0, \\ 0, & \text{otherwise}, \end{cases}
\]

and

\[
\delta(j,k) = \begin{cases} 1, & j = i, \\ -1, & k = i, \\ 0, & \text{otherwise}. \end{cases} \quad (4.10)
\]
**Definition 4.2** We define the $X_0$-curvature representation in $X_0$-currency, and units of $X_0$, by

$$
\sum_{i \neq 0} \min_{R \in (-RW,RW)} \{ F_i^{X_0}(R) - S_i^{X_0} R \} [X_i].
$$

(4.11)

Similarly, in $X_1$-currency, the $X_1$-curvature representation in units of $X_0$ is given by

$$
\sum_{i \neq 1} \min_{R \in (-RW,RW)} \{ F_i^{X_1}(R) - S_i^{X_1} R \} [X_i].
$$

(4.12)

We have already demonstrated that the sensitivities and the Delta representation as functions of the exchange rates depend only on the underlying rates, and now (4.9) shows that curvature representation is also functionally dependent on the underlying rates. We have therefore proved the following.

**Theorem 4.3**

1. The functional form of the curvature representation in any fixed reporting currency, and the curvature capital charge, when converted to the same units, depends only on the underlying rates. The anomaly is therefore due to the aggregation mechanism being noninvariant.

2. There is no value of $\gamma'$ for which the curvature capital charge is invariant under reporting currency change.

3. For uniform risk weights for all currency pairs (with value 0.3) and the regulatory $\gamma' = 0.36$, there exist a portfolio $V$ and two reporting currencies in which the curvature capital charge is approximately 100% higher when compared in the same units. With the same risk weights but with exceptional currency pairs as in Basel Committee on Banking Supervision (2016), and the regulatory $\gamma' = 0.36$, there exist a portfolio $V$ and two reporting currencies in which the curvature capital charge is approximately 300% higher in one reporting currency than the other when compared in the same units.

5 **RESCURING INVARINANCE WITHOUT INCREASING RISK**

We start by giving the standard remedy for noninvariance. The idea is to add symmetry, and we wish to add such symmetry without losing the regulatory framework altogether. In particular, we do not wish to alter the aggregation formula directly. Suppose we have $n$ different reporting currencies (which can, for example, be actual reporting currencies, all currencies or even any numéraire). Take any symmetric function $f(x_1, \ldots, x_n)$, and substitute the different charges for $x_i$. We can define the invariant capital charge function as a constant, equal to the value of $f$, on all reporting currencies. This immediately becomes invariant on the collection of reporting currencies.
we have chosen. We give examples, noting that the mechanism applies to Delta or curvature charges equally. Consider a fixed portfolio $V$ with zero MtM. Recall, as we explained earlier, that there is no loss of generality. Indeed, if a bank has a portfolio $\tilde{V}$ with nonzero MtM, then we write

$$\tilde{V} = V_0 + V,$$

(5.1)

where $V_0$ is the MtM of the portfolio in the reporting currency of the bank. It is the dynamic part, $V$ (also expressed in units of the reporting currency), for which we seek invariance. $V_0$, however, has no capital charge required in either curvature or Delta charges. Indeed, it is purely a constant here and can be ignored. $V$ is the part of the portfolio that can only be nonzero if the FX rates change. While mathematically this is sufficient, we wish to highlight to the reader an intuitive idea here. Any portfolio of zero MtM has zero MtM in any other currency. Suppose, then, that we have a good way of calculating the expected, maximum or quantile loss (or any reasonable risk measure) in some other reporting currency $X$ for this portfolio of a bank that reports in currency $Y$. Suppose the value of the portfolio at the loss scenario is $V$. Then, the value of that loss in the portfolio in currency $Y$ is

$$-V \times p_0[X, Y],$$

(5.2)

where $V$ here is valued in units of $X$. With the same regulatory conservatism, this is bounded by

$$-V \times p_0[X, Y] \times (1 + RW),$$

(5.3)

where $p_0$ represents the base value of the rate. This shows that we can calculate the loss in any reporting currency we choose (no less conservative than the regulations) if we just multiply by $(1 + RW)$ at the very end. In fact, if the bank holds the appropriate capital in the currency in which they calculate the appropriate capital charge, there would be no need for this factor at all.

**Option 1.** We can consider the average of the charges (so all FX books get the same charge, which would be set equal to the average). We have already demonstrated that all the computations are done on the underlying rates directly, and we can compute the different charges in different currencies, convert them to the bank’s currency and average them.

**Option 2.** We can consider the maximum of all charges.

**Option 3.** We can consider the minimum such value over all reporting currencies in our collection.
Option 4. We can consider the minimum over all possible currencies. In this case, the astute reader will note that we will be throwing into the mix a currency that is not necessarily real (or can be a real currency but not necessarily a reporting currency) or can be any numéraire, such as precious metals or any other asset for that matter. We observe, however, that our analysis actually showed that for two currencies not referenced in the underlying rates at all the charges would be equal (all terms in the computations would end up being the same (using equal risk weights here)). Therefore, throwing one such currency into the mix and computing the minimum is the same as throwing in a whole universe of them.

Option 5. An interesting choice is to take a universal currency (say, gold) and compute all the charges with respect to that numéraire to all banks. We like to call this currency Ω or Ω-currency. As we remarked, no rate referencing Ω need be computed at all. This is entirely an algebraic construct to produce different coefficients, and to unify the charge for all reporting currencies.

Option 6. We can use any fixed reporting currency, compute the charges and use their values as our constant values for all reporting currencies. The most natural reporting currency for this option would be US dollars, as we would expect most portfolios to be heavily tilted toward deals on XXX/USD rates.

Our preferred choice, which we view as optimal and justifiable, is the minimum over all reporting currencies. The reason is that the regulatory risk weights were assigned as appropriate for real reporting currencies. Choosing a purely hypothetical currency, while mathematically appealing, deviates from that framework (we could not easily justify the risk weights, for example). So, as long as we consider reporting currencies and take the minimum charge, we are not changing the regulatory framework. This is our optimal proposed solution, though all others would at least achieve the goal of leveling the playing field.

6 PROPOSED SOLUTION

Our final proposal, which we consider economically optimal yet fully regulatory conservative, is as follows.

Option 6 (conservative version). Take the minimum of each (Delta or curvature) charge individually, over all reporting currencies. If a bank reports in the currency for which the charge is minimal, that would be the appropriate charge. If the bank reports in another currency, multiply the charge by $\frac{1}{C_{RW}}$, where RW is the minimum allowed risk weight for the rate between any of the currencies on which the minimum is achieved and the bank’s reporting currency. It is likely that
the US dollar is the most reasonable choice to fit this bill; it also has the additional advantage that the exceptional currency pairs are predominantly of the form XXX/USD. We would therefore suspect that it is optimal (at least for most banks) to compute the FX charges as if they are reporting in US dollars, and then multiply by \((1 + 0.3/\sqrt{2}) \approx 1.21\). That would bring all anomalies to be contained under this (more reasonable) bound (namely 21\%). This is certainly more benign than 300\%. An even better option is to hold the capital for the charge in the currency of the calculation, to avoid the \((1 + \text{RW})\) altogether.

Note that Basel Committee on Banking Supervision (2016, Section A, Paragraph 4, p. 5) gives the impression that the national supervisor may consider the hedging of the required capital against movement in the FX rate to the reporting currency to be exempt from further market risk charges under certain conditions. The context of that paragraph is different from ours, but the principle is related. In the author’s opinion, this situation, if approved, should allow the factor \((1 + \text{RW})\) to be dropped, even if the capital is held in the reporting currency. We believe this is an acceptable situation, but we do not pursue it further, as it is clearly dependent on the national supervisor’s views.

The proposal above relies only on two simple assumptions: the regulatory framework is adequate for all portfolios and adequate for all reporting currencies used in the minimization. These assumptions seem to us to be naturally implied by the regulations. The rest of the proposal simply turns those assumptions into an invariant (or almost invariant) formalism depending on the version of Option 6 used. All the acute anomalies we exposed are no longer present. All the correlations prescribed by the regulators are kept as they are, and the entire aggregation mechanism is preserved.

Finally, for all the required symmetrized functions in any of the options we proposed, we have all the ingredients to compute the Delta and curvature representations using the original form of the valuation as a function only of the underlying rates. Once those coefficients have been computed, the rest can be done on a spreadsheet and no system changes are required.

Before we move on to the next section, we give a simple example of our proposal when applied to a very simple cash portfolio. The technique is the same no matter how complicated the portfolio happens to be. Let us start with a bank reporting in Canadian dollars and having a portfolio of £1. We wish to demonstrate to the reader how this bank can still compute their Delta charge in sterling. This is not a priori obvious, as we certainly do not expect that £1 has FX risk in sterling reporting. The key is the decomposition of the portfolio into static and dynamic parts, as we explained earlier. In our terminology, the CAD-currency representation in units of CAD (which a Canadian-dollar-reporting bank has to compute) is

\[
p_0[\text{GBP, CAD}] + 1 \times (p[\text{GBP, CAD}] - p_0[\text{GBP, CAD}]), \quad (6.1)
\]
where we have decomposed the portfolio into a static part (the first summand) and a dynamic part (the second summand) according to our methodology. We no longer need the static part from this point on. The risk factor in this representation, according to the Basel rules, is \( p[\text{GBP, CAD}] \). The Delta charge in CAD-currency and units of Canadian dollars is now

\[
p_0[\text{GBP, CAD}] \times \text{RW},
\]

where RW must be the risk weight for the GBP/CAD rate. Now, in GBP-currency and units of sterling, our sterling representation of the dynamic part of the portfolio is

\[
p[\text{CAD, GBP}] \times (p[\text{GBP, CAD}] - p_0[\text{GBP, CAD}]) = 1 - p[\text{CAD, GBP}] \times p_0[\text{GBP, CAD}].
\]

This portfolio is now looked at as a portfolio dependent on \( p[\text{CAD, GBP}] \) as a risk factor. The Delta charge in units of GBP is then easily computed as RW, which is the same risk weight for \( p[\text{GBP, CAD}] \). If we convert this to units of Canadian dollars, we get

\[
p_0[\text{GBP, CAD}] \times \text{RW},
\]

which is exactly the same answer as if we compute it in the Canadian-dollar-reporting currency. Note that we do not always expect the answer to be the same, and our proposal is to take the minimum over all representations, but the purpose of the demonstration above is to highlight the fact that this technique does not hide any risk. In particular, we did not lose the risk of the sterling position in Canadian dollar reporting by computing it in sterling reporting currency. The key step is to remove the static part of the portfolio before we proceed. This is also key to the intuitive version we presented in Section 5, demonstrating that the loss of a portfolio of zero MtM can be computed in any currency and then converted appropriately. The fact that the MtM was zero is relevant there.

In summary, the technique works for both curvature and Delta. First, we must express the portfolio in the bank’s reporting currency and remove the static part of the portfolio, leaving a zero-MtM dynamic part. Second, we express the dynamic portfolio in any other reporting currency, \( X \). The risk factors in the latter case need to be expressed as \( p[Y, X] \). Third, we compute the charge of interest in \( X \)-currency representation and units of \( X \). Then, we convert that charge to the bank’s reporting currency using the spot base rate. Finally, we take the minimum of the charge computed over all allowed \( X \). That is the invariant charge for this portfolio. Whether or not a multiplication by \((1 + \text{RW})\) is required, or whether the bank can just keep the capital in the currency that produced the minimum charge to avoid this factor, can be discussed with the national supervisor. Also, in practice, it is likely that it is enough
to compute the optimal \( X \) for a given bank on a regular but infrequent basis, and no daily minimization is required. In fact, it is likely that US dollars is the optimal currency for this purpose anyway, but there can be exceptions to this educated guess. Our goal of regaining a level playing field has now been achieved from within the rules themselves.

7 CONCLUDING REMARKS

We developed a framework in which we were able to fully characterize the invariance of the Delta capital charge for the FX book under a change in reporting currency. We demonstrated that for the current prescribed values of the correlation parameters we do not have invariance, and that, even for a nonexotic FX book, we can have a Delta charge up to 58% higher in one reporting currency than another currency for the same portfolio. We also showed that the curvature capital charge fails invariance by construction for all values of the correlation parameter. We demonstrated that for the current prescribed value of the correlation parameter we can get a curvature charge in one reporting currency that is more than 300% higher than another currency for the same portfolio. We then went on to show how to restore the invariance without dropping the regulatory conservatism. We gave a whole family of options/proposals to correct the anomaly by correcting the aggregation parameters (via symmetrization) in a way that is also taken from the regulatory framework. We also demonstrated that the hope of having a capital charge function that depends on the underlying rates of a portfolio rather than rates referencing the reporting currency has already been realized, but, in a sense, this function lacks symmetrization under a change in reporting currency. During the process of our analysis, we also produced transformation rules that can help the practitioner to compute the Delta and curvature representations in any reporting currency, and similarly for the corresponding charges. We hope that our contribution will add insights into the mathematical framework of the SA and its capital implications. We also hope that it offers to regulators a practical proposal, which we have endeavored to keep in line with the conservatism underlying the regulations. It is rather satisfying to see that it is possible to maintain a level playing field for all banks and yet support a sound and conservative regulatory framework.

DECLARATION OF INTEREST

This work represents the intellectual views of the author and not his employer; any errors are his responsibility. The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.
ACKNOWLEDGEMENTS

The author is grateful to Youngsuk Lee and Lucia Cipolina Kun for useful comments and feedback, and to the anonymous referee for very good suggestions.

REFERENCES


Research Paper

How risk managers should fix tracking error volatility and value-at-risk constraints in asset management

Luca Riccetti

Department of Economics and Law, Università degli Studi di Macerata, Piazza Strambi 1, 62100 Macerata, Italy; email: luca.riccetti@unimc.it

(Received January 18, 2016; revised August 12, 2016; accepted August 12, 2016)

ABSTRACT

Investors usually assign part of their funds to asset managers, who are given the task of beating a benchmark. Asset managers usually face a constraint on maximum tracking error volatility (TEV), which is imposed by the risk management office. In the mean–variance space, Jorion, in his 2003 paper “Portfolio optimization with tracking-error constraints”, shows that this constraint determines an ellipse containing all admissible portfolios. However, many admissible portfolios have problems in mean–variance terms, for example, because of an overly high variance. To overcome this problem, Jorion also fixes a constraint on variance, while, in their 2008 paper “Active portfolio management with benchmarking: adding a value-at-risk constraint”, Alexander and Baptista fix a constraint on value-at-risk (VaR). In this paper, I determine an optimal value for a set of limits composed of the lower limit on TEV, the upper limit on TEV and the upper limit on VaR. To fix the upper limit on VaR, I use the TEV constrained efficient frontier developed in Palomba and Riccetti’s 2013 paper “Asset management with TEV and VAR constraints: the constrained efficient frontiers”, which is the set of portfolios that is on Jorion’s ellipse and not dominated from the mean–variance
perspective. In particular, I develop a strategy to impose on asset managers a set of portfolios that contains as many TEV constrained efficient portfolios and as few inefficient portfolios as possible. Moreover, I show that a limit on maximum VaR is usually better than a limit on maximum variance.

**Keywords:** risk management; active portfolio management; benchmarking; tracking error; value-at-risk (VaR); mean–variance dominance.

## 1 INTRODUCTION

Many investors assign part of their wealth to asset managers (e.g., mutual funds). These investors choose a benchmark but also ask for active management strategies in order to beat it. The active management approach is very different from the passive one, in which the index fund or the exchanged traded fund (ETF) replicates the benchmark’s composition and performance. Instead, active management creates two requirements: (i) keeping the risk of the portfolio (relatively) close to that of the selected benchmark, and (ii) beating the benchmark and maximizing the investors’ utility moving away from the benchmark.

To meet the first requirement, the risk management office usually imposes a maximum value on tracking error volatility (TEV). However, Roll (1992) shows that such portfolios are generally suboptimal, because they do not belong to the mean–variance frontier (MVF; Markowitz (1952)) and are overly risky. Indeed, the so-called mean–TEV frontier (MTF), that is, the set of portfolios with a given expected return and the smallest TEV, is usually far from the efficient frontier. A portfolio’s efficiency loss is defined as the portfolio’s variance minus the variance of the portfolio on the MVF with the same expected return.

Some methods try to mitigate the portfolio efficiency loss by imposing limits on the amount of risk that asset managers can take, using different measures of risk:

---

1 However, investors often face additional sources of risk, such as those arising from labor income and real estate. These risks are usually called “background risk” (see, for example, Gollier 2001). In the presence of background risk, the optimal portfolio is often mean–variance inefficient (Baptista 2008; Flavin and Yamashita 2002). Indeed, the presence of background risk makes individuals less willing to bear other risks (see, for example, Gollier 2001). Therefore, even if TEV optimization is suboptimal for an investor with mean–variance preferences when background risk is absent, TEV optimization can be optimal for the investor when this risk is present. Baptista (2008) provides conditions under which the use of a mean–TEV objective function by the portfolio manager is optimal from the perspective of an investor who faces background risk.

2 Jorion (2002) provides empirical evidence, based on US stock-based funds, that portfolios subject to a TEV limit end up with a variance higher than their benchmarks.
Roll (1992) imposes a limit on beta; Jorion (2003) adds a limit on portfolio variance in a framework with a given TEV ($T$); Alexander and Baptista (2008) impose a limit on value-at-risk (VaR); and Alexander and Baptista (2010) propose a target on ex ante alpha, while still minimizing TEV.

In Palomba and Riccetti (2012), the authors analyze all the interactions among some of the proposed portfolio frontiers. They find an area in the mean–variance space, delimited by the fixed VaR–TEV frontier (FVTF), in which an asset manager can choose portfolios that simultaneously satisfy both TEV and VaR limits and present a trade-off between minimizing relative risk (measured with TEV) and minimizing absolute risk (measured with overall portfolio variance) for any given level of expected return. Moreover, Palomba and Riccetti (2013) draw attention to the mean–variance efficient subset of Jorion’s ellipse (2003), determining the so-called efficient constrained TEV frontier (ECTF). For a more detailed review of these papers, see Section 2.

Concerning the second requirement reported above, Riccetti (2012) proposes applying a lower limit on TEV to force the asset manager to have an active strategy. Indeed, the recent growth of ETFs has increased the importance of beating the benchmark with an active management focused on the maximization of the utility and, thus, on the reduction of the absolute risk with a constraint on the relative risk, as in Jorion (2003). Indeed, investors (especially those who ask for mutual funds with active management strategies) are not interested in minimizing relative risk, but are interested in maximizing their utility, which is a function of portfolio mean, variance and tracking error, as explained by Chow (1995). This author affirms that, due to the uncertainty of absolute returns, investors may decide to compare the performance against a benchmark, but they are “still concerned with the prospect of losing money”; therefore, they “seek portfolios with high return, low standard deviation and low tracking error”. This idea is supported by empirical evidence: most practitioners use both total and relative risk measures.

Therefore, in this paper, I will propose a consistent set of limits that risk management offices should fix in order to help asset managers achieve the two requirements of investors who choose a mutual fund with an active management strategy.

First of all, I will use the lower limit on TEV developed in Riccetti (2012). Beside the lower limit on TEV, I will set a range of optimum values for the upper limit on TEV, using some TEV properties found by Jorion (2003); I will also focus on the

---

3 At the end of the first 2016 quarter, global ETF industry assets reached a new record of over US$3 trillion (www.etf.com). ETFs often take advantage of lower fees, better liquidity and better disclosure compared with mutual funds.
ECTF of Palomba and Riccetti (2013) in order to determine the optimum range for the upper limit on the VaR constraint, with two purposes:

- to reduce the number of efficient constrained TEV portfolios that cannot lie among feasible portfolios if other constraints are used;
- to reduce the amount of feasible portfolios far from the ECTF that an asset manager can choose.

The proposed framework can be applied irrespective of the method used to estimate the required inputs (the expected return vector and variance–covariance matrix), even if, obviously, the method exploited is very relevant for the out-of-sample performance of the allocation. For a good review of the different methods used to estimate expected returns and the variance–covariance matrix, see, for example, DeMiguel et al (2009).

This analysis is significant both for practitioners and for academics. In the former group, it can be suitable for a risk management office and, indirectly, for the allocation chosen by the asset managers. However, in this paper, I only define useful values for some risk management constraints. I do not select a single “optimal” allocation; that is the task of the asset managers, who can freely choose among the portfolios that comply with the constraints. Therefore, as already said, this paper is also indirectly relevant for asset managers, because a good set of constraints can help asset managers to achieve a good performance, but this performance depends on the exact portfolio choice (which is also due to many other factors, such as the goodness of the estimates of the inputs, as explained above). For a numerical practical application, see Section 5.

Instead, this paper contributes to the academic literature because it performs a clear and explicit comparison of variance and VaR limits in order to select which one should be chosen in various contexts. It also adds a range of optimal values for TEV and VaR constraints, linking the limits in order to keep them compatible.

This paper proceeds as follows. Section 2 reviews the approaches of Jorion (2003), Alexander and Baptista (2008) and Palomba and Riccetti (2013). Section 3 explains how lower and upper limits on TEV should be set, while Section 4 explains how a risk management office should set the upper limit on the VaR constraint, given the upper limit on TEV. Section 5 illustrates the results with a simple practical application. Section 6 shows the analytical solutions of the proposed limits. Finally, Section 7 concludes.

2 REVIEW OF SOME PORTFOLIO FRONTIERS

In this section, I review in a nontechnical way (for analytical details, see Sections 6.1, 6.2 and 6.3) some portfolio frontiers that have been proposed in the literature, and which will be useful in this paper. The analysis is conducted in variance–expected
return space \((\sigma^2, \mu)\), but all figures are presented in the usual standard deviation–expected return space \((\sigma, \mu)\). I assume that asset returns have a multivariate normal distribution, and that short sales are allowed. I refer to Palomba and Riccetti (2012) for a deeper analytical analysis of these frontiers.

I analyze the choice of an asset manager who faces an optimization with a constraint on maximum TEV. So, feasible portfolios for the asset manager are inside an elliptical (for relatively low TEV values) area that contains the benchmark, as shown by Jorion (2003). I also call this ellipse the constrained TEV frontier (CTF), as in Palomba and Riccetti (2012).

Moreover, Jorion (2003) shows that, with an upper limit on TEV, in order to avoid overly risky portfolios it is reasonable to choose the portfolio on the CTF at the same level of variance as the benchmark (obviously, the one with the highest expected return, in the upper part of the CTF). This portfolio forces the overall risk (the variance) to be at the same level as the benchmark, and it does not usually have a return that is much smaller than the portfolio on the CTF with the highest expected return, because the top part of the ellipse can be rather flat. Further, Jorion (2003) also characterizes conditions under which this constraint is most useful (for instance, if the benchmark is relatively inefficient).

In Alexander and Baptista (2008), the authors insert a VaR constraint into the framework of Roll (1992). In the \((\sigma, \mu)\) space, the VaR constraint is represented by a line with intercept \(-V\), where \(V\) is the VaR limit, and slope \(z_{\theta}\), where \(z_{\theta}\) is the critical value obtained from the inverse cumulative distribution of a standardized normal with confidence level \(0.5 < \theta < 1\). Portfolios that satisfy the VaR constraint lie on or above the line.

Alexander and Baptista (2008) improve the MTF with the constrained mean–TEV frontier (CMTF): a portfolio is on the CMTF if it satisfies the VaR constraint and there is no other portfolio with the same expected return and a smaller TEV. In Figure 1, CMTF is composed of segment \(M_1R_1\), arc \(R_1R_2\) and segment \(R_2M_2\). Then, CMTF reduces the efficiency loss compared with MTF: for a given expected return, CMTF dominates MTF in mean–variance terms, that is, portfolios on CMTF have a smaller or equal variance (graphically, portfolios on CMTF are at the left of portfolios on MTF or, at most, are the same). Moreover, Alexander and Baptista (2008) provide conditions for the VaR limit, such that the portfolio on the CMTF with a given expected return is also on the CTF determined by a given TEV value.

\(^4\) The area is elliptical if it does not “touch” the mean–variance frontier. Moreover, Jorion (2003) fixes the TEV value and not an upper limit on TEV. Fixing the TEV, Jorion (2003) shows extreme cases in which the required TEV is so high that the benchmark is outside the area delimited by the portfolios with the required TEV. However, this is not the case for the maximum TEV analysis.
Even if the CMTF can select (for a given expected return) the same portfolio determined by Jorion (2003) (for a given level of TEV and standard deviation), the two frontiers are not directly related, and intersections could not exist; indeed, the VaR limit is not related to the benchmark (and even the benchmark can be out of the portfolios that satisfy the VaR limit), and the ellipse could not contain portfolios that satisfy the VaR limit. Palomba and Riccetti (2012) widely discuss all interactions between the cited frontiers. In the current paper, I select the limits on VaR and TEV in a unified framework that is also able to prevent the risk of two incompatible limits.

Palomba and Riccetti (2012) also emphasize that a VaR limit can improve the portfolio allocation because the constrained optimal portfolios are closer to the MVF, but this could have a cost in terms of TEV; indeed, the constrained optimal portfolios can be further away from the MTF. The authors face this issue in the presence of a TEV limit too. Indeed, they introduce the concept of “fixed VaR–TEV frontier” (FVTF): this is composed of the lines that delimit the area of feasible portfolios for the asset manager (which satisfy both VaR and maximum TEV constraints), and that cannot have, for any given level of expected return, other portfolios with both a smaller variance and a smaller TEV. In practice, this area is inside the ellipse, at the left of the CMTF (in Figure 1, among the left arc $K_1K_2$ on the ellipse, segment $K_1R_1$, arc $R_1K_2$ and segment $R_2K_2$). Portfolios inside FVTF face a trade-off between absolute risk (variance) reduction and relative risk (TEV) reduction; indeed, moving toward the left, the variance is reduced and the TEV is increased, and vice versa. A similar issue holds for the limit on variance.

*FIGURE 1* Mean–variance frontier (MVF), mean–TEV frontier (MTF), constrained–TEV frontier (CTF) and constrained mean–TEV frontier (CMTF, in bold), in $(\sigma, \mu)$ space.
How TEV and VaR constraints should be fixed

FIGURE 2  The feasible subsection of the efficient constrained TEV frontier, \( K_2 K_1 \), in \((\mu, \sigma)\) space.

In this case, arc \( J_2 K_2 \) is not admitted.

Palomba and Riccetti (2013) focus on reducing the absolute risk, because, in a classical mean–variance context, it influences the investors’ utility, which is not directly related to TEV. For this reason, the authors introduce the concept of “efficient constrained TEV frontier” (ECTF), defined as the set of portfolios that are on the “constrained TEV frontier” (the ellipse) and are not dominated on the mean–variance criterion.

In the following figures (see, for instance, Figure 2), the ECTF is composed of portfolios on the upper arc \( J_2 J_1 \), where

- \( J_2 \) is the portfolio on the ellipse that has the minimum variance,

- \( J_1 \) is the intersection between the ellipse and the MTF, that is, the portfolio on the ellipse with the highest expected return.

The ECTF definition is unchanged in spite of the slope of the ellipse, and it is also almost unchanged in spite of the level of the upper limit on TEV. Indeed, the ECTF is always the mean–variance efficient part of the set of portfolios feasible after the TEV constraint imposition. Obviously, the ECTF is not an arc on the ellipse in extreme cases, such as the upper limit on TEV being equal to zero (in which case the only admissible portfolio is the benchmark), or the upper limit on TEV being above a threshold \( T_T \), at which the ellipse is tangent to the Markowitz frontier.
3 LOWER AND UPPER LIMIT ON TEV

First of all, I want to fix a range for the TEV, exploiting the analytical results that are derived in Jorion (2003) and Riccetti (2012), and reported in Sections 6.4 and 6.5.

The topic of setting a lower limit on TEV $T_{\text{min}}$ has already been dealt with by Riccetti (2012), who analytically determined the TEV that gives the asset manager the possibility of obtaining an excess return over the benchmark equal to the management fees required from the investors. Indeed, if the asset manager exactly duplicated the benchmark $B$, the investor would obtain a return that would be equal to the benchmark’s return less the fees, with a portfolio $B_{\text{com}}$ that would present the same standard deviation of the benchmark. This could be a relatively inefficient strategy if the investor could replace the passive asset manager with an ETF that would require a lower fee. Therefore, as already explained in the introduction, the large growth of ETFs has increased the importance of beating the benchmark with an active portfolio allocation. Then, the risk management office has to force the asset manager to perform an active strategy to beat the benchmark, imposing a lower limit on TEV.\(^5\)

The choice of the upper limit on TEV $T$ is free to be made by the risk management office, given the characteristics with which the mutual fund wants to comply, in order to attract investors. However, I can give some insights in order to establish a reasonable range for this constraint.

I distinguish two cases, depending on whether the expected return of the benchmark $B$ is above or below the expected return of the global minimum variance portfolio on the Markowitz efficient frontier, which will be called $C$.

In the first case, that is, if the expected return of the benchmark $\mu_B$ is above the expected return of the global minimum variance portfolio $\mu_C$, the upper limit on TEV $T$ should not be above a value $T_M$, found by Jorion (2003), at which portfolio $J_2$ coincides with portfolio $C$. An upper limit on TEV above this value could be dangerous; indeed, Jorion (2003) shows that portfolios with a high TEV are on a line that moves to the right when the TEV increases. In order to reach a very high TEV, the asset manager has to choose a portfolio with a very high overall variance. Obviously, the asset manager is not forced to use all the TEV, but the risk management office has to prevent this possible event. When $T$ increases above $T_M$, the ECTF only enlarges with portfolios that are on the Markowitz frontier with a high variance level; at the same time, a large number of very risky dominated portfolios enter the feasible set of the asset manager. Instead, below $T_M$ a TEV increase makes the ECTF improve, reaching portfolios ever closer to the Markowitz frontier till the TEV value at which the ellipse has a contact point with the Markowitz efficient frontier (analytically found

\(^5\) Riccetti (2012) also faces some peculiar cases, such as very high fees or a benchmark on (or very near) the Markowitz efficient frontier; in the latter case, choosing an ETF that reproduces the benchmark (or a mutual fund with a passive management strategy) is obviously best.
by Jorion (2003)); moreover, between this value and $T_M$, ECTF touches additional portfolios on the efficient frontier both with lower and higher variance compared with the portfolio at the tangency point.

If $T_M$ is a reasonable maximum for the upper limit on TEV, I now want to set a minimum in order to have a range for the upper limit on the TEV constraint. This minimum should be above the lower limit on TEV for an amount large enough to give asset managers a nonempty and not-too-narrow set of feasible portfolios. Therefore,

$$T_{\text{min}} < T \leq T_M. \quad (3.1)$$

In the opposite case, that is, if $\mu_B \leq \mu_C$, the maximum of the range should be enlarged. Indeed, the benchmark could be very inefficient, because it could be far from the efficient part of the Markowitz frontier, that is, the part with a return above $\mu_C$. Then, the asset manager can largely improve the benchmark performance with a very active strategy that requires a lot of “bets”; this moves the managed portfolio allocation away from the benchmark, with a consequent relatively high TEV. Therefore, the further the benchmark is from portfolio $C$, the more the maximum of the range should be enlarged. Subsequently, $T_M$ should be multiplied for a number $\alpha > 1$ that is an increasing function of both $\mu_C - \mu_B$ and $\sigma_B^2 - \sigma_C^2$.

Given that the minimum of the range is unchanged, the range is now

$$T_{\text{min}} < T \leq \alpha T_M. \quad (3.2)$$

In both cases, if the minimum of the range $T_{\text{min}}$ is above a reasonable level for the maximum of the range, this is a signal that the fees (which determine the value of $T_{\text{min}}$) are probably too high.

4 UPPER LIMIT ON VALUE-AT-RISK IN JORION’S FRAMEWORK

Given the upper limit on TEV (which determines the positions of portfolios $J_2$ and $J_1$), I can set an upper limit on VaR. Palomba and Riccetti (2013) insert the VaR limit in the framework of Jorion (2003). Indeed, instead of constraining the standard deviation $\sigma$ to be no more than the standard deviation $\sigma_B$ of a reference portfolio $B$ (for instance, the benchmark’s variance as in Jorion (2003)), the VaR $V$ could be constrained to be no more than the VaR $V_B$ of the reference portfolio $B$. Palomba and Riccetti (2013) determine the subsection of the ECTF that is available for asset managers after the imposition of the limit on standard deviation or VaR. Moreover, they analyze all the possible relationships among these frontiers, showing under which analytical conditions the VaR or standard deviation cut a different subset of efficient constrained TEV portfolios.

Starting from that analysis, I want to determine an optimal range for the VaR limit. This optimal range has to leave as many efficient constrained TEV portfolios in the
set of feasible portfolios for the asset manager as possible and, at the same time, reduce (as much as possible) the number of portfolios that are not efficient. In this discussion, I will also show that the decision to set a limit on the portfolio variance, as in Jorion (2003), is usually less efficient than the decision to set the VaR constraint in the optimal range that I will determine. I will explain all these concepts graphically; I refer to Section 6 for the analytical framework.

First of all, I have to distinguish two cases.

1. The case in which the VaR of portfolio \( J_2 \), \( V_{J_2} \), is lower than the VaR of portfolio \( J_1 \), \( V_{J_1} \), which is usual for the commonly employed confidence level values used for the VaR (above 90%): \( V_{J_2} \leq V_{J_1} \);

2. The opposite and peculiar case in which \( V_{J_1} < V_{J_2} \).

For the analytical values of these VaR limits, see Section 6.6 for the case of \( V_{J_2} \leq V_{J_1} \) and Section 6.7 for the opposite case.

### 4.1 Case \( V_{J_2} \leq V_{J_1} \)

When the VaR of portfolio \( J_2 \) is lower than the VaR of portfolio \( J_1 \),

1. the upper limit on VaR \( V \) has to be set above \( V_{J_2} \) and below \( V_{J_1} \): \( V_{J_2} \leq V \leq V_{J_1} \);

2. the previous limit on VaR is “better” than a limit on variance, that is, it can leave more efficient constrained TEV portfolios in the set of feasible portfolios for the asset manager and, at the same time, reduce the number of inefficient ones.

The reason for setting \( V_{J_2} \) as the minimum value of the range for the upper limit on VaR is that a VaR below this threshold cuts some portfolios near \( J_2 \) from the ECTF (the arc \( JK_2 \) in Figure 2) or cannot intersect the ellipse, leaving the asset manager with an empty set of admissible portfolios. In other words, \( V \leq V_{J_2} \) does not give asset managers the possibility of working and choosing the part of the efficient subsection of the ellipse with the lowest standard deviation (arc \( JK_2 \)).

The maximum value of the range for the upper limit on VaR has to be set below or equal to \( V_{J_1} \). If I fix the VaR limit at the level of portfolio \( J_1 \), that is, the portfolio with maximum expected return of the ellipse, the feasible set of portfolios includes all the efficient constrained portfolios, as shown in Figure 3. Setting a VaR above this value only adds, in the feasible set, portfolios dominated by the portfolios on the ECTF.

Regarding the second issue stated above, it is straightforward to show graphically that the constraint on VaR is “better” than the constraint on portfolio variance. Indeed, the VaR constraint eliminates from the admissible portfolios the inefficient ones in the
area among segment $J_1K_2$, segment $J_1K_3$ and arc $K_2K_3$ in Figure 3. The reason for this improvement is that the VaR constraint is always less inclined than the vertical variance constraint.

However, $V_{J_1}$ can sometimes represent quite a high level of risk compared with the risk of a reference portfolio (usually the benchmark). Then, I can distinguish two cases:

1. the VaR of the benchmark portfolio $V_B$ is above the VaR of portfolio $J_1$, that is, $V_B > V_{J_1}$;
2. the VaR of the benchmark portfolio $V_B$ is below the VaR of portfolio $J_1$, that is, $V_B < V_{J_1}$.

The first case does not present the previously explained problem; the maximum level of admitted VaR is even below the VaR of the reference portfolio.

The second case is represented in Figure 4. It is easy to show that a limit on VaR is, again, better than a limit on variance. Suppose, for instance, that the upper limit on variance or the upper limit on VaR is set at the benchmark (portfolio $B$) level. Admissible portfolios for the asset manager are the following:

- portfolios in the left area, between segment $B_1B_2$ and arc $B_1B_2$, if the risk management office sets an upper limit on TEV and a variance constraint at the benchmark level;
- portfolios in the left area, between segment $K_1K_2$ and arc $K_1K_2$, if the risk management office sets an upper limit on TEV and a VaR constraint at the benchmark level.
A variance constraint cuts off from the feasible portfolios those on arc $B_1J_1$ that are efficient constrained. Instead, a VaR constraint saves efficient constrained portfolios on arc $B_1K_1$, while it cuts off inefficient ones in the area among segment $BK_2$, segment $BB_2$ and arc $K_2B_2$.

This result is obtained with a positive slope of the ellipse, but it could be generalized to the cases of horizontal or negative slope.

**Summary**

The upper limit on VaR $V$ has to be set in the following way: $V_{J_2} \leq V \leq V_{J_1}$. The use of VaR is preferable to the use of variance with regard to constraining the absolute risk, because it can eliminate several portfolios that are surely not optimal (in the low part of the ellipse) from the feasible set, while inserting portfolios that could be of interest in mean–variance terms. Moreover, in order to give asset managers a not-too-narrow feasible set of portfolios, the upper limit on VaR should be set as follows (see Section 6.6.3):

- if $V_B > V_{J_1}$, then the optimal choice for risk management is to couple the upper limit on TEV with an upper limit on VaR set to pass through $J_1$: $V = V_{J_1}$;

- if $V_B \leq V_{J_1}$, then the optimal choice for risk management is to couple the upper limit on TEV with an upper limit on VaR set at the same level as the benchmark: $V = V_B$. 
In this way,

- compared with Alexander and Baptista (2008), I solve the problem of a VaR constraint that is not related to the benchmark and could create empty FVTF (as explained in Palomba and Riccetti (2012));

- compared with Jorion (2003), I improve the set of feasible portfolios, substituting the constraint on variance with the constraint on VaR (reducing or avoiding the removal of efficient constrained portfolios and, at the same time, reducing the number of inefficient portfolios).

4.2 Case $V_{J_1} < V_{J_2}$

This very special case occurs when the slope of the VaR straight line is flatter than the slope of the straight line that passes through points $J_2$ and $J_1$.\(^6\)

However, this is the only situation that can really create a trade-off between the variance constraint and VaR constraint in terms of efficient constrained portfolios. For example, in Figure 5, the VaR line that passes through the benchmark, compared with the constraint on variance, adds to the feasible set all portfolios between $B_1$ and $J_1$; however, it cuts the efficient constrained portfolios with the lowest variance, which are in the arc $JK_2$. Moreover, the VaR constraint cuts off some inefficient portfolios with low variance (under segment $K_2B$ and at the left of segment $BB_2$), but it adds some inefficient portfolios with high variance (above $BK_1$ and at the right of $B_1B$).

It is possible to set the VaR to pass through $J_1$ in order to avoid extremely risky portfolios, but the feasible set could be quite small (between segment $K_2J_1$ and arc $K_2J_1$ in Figure 6); however, the less risky efficient constrained TEV portfolios (arc $J_2K_2$) are excluded from the feasible portfolios.

To summarize, in this case, the choice between VaR and variance constraint, as well as the level of the constraint, is a free choice for the risk management office; however, the choice of VaR implies a riskier set of feasible portfolios in variance terms.

5 AN EMPIRICAL APPLICATION

A short numerical application is reported in order to show the practical simplicity of the method illustrated above.

I use a small portfolio composed of eight of the largest eurozone banking groups: Banco Bilbao Vizcaya Argentaria SA (BBVA), Banco Santander SA, BNP Paribas

\(^6\) This case occurs if the confidence level used to determine VaR is sufficiently low. Moreover, if the confidence level $\theta$ is not high enough, it could happen that the intercept $-\bar{V}$ of the VaR straight line is on the positive side of the $y$-axis. Here, it is not VaR that is a loss by definition. To avoid this possible problem, I define VaR as the “worst expected return” case with probability $(1 - \theta)$. 
FIGURE 5  VaR and variance constraints at benchmark level, in special case 2, in \((\sigma, \mu)\) space.

FIGURE 6  VaR constraint at portfolio \(J_1\) level, in special case 2, in \((\sigma, \mu)\) space.

SA, Crédit Agricole SA, Deutsche Bank AG, ING Group NV, Société Générale SA and UniCredit SpA. The chosen benchmark is the STOXX Europe 600 Banks (which is composed of a subset of forty-eight banking groups included in the STOXX Europe 600 Index). I use daily returns (for simplicity, continuously computed as \(\log(P_t/P_{t-1})\)) of the last two years (2014 and 2015), with 506 observations for each time series. In the considered period, the eurozone banking sector shows no trend, with a flat evolution, as shown by the benchmark; this presents a daily average
How TEV and VaR constraints should be fixed

**TABLE 1** Descriptive statistics (in %) of the assets used in the empirical example.

<table>
<thead>
<tr>
<th>Asset</th>
<th>Mean</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>BBVA</td>
<td>−0.036</td>
<td>−0.005</td>
<td>−6.208</td>
<td>5.547</td>
<td>1.612</td>
</tr>
<tr>
<td>Banco Santander</td>
<td>−0.036</td>
<td>0.044</td>
<td>−15.186</td>
<td>5.763</td>
<td>1.788</td>
</tr>
<tr>
<td>BNP Paribas</td>
<td>−0.011</td>
<td>0.122</td>
<td>−5.037</td>
<td>4.931</td>
<td>1.656</td>
</tr>
<tr>
<td>Crédit Agricole</td>
<td>0.032</td>
<td>0.044</td>
<td>−10.725</td>
<td>7.324</td>
<td>1.925</td>
</tr>
<tr>
<td>Deutsche Bank</td>
<td>−0.072</td>
<td>−0.092</td>
<td>−8.036</td>
<td>4.883</td>
<td>1.757</td>
</tr>
<tr>
<td>ING Group</td>
<td>0.044</td>
<td>0.016</td>
<td>−5.774</td>
<td>6.857</td>
<td>1.800</td>
</tr>
<tr>
<td>Société Générale</td>
<td>0.006</td>
<td>0.006</td>
<td>−6.905</td>
<td>7.599</td>
<td>1.855</td>
</tr>
<tr>
<td>UniCredit</td>
<td>−0.002</td>
<td>0.000</td>
<td>−7.386</td>
<td>6.667</td>
<td>2.193</td>
</tr>
<tr>
<td>Benchmark</td>
<td>−0.016</td>
<td>0.064</td>
<td>−6.014</td>
<td>5.382</td>
<td>1.616</td>
</tr>
</tbody>
</table>

SD denotes standard deviation.

return equal to −0.016% and a median return equal to +0.064%. However, in the same period, many banks show a high return volatility. Table 1 reports more detailed summary statistics.

I also assume that the fees required by the asset manager are equal to a fixed yearly com = 1.50% (daily 0.006%), and that the risk management office computes the VaR with a confidence interval $\theta = 99\%$.

First of all, I compute the minimum variance portfolio $C$ feasible with the eight stocks included in the portfolio. Portfolio $C$ presents a daily return equal to −0.034%, with a standard deviation of 1.454%. Following Riccetti (2012) (in particular, applying the simplified equation (6.5)), I find that the daily TEV should be above 0.004%.

Now, I can compute the range for the upper limit on TEV. I observe that the expected return of the benchmark $\mu_B$ is above the expected return of the global minimum variance portfolio $\mu_C$, therefore the upper limit on TEV should not be above the value $T_M$ found by Jorion (2003), at which portfolio $J_2$ coincides with portfolio $C$. This maximum (equation (6.6)) is equal to 0.497%. Therefore, the upper limit on TEV $T$ should be in the following range: $0.004% < T \leq 0.497%$. As already said, $T$ is a free choice for the risk management office, given the characteristics with which the mutual fund wants to comply. In this case, I decide to set the upper limit on TEV to be in the middle of the previous range, that is, $T = 0.250\%$.

After setting $T$, it is possible to compute the characteristics of the relevant portfolios $J_2$ and $J_1$, reported in Table 2 together with the benchmark. For mean and variance equations, see Table 4; the VaR can be computed with (6.2) or (6.8) and (6.9); the Sharpe ratio is computed assuming a daily risk-free rate equal to zero (therefore, it is simply the expected return divided by the standard deviation); and the information ratio is calculated as the expected return of the portfolio less the expected return of the benchmark, divided by the TEV.
In order to define the range for the upper limit on VaR, I have to check the VaR of portfolios $J_2$ and $J_1$. In this case, $V_{J_2} < V_{J_1}$; therefore, the minimum of the range for the upper limit on VaR is $V_{J_2} = 3.445\%$, and the maximum of the range for the upper limit on VaR is $V_{J_1} = 4.026\%$, that is, $3.445\% < V < 4.026\%$. Inside this range, observing that $V_B < V_{J_1}$, a possible choice for the upper limit on VaR is $V = V_B = 3.775\%$. Graphically, this case is similar to that reported in Figure 4. Moreover, as already explained, in this case an upper limit on VaR “dominates” an upper limit on standard deviation/variance.

Table 3 (that is, the empirical application of Table 5) summarizes the defined daily limits on minimum TEV, maximum TEV and maximum VaR for the analyzed portfolio.

6 ANALYTICAL SOLUTIONS OF THE LIMITS

To analytically describe all the selected limits, as is usual in the literature, some notation has to be provided.

6.1 TEV

First of all, the TEV measures the volatility of the deviations of active portfolio returns from the benchmark portfolio returns. Using the following notation,

- $q_b$ (vector of benchmark asset weights),
- $q_p$ (vector of portfolio asset weights),
- $x = q_p - q_b$ (vector of deviations from benchmark),
- $\Omega$ (variance–covariance matrix of asset returns),

the tracking error variance $T$, also called tracking error volatility (TEV), can be written as

$$T = x' \Omega x. \quad (6.1)$$
6.2 VaR

Alexander and Baptista (2008) insert a VaR constraint, which is the maximum loss at a given confidence level that the portfolio can suffer over a period of time. In the usual mean–variance framework, it is assumed that asset returns have a multivariate normal distribution. Then, the portfolio VaR $V_P$ at the $\theta$ confidence level (with $0.5 < \theta < 1$) is

$$V_P = z_\theta \sigma_P - \mu_P,$$

(6.2)

where $z_\theta = \Phi^{-1}(\theta)$, $\Phi(\cdot)$ is the standard normal cumulative distribution function, $\sigma_P$ is the standard deviation of portfolio $P$ returns and $\mu_P$ is its expected return.\footnote{Some academics, such as Maspero and Saita (2005), and some practitioners define as relative VaR (RVaR) the opposite of what I call VaR.}

Inserting a constraint on maximum VaR $V_P \leq V$ in (6.2), I find that

$$E(\mu_P) \geq -V + z_\theta \sigma_P,$$

(6.3)

which represents, in the $(\sigma, \mu)$ space, a straight line with intercept $-V$ and slope $z_\theta$. Portfolios that satisfy the VaR constraint lie to the left/above the half-plane generated by the line represented by (6.3).

6.3 Efficient constrained TEV frontier (ECTF)

Palomba and Riccetti (2013) define the efficient constrained TEV frontier (ECTF) as the set of portfolios that are on the constrained TEV frontier (that is, on the ellipse if the TEV is not too large) and are not dominated in mean–variance terms. The ECTF is composed of portfolios on the small arc $J_2 J_1$ on the ellipse. To describe the two portfolios $J_2$ and $J_1$, further notation has to be provided: given $n$ assets, the $n$-dimensional column vector $\mu$ contains the expected returns (while the squared $n \times n$ matrix $\Omega$ represents the variance–covariance matrix, as already explained). The following constants are defined: $a = i' \Omega^{-1} i$, $b = i' \Omega^{-1} \mu$, $c = \mu' \Omega^{-1} \mu$ and $d = c - b^2/a$, where $i$ is an $n$-dimensional column vector in which each element is 1.

The minimum variance portfolio of the mean–variance frontier $C$ has expected return $\mu_C = b/a$ and variance $\sigma_C^2 = 1/a$.\footnote{Some academics, such as Maspero and Saita (2005), and some practitioners define as relative VaR (RVaR) the opposite of what I call VaR.}
TABLE 4  Relevant portfolios.

<table>
<thead>
<tr>
<th>Portfolio</th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B$: benchmark</td>
<td>$\mu_B$</td>
<td>$\sigma_B^2$</td>
</tr>
<tr>
<td>$J_2$: ellipse portfolio with minimum variance</td>
<td>$\mu_B - \frac{\Delta_1 \sqrt{T}}{\sqrt{\Delta_2}}$</td>
<td>$\sigma_B^2 + T - 2\sqrt{T\Delta_2}$</td>
</tr>
<tr>
<td>$J_1$: ellipse portfolio with maximum return</td>
<td>$\mu_B + \sqrt{dT}$</td>
<td>$\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d}$</td>
</tr>
</tbody>
</table>

All these values are independent from managers’ strategies because they are derived exclusively from the available data.

The benchmark $B$ can be defined as the reference portfolio with expected return $\mu_B$ and variance $\sigma_B^2$.

Last, I define

\[
\Delta_1 = \mu_B - \mu_C, \\
\Delta_2 = \sigma_B^2 - \sigma_C^2.
\]

$\Delta_1$ determines the slope of the ellipse: if $\Delta_1$ is positive, the slope is positive, and vice versa.

Now I can determine portfolios $J_2$ and $J_1$. $J_2$, which has expected return $\mu_B - (\Delta_1 \sqrt{T} / \sqrt{\Delta_2})$ and variance $\sigma_B^2 + T - 2\sqrt{T\Delta_2}$, is the portfolio of the ellipse that has the minimum variance.

$J_1$, which has expected return $\mu_B + \sqrt{dT}$ and variance $\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d}$, is the intersection between the ellipse and the MTF, that is, the portfolio on the ellipse with the highest return.

Analytical details about these portfolios are in Jorion (2003). Table 4 summarizes the characteristics of the three portfolios that are most relevant in determining the upper limit on VaR.

6.4 Lower limit on TEV

Riccetti (2012) deals with the topic of analytically determining the minimum TEV that must be requested for asset managers to be considered as actively managing. In particular, the paper requires the minimum effort to beat the benchmark covering the fees requested from the investors. In other words, Riccetti (2012) derives the lower limit on TEV that gives the asset manager the possibility of obtaining a gain equal to the requested fees.

Defining “com” as the amount of fees paid by the investors to the asset manager, the equations that a risk management office has to apply to understand if the asset manager has done enough activity to beat the benchmark are the following.
• If com > \(-2\Delta_1\), the lower limit on tracking error variance is

\[
T_{\min} = 2\Delta_2 - \frac{2\Delta_1(\Delta_1 + \text{com})}{d} - \frac{2\Delta_2}{d} \sqrt{\frac{\Delta_4 + 2\text{com}\Delta_1^3 + \text{com}^2\Delta_1^2}{\Delta_2^2} - \frac{2d\Delta_1^2 + 2\text{com}\Delta_1 + \text{com}^2d}{\Delta_2}} + d^2.
\]  

(6.4)

• Otherwise (that is, if com ≤ \(-2\Delta_1\)), the lower limit on tracking error variance is

\[
T_{\min} = \frac{\text{com}^2}{d}.
\]  

(6.5)

Even if (6.5) should only be used with an inefficient benchmark (negative \(\Delta_1\), that is, \(\mu_B < \mu_C\)), as shown by Riccetti (2012), it is much simpler than (6.4). In addition, it always implies a lower limit on TEV smaller than that obtained by (6.4), so it can be set as a lower limit for tracking error variance.

6.5 Upper limit on TEV

6.5.1 Upper limit on TEV: maximum of the range

The upper limit on TEV should not be above a value \(\alpha T_M\), where \(T_M\) is the value at which portfolio \(J_2\) coincides with \(C\) (the global minimum variance portfolio on the Markowitz efficient frontier), found by Jorion (2003). Jorion analytically derives this value, finding

\[
T_M = \Delta_2.
\]  

(6.6)

Further, \(\alpha \geq 1\) should be a decreasing function of \(\Delta_1\) and an increasing function of \(\Delta_2\). In particular, \(\alpha\) should be set at 1 if \(\Delta_1 > 0\), and above 1 when \(\Delta_1 \leq 0\).

6.5.2 Upper limit on TEV: minimum of the range

The minimum of the range for the upper limit on TEV has to be above the lower limit on TEV defined in Section 6.4.

Therefore, the upper limit on TEV should be in the following range:

\[
\frac{\text{com}^2}{d} < T \leq \alpha\Delta_2.
\]  

(6.7)

6.6 Upper limit on VaR: the \(V_{J_2} \leq V_{J_1}\) case

6.6.1 Upper limit on VaR: minimum of the range

To fix a VaR constraint smaller than \(V_{J_2}\) causes the loss of the efficient constrained TEV portfolios with the smallest variance. This is due to a trade-off between VaR
and standard deviation of returns for the portfolios included in arc $J_2K$, where $K$ is the tangency point between the VaR straight line and the CTF (see Palomba and Riccetti (2012) to determine portfolio $K$). Among these portfolios, $J_2$ presents the minimum standard deviation and the maximum VaR, while $K$ has the maximum standard deviation and the minimum VaR; all portfolios inside this arc face a trade-off between VaR and variance reduction. For a discussion on this trade-off, see Palomba and Riccetti (2013) or Alexander (2009). ⁸

Therefore, the minimum of the range for the upper limit on VaR is

$$V_{J_2} = z_\theta (\sigma_B^2 + T - 2\sqrt{T\Delta_2})^{1/2} - \mu_B + \frac{\Delta_1 \sqrt{T}}{\sqrt{\Delta_2}}. \quad (6.8)$$

### 6.6.2 Upper limit on VaR: maximum of the range

In general, the maximum of the range for the upper limit on VaR is $V_{J_1}$:

$$V_{J_1} = z_\theta (\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d})^{1/2} - \mu_B - \sqrt{dT}. \quad (6.9)$$

### 6.6.3 Upper limit on VaR: a possible choice

As explained in Section 4, when selecting the upper limit on VaR, two cases are of interest.

#### The $V_{J_2} \leq V_{J_1} < V_B$ case

In this case, the optimal VaR level is reported in (6.9), that is, the one that passes through the portfolio with the highest return of the ellipse ($J_1$). It makes all efficient constrained TEV portfolios feasible, and, compared with the variance constraint at the $J_1$ level, this constraint always cuts off some inefficient portfolios in the lower part of the ellipse.

$V_{J_1} < V_B$ means that the value of the VaR constraint for the VaR straight line that passes through the benchmark is higher than the value for the VaR straight line that passes through the portfolio with the highest return of the ellipse (portfolio $J_1$). The former is equal to

$$V_B = z_\theta \sigma_B - \mu_B. \quad (6.10)$$

while the latter is in (6.9).

So, the first is higher than the second if

$$\sigma_B > \left(\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d}\right)^{1/2} - \frac{\sqrt{dT}}{z_\theta}. \quad (6.11)$$

⁸ Substituting the concept of mean–variance efficiency with the concept of mean–VaR efficiency (Alexander 2009), this problem is obviously overcome. The limit for setting the minimum acceptable VaR is $V_K$, even if in this extreme case the only feasible portfolio is $K$, which is a too-restricted and practically unreasonable space for asset managers’ choices.
TABLE 5  Optimal limits.

<table>
<thead>
<tr>
<th>Minimum TEV</th>
<th>$T_{\text{min}} = \frac{\sigma^2}{d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper limit on TEV (range)</td>
<td>$\frac{\sigma^2}{d} &lt; T \leq \alpha \Delta_2$</td>
</tr>
<tr>
<td>Upper limit on VaR, $V_{J_2} \leq V_{J_1} &lt; V_B$ case</td>
<td>$V = z_{\theta} (\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d})^{1/2} - \mu_B - \sqrt{dT}$</td>
</tr>
<tr>
<td>Upper limit on VaR, $V_{J_2} \leq V_B \leq V_{J_1}$ case</td>
<td>$V = z_{\theta} \sigma_B - \mu_B$</td>
</tr>
</tbody>
</table>

As explained in Palomba and Riccetti (2013), a necessary and sufficient condition for $V_B > V_{J_1}$ is that the slope of the straight line that passes through $B$ and $J_1$ is negative or vertical, or is positive and higher than $z_{\theta}$ (that is, $(\mu_B - \mu_{J_1})/(\sigma_B - \sigma_{J_1}) > z_{\theta}$).

The $V_{J_2} \leq V_B \leq V_{J_1}$ case

This is contrary to the previous case. In this case, it may be better to set the upper limit on VaR at the $V_B$ level (6.10).

6.7 Upper limit on VaR: the $V_{J_1} < V_{J_2}$ case

As explained in Palomba and Riccetti (2013), a necessary and sufficient condition for $V_{J_1} < V_{J_2}$ is that the slope of the VaR straight line is flatter than the slope of the straight line that passes through points $J_2$ and $J_1$. Since the straight line passing through $J_2$ and $J_1$ has the slope

$$\hat{z} = \frac{\mu_{J_1} - \mu_{J_2}}{\sigma_{J_1} - \sigma_{J_2}}, \quad (6.12)$$

where $\sigma_{J_2}$ and $\sigma_{J_1}$ are the standard errors associated with portfolios $J_2$ and $J_1$, then $V_{J_1} < V_{J_2}$ when $z_{\theta} < \hat{z}$.

With a further decomposition of the equation, I also obtain

$$z_{\theta} < \frac{\sqrt{d} \left( \sqrt{\sigma_B^2 + T + 2\Delta_1 \sqrt{T/d}} + \sqrt{\sigma_B^2 + T - 2\sqrt{T\Delta_2}} \right)}{2\sqrt{\Delta_2}}. \quad (6.13)$$

In this very peculiar case, the risk management office might prefer a limit on variance to a limit on VaR.

6.8 Summary

Table 5 summarizes the proposed constraints.
7 CONCLUSIONS

Asset managers that follow active management strategies usually face an upper limit on the value of TEV; this is imposed by the risk management office to avoid asset managers holding a portfolio that is much riskier than the selected benchmark. In the mean–variance space, Jorion (2003) shows that this constraint determines an ellipse that contains all the admissible portfolios. However, many feasible portfolios are not a good choice for investors in mean–variance terms. For example, many of these portfolios are overly risky. To overcome this problem, Jorion (2003) adds to the TEV constraint a variance constraint, while Alexander and Baptista (2008) add a VaR constraint.

In this context, Palomba and Riccetti (2013) define “efficient constrained TEV frontier” (ECTF) as the set of portfolios that is on the “constrained TEV frontier” (Jorion’s ellipse) and not dominated in mean–variance terms.

In this paper, I choose a strategy that imposes on asset managers a set of (constrained TEV) feasible portfolios, which contains as many efficient constrained portfolios and, at the same time, as few inefficient portfolios as possible. With this aim, the choice of setting a VaR constraint in the Jorion framework is better than the choice of setting a variance constraint; in other words, a good choice for risk management is to fix an upper limit on both TEV and VaR. In particular, the VaR can be chosen in this way:

- if the value of the VaR constraint that passes through the benchmark is lower than that which passes through the portfolio with the highest expected return of the ellipse, then a reasonable choice is to set the VaR at the same level as the benchmark’s VaR;

- if the value of the VaR constraint that passes through the benchmark is higher than (or equal to) that which passes through the portfolio with the highest expected return of the ellipse, then the optimal VaR is set so that the straight line passes through the portfolio with the highest expected return of the ellipse.

Indeed, on the one hand, the variance constraint keeps several inefficient portfolios in the feasible area in the lower part of the ellipse, while the VaR constraint cuts part of them. On the other hand, the VaR constraint can add some efficient constrained portfolios that a variance constraint might cut. Moreover, in this way, I link the VaR constraint to the TEV constraint; thus, I solve the Alexander and Baptista (2008) problem of a VaR not related to the benchmark.

The only peculiar exception, in which this strategy could be a non-optimal way of fixing constraints for a risk-averse risk management office, is the case in which the slope of the VaR straight line is flatter than the slope of the straight line that passes through points \( J_2 \) and \( J_1 \).
The choice of the optimal VaR limit derives from the choice of the upper limit on TEV. Using a result obtained by Jorion (2003) and the lower limit on TEV calculated by Riccetti (2012), I also determine a range in which to set the upper limit on TEV. Moreover, as already said, I recall the lower limit on TEV established by Riccetti (2012) to force asset managers to have an active strategy.

All in all, I determine a consistent set of limits that risk management offices should impose on asset managers, composed of

1. a lower limit on TEV,
2. an upper limit on TEV,
3. an upper limit on VaR.

Further developments of this analysis could be the goal of future research. First, this framework can be extended to methodologies, such as the Black–Litterman model, commonly used in the mutual fund industry. Moreover, the most important advancements are (i) studying the combination of TEV, VaR and variance constraints, with constraints on portfolio weights (see, for example, Bajeux-Besnainou et al. 2011) as short sales prohibition, and (ii) studying more complex utility functions that also consider higher moments (such as skewness and kurtosis) of the returns distribution.

DECLARATION OF INTEREST

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

I am grateful to an anonymous referee for helpful comments and suggestions, to Giulio Palomba and to participants in the “ADEIMF 2014 Summer Conference” (organized at Università Cattolica delSacro Cuore, Milan, September 5–6, 2014) and, in particular, to Paolo Cucurachi.

REFERENCES


