Impact of nonstationarity on estimating and modeling empirical copulas of daily stock returns
Marcel Wollschläger and Rudi Schäfer

Path-consistent wrong-way risk: a structural model approach
Markus Hofer

Decomposition of portfolio risk into independent factors using an inductive causal search algorithm
Brian D. Deaton

Optimal asset management for defined-contribution pension funds with default risk
Shibo Bian, James Cicon and Yi Zhang

A fuzzy data envelopment analysis model for evaluating the efficiency of socially responsible and conventional mutual funds
I. Baeza-Sampere, V. Coll-Serrano, B. M’Zali and P. Méndez-Rodríguez
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 corpsubs@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. BÁLı Georgetown University
JEREMY BERKOWITZ University of Houston
PETER CARR Morgan Stanley
New York University
JON DANIELSSON London School of Economics
ROBERT F. ENGLE New York University
JOHN HULL University of Toronto
ROBERT JARROW Cornell University

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 Incisive Risk Information Limited, Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK. Subscriptions are available on an annual basis, and the rates are set out in the table below.

<table>
<thead>
<tr>
<th></th>
<th>UK</th>
<th>Europe</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>Risk.net Journals</td>
<td>£1945</td>
<td>€2795</td>
<td>$3095</td>
</tr>
<tr>
<td>Print</td>
<td>£890</td>
<td>€1265</td>
<td>$1595</td>
</tr>
<tr>
<td>Risk.net Premium</td>
<td>£2750</td>
<td>€3995</td>
<td>$4400</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: Incisive Media (c/o CDS Global), Tower House, Sovereign Park, Market Harborough, Leicestershire, LE16 9EF, UK. Tel: 0870 787 6822 (UK), +44 (0)1858 438 421 (ROW); fax: +44 (0)1858 434958

US & Canada Office: Incisive Media, 55 Broad Street, 22nd Floor, New York, NY 10004, USA. Tel: +1 646 736 1888; fax: +1 646 390 6612

Asia & Pacific Office: Incisive Media, 20th Floor, Admiralty Centre, Tower 2, 18 Harcourt Road, Admiralty, Hong Kong. Tel: +852 3411 4888; fax: +852 3411 4811

Website: www.risk.net/journal  E-mail: incisivehv@subscription.co.uk

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@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 \TeX version and a PDF file. \TeX 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, Incisive Media,
Haymarket House, 28–29 Haymarket, London SW1Y 4RX, UK
Tel: +44 (0)20 7004 7531; Fax: +44 (0)20 7484 9758
E-mail: journals@incisivemedia.com

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
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 Risk.net/subscribe or email corpsubs@risk.net
CONTENTS

Letter from the Editor-in-Chief vii

RESEARCH PAPERS

Impact of nonstationarity on estimating and modeling empirical copulas of daily stock returns 1
Marcel Wollschläger and Rudi Schäfer

Path-consistent wrong-way risk: a structural model approach 25
Markus Hofer

Decomposition of portfolio risk into independent factors using an inductive causal search algorithm 43
Brian D. Deaton

Optimal asset management for defined-contribution pension funds with default risk 63
Shibo Bian, James Cicon and Yi Zhang

A fuzzy data envelopment analysis model for evaluating the efficiency of socially responsible and conventional mutual funds 77
I. Baeza-Sampere, V. Coll-Serrano, B. M’Zali and P. Méndez-Rodríguez
LETTER FROM THE EDITOR-IN-CHIEF

Farid AitSahlia
Warrington College of Business Administration,
University of Florida

This issue of The Journal of Risk contains papers that focus on modeling dependence via copulas and on identifying independent risk factors for better portfolio risk attribution. It also includes an extension of a classical portfolio selection problem to account for bond default, as well as an approach to compare mutual funds on the basis of hard-to-measure social and environmental criteria.

In the first paper of this issue, “Impact of nonstationarity on estimating and modeling empirical copulas of daily stock returns”, Marcel Wollschläger and Rudi Schäfer conduct an empirical study that systematically distinguishes between local and global behavior of correlated time series. Their findings suggest copulas that best capture asymmetry in the tails when dependence is not stationary, especially during periods of high volatility.

In “Path-consistent wrong-way risk: a structural model approach”, the second paper in the issue, Markus Hofer considers wrong way risk, which is present when the probability of default of a counterparty increases with the value of a credit portfolio. Here, too, the author uses copulas to capture dependence in a way that is useful for credit valuation adjustment, satisfying a path-consistent weight requirement, which in turn provides a framework for model calibration.

In our third paper, “Decomposition of portfolio risk into independent factors using an inductive causal search algorithm”, Brian Deaton applies a search method to identify the set of independent factors affecting the returns of a portfolio. As a consequence, risk attribution and decomposition in portfolio strategies is better conducted.

The issue concludes with two short notes.

In “Optimal asset management for defined-contribution pension funds with default risk”, Shibo Bian, James Cicon and Yi Zhang extend the classical portfolio selection problem involving a bond and a stock in the presence of a risk-free rate. They provide closed-form solutions for non-self-financing strategies in the presence of defaultable bonds. They show, in particular, that investment in the defaultable bond increases with the jump-risk premium and time to maturity.

In the second note, “A fuzzy data envelopment analysis model for evaluating the efficiency of socially responsible and conventional mutual funds”, I. Baeza-Sampere, V. Coll-Serrano, B. M’zali and P. Méndez-Rodríguez show how data envelopment analysis can be used to assess the performance of mutual funds along a dimension that goes beyond standard measures: namely, that of a social and environmental index, for which fuzzy set theory is useful given the difficulty in measuring the index inputs.

www.risk.net/journal

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
Research Paper

Impact of nonstationarity on estimating and modeling empirical copulas of daily stock returns

Marcel Wollschlorger¹,² and Rudi Schäfer¹,³

¹Faculty of Physics, University of Duisburg-Essen, Lotharstrasse 1, 47048 Duisburg, Germany
²Chair for Energy Trading and Finance, University of Duisburg-Essen, Universitätsstrasse 12, 45141 Essen, Germany; email: marcel.wollschaeger@uni-duisburg-essen.de
³Research and Prototyping, Arago GmbH, Eschersheimer Landstraße 526–532, 60433 Frankfurt am Main, Germany; email: rschaefer@arago.de

(Received July 3, 2015; accepted November 30, 2015)

ABSTRACT

All too often, measuring statistical dependencies between financial time series is reduced to the study of a linear correlation coefficient. However, this may not capture all facets of reality. We study empirical dependencies of daily stock returns by their pairwise copulas. Here, we investigate in particular the extent to which the nonstationarity of financial time series affects both the estimation and the modeling of empirical copulas. We estimate empirical copulas from the nonstationary, original return time series and stationary, locally normalized ones, and we are thereby able to explore the empirical dependence structure on two different scales: globally and locally. Additionally, the asymmetry of the empirical copulas is emphasized as a fundamental characteristic. We compare our empirical findings with a single Gaussian copula, with a correlation-weighted average of Gaussian copulas, with the \( K \)-copula directly addressing the nonstationarity of dependencies as a model parameter and with the skewed Student \( t \)-copula. The \( K \)-copula covers the empirical dependence...
structure on the local scale most adequately, whereas the skewed Student $t$-copula best captures the asymmetry of the empirical copula found on the global scale.

Keywords: copulas; financial time series; nonstationarity; asymmetry; multivariate mixture; $K$-copula.

1 INTRODUCTION

The study of empirical dependencies of financial time series is not only important for risk management and portfolio optimization, but a prerequisite for a deeper understanding. It is common practice to reduce the question of statistical dependence to the study of the Pearson correlation coefficient; the most notable works in this direction include Markowitz (1952), Martens and Poon (2001), Pelletier (2006) and Engle and Colacito (2006). The correlation coefficient, however, only measures the degree of linear dependence between two random variables. Nonlinear dependencies or more complicated dependence structures are not captured appropriately. Here, we choose a copula approach to investigate the full range of statistical dependencies. In contrast with the multivariate distribution, which also contains the marginal distributions, the copula introduced by Sklar (1973) transforms these marginal distributions to uniform distributions. This allows for statistical dependencies to be studied directly. Copulas find application in many fields, eg, civil engineering (Kilgore and Thompson 2011), medicine (Onken et al 2009), climate and weather research (Schölzel and Friederichs 2008) and the generation of random vectors (Strelen 2009). In finance, copulas are primarily used in risk and portfolio management (see, for example, Brigo et al 2010; Embrechts et al 2003, 2002; Low et al 2013; McNeil et al 2005; Meucci 2011; Rosenberg and Schuermann 2006). A compact survey of the growing literature of copula models by Patton (2012a,b) provides more detail. For a more general introduction to copulas, the reader is referred to Joe (1997) and Nelsen (2006).

In all these contexts, easily tractable analytical copulas are used as a building block in the multivariate distributions. Empirical studies that do not assume an analytical dependence structure a priori are few and far between (see, for example, Münnix et al 2012; Ning et al 2008). This is where our contribution fits in. We study empirical dependencies of daily returns of S&P 500 stocks. To achieve a good statistical significance for the estimation of a complete copula, a very large amount of data is necessary. Therefore, we restrict ourselves to the bivariate case and average over many pairwise copulas. In addition, we consider a rather long time horizon. In considering a long time horizon we are confronted with two problems: the nonstationarity of the individual time series and their dependencies. The former (ie, time-varying trends and volatilities) must be taken into account when we estimate empirical copulas. To this end we apply the local normalization to the empirical return time series, which rids...
them of nonstationary trends and volatilities. The local normalization was introduced by Schäfer and Guhr (2010) and yields stationary, standard normally distributed time series while preserving dependencies between different time series. In particular, it allows us to study the empirical dependence structure on a local scale, whereas the copula of original returns reveals the dependence structure on a global scale.

We compare our empirical findings with different analytical copulas, addressing in particular the question of how far the assumption of a Gaussian dependence structure may carry. Since we average over many stock pairs, it would be naive to assume that the resulting average pairwise copula could be described by a single Gaussian copula where only the average correlation is included. And, indeed, we find rather poor agreement, especially on the global scale. A correlation-weighted average of Gaussian copulas, which takes into account the different pairwise correlations, provides only a slightly better description. We have to consider the time-varying dependencies as well. These can be addressed by an ensemble average over random correlations (see Schmitt et al 2013). This ansatz yields good agreement for multivariate return distributions. Here, we derive the pairwise $K$-copula resulting from this random matrix approach and compare it with our empirical findings. Overall, we find rather good agreement, especially on the local scale. However, the empirical copulas show an asymmetry in the tail dependence, which our model cannot account for. As a model that allows for an asymmetry in the dependence structure, the skewed Student $t$-distribution, first introduced by Hansen (1994), and its corresponding copula, introduced by Demarta and McNeil (2005), are steadily gaining popularity, particularly with regard to credit risk management (see, for example, Dokov et al 2008; Hu and Kercheval 2006) and the study of asymmetric dependencies (see, for example, Sun et al 2008). On the global scale, where the asymmetry is more pronounced, the skewed Student $t$-copula yields improved agreement with our empirical copulas.

The paper is structured as follows. In Section 2, we briefly present our data set and the theoretical background for the local normalization, correlations and copulas. Additionally, we derive the $K$-copula. We present our results in Section 3 and conclude in Section 4.

2 DATA SET AND THEORETICAL BACKGROUND

2.1 Prices and returns

In our empirical study, we consider the daily closing prices of stocks in the S&P 500 stock index, which have been adjusted for splits and dividends. Our observation period ranges from March 1, 2006 to December 31, 2012. The data is obtained from http://finance.yahoo.com/. From the price time series we calculate the returns (later referred
to as “original” returns) as

$$r_k(t) = \frac{S_k(t + \Delta t) - S_k(t)}{S_k(t)},$$

(2.1)

where $S_k(t)$ is the price at time $t$ of a stock $k$ and $\Delta t = 1$ trading day. We finally arrive at $T = 1760$ daily returns for $K = 460$ stocks that were continuously traded and listed in the S&P 500 during the entire observation period. It is well known that original returns show strongly nonstationary behavior.

### 2.2 Local normalization

To correct for nonstationary trends and volatilities in the original returns, we employ a method called local normalization, introduced by Schäfer and Guhr (2010). Here, the locally normalized return is defined as

$$\rho_k(t) = \frac{r_k(t) - \mu_k(t)}{\sigma_k(t)},$$

(2.2)

with local average $\mu_k(t)$ and local volatility $\sigma_k(t)$ of a stock $k$, which are both estimated over a time window of the preceding thirteen trading days. As has been shown in Schäfer and Guhr (2010), this interval of thirteen trading days for the local average and volatility yields the best approximation of stationary, standard normally distributed time series.

### 2.3 Correlations

To measure the statistical dependence of two random variables $X, Y$, the Pearson correlation coefficient,

$$C_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y},$$

(2.3)

is often considered, where $\text{Cov}(X, Y)$ is the covariance of both random variables and $\sigma_X, \sigma_Y$ are their respective standard deviations. In our case, the random variables $X, Y$ are returns $r_k, r_l$. Nonstationarity of the individual time series is a fundamental problem for the estimation of correlation coefficients. For return time series, the time-varying volatilities lead to estimation errors in (2.3) (see, for example, Münnix et al 2012; Schäfer and Guhr 2010). This can be taken into account by the local normalization introduced above. As already mentioned in Section 1, the correlation coefficient only measures the linear dependence between two random variables. Therefore, it is not well suited to measuring arbitrary dependencies. For the latter purpose we consider copulas in the following section.
2.4 Definition of copulas

In our short introduction to copulas, we shall confine ourselves to the bivariate case, since we only study empirical pairwise copulas. All the information about the statistical dependence of two random variables is certainly contained in their bivariate distribution. However, it is difficult to compare bivariate distributions if the marginal distributions are not the same for all random variables. To achieve comparability we can transform each marginal distribution into a uniform distribution. The bivariate distribution of such transformed variables is then called a copula.

We consider two continuous random variables $X, Y$ with a joint probability density function $f_{X,Y}(x, y)$. Then, the joint cumulative distribution function is given by

$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(x', y') \, dx' \, dy'.$$  \hfill (2.4)

We denote the marginal distribution functions by $F_X(x)$, $F_Y(y)$, and the corresponding probability density functions by

$$f_X(x) = \frac{d}{dx} F_X(x), \quad f_Y(y) = \frac{d}{dy} F_Y(y).$$

According to Sklar (1973), we define the copula by

$$F_{X,Y}(x, y) = \text{Cop}_{X,Y}(F_X(x), F_Y(y)) \quad \iff \quad \text{Cop}_{X,Y}(u, v) = F_{X,Y}(F_X^{-1}(u), F_Y^{-1}(v)), \quad (2.5)$$

where we used the transformations $u = F_X(x)$, $v = F_Y(y)$, and $F_X^{-1}, F_Y^{-1}$ indicate the inverse marginal distribution functions, also called quantile functions. According to the usual definition of a probability density function, the copula density is defined by the partial derivatives with respect to $u$ and $v$:

$$\text{cop}_{X,Y}(u, v) = \frac{\partial^2 \text{Cop}_{X,Y}(u, v)}{\partial u \partial v} \frac{\partial F_Y^{-1}(v)}{\partial v}$$

$$= \frac{\partial^2 F_{X,Y}(x, y)}{\partial x \partial y} \bigg|_{x=F_X^{-1}(u), \ y=F_Y^{-1}(v)} \frac{\partial F_X^{-1}(u)}{\partial u}$$

$$= \frac{f_{X,Y}(x, y)}{f_X(x) f_Y(y)} \bigg|_{x=F_X^{-1}(u), \ y=F_Y^{-1}(v)} \quad (2.6)$$

where we used the inverse function rule in the last step.

2.5 Empirical pairwise copula

In order to calculate the empirical pairwise copula of two return time series $r_k, r_l$, we have to transform their marginal distributions to uniformity. We then refer to the
transformed time series as \( u = u_k, v = u_l \). To achieve uniform distributed time series, we utilize the empirical distribution function

\[
    u_k(t) = F_k(r_k(t)) = \frac{1}{T} \sum_{\tau=1}^{T} \mathbf{1}\{r_k(\tau) \leq r_k(t)\} - \frac{1}{2T}, \tag{2.7}
\]

where \( \mathbf{1} \) is the indicator function. The term \(-\frac{1}{2}\) in definition (2.7) simply ensures that all values \( u_k(t) \) lie strictly in the interval \((0, 1)\). Finally, we arrive at the empirical pairwise copula density as a two-dimensional histogram of data pairs \((u_k(t), u_l(t))\).

### 2.6 Bivariate Gaussian copula

The Gaussian copula is the dependence structure that arises from a normal distribution. We consider two random variables \( X, Y \), which follow a joint normal distribution with correlation coefficient \( c \). The bivariate Gaussian copula is then given by

\[
    \text{Cop}_c^G(u, v) = \Phi_c(\Phi^{-1}(u), \Phi^{-1}(v)), \tag{2.8}
\]

where \( \Phi_c \) describes the cumulative distribution function of a bivariate standard normal distribution, and \( \Phi^{-1} \) describes the inverse cumulative distribution function of a univariate standard normal distribution:

\[
    \Phi_c(x, y) = \int_{-\infty}^{x} dx' \int_{-\infty}^{y} dy' \frac{1}{2\pi \sqrt{1-c^2}} \exp \left( \frac{-(x')^2 - 2cx'y' + (y')^2}{2(1-c^2)} \right), \tag{2.9}
\]

\[
    \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{(x')^2}{2} \right) dx'. \tag{2.10}
\]

Partial differentiation yields the bivariate Gaussian copula density,

\[
    \text{cop}_c^G(u, v) = \frac{1}{\sqrt{1-c^2}} \exp \left( \frac{-c^2\Phi^{-1}(u)^2 - 2c \Phi^{-1}(u)\Phi^{-1}(v) + c^2\Phi^{-1}(v)^2}{2(1-c^2)} \right). \tag{2.11}
\]

In Figure 1, we show as an example two Gaussian copula densities \( \text{cop}_c^G(u, v) \) for a positive and a negative correlation \( c \), respectively. Here, we can distinguish the inherently symmetric character as a fundamental property. Gaussian copula densities with positive correlations describe strong dependencies between events in the same tail in each marginal distribution, while negative correlations describe strong dependencies between events in the opposite tails.
FIGURE 1 Gaussian copula density \( \text{cop}_c^G(u, v) \) for different correlations \( c \): (a) \( c = 0.8 \) and (b) \( c = -0.5 \).

2.7 Bivariate \( K \)-copula

Correlations between financial time series are time dependent (see, for example, Cappiello et al 2006; Engle 2002; Münnix et al 2012; Schäfer and Guhr 2010; Tse and Tsui 2002). To take this empirical fact into account, we consider a random matrix model, introduced by Schmitt et al (2013), which assumes, for each return vector \( \mathbf{r}(t) = (r_1(t), \ldots, r_K(t)) \) at each time \( t = 1, \ldots, T \), a multivariate normal distribution

\[
g(\mathbf{r}(t), \Sigma_t) = \frac{1}{\sqrt{\det(2\pi \Sigma_t)}} \exp \left( -\frac{\mathbf{r}(t)^\top \Sigma_t^{-1} \mathbf{r}(t)}{2} \right)
\]

with a time-dependent covariance matrix \( \Sigma_t \). Each \( \Sigma_t \) is then modeled by a random matrix \( \Sigma_t \rightarrow AA^\top/N \) drawn from a Wishart distribution (see Wishart 1928). To this end we choose a multivariate normal distribution,

\[
w(A, \Sigma) = \frac{1}{\det^{N/2}(2\pi \Sigma)} \exp \left( -\frac{\text{tr} A^\top \Sigma^{-1} A}{2} \right),
\]

for the matrix elements of \( A \), which fluctuate around the empirical average covariance matrix \( \Sigma = (\Sigma_t)_{t=1,\ldots,T} \). The distribution of a sample of return vectors \( \mathbf{r} \) is then given by a multivariate normal–Wishart mixture: we are averaging a multivariate normal distribution over an ensemble of Wishart-distributed covariances, which finally results...
in a $K$-distribution,

$$\langle g \rangle (r, \Sigma, N) = \int d[A] g \left( r, \frac{1}{N} A A^\dagger \right) w(A, \Sigma)$$

$$= \frac{1}{2^{N/2+1} \Gamma(N/2) \sqrt{\det(2\pi \Sigma/N)}} \frac{\mathcal{K}_{(K-N)/2}(\sqrt{N} r^\dagger \Sigma^{-1} r)}{\sqrt{N} r^\dagger \Sigma^{-1} r^{(K-N)/2}}$$

$$= \frac{1}{(2\pi)^K \Gamma(N/2) \sqrt{\det \Sigma}}$$

$$\times \int_0^\infty dz z^{(N/2)-1} e^{-z} \sqrt{\frac{\pi N}{z}} \exp \left( -\frac{N}{4z} r^\dagger \Sigma^{-1} r \right), \quad (2.14)$$

where $\int d[A]$ denotes the integral over all matrix elements of $A$, and $\mathcal{K}_\lambda$ is the modified Bessel function of the second kind, of order $\lambda$. The $K$-distribution (2.14) contains only two parameters: the empirical average covariance matrix $\Sigma$, and a free parameter $N$, which characterizes the fluctuations of covariances around the empirical average $\Sigma$. In this manner, the empirically observed nonstationarity of covariances enters directly into the random matrix model. Note that the $K$-distribution (2.14) is a special case of the multivariate generalized hyperbolic distribution (see, for example, Aas and Haff 2006; Browne and McNicholas 2014; McNeil et al 2005; Necula 2009; Schmidt et al 2006; Socgnia and Wilcox 2014; Vilca et al 2014).

By considering the $K$-distribution (2.14) for the bivariate case, we can now derive the bivariate $K$-copula density via (2.6). In our case, the probability density function is the bivariate $K$-distribution, $f_{X,Y}(x, y) = \langle g \rangle (r, \Sigma, N)$ with $K = 2$. Since the copula density is independent of the marginal distributions, we can choose the standard deviations $\sigma_X = \sigma_Y = 1$, leading to the covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_X^2 & \sigma_X \sigma_Y c \\ \sigma_X \sigma_Y c & \sigma_Y^2 \end{pmatrix} = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}, \quad (2.15)$$

which is now merely a correlation matrix with the empirical average correlation coefficient $c$. Thus, the parameter $N$ simply characterizes the fluctuation strength of the correlations around their mean value. The smaller the value of $N$, the larger the fluctuations. In the limit $N \to \infty$ the fluctuations vanish and we arrive at a normal distribution and a Gaussian copula, respectively. The marginal probability density functions of the $K$-distribution (2.14) are identical ($f_X(\cdot) = f_Y(\cdot)$) with

$$f_X(x) = \int_{-\infty}^\infty dy f_{X,Y}(x, y)$$

$$= \frac{1}{\Gamma(N/2)} \int_0^\infty dz z^{(N/2)-1} e^{-z} \sqrt{\frac{N}{4\pi z}} \exp \left( -\frac{N}{4z} x^2 \right), \quad (2.16)$$
FIGURE 2  $K$-copula densities $c_{K,c,N}(u, v)$ for average correlation $c$ and parameter $N$: (a) $c = 0$, $N = 5$; (b) $c = 0.2$, $N = 4$.

and the marginal cumulative distribution functions, $F_X(\cdot) = F_Y(\cdot)$, with

$$F_X(x) = \int_{-\infty}^{x} d\xi f_X(\xi)$$

$$= \frac{1}{\Gamma(N/2)} \int_{0}^{\infty} dz z^{(N/2)-1}e^{-z} \int_{-\infty}^{x} d\xi \sqrt{\frac{N}{4\pi z}} \exp\left(-\frac{N}{4z} z^2 \right).$$  \hspace{1cm} (2.17)

have to be calculated and inverted numerically. Insertion into (2.6) then yields the $K$-copula density $c_{K,c,N}(u, v)$. For illustration, Figure 2 shows the $K$-copula density for an average correlation $c = 0$ and $N = 5$ on the one hand, and for $c = 0.2$ and $N = 4$ on the other. We observe that the $K$-copula density contains both positive and negative correlations, in contrast to the Gaussian copula density. In particular, negative correlations are also covered if the average correlation $c$ is positive but the fluctuations around it are large enough. However, the $K$-copula density is still symmetric with respect to the facing corners.

2.8 Bivariate skewed Student $t$-copula

We consider a bivariate skewed Student $t$-distributed random variable $Z = (X, Y)$ represented by the normal mean–variance mixture

$$Z = \mu + \gamma W + Q\sqrt{W},$$  \hspace{1cm} (2.18)

where $\mu = (\mu_1, \mu_2)$ is a location parameter vector; $W \sim IG(v/2, v/2)$ follows an inverse gamma distribution with $v$ degrees of freedom; $Q \sim N(0, \Sigma)$ is independent
of \( W \) and drawn from a bivariate normal distribution with zero mean and covariance matrix \( \Sigma \); and \( \mathbf{y} = (\gamma_1, \gamma_2) \) is the skewness parameter vector. Since we are aiming at deriving a skewed Student \( t \)-copula density from its corresponding distribution following (2.6), we are able to drop the location parameter vector \( \mathbf{\mu} \) in the stochastic representation (2.18), ie, \( \mathbf{\mu} = 0 \), and set the standard deviations to 1 such that \( \Sigma \) conforms to the correlation matrix (2.15) again. In this case, the bivariate skewed Student \( t \) probability density function reads

\[
f_Z(z) = \frac{1}{2^{v/2} \Gamma(v/2) \pi \sqrt{\det \Sigma}} \exp(z^\top \Sigma^{-1} \mathbf{y}) \times \frac{\mathcal{K}_{v/2+1}(\sqrt{(v + z^\top \Sigma^{-1} z) \mathbf{y}^\top \Sigma^{-1} \mathbf{y}})}{\left(\sqrt{(v + z^\top \Sigma^{-1} z) \mathbf{y}^\top \Sigma^{-1} \mathbf{y}}\right)^{(v/2+1)}},
\]

where \( \mathcal{K}_\lambda \) is the modified Bessel function of the second kind, of order \( \lambda \). The univariate marginal probability density functions are identical \( (f_X(\cdot) = f_Y(\cdot)) \) with

\[
f_X(x) = \frac{1}{2^{(1-v)/2} \Gamma(v/2) \sqrt{\pi} \sqrt{\det \Sigma}} \exp(x \gamma_1) \times \frac{\mathcal{K}_{(v+1)/2}(\sqrt{(v + x^2) \gamma_1^2})}{\left(\sqrt{(v + x^2) \gamma_1^2}\right)^{(v+1)/2}},
\]

and the marginal cumulative distribution functions, \( F_X(\cdot) = F_Y(\cdot) \), with

\[
F_X(x) = \int_{-\infty}^{x} d\xi f_X(\xi),
\]

have to be calculated and inverted numerically. Bringing it all together corresponding to (2.6) with \( f_{X,Y}(x, y) = f_Z(z) \) then yields the bivariate skewed Student \( t \)-copula density: \( \text{cop}^t_{c,v,\mathbf{y}}(u, v) \). In the limit \( v \to \infty \) we obtain the Gaussian copula density.

### 3 EMPIRICAL RESULTS AND MODEL COMPARISON

In Section 3.1 we present the empirical copula densities for original returns and locally normalized returns, respectively. In Section 3.2 we discuss the asymmetry in the tails of the empirical copula densities in more detail. We then compare the empirical copulas to the Gaussian copula with mean correlation in Section 3.3, to a correlation-weighted Gaussian copula in Section 3.4, to the \( K \)-copula in Section 3.5 and to the skewed Student \( t \)-copula in Section 3.6.
3.1 Empirical copula densities

We consider empirical pairwise copula densities averaged over all $K(K - 1)/2$ stock pairs

$$
\text{cop}(u, v) = \frac{2}{K(K - 1)} \sum_{k,l=1,l>k}^K \text{cop}_{k,l}(u, v),
$$

(3.1)

where $\text{cop}_{k,l}(u, v)$ is calculated as a two-dimensional histogram of the data pairs $(u_k(t), u_l(t))$, according to (2.7). For the bin size of these histograms we choose $\Delta u = \Delta v = 1/20$. In the following, we denote the empirical copula density by $\text{cop}^{(\text{glob})}(u, v)$ for original returns, and by $\text{cop}^{(\text{loc})}(u, v)$ for locally normalized returns.

As mentioned above, the main difference between these two cases is the consideration of nonstationarity: the original returns exhibit time-varying trends and volatilities; thus, we might argue the return time series are nonstationary. The locally normalized returns, on the other hand, show stationary behavior. We compare the results for both cases because each has its merits. The copula for the original returns provides the statistical dependence over the full time horizon, ie, on a global scale, while the copula for the locally normalized returns reveals the statistical dependence on a local scale.

In Figure 3, we compare the empirical copula densities. At first sight, the dependence structure seems very similar for both cases. Deviations exist mainly in the corners. Overall, the statistical dependence is preserved rather well when we strip the time series from time-varying trends and volatilities. For the local normalization,
Schäfer and Guhr (2010) showed that correlations are preserved under this procedure. Apparently, this also holds for the copula to some degree. For the original returns, the two peaks in the corners are higher. This can be explained as follows: the returns of periods with high volatility are more likely to end up in the lowest or highest quantile, i.e., \( u_k(t) \) and \( u_l(t) \) are close to 0 or 1. Therefore, periods with high volatility contribute strongly to the corners of this copula. Since high volatility typically coincides with strong correlations in the market, the corners exhibit a stronger dependence.

Qualitatively, the empirical results resemble Gaussian copula densities except for the corners, where an asymmetry is observed. This empirical asymmetry implies that large negative returns of two stocks show stronger dependence than large positive returns. Although the asymmetry cannot be captured either by the Gaussian copula, a Gaussian mixture or by the \( K \)-copula, we shall investigate how well these analytical copulas can approximate the overall dependence structure of empirical data. In addition, we shall compare our empirical copulas to the skewed Student \( t \)-copula, which is able to model asymmetries in the dependence structure. First, however, we take a closer look at the empirical asymmetry of the tail dependence itself.

### 3.2 Asymmetry of the tail dependence

We now elaborate on the features of the empirical copula densities. Their essential characteristic is the asymmetry between the lower–lower and upper–upper corners: the dependence between large negative returns is always stronger than that between large positive returns. Consequently, we observe a stronger dependence between coinciding downside movements than between coinciding upside movements. In other words, we distinguish an asymmetry between the bearish and bullish markets. This asymmetry is particularly evident for the original returns. Ang and Chen (2002), studying the dependence between US stocks and the US market, observed the same asymmetry in correlation.

To quantify this empirical asymmetry, we estimate the tail dependence by integrating over the \( 0.2 \times 0.2 \) area in all four corners for each of the \( K(K-1)/2 \) empirical pairwise copula densities. We obtain the tail-dependence asymmetries by subtracting the integrated areas in the opposite corners of the empirical copula densities:

\[
p_{k,l} = \int_{0.2}^{1} du \int_{0.2}^{1} dv \, \text{cop}_{k,l}(u,v) - \int_{0}^{0.2} du \int_{0}^{0.2} dv \, \text{cop}_{k,l}(u,v),
\]

\[
q_{k,l} = \int_{0}^{0.2} du \int_{0.2}^{1} dv \, \text{cop}_{k,l}(u,v) - \int_{0}^{1} du \int_{0}^{0.2} dv \, \text{cop}_{k,l}(u,v).
\]

We call \( p_{k,l} \) the positive tail-dependence asymmetry, and \( q_{k,l} \) the negative tail-dependence asymmetry. The histograms of these tail-dependence asymmetries,
f(p_{k,l}) and f(q_{k,l}), respectively, are shown in Figure 4 for both data sets. While the negative tail-dependence asymmetries \( q_{k,l} \) are centered around zero for both original and locally normalized returns, we observe a distinct negative offset for the positive tail-dependence asymmetries \( p_{k,l} \). Hence, on average there is no asymmetry in the negative tail dependence, i.e., between coinciding large positive and large negative movements, but there is an asymmetry in the positive tail dependence, i.e., between coinciding large negative movements and coinciding large positive movements. We notice additionally that the locally normalized returns show a much weaker asymmetry in the positive tail dependence than the original returns. This empirical finding can be interpreted as follows: for the original return time series, events in the tails of the distribution reflect periods of high volatility. And since high volatility often occurs simultaneously in most stocks, the tail dependence also reflects these periods in particular. In contrast, the locally normalized returns are stationary with a constant volatility of 1. Therefore, the tail dependence in this case reflects the average behavior over the whole time period. Thus, our empirical results imply that the asymmetry in the tail dependence is not stationary, but it is particularly strong in periods with high

---

**FIGURE 4** Histograms of the tail-dependence asymmetries.

(a), (c) Asymmetry for the positive tail dependence, \( p_{k,l} \). (b), (d) Asymmetry for the negative tail dependence, \( q_{k,l} \). (a), (b) Original returns. (c), (d) Locally normalized returns.
volatility. In fact, periods with high volatility coincide both with strong correlations and strong asymmetry in the dependence structure.

3.3 Comparison with the Gaussian copula

At first sight, our empirical copula densities seem to roughly resemble the Gaussian copula. How good is this agreement quantitatively? In Figure 5, we compare the empirical copula densities with the respective analytical Gaussian copula density $\text{cop}_G(u, v)$. Here, we set the correlation coefficient in each Gaussian copula to the empirical average correlations: $\overline{\varepsilon^{(\text{glob})}} = 0.44$ for original returns and $\overline{\varepsilon^{(\text{loc})}} = 0.39$ for locally normalized returns. In Figure 5, we observe clear deviations from the respective Gaussian copula densities. This is to be expected since the empirical copula
TABLE 1  Least mean squares between analytical and empirical copula densities.

<table>
<thead>
<tr>
<th>Analytical density</th>
<th>Original returns</th>
<th>Locally normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gaussian copula</td>
<td>27.52</td>
<td>5.26</td>
</tr>
<tr>
<td>Correlation-weighted Gaussian copula</td>
<td>26.42</td>
<td>5.11</td>
</tr>
<tr>
<td>K-copula</td>
<td>6.26</td>
<td>2.27</td>
</tr>
<tr>
<td>Skewed Student t-copula</td>
<td>0.65</td>
<td>2.87</td>
</tr>
</tbody>
</table>

FIGURE 6  Histograms of correlations for (a) original returns, $f(C_{k,l}^{(\text{glob})})$, and (b) locally normalized returns, $f(C_{k,l}^{(\text{loc})})$.

densities are in fact an average over $K(K-1)/2$ pairwise copula densities with different correlations. Hence, a comparison with a single Gaussian copula with average correlation cannot yield suitable results. Indeed, for the original returns we find considerable deviations not only in the corners but over the entire dependence structure. For the locally normalized returns, the corners are rather well described. Nonetheless, deviations over the whole dependence structure are clearly visible. Table 1 summarizes the deviations between empirical and analytical copula densities in the form of least mean squares. For the original returns we find a least mean square of 27.52, while the locally normalized returns yield a smaller 5.26.

3.4 Comparison with a correlation-weighted Gaussian copula

Aiming for a more suitable description of the empirical copulas, we now consider a weighted average of Gaussian copulas for different correlation coefficients. This takes into account the empirical average over different stock pairs (see (3.1)). Figure 6 shows the histograms of empirical correlation coefficients $f(C_{k,l})$ for original and locally
(a), (c) Correlation-weighted Gaussian copula densities \( \text{cop}(u, v)^{\text{CWG}} \). (b), (d) Differences between the empirical copula density and the correlation-weighted Gaussian copula density, \( \text{cop}^{(\text{glob})}(u, v) - \text{cop}(u, v)^{\text{CWG}} \) and \( \text{cop}^{(\text{loc})}(u, v) - \text{cop}(u, v)^{\text{CWG}} \), respectively. (a), (b) Original returns. (c), (d) Locally normalized returns.

Normalized returns estimated on the entire time series. The bin size is \( \Delta c = 0.02 \). In our data sets we only observe positive correlations. We now obtain a Gaussian mixture by weighting each Gaussian copula density with the value of the probability density function of the respective correlation, \( h(C_{k,l}) = f(C_{k,l}) \Delta c \). This yields the correlation-weighted Gaussian copula density

\[
\text{cop}^{\text{CWG}}(u, v) = \sum_{C_{k,l} = -1}^{1} h(C_{k,l}) \text{cop}_{C_{k,l}}^{G}(u, v),
\]

which is compared with the empirical copula densities in Figure 7. For both data sets we find only slight improvement over the single Gaussian copula, as depicted in Figure 5. The least mean squares are only slightly smaller as well (see Table 1).
Table 2: Parameter values for the $K$-copula densities.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original returns</th>
<th>Locally normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>$c$</td>
<td>0.44</td>
<td>0.39</td>
</tr>
<tr>
<td>$N$</td>
<td>3.2</td>
<td>7.8</td>
</tr>
</tbody>
</table>

This can be attributed to the fact that correlations also vary in time (see, for example, Münix et al. 2012; Schäfer and Guhr 2010). Thus, it is not sufficient to take the average over different correlations into account. For a better approximation of the empirical copula densities, the correlations would have to be estimated on shorter time intervals. This, however, leads to increasing estimation errors for shorter estimation intervals. Consequently, the reliable attainment of time-dependent correlations is problematic. This is where the $K$-copula, discussed in Section 2.7, comes into play.

3.5 Comparison with the $K$-copula

We now consider the $K$-copula density $\text{cop}^K_{c,N}(u, v)$, which takes inhomogeneous and time-varying correlations into account. Here, the fluctuations of correlations around their mean value $c$ are characterized by the free parameter $N$. We calculate the mean correlation coefficient for the original returns and the locally normalized returns. As noted above, we find $\bar{c}^{(\text{glob})} = 0.44$ and $\bar{c}^{(\text{loc})} = 0.39$, respectively. The free parameter $N$ is fitted to the empirical copula densities by the method of least squares. We find $N^{(\text{glob})} = 3.2$ for original returns and $N^{(\text{loc})} = 7.8$ for locally normalized returns. The parameter values for $\bar{c}$ and $N$ are also summarized in Table 2. The lower value for the original returns reflects the fact that the locally normalized returns have a constant variance of 1. Hence, there are smaller fluctuations in the covariances, which are in this case simply the correlations. In the $K$-copula, smaller fluctuations of the covariances or correlations are described by larger values of $N$.

In Figure 8, the resulting $K$-copula densities are compared with the empirical copula densities. In both cases we find improved agreement with the empirical results. This is also reflected in much lower values for the least mean squares (see Table 1). However, for the original returns, a large deviation between empirical and analytical results remains. The overall structure of the empirical dependence is not captured very well by the $K$-copula. For the locally normalized returns on the other hand, the $K$-copula yields a very good description of the empirical dependence. Only slight deviations persist in the lower–lower and upper–upper corners. The asymmetry of the empirical copula densities cannot be captured by the $K$-copula density due to its
FIGURE 8 Differences in $K$-copula densities.

(a), (c) $K$-copula densities $\text{cop}_{c,N}^K(u,v)$. (b), (d) Differences between the empirical copula density and the $K$-copula density, $\text{cop}^{(\text{glob})}(u,v) - \text{cop}_{c,N}^K(u,v)$ and $\text{cop}^{(\text{loc})}(u,v) - \text{cop}_{c,N}^K(u,v)$, respectively. (a), (b) Original returns, $c^{(\text{glob})} = 0.44$, $N^{(\text{glob})} = 3.2$. (c), (d) Locally normalized returns $c^{(\text{loc})} = 0.39$, $N^{(\text{loc})} = 7.8$.

TABLE 3 Parameter values for the skewed Student $t$-copula densities.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Original returns</th>
<th>Locally normalized</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\tilde{c}$</td>
<td>0.44</td>
<td>0.39</td>
</tr>
<tr>
<td>$\nu$</td>
<td>3.3</td>
<td>8.0</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.06</td>
<td>0.04</td>
</tr>
</tbody>
</table>

symmetric character. This asymmetry is only weakly present in the case of locally normalized returns, which leads to a better agreement between the $K$-copula and the empirical copula.
3.6 Comparison with the skewed Student $t$-copula

The skewed Student $t$-copula allows for an asymmetry in the dependence structure. Hence, it is a natural candidate with which to compare our empirical findings. Since we observe no asymmetry on average in the negative tail dependence of the empirical copula densities (see histograms (b) and (d) in Figure 4), we choose the same value for both components of the skewness parameter vector $\gamma$, $\gamma_1 = \gamma_2 = \gamma$. This leaves only an asymmetry in the positive tail dependence of the skewed Student $t$-copula and reduces the number of parameters. The parameter values for $\nu$ and $\gamma$ are fitted to the empirical copula densities by the method of least squares, whereas the mean correlation $c$ is empirically determined by $\bar{c}^{(\text{glob})} = 0.44$ and $\bar{c}^{(\text{loc})} = 0.39$, respectively. We
find  \( \nu^{(\text{glob})} = 3.3, \gamma^{(\text{glob})} = 0.06 \) for original returns and  \( \nu^{(\text{loc})} = 8.0, \gamma^{(\text{loc})} = 0.04 \) for locally normalized returns. The parameter values are summarized in Table 3.

Figure 9 illustrates the resulting skewed Student \( t \)-copula densities and their differences from the empirical copula densities. For original returns, we find a remarkable agreement. The skewed Student \( t \)-copula is able to capture the overall dependence structure very well, including the asymmetry in the positive tail dependence. There are only small deviations in the corners: in particular a mild overestimation for the negative tail dependence of extreme events. For locally normalized returns, the skewed Student \( t \)-copula provides a slightly worse fit of the empirical data than the \( K \)-copula. There is not much asymmetry to capture in the empirical copula density. Hence, the skewness parameter plays only a minor role in this case. Further, the positive tail dependence is more pronounced for the skewed Student \( t \)-copula than for the \( K \)-copula. On a local scale, however, this is not reflected in the empirical dependence structure.

4 CONCLUSION

We presented an empirical study on the statistical dependencies between daily stock returns. To this end we estimated empirical pairwise copulas for the original returns and for locally normalized returns. Considering the former allowed us to study the dependence structure on a global scale. In the latter case, the nonstationary characteristics of return time series, ie, time-varying trends and volatilities, were removed. This provided us with the dependence structure on a local scale.

How far does the concept of Gaussian dependence carry in light of our empirical findings? To answer this question, we compared the empirical results not only with a single Gaussian copula, but also with a correlation-weighted average of Gaussian copulas and with the \( K \)-copula. The latter arises from a random matrix approach that models time-varying covariances with a Wishart distribution. This yields a \( K \)-distribution that adequately describes multivariate returns. We derived the resulting \( K \)-copula for the bivariate case and found very good agreement with empirical pairwise dependencies of the locally normalized returns. Thus, we arrived at a consistent picture within the random matrix model: the \( K \)-distribution is able to describe the tail behavior of the marginal return distributions, and the \( K \)-copula captures the overall empirical dependence structure. This implies that Gaussian statistics, and thus also a Gaussian dependence structure, provide a good description on a local scale. However, on a global scale, ie, when the empirical distribution function, which is involved in the estimation of the copula, is applied to the original return time series instead, we find rather significant deviations from the \( K \)-copula. In particular, we observe a pronounced asymmetry in the positive tail dependence. Therefore, we also compared our empirical findings with a model that explicitly allows for such an asymmetry: the
skewed Student $t$-copula. And indeed, we found a rather compelling agreement with the empirical dependence structure of original returns. For the local scale, however, the empirical copula exhibited only a mild asymmetry and is overall better described by the $K$-copula. How can we understand this? For the original returns, the tail dependence reflects periods with high volatility, while for the locally normalized returns all periods contribute equally to the tail dependence. Thus, our results imply that the asymmetry in the tail dependence is not stationary, and it is stronger in periods with high volatility.

**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 thank Desislava Chetalova for helpful discussions.

**REFERENCES**


Research Paper

Path-consistent wrong-way risk: a structural model approach

Markus Hofer

Quantitative Analytics, ING Bank, ING Financial Markets, Bijlmerdreef 109, 1102 BW Amsterdam, The Netherlands; email: markus.hofer@ing.nl

(Received June 26, 2015; revised December 15, 2015; accepted January 11, 2016)

ABSTRACT

We present a general and path-consistent wrong-way risk (WWR) model, which does not require simulation of credit and market variables simultaneously. Although similar so-called copula models are well known, our approach is novel in several ways. First, our method can model a wide range of dependence structures while always guaranteeing path consistency of default probabilities (the possibility of path-inconsistencies in copula models was highlighted in a recent article). Second, we place special emphasis on the difficult task of calibrating the underlying dependence structure. In particular, we consider a default correction of the dependence structure. Third, our model serves as a bridge between structural model approaches, where dependence between exposure and equity price is modeled, and copula models, where exposure is directly correlated to default time. Finally, we apply our method in realistic situations and show that we can achieve a wide range of WWR impacts.

Keywords: credit valuation adjustment; wrong-way risk; copula; structural credit model; model calibration.
1 INTRODUCTION

In recent years, credit valuation adjustment (CVA) has received enormous attention, focused on both the computational and modeling aspects. (Unilateral) CVA is defined as the expected loss in a portfolio at default of the corresponding counterparty; more precisely,

\[ \text{CVA} = \mathbb{E}[(1 - \text{RR})V^+_\tau 1(\tau < T)], \]

where RR denotes the so-called recovery rate, \( V^+_\tau \) denotes the discounted, positive exposure of the portfolio at the time of default \( \tau \) of the counterparty, and \( T \) is the maximal maturity of a derivative within the portfolio.

In this context, an important challenge is the modeling of the dependencies between \( V_t \) and \( \tau \) or, more generally, between exposure and probability of default. Since there are almost no liquidly traded products that depend on market and credit variables simultaneously, modeling and calibration of these dependencies is extremely difficult.

A special case of dependence, usually denoted wrong-way risk (WWR), is present if the probability of default of the counterparty is positively correlated with the value of the portfolio. In particular, in times of crisis the increase in CVA due to WWR can be significant. A typical example of very strong WWR is portfolios with emerging market counterparties, where the portfolio value depends heavily on foreign exchange (FX) rates involving the local currency.

Hull and White (2012) were among the first to publish a consistent WWR framework. They model WWR as (deterministic) dependence between hazard rates and exposure. This approach has the advantage of requiring only small changes to standard methods for calculating CVA without WWR.

Another popular modeling approach was followed by a number of authors (see, for example, Pykthin 2012; Rosen and Saunders 2012; Sokol 2010). In Sokol (2010) the dependence between default time and exposure is modeled by a copula, where the choice of copula parameters depends on the correlation between the counterparty’s creditworthiness and market risk factors. Recently, Böcker and Brunnbauer (2014) showed that these so-called copula models are prone to losing the path consistency of default probabilities if the model is not set up properly. Böcker and Brunnbauer give a characterization of path-consistent WWR models and show, for instance, that the Rosen–Saunders model and the above-mentioned Hull–White model belong to the class of path-consistent WWR models.

In contrast to the above examples, where dependence between credit and market variables is directly modeled, it is also possible to introduce dependencies between exposure at default and time of default, implicitly, by using a structural credit model; see, for instance, Buckley et al (2011) in which default occurs if the equity process of the counterparty hits a specific barrier, and exposure dynamics are modeled as a stochastic process correlated with the equity process.
Finally, a number of articles focus on particular aspects of WWR such as systemic wrong-way risk (Pykthin and Sokol 2013; Turlakov 2013), econometric investigations of the calibration of WWR models (Ruiz et al 2015) or modeling of exposure jumps at default (Pykthin and Sokol 2013). Furthermore, Glasserman and Yang (2015) compute model-independent upper bounds for WWR using linear programming. See Hofer and Iacó (2014) for a general numerical method to compute worst case bounds when the dependence structure is unknown.

Here we pick up the ideas in Böcker and Brunnbauer (2014), Rosen and Saunders (2012), Buckley et al (2011) and Sokol (2010), to define a WWR model defined on a discrete-time grid, which is path consistent in the sense of Böcker and Brunnbauer (2014). The contribution of this paper is threefold. First, we present a general, realistic WWR model in line with the path-consistency definition given in Böcker and Brunnbauer (2014). WWR is explicitly modeled by the joint distribution of the exposure and the equity value of the counterparty. Second, we motivate and provide a calibration strategy, since calibration of WWR models has not received much attention in the literature. Finally, we provide a link between WWR models that define dependence between exposure and default time directly and those that model the dependence indirectly via a structural model.

The structure of the paper is as follows. In Section 2 we introduce our model setup and prove our main theoretical contribution: the path consistency of our model. In Section 3 we present a calibration methodology for our model, and in Section 4 we apply our model in realistic situations and show that it is able to capture a wide range of WWR impacts.

2 A PATH-CONSISTENT WRONG-WAY RISK MODEL

In this section, we consider only unilateral CVA, ie, we only model the default of the counterparty. The main result of this section is Theorem 2.5, in which we show that our method defines path-consistent exposure weights in the sense of Böcker and Brunnbauer (2014).

The drift of the portfolio value can smoothly mimic sudden changes such as payments or expiration of parts of the portfolio. Note that $V_t$ depends only on the market value of the portfolio and not on the default of the counterparty; thus, we call $Y_t$ the default-risk-free portfolio value.

We model default in a similar way to structural credit models (see, for example, Merton 1974). For a given time grid $t_0 = 0 < t_1 < \cdots < t_M = T$, we assume the counterparty defaults at $t_j$ if its stock price $\tilde{S}_{t_j}$ is smaller than a barrier level $B_j$. In this setting, default is only possible at a grid point $t_j$. In our model the barrier levels $B_j$ are defined such that the model default probabilities

$$p_j = \mathbb{P}(\tilde{S}_{t_j} < B_j, \tilde{S}_{t_k} > B_k \forall k < j)$$
match market-implied values in the following sense:

$$p_j = \mathbb{P}_{\text{market implied}}(\text{default in } [t_{j-1}, t_j] \mid \text{no default before } t_{j-1}).$$  \tag{2.1}$$

Since we work on a discrete-time grid, we are able to divide the total CVA into contributions from each monitoring point \( t_j \). More precisely, we can write

$$\text{CVA} = \sum_{j=1}^{M} \text{CVA}_j,$$

which will simplify notation in the rest of this section.

Assume \( \text{RR} = 0 \). We can compute the expected positive exposure at default at time \( t_1 \) in a straightforward way. Furthermore, we denote the value of the underlying portfolio at time \( t \) as \( \tilde{V}_t \). More precisely, we have \( \tilde{V}_t = V_t \mid S_{t_j}^{\text{CP}} > B_j \) for all \( t_j < t \).

In the same way, we have to consider the equity process conditionally on no-default in the past:

$$\tilde{S}_t^{\text{CP}} = S_{t_j}^{\text{CP}} \mid S_{t_j}^{\text{CP}} > B_j \quad \forall t_j < t.$$

Note that \( \tilde{V}_t \) depends only on the market value of the portfolio and not on the default of the counterparty. Thus, we call \( \tilde{V}_t \) the default-risk-free portfolio value. Since the dynamics of \( \tilde{V}_t \) are usually not available in closed form, we have to rely on Monte Carlo simulations. Assume that we have simulated \( N \tilde{V}_t \)-paths on the time grid \((t_j)_{0 \leq j \leq M}\) and the resulting values are denoted by \( \tilde{V}_{ij}^t \) for \( i = 1, \ldots, N \) and \( j = 0, \ldots, M \).

When exposure and probability of default are not independent, we have to consider risky portfolio values \( V_t \). More precisely we correct the portfolio values \( \tilde{V}_t \) for probability at default at \( s < t \). Similarly, we will denote the default-adjusted equity process by \( S_{t_j} \). The distributions of \( V_t \) and \( \tilde{V}_t \) can differ significantly, as we show in the following simple example.

**Example 2.1** Let the dynamics of the portfolio value be given by

$$\tilde{V}_{t_j} = \begin{cases} 
0 & \text{with probability 0.9 if } \tilde{V}_{ij-1} = 0, \\
1 & \text{with probability 0.1 if } \tilde{V}_{ij-1} = 0, \\
1.000 & \text{with probability 1 if } \tilde{V}_{ij-1} = 1,
\end{cases}$$

where \( \tilde{V}_0 = 0 \) and \( p_j = 0.1 \) for all \( j \). If we assume independence between the portfolio value and default events we get

$$V_{t_j} = \begin{cases} 
0 & \text{with probability 0.9 if } V_{ij-1} = 0, \\
1 & \text{with probability 0.1 if } V_{ij-1} = 0, \\
1.000 & \text{with probability 1 if } V_{ij-1} = 1.
\end{cases}$$
However, if we assume comonotonicity between the default and the portfolio value, meaning that at every \( t_j, j > 1 \), the top \( p_j \)-quantile of \( V_{t_j} \) defaults and the rest has a 0% probability of default, then we have

\[
V_{t_j} = \begin{cases} 
0 & \text{with probability } 0.9 \text{ if } V_{t_{j-1}} = 0, \\
1 & \text{with probability } 0.1 \text{ if } V_{t_{j-1}} = 0,
\end{cases}
\]

where \( V_{t_0} = 0 \). In this case the highest portfolio value, 1.000, can never be reached, although portfolio value and default are strongly positive correlated.

In our model, the dependence structure between market and credit variables is induced by the dependence structure between the portfolio value and the counterparty’s equity value. More precisely, we assume that the joint distribution of \( V_{t_j} \), the portfolio value at \( t_j \), and \( S_{t_j} \), the counterparty’s equity value, is defined as

\[
F_{V_{t_j}, S_{t_j}}(v, s) = C_j(F_{V_{t_j}}(v), F_{S_{t_j}}(s)),
\]

where \( C_j \) is a copula and \( F_{V_{t_j}} \) is the empirical distribution of the portfolio value at \( t_j \).

In the following equations we denote the per-time-slice ordered exposures by \( U_{t_{j}}^i \), i.e., \( U_{t_{j}}^1 \leq \cdots \leq U_{t_{j}}^N \), and the position of \( V_{t_j}^i \) in the ordered set by \( g(j, i) \); thus,

\[
V_{t_j}^i = U_{t_{j}}^{g(j,i)}.
\]

Under the above assumptions, the CVA contribution at time \( t_1 \) can be written as

\[
CVA_1 = (1 - RR) \int_{(-\infty, \infty) \times [0, B_1)} \max(v, 0) \, dF_{V_{t_1}, S_{t_1}}(v, s) \\
= (1 - RR) \sum_{i=1}^{N} \max(U_{t_1}^i, 0) \int_{[(i-1)/N, i/N) \times [0, F_{S_{t_1}}(B_1))]} dC_1(x, y) \\
= (1 - RR) \sum_{i=1}^{N} \max(U_{t_1}^i, 0) w_1^i p_1,
\]

(2.2)

where

\[
0 \leq w_1^i \leq \min\left(\frac{1}{Np_1}, 1\right) \quad \forall i
\]

and

\[
\sum_{i=1}^{N} w_1^i = 1,
\]

which follows from copula properties and the Fréchet–Hoeffding bound (see, for example, Nelsen 2006). Note that \( w_1^i = 1/N \) for all \( i \) (which replicates the no-WWR case) holds if, for instance, \( C_1(x, y) \) is the independence copula or if \( p_1 = 1 \).
Furthermore, in the above setting, no knowledge about the dynamics of $S_t$, except for the parameters of the copula $C_1$ and the counterparty’s default probabilities, is needed to compute $CVA_1$. Moreover, note that $V_{t_1} = \tilde{V}_{t_1}$ and $S_{t_1} = \tilde{S}_{t_1}$.

Next we compute $F_{V_{t_2}}$, i.e., we have to adjust $F_{\tilde{V}_{t_2}}$ for a default at $t_1$. For this reason we define survival probabilities $P_{t_j}^i$ for every simulated exposure path $(\tilde{V}_{t_j})_{0 \leq j \leq M}$, where $P_{t_j}^i$ denotes the probability that the counterparty did not default before $t_j$ conditional on having the exposures $\tilde{V}_{t_k}$ for $0 \leq k < j$. Obviously, $P_{t_1}^i = 1$ for all $i$, and, following (2.2), the $P_{t_2}^i$ can be computed as

$$P_{t_2}^i = P_{t_1}^i \left(1 - N w_{1}^g(1,i) p_1\right). \quad (2.3)$$

In (2.3), we used the fact that at $t_1$ every simulated exposure path accounts for $1/N$ of the empirical distribution. We present a general version of (2.3) below.

To further simplify notation, we use aggregated survival probabilities $\tilde{P}_{t_j}^i$, ordered according to the corresponding simulated exposures. More precisely,

$$\tilde{P}_{t_j}^i = \sum_{k=1}^{i} P_{t_j}^{g^{-1}(j,k)},$$

where $g^{-1}$ is defined such that $i = g^{-1}(g(j,i))$ for all $i$, $j$ and $\tilde{P}_{t_j}^0 = 0$ for all $j$.

Now we can compute $CVA_2$:

$$CVA_2 = (1 - RR) \int_{(-\infty,\infty) \times [0,B_2)} \max(v,0) \, dF_{V_{t_2},S_{t_2}}(v,s)$$

$$= (1 - RR) \sum_{i=1}^{N} \max(U_{t_2}^i,0) \int_{[\tilde{P}_{t_2}^{i-1}/\tilde{P}_{t_2}^{N},\tilde{P}_{t_2}^{i}/\tilde{P}_{t_2}^{N}] \times [0,F_{S_{t_2}}(B_2)]} dC_2(x,y)$$

$$= (1 - RR) \sum_{i=1}^{N} \max(U_{t_2}^i,0) w_2^i p_2,$$

and, in general,

$$CVA_j = (1 - RR) \int_{(-\infty,\infty) \times [0,B_j)} \max(v,0) \, dF_{V_{t_j},S_{t_j}}(v,s)$$

$$= (1 - RR) \sum_{i=1}^{N} \max(U_{t_j}^i,0) \int_{[\tilde{P}_{t_j}^{i-1}/\tilde{P}_{t_j}^{N},\tilde{P}_{t_j}^{i}/\tilde{P}_{t_j}^{N}] \times [0,F_{S_{t_j}}(B_j)]} dC_j(x,y)$$

$$= (1 - RR) \sum_{i=1}^{N} \max(U_{t_j}^i,0) w_j^i p_j, \quad (2.4)$$

where

$$P_{t_{j+1}}^i = P_{t_j}^i \left(1 - \frac{\tilde{P}_{t_j}^N}{P_{t_j}^i} w_{j}^g(j,i) p_j\right). \quad (2.5)$$
Note that, since $\tilde{P}_{ti}^i = i$ and $\tilde{P}_{ti}^N = N$, we have that (2.4) is equivalent to (2.2) for $j = 1$, and (2.5) is equivalent to (2.3) for $j = 1$ as $P_{ti}^i = 1$ for all $i$.

The following lemma proves the nonnegativity of $P_{ti}^i$.

**Lemma 2.2** $P_{ti}^i \geq 0$ for all $i, j$.

**Proof** First assume that $P_{ti}^i = 0$. Then $w_j^{g(j,i)} = 0$, since the corresponding integration domain in (2.4) is empty.

Next we assume that $P_{ti}^i > 0$. Then

$$P_{ti+1}^i = P_{ti}^i \left( 1 - \frac{\tilde{P}_{ti}^N}{P_{ti}^i} w_j^k p_j \right),$$

where $k = g(j,i)$. Thus, the statement is proved if

$$\left( 1 - \frac{\tilde{P}_{ti}^N}{P_{ti}^i} w_j^k p_j \right) \geq 0.$$

We have

$$\frac{\tilde{P}_{ti}^N}{P_{ti}^i} w_j^k p_j = \frac{\tilde{P}_{ti}^N}{P_{ti}^i} \int_{[\tilde{p}_{ti}^k/\tilde{P}_{ti}^N, \tilde{p}_{ti}^k/\tilde{P}_{ti}^N] \times [0, F_{St_j}(B_j)]} dC_j(x, y) \leq \frac{\tilde{P}_{ti}^N}{P_{ti}^i} \min \left( p_j, \frac{\tilde{p}_{ti}^k - \tilde{p}_{ti}^{k-1}}{\tilde{p}_{ti}^N} \right) \leq 1,$$

where the first inequality follows from the upper Fréchet–Hoeffding bound,

$$\int_{[x_1, x_2] \times [y_1, y_2]} dC(x, y) \leq \min(x_2 - x_1, y_2 - y_1),$$

which holds for every copula $C$. □

To prove the path consistency of our model, we recall the notion of path-consistent exposure weight functions characterized in Böcker and Brunnbauer (2014, Theorem 9). Let $\Omega$ be the path space of market variables and let $F(t)$ be the distribution function of the default time $\tau$ of the counterparty. Then the CVA can be written as

$$\text{CVA} = (1 - \text{RR}) \mathbb{E} \left[ \int_0^\infty \tilde{V}^+ (t, \omega) dq(t, \omega) \right],$$

where $\tilde{V}^+$ represents the positive part of the stochastic process $\tilde{V}$.
where $\omega \in \Omega$ and
\[ q(t, \omega) = \mathbb{P}(\tau(\omega) \leq t). \]

We define a continuous weight function $w_t(\omega)$ as
\[ w : \Omega \times [0, \infty) \to \mathbb{R}, \quad w_t(\omega) = \frac{q'(t, \omega)}{F'(t)}. \]

Using this notation, we can write the CVA as
\[ \text{CVA} = (1 - \text{RR}) \int_0^\infty \mathbb{E}[\bar{V}^+(t, \omega) w_t(\omega)] \, dF(t). \quad (2.6) \]

Furthermore, we denote the infinitesimal CVA contribution by
\[ \text{CVA}^{\text{Cont}}_t = (1 - \text{RR})\mathbb{E}[V^+(t, \omega) w_t(\omega)] \, dF(t). \quad (2.7) \]

**Definition 2.3** (Path-consistent exposure weight function) The exposure weight function $w_t(\omega)$ in (2.6) is called path consistent if and only if

1. $\mathbb{E}[w_t] = 1$ for all $t \geq 0$, and
2. the following holds for all $\omega \in \Omega$:
\[ \int_0^\infty w_t(\omega) \, dF(t) = 1. \]

We now discuss path consistency in our discrete setting, following the above approach. Thus, we introduce the notion of discrete path-consistent exposure weights. As shown in (2.4), in our model the CVA contribution in the time interval $[t_{j-1}, t_j)$ is given as
\[ \text{CVA}_j = (1 - \text{RR}) \sum_{i=1}^N \max(U^i_{t_j}, 0) w^i_j p_j, \]
where we call $w^i_j, j = 1, \ldots, M, i = 1, \ldots, N,$ discrete exposure weights.

**Definition 2.4** (Discrete path-consistent exposure weights) We call the discrete exposure weights $(w^i_j)_{1 \leq j \leq N, 0 \leq i \leq M}$ path consistent if and only if

1. $\mathbb{E}[w^i_j] = 1/N$ for all $i, j$, and
2. for every given exposure path realization, the corresponding infinite time default probability is 1.

In the following main result of this section we show that our method always includes path-consistent exposure weights, which are the discrete analog of the path-consistent exposure weight functions in Definition 2.3.
THEOREM 2.5 The weights \( (w^j_i)_{1 \leq j \leq N, 0 \leq i \leq M} \) are discrete path-consistent exposure weights.

PROOF Condition (1) in Definition 2.4 is equivalent to

\[
\sum_{i=1}^{N} w^j_i = 1
\]

for all \( j \), which follows from (2.4) and the definition of a copula: more precisely, from \( C(1, x) = x \) for all \( x \in [0, 1] \) and since

\[
\frac{\tilde{P}^i_{t_j}}{P^N_{t_j}} = \begin{cases} 
0 & \text{for } i = 0, \\
1 & \text{for } i = N.
\end{cases}
\]

To prove condition (2) in Definition 2.4 we first assume that the distribution function of the underlying default time has compact support, i.e., there exists \( t^* < \infty \) such that \( F(t^*) = 1 \), and we set \( t_M = t^* \). Then for a fixed path with index \( i \) we have that \( p_M = 1 \), and thus

\[
\begin{align*}
p^i_{t_M} &= p^i_{t_{M-1}} \left( 1 - \frac{\tilde{P}^N_{t_{M-1}} w^g_{M-1,i} p_{M-1}}{P^N_{t_{M-1}}} \right) \\
&= p^i_{t_{M-1}} \left( 1 - \frac{\tilde{P}^N_{t_{M-1}} p^i_{t_{M-1}}}{P^N_{t_{M-1}}} \right) \\
&= 0.
\end{align*}
\]

If \( F \) has support on \( \mathbb{R}^+ \), we define a sequence of distribution functions \( F_k \), with

\[
F_k(t) = \begin{cases} 
F(t) & \text{for } 0 \leq t \leq \tau_k, \\
1 & \text{for } t > \tau_k,
\end{cases}
\]

where for every \( k \) we choose \( \tau_k = t_M \). Then we have that, for every \( k \) and for every path survival, the probabilities at \( t_M = \tau_k \) are 0. With \( \tau_k \to \infty \) we can approximate \( F \) arbitrarily close with \( F_k \), which proves the statement. \( \square \)

We close this section with three remarks, in which we emphasize the connection between our model and recent literature.

REMARK 2.6 A frequently stated requirement for the WWR model is that the CVA should be higher in the WWR case than in the no-WWR case. However, it was observed and analyzed in Ghamami and Goldberg (2014) that the presence of WWR can decrease CVA. This somewhat nonintuitive effect can also be seen in Example 2.1: although the highest present exposures always have the highest probability of default, the CVA is largest in the case of independence between market variables and credit events. In our WWR model this effect is present, which makes the model realistic.
Remark 2.7 A similar model to that presented in this section can also be defined by modeling the dependence between risky exposure, $V_t$, and conditional default time $(\tau \mid \tau > t_{j-1})$ in the following way:

$$F_{V_{tj}, \tau > t_{j-1}}(v, s) = C_j(F_{V_{tj}}(v), F_{\tau > t_{j-1}}(s)),$$

where $C_j$ is a copula, $F_{V_{tj}}$ is the empirical distribution of the portfolio value at $t$ and $F_{\tau > t_{j-1}}$ is the distribution of the default time conditional on no-default before $t_{j-1}$. When applied in a similar way to that in our model, this joint distribution will also lead to path-consistent exposure weights and to formulas very similar to (2.4). However, in this case the difficult (in our opinion) task of calibrating dependence structures involving default time arises. Nevertheless, there is a close relationship between our (structural) model and, for instance, the approach in Sokol (2010).

Remark 2.8 Rosen and Saunders (2012) propose modeling WWR as joint distribution of the exposure and a so-called creditworthiness indicator, which is a two-factor latent variable consisting of a systematic effect and an idiosyncratic effect. If we omit the systematic factor and define the idiosyncratic effect as a normalized asset return, we arrive at a similar model to the one in this paper. However, note that their approach is only path consistent if the creditworthiness indicator is a unilateral variable, whereas in our approach creditworthiness can change over time.

3 CALIBRATION

In this section we propose a calibration procedure for our WWR model. To better illustrate the calibration, we apply the following simplifying assumptions:

1. the copulas $C_j$ are Gaussian copulas with correlation parameters $\rho_j$;
2. historical dynamics of $\widetilde{V}_t$ can be observed by repricing today’s composition of the underlying portfolio using historical market data;
3. $\log(\tilde{S}_t)$ is a Brownian motion with zero drifts and constant volatility $\sigma_s$;
4. we assume that the process $\widetilde{V}_t$ has independent, homogeneous increments, ie, $\widetilde{V}_{t_2} - \widetilde{V}_{t_1}$ is stochastically independent of $\widetilde{V}_{s_2} - \widetilde{V}_{s_1}$ for all $t_2 \geq t_1 \geq s_2 \geq s_1$ and the distribution of $\widetilde{V}_{t_2 + t} - \widetilde{V}_{t_1 + t}$ is the same for all $t$ for fixed $t_2 > t_1$.

We define a set of historical market dates for which we compute the mark-to-market value of the underlying portfolio according to the current portfolio composition. Then we calculate the Pearson correlation estimator $\tilde{\rho}$ between the resulting exposure time series and the corresponding historical counterparty logarithmic equity time series.
The main idea in this section is to introduce a latent variable process $Z_{tj}$, which has normally distributed increments. More precisely, we define $Z_{tj}$ as

$$Z_{tj} = Z_{tj-1} + \sigma_Z(t_j - t_{j-1})\Phi^{-1}(F_{V_{tj}} - V_{tj-1} - (V_{tj} - V_{tj-1})),$$

where $\Phi^{-1}$ is the inverse distribution function of a standard normal distribution and $F_{V_{tj}} - V_{tj-1}$ is the distribution function of $V_{tj} - V_{tj-1}$. In this case $Z_{tj}$ is a latent variable process for the portfolio value with independent normally distributed increments.

Since there is no default up to $t_1$ in our model, we set the correlation parameter of the copula $C_1$ to be

$$\rho_1 = \tilde{\rho}.$$

Now we want to investigate how taking default into account changes the correlation between the exposure process, $V_t$, and the log equity process, $S_t$. To simplify notation we set $s_t = \log(S_t)$, $\tilde{s}_t = \log(\tilde{S}_t)$ and $b_i = \log(B_i)$. We have

$$\text{Cov}(Z_{t_2}, s_{t_2}) = \text{Cov}(\tilde{Z}_{t_2}, \tilde{s}_{t_2} \mid \tilde{s}_{t_1} > b_1)$$

$$= \mathbb{E}[(\tilde{Z}_{t_2} - \tilde{Z}_{t_1} - \mathbb{E}[\tilde{Z}_{t_2} - \tilde{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1]) \times (\tilde{s}_{t_2} - \tilde{s}_{t_1} - \mathbb{E}[\tilde{s}_{t_2} - \tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1]) \mid \tilde{s}_{t_1} > b_1]$$

$$= \text{Cov}(\tilde{Z}_{t_2} - \tilde{Z}_{t_1}, \tilde{s}_{t_2} - \tilde{s}_{t_1}) + \text{Cov}(\tilde{Z}_{t_1}, \tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1).$$

(3.1)

Furthermore, by Cholesky decomposition, $\tilde{Z}_t$ can be written as

$$\tilde{Z}_t = \rho \frac{\sigma_Z}{\sigma_S} \tilde{s}_t + \sqrt{1 - \rho^2} \sigma_Z \sqrt{t} U,$$

where $U$ is a standard normal random variable independent of $\tilde{s}_t$. Thus, we can simplify the second term in (3.1) as

$$\text{Cov}(\tilde{Z}_{t_1}, \tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)$$

$$= \mathbb{E}[\tilde{Z}_{t_1} \tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1] - \mathbb{E}[\tilde{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1] \mathbb{E}[\tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1]$$

$$= \rho \frac{\sigma_Z}{\sigma_S} \text{Var}(\tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1).$$

Moreover, we have

$$\text{Var}(\tilde{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1) = \mathbb{E}[\tilde{Z}_{t_1}^2 \mid \tilde{s}_{t_1} > b_1] - \mathbb{E}[\tilde{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1]^2$$

$$= \left(\rho \frac{\sigma_Z}{\sigma_S}\right)^2 \text{Var}(\tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1) + (1 - \rho^2)\sigma_Z^2 t_1 \text{Var}(U)$$

$$= \sigma_Y^2 \left(\left(\rho \frac{\sigma_Z}{\sigma_S}\right)^2 \text{Var}(\tilde{s}_{t_1} \mid \tilde{s}_{t_1} > b_1) + (1 - \rho^2) t_1\right).$$

(3.2)
Now we can compute the required correlation

\[
\text{Corr}(Z_{t_2}, s_{t_2}) = \frac{\text{Cov}(\tilde{Z}_{t_2-t_1}, \tilde{s}_{t_2-t_1}) + \text{Cov}(\hat{Z}_{t_1}, \hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)}{\sqrt{\text{Var}(Z_{t_2})} \sqrt{\text{Var}(s_{t_2})}}
\]

\[
= \frac{\rho \sigma_Z \sigma_S (t_2 - t_1) + \rho (\sigma_Z / \sigma_S) \text{Var}(\hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)}{\sqrt{\text{Var}(Z_{t_2})} \sqrt{\text{Var}(s_{t_2})}}
\]

\[
= \frac{\rho \sigma_Z \sigma_S (t_2 - t_1) + \rho (\sigma_Z / \sigma_S) \text{Var}(\hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)}{\sqrt{\text{Var}(\tilde{Z}_{t_2-t_1} + \hat{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1)}} \frac{\sqrt{\text{Var}(\hat{s}_{t_2-t_1} + \hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)}}{\sqrt{\sigma_Z^2 (t_2 - t_1) + \text{Var}(\tilde{Z}_{t_1} \mid \tilde{s}_{t_1} > b_1)}} \left(\sigma_S^2 (t_2 - t_1) + \text{Var}(\hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)\right)
\]

(3.3)

where \(\text{Var}(\hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1)\) is the only unknown quantity if we use (3.2).

This variance term can be computed in closed form, since the variance of a truncated normal distribution is given by

\[
\text{Var}(\hat{s}_{t_1} \mid \tilde{s}_{t_1} > b_1) = \sigma^2_{S t_1} (1 - \delta(\alpha)),
\]

where

\[
\alpha = \frac{b_1}{\sigma_S \sqrt{t_1}}, \quad \lambda(\alpha) = \frac{\phi(\alpha)}{1 - \Phi(\alpha)} \quad \text{and} \quad \delta(\alpha) = \lambda(\alpha)(\lambda(\alpha) - \alpha),
\]

\(\phi\) and \(\Phi\) are the probability distribution function and cumulative distribution function of the standard normal distribution (see, for example, Johnson et al (1994) for more details).

Following the same approach as above, we can derive formulas similar to (3.3) for \(\text{Corr}(Z_{s_j}, s_{s_j}), j > 2\). However, since the main ingredient \(\text{Var}(\hat{s}_{s_j} \mid \tilde{s}_{s_1} > b_1, \ldots, \tilde{s}_{s_j} > b_j)\) is not analytically available for \(j > 1\), the correlation is not given in explicit form. An alternative is Monte Carlo simulation, which we will use in the next section.

**Remark 3.1** The parameter \(\tilde{\rho}_t\), denoting the correlation between \(\tilde{Z}_t\) and \(\tilde{s}_t\), can be further enhanced by making it time dependent. By doing this we can, for instance, account for the so-called aging of the portfolio, which means that the dynamics of the portfolio value change as derivatives in the portfolio expire.

### 4 Numerical Results

In this section, we test our WWR model in two realistic situations: a EUR/RUB FX forward traded with a Russian counterparty (Sberbank in our example) and a
euro interest rate swap (IRS) with extremely long maturity, traded with the Dutch government. In this context the question of what we mean by the “equity of a state” arises. For various reasons, e.g., an obvious correlation with the local economy, we propose to use the major equity index in the underlying country as a proxy of the government’s “equity”. In the case of the Dutch government this is the AEX.

For our exposure simulations we use a geometric Brownian motion to model the EUR/RUB FX rate, and a one-factor Hull–White model for each interest rate involved. Furthermore, we use an intensity model with time-varying deterministic intensities to model default probabilities. All models including correlations are calibrated to the market.

4.1 Correlation calibration

The FX forward is par and has a maturity $T = 2$ years. The notional of the EUR leg is $N_{EUR} = €10$ million. For the market data of June 5, 2015 this gives a notional $N_{RUB}$ of approximately Rb790 million. Figure 1 shows the Sberbank equity time series and the artificial exposure of the FX forward trade at time $t = 0$, given by

$$\text{Exposure}_{FX}(t) = D_{EUR}(t, T)N_{EUR} - D_{RUB}(t, T)N_{RUB} \cdot S_{EUR/RUB}(t), \quad (4.1)$$

where $D_{CCY}(t, T)$ is the discount factor in currency CCY and $S_{EUR/RUB}(t)$ is the EUR/RUB spot FX rate. Note that the parameters of the trade are fixed. Thus, the exposure is in general only zero on June 5, 2015.

The daily changes of the above two-year time series give a correlation $\hat{\rho}$ of about $-52\%$.

In our second scenario, the euro IRS is par issued at June 5, 2015 with a maturity of thirty years, yearly payments of a floating rate $r_t$ against a fixed rate $r$, and a notional $N$ of €10 million. Historical exposures of the IRS and the AEX index are given in Figure 2. This data set gives a correlation of $-12\%$.

In our two simple examples it makes sense to assume no aging of the portfolio, as mentioned in the previous section, since we only have a single contract per portfolio. However, we compute the adjustment to the correlation parameter when taking default into account. To do this we have to specify the observation dates $t_j$ as given in (2.1).

We choose the intervals between observation dates to be very small (weekly) for the first month, then monthly up to two years, quarterly up to five years, semiannually up to ten years and annually thereafter.

With this setting we compute $\text{Var}(\tilde{s}_{t_j} \mid \tilde{s}_{t_1} > b_1, \ldots, \tilde{s}_{t_j} > b_j)$ for all $j$ using Monte Carlo simulations and then $\rho_j$ by (3.2) and (3.3). The results are displayed in Figures 3 and 4. Note that the spikes are due to the change in the interval lengths and uncertainty in Monte Carlo estimators. We plot correlation values only up to the maturity of the corresponding portfolio.
FIGURE 1  Time series of historical FX forward exposure at $t = 0$ and Sberbank from June 6, 2013 to June 5, 2015.

FIGURE 2  Time series of historical IRS exposure at $t = 0$ and AEX from June 6, 2013 to June 5, 2015.
4.2 CVA computations

In Tables 1 and 2, we present CVA numbers for the calibrated and default-adjusted two-year historical correlation and four generic cases of low and high, wrong- and
right-way risk (RWR) to relate the impact of correlation parameters on CVA results. The corresponding expected positive exposures are given in Figures 5 and 6. Note that the (first-year) annual default probability for the Dutch government, according
TABLE 1  CVA results of an EUR/RUB FX forward deal with Sberbank for different correlation parameters.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\rho}$ (%)</th>
<th>CVA (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No WWR</td>
<td>0</td>
<td>77.670</td>
</tr>
<tr>
<td>Calibrated correlation</td>
<td>-52</td>
<td>223.832</td>
</tr>
<tr>
<td>High WWR</td>
<td>-50</td>
<td>217.681</td>
</tr>
<tr>
<td>Low WWR</td>
<td>-5</td>
<td>88.247</td>
</tr>
<tr>
<td>Low RWR</td>
<td>5</td>
<td>65.801</td>
</tr>
<tr>
<td>High RWR</td>
<td>50</td>
<td>6.448</td>
</tr>
</tbody>
</table>

TABLE 2  CVA results of a euro IRS deal with the Dutch government for different correlation parameters.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\hat{\rho}$ (%)</th>
<th>CVA (€)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No WWR</td>
<td>0</td>
<td>76.389</td>
</tr>
<tr>
<td>Calibrated correlation</td>
<td>-12</td>
<td>115.218</td>
</tr>
<tr>
<td>High WWR</td>
<td>-50</td>
<td>327.664</td>
</tr>
<tr>
<td>Low WWR</td>
<td>-5</td>
<td>88.740</td>
</tr>
<tr>
<td>Low RWR</td>
<td>5</td>
<td>58.339</td>
</tr>
<tr>
<td>High RWR</td>
<td>50</td>
<td>2.478</td>
</tr>
</tbody>
</table>

to CDS market data of June 5, 2015, is about 0.2%, and the corresponding value for Sberbank is about 5.6%.

From these figures we draw two conclusions:

(1) our model can show a wide range of CVA impacts; and

(2) the $-52\%$ as in the Sberbank example is already an extreme case for empirical data, which is in line with our understanding of WWR.

DECLARATION OF INTEREST

The author reports no conflicts of interest. The author alone is responsible for the content and writing of the paper. The views expressed here in this paper those of the author and do not necessarily reflect the position of his employer.

ACKNOWLEDGEMENTS

The author is grateful to two anonymous referees, whose suggestions helped to improve the paper.
REFERENCES


Get the information you need straight to your inbox

We also offer daily, weekly and live news updates
Visit Risk.net/alerts now and select your preferences

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
Research Paper

Decomposition of portfolio risk into independent factors using an inductive causal search algorithm

Brian D. Deaton

Department of Agricultural Economics, Texas A&M University, 2124 TAMU, College Station, TX 77843-2124, USA; email: deaton.brian@mcm.edu

(Received September 29, 2014; revised September 25, 2015; accepted September 28, 2015)

ABSTRACT

A method is presented to estimate and decompose a portfolio’s risk along independent factors. This decomposition is based upon a market’s underlying independent risk factors, which are found empirically by using an inductive causal search algorithm that is based on independent component analysis. Since independent risk factors can be understood to always add risk to a portfolio, a portfolio manager can use them to better understand and budget risk. In contrast, portfolio management using the classic marginal analysis is confusing because adding a risky security to a portfolio might actually reduce the portfolio’s risk. In a small application using the six most widely traded currencies (the Australian dollar, Canadian dollar, euro, sterling, Japanese yen and US dollar), independent-factor risk contributions are constrained during portfolio optimizations, and the internal risk characteristics of the resulting portfolios are found to compare favorably with those created by using constraints on the risk contributions of the original assets.

Keywords: risk management; risk attribution; portfolio optimization; portfolio risk decomposition; inductive causation.
1 INTRODUCTION

This paper presents a new way to estimate and decompose a portfolio’s risk into independent factors. This risk decomposition is based upon a market’s underlying independent risk factors, which are found empirically using a method based on independent component analysis. When the risk of a portfolio is decomposed into independent factors, the risk contribution of any factor is nonnegative. Nonnegative risk contributions are convenient for portfolio managers because they allow them to create a risk budget, which is impossible to do when factors are correlated, because the marginal risk contributions are relative to an existing portfolio and can be negative. In an application, it is shown that the internal risk characteristics of portfolios can be controlled by placing constraints on the risk contributions from the independent factors.

The analysis is built on the framework of Meucci (2007) for analyzing risk contributions from a set of arbitrary risk factors. In this framework, a portfolio’s profit and loss $\Pi$ is represented by the product of a vector of risk factors $F$ and a vector of their respective exposures $b$:

$$\Pi = b^T F,$$

(1.1)

where the superscript “T” denotes the transpose of a vector or matrix. The vectors $b$ and $F$ can be defined in a variety of ways, including the case where $b \equiv \omega$ is a vector of relative portfolio weights and $F \equiv R$ is a vector of security returns.

It is often desirable to analyze a portfolio in terms of a set of new risk factors $\tilde{F}$ that completely span the market. In this case, there is an invertible matrix $P$ that linearly transforms the set of original risk factors into the set of new risk factors

$$\tilde{F} = PF.$$

(1.2)

The matrix $P$ can be interpreted as a “pick” matrix containing rows that transform each original factor into its respective new factor. The exposures to the new factors $\tilde{b}$ can then be computed in terms of the exposures to the original factors $b$ as follows:

$$\Pi = b^T P^{-1} PF = \tilde{b}^T \tilde{F},$$

(1.3)

so that

$$\tilde{b} \equiv P^{-T} b,$$

(1.4)

where the superscript “$-1$” denotes the inverse of a matrix.

2 PORTFOLIO THEORY

For simplicity, only mean–variance analysis will be considered here; the methods presented below can easily be used with expected utility optimization or with alternative risk and reward measures. In the mean–variance portfolio theory of Markowitz
Decomposition of portfolio risk using an inductive causal search algorithm

(1952), an investor only cares about the first two moments of the distribution of returns. The goal of the investor is to create a portfolio of risky assets that maximizes their expected return for a given level of risk, which is measured by the variance or the standard deviation of the portfolio return. Canonical mean–variance optimization is performed for an investment horizon of one period so that a reference to time is unnecessary.

The risk factors $F$ of the previous section are defined as the returns $R$ on a set of $N$ securities:

$$R = \begin{pmatrix} R_1 \\ \vdots \\ R_N \end{pmatrix},$$

(2.1)

where each $R_i$ is the return on security $i$. Note that many portfolios contain a risk-free security; when this is the case, the risk-free security is simply included in this set of $N$ securities. The expected returns for the investment horizon are

$$\mathbb{E}(R) = \mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_N \end{pmatrix},$$

(2.2)

where each $\mu_i$ is the expected return of security $i$. The relationships between the returns are summarized in a covariance matrix:

$$\text{cov}(R) = \Omega.$$

(2.3)

The exposures $b$ to the set of returns $R$ are defined as a set of relative portfolio weights $\omega$. A portfolio is a weighted average of $N$ risk factors with relative weight $\omega_i$ assigned to factor $i$:

$$\omega = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_N \end{pmatrix}.$$

(2.4)

Relative portfolio weights can be either positive for a long position or negative for a short position but must sum to 1:

$$\sum_{i=1}^{N} \omega_i = 1.$$ 

The portfolio’s expected return is a weighted average of the security returns,

$$R_p = \omega^T \mu = \sum_{i=1}^{N} \omega_i \mu_i,$$

(2.5)
and the variance of the portfolio’s return is

\[ \sigma_p^2 = \omega^T \Omega \omega = \sum_{i=1}^{N} \sum_{j=1}^{N} \omega_i \omega_j \Omega_{i,j}, \] (2.6)

where \( \Omega_{i,j} \) is the covariance between factor \( i \) and factor \( j \).

The canonical mean–variance model is

\[
\begin{align*}
\text{minimize} & \quad \lambda \sigma_p^2 - (1 - \lambda) R_p \\
\text{subject to} & \quad \sum_{i=1}^{N} \omega_i = 1,
\end{align*}
\] (2.7)

where \( \lambda \) is a risk aversion parameter; a risk-neutral investor has \( \lambda = 0 \), while a highly risk-averse investor has \( \lambda = 1 \) (Zenios 2007). This model can be used to trace an efficient frontier by varying \( \lambda \) across the interval \([0, 1]\); every \( \lambda \) in this interval corresponds to an efficient portfolio that has minimum variance for a given expected return.

Additional constraints can be added to the mean–variance model to achieve certain objectives. It is well known that the mean–variance model tends to allocate too much relative weight to certain assets (Green 1992; Gerard and Tutuncu 2007). Thus, a common addition is to put constraints on the relative portfolio weights so that the portfolio wealth is not focused on any one asset:

\[ \omega_i \leq m \quad \text{for} \quad i = 1, \ldots, N. \] (2.8)

This idea can be generalized so that limits are placed on arbitrary factor exposures \( b \):

\[ b_i \leq m \quad \text{for} \quad i = 1, \ldots, N. \] (2.9)

### 3 MARGINAL RISK CONTRIBUTIONS

In the risk attribution literature, statistical measures of risk are decomposed by using the fact that they are homogeneous of degree 1. For instance, popular risk measures such as standard deviation, value-at-risk and expected shortfall all share this property. When the portfolio exposures \( b \) are scaled by a factor \( k \), a risk measure \( \mathcal{R} \) that is homogeneous of degree 1 increases by that same factor:

\[ \mathcal{R}(k b) = k \mathcal{R}(b). \] (3.1)

Risk measures that behave this way can be decomposed as a summation by taking the derivative of each side of (3.1) with respect to \( k \):

\[ \sum_{i=1}^{N} b_i \frac{\partial \mathcal{R}(k b)}{\partial k b_i} = \mathcal{R}(b). \] (3.2)
Now, set \( k = 1 \) to obtain

\[
\mathcal{R}(\mathbf{b}) = \sum_{i=1}^{N} b_i \frac{\partial \mathcal{R}(\mathbf{b})}{\partial b_i}.
\] (3.3)

This equation states that the risk of the portfolio, which is a scalar quantity, is the sum of the risk contributions from each individual factor. The risk from each individual factor is expressed as a product of the factor exposure \( b_i \) and the marginal rate of change in risk per unit change in the factor exposure \( \frac{\partial \mathcal{R}(\mathbf{b})}{\partial b_i} \). Each risk contribution can also be written in terms of a percentage contribution to portfolio risk as

\[
\frac{1}{\mathcal{R}(\mathbf{b})} \left( b_i \frac{\partial \mathcal{R}(\mathbf{b})}{\partial b_i} \right) \times 100,
\] (3.4)

so that the sum of the percentage contributions is 100.

As Litterman (1996) notes, this is a marginal analysis and useful only in analyzing risk relative to an existing portfolio. For example, eliminating a position that accounts for half the risk will not reduce the portfolio’s risk by half. As a position is reduced in size, its marginal contribution to risk will also be reduced. Additionally, some positions have a negative marginal contribution to risk, so increasing the size of such a position will reduce the risk of the portfolio at the margin.

It is straightforward to obtain a risk metric decomposition in terms of a new set of factors via the risk decomposition computations that use the original set of factors. The marginal rates of change in risk per unit change in the new factor exposures are linear translations of the originals (Meucci 2007):

\[
\frac{\partial \mathcal{R}(\tilde{\mathbf{b}})}{\partial \tilde{\mathbf{b}}} = \mathbf{P} \frac{\partial \mathcal{R}(\mathbf{b})}{\partial \mathbf{b}};
\] (3.5)

so the risk measure in terms of the new factor exposures is

\[
\mathcal{R}(\tilde{\mathbf{b}}) = \sum_{i=1}^{N} \tilde{b}_i \frac{\partial \mathcal{R}(\tilde{\mathbf{b}})}{\partial \tilde{b}_i}.
\] (3.6)

For mean–variance analysis, the portfolio’s standard deviation will be used as the measure of risk,

\[
\mathcal{R}(\mathbf{b}) = \sqrt{\mathbf{b}^\top \text{cov}\{\mathbf{F}\} \mathbf{b}},
\] (3.7)

where \( \text{cov}\{\mathbf{F}\} \) is the risk-factor covariance matrix. The partial derivatives of the portfolio standard deviation are

\[
\frac{\partial \mathcal{R}(\mathbf{b})}{\partial \mathbf{b}} = \frac{\text{cov}\{\mathbf{F}\} \mathbf{b}}{\sqrt{\mathbf{b}^\top \text{cov}\{\mathbf{F}\} \mathbf{b}}}.
\] (3.8)
In the application below, $b$ is the vector of relative weights on the original currencies $b \equiv \omega$ (with $F \equiv R$), and $\tilde{b}$ is the vector of relative weights on the independent factors $\tilde{b} \equiv \tilde{\omega}$ (with $\tilde{F} \equiv \tilde{R}$).

When factors are independent, each marginal risk contribution is nonnegative and can be interpreted as having a potentially adverse effect on the portfolio (in contrast to the potentially negative contributions when factors are not independent). In this case, the portfolio manager is able to create a risk budget for the independent factors that sums to 100% of the portfolio’s risk. Risk contributions are no longer relative to an existing portfolio in this analysis, and it is no longer necessary to resort to interpreting risk contributions as potential loss contributions (as in Qian (2006)).

4 INDEPENDENT COMPONENT ANALYSIS

Independent component analysis (ICA) can be used to find independent factors, and it is a primary component of the algorithm discussed in the next section. It is based on the premise that $N$ observed variables $x_1, \ldots, x_N$ are linear combinations of underlying, statistically mutually independent source variables $s_1, \ldots, s_N$,

$$x_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{iN}s_N \quad \text{for all } i = 1, \ldots, N,$$

(4.1)

where the $a_{ij}$ are mixing coefficients. This basic ICA model is written in vector–matrix form as

$$x = As.$$

(4.2)

The observed variables $x$ are used to estimate both the unknown mixing coefficient matrix $A$ and the unobserved independent components $s$. The observed variables $x$ and the independent components $s$ are both assumed to have zero mean. If this assumption does not hold, then the original observed variables, denoted by $x_o$, can be centered in a preprocessing step

$$x = x_o - \mathbb{E}(x_o).$$

(4.3)

This preprocessing also forces the independent components to have zero mean since

$$\mathbb{E}(s) = A^{-1}\mathbb{E}(x).$$

(4.4)

Estimation of the basic ICA model relies on the following assumptions (Hyvärinen et al 2001).

(1) The independent components are assumed to be statistically independent, but in application some dependency is acceptable.

(2) The mixing matrix $A$ is assumed to be square and invertible for convenience and simplicity.

(3) The independent components must have non-Gaussian distributions.
Some ICA models are slightly different from the basic ICA model and have their own assumptions (see Hyvärinen et al (2001) for details).

The independent components $s$ are more than just uncorrelated: they are as statistically independent as possible. Achieving this requires more information than is available in a correlation matrix unless all of the variables are normally distributed, in which case zero correlation is equivalent to independence. The estimation of independent components uses higher-order moments or other information such as the autocovariance structure for time series variables in addition to correlation information.

It is always possible to linearly transform the observed random variables $x$ into uncorrelated variables. It is also often desirable that the transformed variables have variances equal to unity. The process called whitening transforms zero-mean variables $x$ into uncorrelated variables $z$ that have unit variances. The result of whitening is a matrix $V$ that decorrelates the observed data vector,

$$z = Vx.$$ (4.5)

The matrix $V$ is computed as

$$V = D^{-1/2} E^T,$$ (4.6)

where $E = (e_1 \cdots e_N)$ is the matrix whose columns are the unit-norm eigenvectors of the covariance matrix $C_x = E\{xx^T\}$, and $D = \text{diag}(d_1 \cdots d_N)$ is the diagonal matrix of the eigenvalues of $C_x$. There are an infinite number of matrices $V$ that can create decorrelated components. This is the reason that estimation of the basic ICA model requires the higher-order moments of non-Gaussian distributions (Hyvärinen et al 2001).

The basic ICA model in (4.2) has the following ambiguities:

1. the variances of the independent components cannot be determined;
2. the order of the independent components cannot be determined.

The first ambiguity occurs because any scalar multiple of one of the independent components $s_i$ can be canceled by dividing the corresponding column of the mixing matrix $A$ by that same multiple. The second ambiguity follows from modifying the model with any permutation matrix $P$ and its inverse,

$$x = AP^{-1}Ps.$$ (4.7)

Now $AP^{-1}$ is the unknown mixing matrix and $Ps$ are the independent components in a different order. In applications that use ICA, neither of these ambiguities is important; for an explanation, see Hyvärinen et al (2001).
5 DYNAMIC DIRECTED GRAPH DISCOVERY (VAR–LiNGAM)

A good way to find independent risk factors is to use a causal discovery algorithm that incorporates ICA internally. The method used here, called VAR–LiNGAM, is an enhancement of a vector autoregressive model with the LiNGAM causal discovery algorithm (Hyvärinen 2008; Hyvärinen et al 2010; Shimizu et al 2006). This method is purportedly the first to fully identify a structural vector autoregressive model with the assumptions that there is acyclicity and there are no latent variables. It uses the non-Gaussian structure of the data, whereas other methods use only the covariance information, which is not always sufficient for identification. The non-Gaussian information is used in the ICA step of the LiNGAM algorithm to find the matrix that transforms a vector of returns into its vector of corresponding independent factors.

VAR–LiNGAM is well suited for use in economics because it works with time series and, importantly, provides parameter estimates. Additionally, the LiNGAM algorithm is very computationally efficient compared with most other causal search algorithms. Because of its computational speed, the VAR–LiNGAM method might still be feasible in cases where a portfolio contains a large number of assets, so long as enough data is available for good statistical estimates. When using the method with a large number of assets, the asset returns should be used directly without an initial transformation, so that the underlying causal structure is not obscured. The number of estimated independent factors will be the same as the number of original assets (see Section 6), but with certain causal structures the number of independent factors could be reduced by ignoring those with very small variances or using some similar criteria; some asset managers may find this reduced dimensionality convenient.

Before introducing the VAR–LiNGAM model, we discuss the two components on which it is built. The first component, a structural equation model (SEM), typically assumes that the observed data is independent and identically distributed; an SEM model does not consider the time series structure in data. A vector of contemporaneous returns $\mathbf{R}$ is modeled in SEM form as

$$\mathbf{R} = \mathbf{B}\mathbf{R} + \mathbf{e},$$

(5.1)

where $\mathbf{e}$ is a vector of disturbances and $\mathbf{B}$ is a matrix of coefficients; the diagonal of $\mathbf{B}$ is defined to be zero. Equation (5.1) can easily be transformed into an ICA model:

$$\mathbf{R} = (\mathbf{I} - \mathbf{B})^{-1}\mathbf{e}.$$  

(5.2)

The VAR–LiNGAM assumptions on the disturbance terms $\mathbf{e}$, enumerated below, allow ICA to be used as part of its estimation procedure.

---

1 LiNGAM: linear, non-Gaussian, acyclic causal models.
The second component of VAR–LiNGAM, an autoregressive model, does consider the time series structure in data. The vector of return observations is denoted as a time series by

$$R(t) = \begin{pmatrix} R_1(t) \\ \vdots \\ R_N(t) \end{pmatrix}, \quad t \in 1, \ldots, T,$$

(5.3)

where \(R_i(t)\) is the return on a particular asset \(i \in 1, \ldots, N\) at time \(t \in 1, \ldots, T\). An autoregressive model of multivariate time series return data \(R(t)\) is written as

$$R(t) = \sum_{\tau=1}^{k} B_\tau R(t - \tau) + e(t),$$

(5.4)

where \(k\) is the number of time delays (lags) of the autoregression, \(B_\tau\) are \(n \times n\) matrixes of coefficients and \(e(t)\) is the innovation process.

The VAR–LiNGAM model is a combination of both the contemporaneous (lag zero) structural equation model (5.1) and the autoregressive model with time delays (5.4). The notation is similar to the previous notation: \(k\) is again the number of time delays (lags) of the autoregression and \(B_\tau\) are the \(n \times n\) matrixes containing the causal effects between returns \(R(t - \tau)\) with time lag \(\tau = 1, \ldots, k\). The \(B_\tau\) matrixes for \(\tau > 0\) correspond to effects from the past to the present, while \(B_0\) corresponds to instantaneous effects. The complete VAR–LiNGAM model is

$$R(t) = \sum_{\tau=0}^{k} B_\tau R(t - \tau) + e(t),$$

(5.5)

where \(e(t)\) are random disturbances. This model is based on three assumptions:

1. \(e(t)\) are mutually independent and temporally uncorrelated, both with each other and over time;
2. \(e(t)\) are non-Gaussian;
3. the matrix \(B_0\) corresponds to an acyclic graph.

The model is estimated in two stages. In the first stage, estimate the following vector autoregressive model using any least-squares method:

$$R(t) = \sum_{\tau=1}^{k} M_\tau R(t - \tau) + n(t),$$

(5.6)

where \(M_\tau\) is the matrix of autoregressive least-squares estimates. Note that \(M_\tau\) contains no contemporaneous parameter estimates, ie, \(\tau = 0\). Then, compute the residuals.
of the model,

\[ \hat{n}(t) = R(t) - \sum_{\tau=1}^{k} M_\tau R(t - \tau). \]  

(5.7)

In the second stage, perform a LiNGAM analysis on the residuals to obtain an estimate of the matrix \( B_0 \), which is the solution to the instantaneous causal model

\[ \hat{n}(t) = B_0 \hat{n}(t) + e(t). \]  

(5.8)

Use \( \hat{B}_0 \) to compute \( \hat{B}_\tau \),

\[ \hat{B}_\tau = (I - \hat{B}_0) M_\tau \quad \text{for } \tau > 0, \]  

(5.9)

and substitute the estimates of these causal effect matrixes into (5.5).

The VAR–LiNGAM model generates forecasts in the same way that a traditional VAR model generates forecasts. Use historical data and parameter estimates on the right-hand side of (5.5) to estimate one-step-ahead forecasts.

### 6 PORTFOLIOS OF INDEPENDENT FACTORS

The source of variance in the VAR–LiNGAM model is the vector of disturbances \( e(t) \) in the model above; each disturbance term can be interpreted as an independent factor. Any portfolio composed of a set of securities that correspond to the VAR–LiNGAM model can alternatively be thought of as a portfolio composed of a set of independent factors. Identifying and controlling the portfolio’s exposure to the independent factors is important because they are the true sources of both risk and return in the VAR–LiNGAM model. As shown in Section 7 on applications, however, controlling a portfolio’s exposure to a particular security does not effectively control the portfolio’s exposure to that security’s underlying independent factor.

The basic VAR–LiNGAM model definition can be written in terms of the returns on a set of independent factors (Hyvärinen et al 2010):

\[ \tilde{R}(t) = (I - B_0) R(t) = \sum_{\tau=1}^{k} B_\tau R(t - \tau) + e(t). \]  

(6.1)

The independent-factor returns \( \tilde{R}(t) \) are composed of a component that contains only information from the past (\( \tau > 0 \)) and a component \( e(t) \), which is a vector of independent disturbance terms at \( \tau = 0 \). When the matrix \( B_0 \) is interpreted as a directed acyclic graph containing nodes and directed edges, each row of the matrix \((I - B_0)\) subtracts the effects of a node’s parents from the effects of the node itself, leaving only the independent factor that drives each node.
Using (6.1), the relative weights of the underlying independent factors can be identified in a portfolio of correlated securities. This can be done in a similar way to finding the new factor exposures in (1.3), by allowing the matrix \((I - B_0)\) to function as the “pick” matrix,

\[
R_P(t) = \omega^T R(t) = \omega^T (I - B_0)^{-1}(I - B_0)R(t) = \tilde{\omega}^T \tilde{R}(t).
\] (6.2)

Thus, the relative weights \(\tilde{\omega}\) on the independent factors are

\[
\tilde{\omega} = (I - B_0)^{-T} \omega.
\] (6.3)

These weights can be used in the process of portfolio optimization to control risk, as shown below.

**7 APPLICATION**

In the remainder of the paper, the VAR–LiNGAM model is used to generate forecasts for the set of the six most widely traded currencies. In addition, estimates from the model are used to find the set of the currencies’ underlying independent factors. These independent factors are used in the risk contribution framework to provide portfolio optimization constraints. Six portfolios constructed using different optimization constraints are compared on a small out-of-sample data set. The purpose of this brief study is not to find the best performing portfolio but to examine the internal properties of the portfolios that are a result of using different constraints during optimization. The constraints are used to prohibit the allocation of too much wealth to any one risk factor in the portfolio, and some constraints do this more effectively than others. A set of plots shows how effective each constraint set is at controlling each measure of risk contribution; all plots are generated using Gnuplot software.²

**8 DESCRIPTION OF THE DATA**

Data was obtained from the Sierra Chart FXCM Forex data service using Sierra Chart software.³ This data is based on transactions between the FXCM foreign exchange dealer and its clients. The data set consists of direct quotations with fifteen-minute periodicity for the Australian dollar (AUD), Canadian dollar (CAD), euro (EUR), sterling (GBP), Japanese yen (JPY) and US dollar (USD), all in terms of the Swiss franc (CHF). In other words, the exchange rates used in the study are the CHF/\(X\) exchange rates, where \(X\) represents one of the six currencies listed above. These currencies are chosen because they had the largest market turnover rates in 2010 according to the

² URL: www.gnuplot.info/.
³ URL: www.sierrachart.com/.
Triennial Central Bank Survey (Bank for International Settlements 2010). Data for the complete 2009 financial year (24,916 observations) is used for model estimation, while data on January 4, 2010 from 00:00 to 12:30 (fifty-one observations) is used for a portfolio performance comparison in which forecasts are made and then portfolios are rebalanced with portfolio optimizations based on these forecasts. Missing data is replaced by the most recent observation in each currency series. Log returns are computed by taking the natural logarithm and then first-differencing the exchange rates. The expected values and covariance matrix of the 2009 log returns are shown in Tables A-1 and A-2 in the online appendix.

9 VAR–LiNGAM ESTIMATION RESULTS

The two-stage estimation technique was used to estimate the VAR–LiNGAM model for the 2009 currency returns. The VAR estimation was performed with SAS software. An order 1 VAR was selected using the Hannan–Quinn information criterion and the Schwarz Bayesian criterion. Both criteria were negligibly better for order 2, but order 1 was chosen for parsimony. The LiNGAM procedure was then performed on the VAR residuals using the MATLAB code provided by Shimizu et al. (2006). The same code was used to prune the $B_0$ matrix at alpha level 0.01 using the Bonferroni correction technique (Shaffer 1995). The expected returns for both the VAR residuals and the independent factors are presented in Table A-1, while their respective covariance matrices are shown in Tables A-3 and A-4. The pruned $B_0$ matrix is shown in Table A-5; this matrix corresponds to the directed acyclic graph shown in Figure A-1. The $B_1$ matrix is shown in Table A-6, and the autoregressive matrix $M_1$ is shown in Table A-7. This procedure was repeated for the log returns of the complete 2010 financial year, and the resulting directed acyclic graph is shown in Figure A-2 for comparison. The directed acyclic graphs in Figures A-1 and A-2 show the instantaneous causal effects between the currencies and were generated using Graphviz software. The edges of the graphs are directed to represent causal flow. For instance, an edge $X \rightarrow Y$ indicates that a movement of currency $X$ causes a movement of currency $Y$.

As evidence that the VAR–LiNGAM non-Gaussian assumption holds, a Kolmogorov–Smirnov test performed on each currency’s corresponding independent factor confirmed that the null hypothesis of normality was rejected, with a $p$-value less than 0.01 for each factor.

---

5 See Appendix A online for the relevant tables and figures.
6 URL: www.graphviz.org/.
10 PORTFOLIO OPTIMIZATION WITH CONSTRAINTS

A variety of different constraint sets can be added to the canonical mean–variance model to achieve varying portfolio objectives. The purpose of the constraints described here is to increase portfolio diversification so that the portfolio’s risk is not concentrated too greatly in any one risk factor. In each case, the metric or attribute being constrained characterizes the diversification differently. Constraints on the following attributes are added to the canonical mean–variance model (2.7): relative portfolio weights, independent-factor relative portfolio weights, percentage marginal risk contributions and independent-factor percentage marginal risk contributions. All optimizations are performed using the AMPL algebraic modeling system.7

The relative portfolio weight constraints and the independent-factor relative-weight constraints are created by normalizing the relative weights with the sum of the risky asset weights (excluding the risk-free asset). This normalization attempts to make the portion of the portfolio invested in risky assets more diversified. The limits on the relative weights are chosen to be 0.25 so that no more than 25% of the risky portion of the portfolio is invested (long or short) in any one risky asset. This leads to the addition of the following constraint set to the canonical mean–variance model:

\[
-0.25 \leq \frac{b_i}{\sum_{i=1}^{6} b_i} \leq 0.25 \quad \text{for } i = 1, \ldots, 6,
\]

where \( b \) is defined as either \( \omega \) or \( \bar{\omega} \). This constraint set is referred to as the relative portfolio weight constraint set (WC) when \( b \) is defined as \( \omega \), and the independent-factor relative-weight constraint set (WIC) when \( b \) is defined as \( \bar{\omega} \).

Limits placed on each factor’s percentage marginal risk contribution in a portfolio optimization prohibit any one factor’s marginal risk contribution from dominating the portfolio. The limits on the percentage marginal risk contributions are chosen to be 0.25 to be consistent with the constraints on the relative portfolio weights above. The resulting constraint set is

\[
-0.25 \leq \frac{1}{\mathcal{R}(b)} \left( b_i \frac{\delta \mathcal{R}(b)}{\delta b_i} \right) \leq 0.25 \quad \text{for } i = 1, \ldots, 6,
\]

where \( b \) is defined as either \( \omega \) or \( \bar{\omega} \). For this application, the portfolio’s standard deviation is chosen as the measure of risk (3.7). This constraint set is referred to as the marginal risk contribution constraint set (MRC) when \( b \) is defined as \( \omega \), and the independent-factor marginal risk contribution constraint set (MRIC) when \( b \) is defined as \( \bar{\omega} \).

7 URL: www.ampl.com.
11 EFFICIENT FRONTIERS

In this section, an efficient frontier is constructed using the unconstrained canonical mean–variance model, and using each of the constraints discussed above appended to the canonical mean–variance model one at a time. To simplify the computation of the efficient frontiers, the risk-free rate is set equal to zero and all estimates are computed using data from the complete 2009 financial year. The portfolio’s expected return is computed using the expected value of the currency log returns,

\[ R_P = \omega^T \mathbb{E}\{ \mathbf{R}(t) \}, \quad t \in 2009, \]  

(11.1)

and the variance of the portfolio’s return is computed using the covariance matrix of the currency log returns,

\[ \sigma_P^2 = \omega^T \text{cov}\{ \mathbf{R}(t) \} \omega, \quad t \in 2009. \]  

(11.2)

The matrix \( B_0 \), which is required to compute the relative weights on the independent factors \( \tilde{\omega} \) via \( (\mathbf{I} - B_0)^{-T} \), was also estimated using data from the complete 2009 financial year. In each of the optimizations the efficient frontier is constructed by varying \( \lambda \) across the interval \([0, 1]\). The resulting frontiers reflect the year 2009 in general rather than any particular point in time.

The set of efficient frontiers is shown in Figure 1. The frontiers presented here are linear because a risk-free asset is present in the portfolio. In contrast, when a risk-free asset is not present, an efficient frontier is the upper edge of a hyperbolic region that surrounds all possible portfolios. The efficient frontier lines in Figure 1 start at the risk-free rate on the \( y \)-axis and continue through a tangency portfolio on the upper edge of their respective hyperbolic regions (not shown) that surround the set of all possible risky portfolios.

Unconstrained frontiers are ubiquitous in finance, and frontiers created with relative portfolio weight constraints are considered occasionally. Frontiers created with marginal risk contribution constraints are rarely used, and the frontiers created with independent-factor constraints are unique to this study.

There is a cost in terms of efficiency of adding constraints to the optimization. The unconstrained frontier is the most efficient, ie, has the greatest portfolio expected return per standard deviation of portfolio return, followed by (in order) the frontiers generated with the following constraint sets: WIC, MRC, MRIC and WC.

All but the WC set produce frontiers that are near to one another and relatively near to the unconstrained frontier. The WC’s large distance from the unconstrained frontier indicates that the cost of this constraint set is much greater than the cost of the others.
The frontiers are generated using constrained mean–variance optimization with data from 2009. Each frontier in the plot is labeled according to the constraint set used in its creation: the relative portfolio weight constraint (WC) set, the independent-factor relative-weight constraint (WIC) set, the marginal risk contribution constraint (MRC) set and the independent-factor marginal risk contribution constraint (MRIC) set.

12 PORTFOLIO PERFORMANCE

The performance of each of these portfolios is compared using data that spans 00:00–12:30 on January 4, 2010; the realized returns for each of the six currencies during this time period are shown in Figure A-3. It is assumed that the risk-free rate is zero and that there are no transaction costs. The VAR–LiNGAM estimation results from 2009 of $\mathbb{E}\{n(t)\}$, $\text{cov}\{n(t)\}$ and $M_1$ are used in the time series model that serves as the base for the portfolio optimizations. The 2009 estimate for $B_0$ is used to compute the relative weights on the independent factors $\tilde{\omega}$. For each period’s portfolio optimization, the most recently observed return $R(t)$ is used to compute the next period’s conditional expected value $\mathbb{E}\{R(t+1) \mid M_1, R(t)\}$, as follows:

$$
\mathbb{E}\{R(t+1) \mid M_1, R(t)\} = \mathbb{E}\{M_1 R(t) + n(t) \mid M_1, R(t)\} = M_1 R(t) + \mathbb{E}\{n(t)\}.
$$  \hspace{1cm} (12.1)

The 2009 covariance matrix of the residuals is used for each portfolio optimization and is assumed to remain constant throughout January 4, 2010.

\hspace{1cm} 

---

8 See Appendix A online for the relevant figures referred to in this section.
FIGURE 2 Portfolio performance comparison on January 4, 2010 from 00:00 to 12:30.

For every point in time in the plot, the VAR–LiNGAM model is used to generate forecasts, and then constrained mean–variance optimization is used to rebalance each portfolio. Initial portfolio wealth is 100 CHF. Each portfolio in the plot is labeled according to the constraint set used in its creation: the relative portfolio weight constraint set (WC), the independent-factor relative-weight constraint set (WIC), the marginal risk contribution constraint set (MRC) and the independent-factor marginal risk contribution constraint set (MRIC).

Each portfolio is created with $\lambda = 0.99$ in its corresponding optimization model for every time period in the simulation; this high degree of risk aversion was chosen so that the portfolios do not become too volatile. Each portfolio starts with an initial wealth in cash (unallocated to other currencies) of 100 CHF. A comparison of the portfolio performance is shown in Figure 2. The cumulative wealth of each portfolio at any particular point in time in Figure 2 can be explained by comparing the portfolio holdings shown in Figure A-4 with the realized returns shown in Figure A-3. A portfolio experiences a large change in value when it has a large position in a currency that has a large realized return. For example, in period 30 the unconstrained portfolio had a substantial loss in value (see Figure 2) because it had a large short (ie, negative) position in the euro (see Figure A-4) and the euro had a large positive return (see Figure A-3). The value of the other portfolios did not change much in period 30 because none of them had any substantial positions.

In period 31, the unconstrained portfolio made up the lost ground from the previous period by having an even larger short position in the euro and a large short position in the yen; the yen’s negative return in period 31 was large, while the euro’s was
modest. The WC, MRC and MRIC portfolios also experienced large gains in period 31 (see Figure 2) by having large short positions in the yen and euro. In period 50, all portfolios lost value because they had long positions in sterling, which had a large negative return. Overall, on this small data set no one portfolio performed markedly better than the others, but the portfolio generated with WIC did seem to lag behind.

Each of the five plots in Figures A-4–A-7 corresponds to a portfolio constructed with one of the constraints discussed above. All of the plots within a figure display the same attribute (eg, percentage marginal risk contributions in Figure A-6) for the positions contained within a portfolio.

Plots of portfolio currency allocations are shown in Figure A-4; each line in the figure represents the total value of a particular currency holding in Swiss francs. At some points in time and for all constraints except the weight constraints, the allocation to a specific currency is disproportionately large. Only with the relative weight constraints are the allocations of all currencies kept within a narrow boundary.

Independent-factor relative portfolio weights over time are displayed in Figure A-5. The WIC and WC sets appear to be the most effective at keeping the independent-factor relative portfolio weights constrained. The other constraint sets allow some of the independent-factor relative portfolio weights to become many times larger than the others; this concentration of wealth in a small number of independent factors is undesirable because it represents a loss of diversification.

Percentage marginal risk contribution plots are shown in Figure A-6, and independent-factor percentage marginal risk contribution plots are shown in Figure A-7. None of the plots in either of these figures indicates that the marginal risk contributions become as divergent as some of the independent-factor relative portfolio weights do in Figure A-5. In this small example, the WC set seems to be almost as effective as the MRC set at controlling the percentage marginal risk contributions in Figure A-6, whereas the MRIC set is the only constraint set that appears truly effective at controlling independent-factor percentage marginal risk contributions in Figure A-7.

13 CONCLUSION

When a portfolio contains a set of correlated assets, a portfolio manager’s life is complicated by the fact that risk can only be analyzed relative to an existing portfolio. However the portfolio is constructed, marginal risk analysis is confusing because changing the size of one position changes not only its marginal risk contribution but also the marginal risk contributions from the other positions. To further complicate matters, some positions have negative marginal risk contributions, so increasing such a position’s size will actually reduce the risk of the portfolio at the margin.

VAR–LiNGAM is presented in this paper as a method of finding a portfolio’s underlying independent risk factors. Such factors can be used to decompose a portfolio’s
risk so that the risk contribution from any independent factor is nonnegative. With this decomposition, the risk contribution from each independent factor can be interpreted as having a potentially adverse effect on the portfolio, and portfolios can be constructed using a risk budget. In the example application, percentage independent-factor risk contributions are constrained to be less than or equal to 25% of the total portfolio’s risk. This is equivalent to a portfolio manager deciding that no more than 25% of the portfolio’s risk should be allocated to any one independent risk factor. This decision takes place before the portfolio is created, and the percentage risk allocations are not relative to an existing portfolio, as they would be when risk factors are not independent.

In Section 7, we found the market structure underlying a set of currencies, and from this we identified the set of independent factors. Portfolios are constructed on a sample data set by constraining the portfolio weights and the percentage marginal risk contributions of both the original assets and the independent factors during mean–variance optimizations. The properties of the portfolios and efficient frontiers created using constraints on the independent risk factors compare favorably with those created using constraints on the original assets. Plots of the independent-factor percentage marginal risk contributions show that the best way to control independent-factor marginal risk contributions is to put constraints on the independent-factor risk marginal contributions themselves.

DECLARATION OF INTEREST

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

ACKNOWLEDGEMENTS

Thanks to David A. Bessler, James W. Mjelde, David J. Leatham and an anonymous reviewer for helpful comments that improved the quality of the paper.

REFERENCES


Research Paper

Optimal asset management for defined-contribution pension funds with default risk

Shibo Bian,¹ James Cicon² and Yi Zhang³

¹School of Finance, Shanghai Lixin University of Accounting and Finance, Shanghai 201620, People's Republic of China; email: bsb@lixin.edu.cn
²Harmon School of Business and Professional Studies, University of Central Missouri, Warrensburg, MO 64093, USA; email: cicon@ucmo.edu
³School of Public Economics and Administration, Shanghai University of Finance and Economics, Shanghai 200433, People's Republic of China; email: zhang.yi@mail.shufe.edu.cn

(Received September 17, 2015; revised December 4, 2015; accepted January 25, 2016)

ABSTRACT

We explore how a defined-contribution pension fund optimally distributes wealth between a defaultable bond, a stock and a bank account, given that a salary is a stochastic process. We assume that the investment objective of the defined-contribution pension fund is to maximize the expected constant relative risk aversion utility of terminal wealth. We thus obtain a closed-form solution to the optimal problem using a martingale approach. We develop numerical simulations, which we graph as illustrations. Finally, we discuss relevant economic insights obtained from our results.

Keywords: defined-contribution pension plan; optimal investment; defaultable bond; stochastic salary; martingale approach.
1 INTRODUCTION

The Latin American debt crisis, the US subprime crisis and the European sovereign debt crisis serve to remind us that default risk is a major issue facing investors. Furthermore, these crises emphasize the importance of incorporating the default feature into portfolio management models. Consequently, interest in the defaultable market-optimal portfolio problem has increased.

Korn and Kraft (2003) studied the optimal portfolio in the context of defaultable bonds and price the defaultable bond using Merton’s (1974) structural approach. They use elasticity and duration as control variables in their optimization problem, expressing the optimal amounts invested in each security as functions of these two variables. The paper by Kraft and Steffensen (2006) extends their work within a “first passage time” framework (Black and Cox 1976).


The crises demonstrate that contagion risk plays a fundamental role in financially distressed situations. The following representative studies analyze contagion issues. Kraft and Steffensen (2008) analyze the impact of joint default risk on the investor’s portfolio decision, where the investor wishes to invest in defaultable bonds under the multiple default framework of Schönbucher (1998). In a subsequent study, Kraft and Steffensen (2009) consider the impact of contagion on investor demand for defaultable bonds. Jiao et al (2013) combine stochastic control and backward stochastic differential equation decomposition. This approach permits them to provide generalized formulas that explain the terminal wealth and the optimal investment strategy with defaultable bonds. Bo and Capponi (2014) use credit default swaps to analyze how contagion risk affects the optimal allocation decision.
These representative studies leave a gap in the literature. They focus on the investor who has a self-financing investment strategy. However, they do not address the fact that many investors in defaultable bonds pursue a non-self-financing investment strategy. Such investors include pension funds, insurers and mutual funds.

We aim to fill this gap in the literature. Our paper shows how a defined-contribution (DC) pension fund can invest optimally in defaultable bonds. We show that the DC pension fund’s investment strategy is not self-financing because it experiences continuous contributions. Some studies use a stochastic framework to investigate the DC pension fund’s optimal investment strategy. Such studies include Boulier et al (2001), Haberman and Vigna (2002), Deelstra et al (2004), Menoncin and Scaillet (2006), Zhang et al (2007) and Gao (2008). To the best of our knowledge, we are the first to study the optimal investment problem of a DC pension fund whose investment opportunity set includes a default security.

In this paper, we consider a unique investment optimization problem: a DC pension fund operating in a financial market that provides a defaultable bond, a risky stock asset and a bank account. The fund manager may dynamically share their wealth between the three aforementioned assets. They seek to maximize terminal wealth under a finite horizon of expected utility. We use the martingale approach to obtain a closed-form solution to this optimization problem.

The paper is organized as follows. In Section 2 we set up the mathematical framework of our financial market model. In Section 3 we introduce our portfolio-optimization problem. In Section 4 we carry out our numerical analysis. Section 5 summarizes our results and concludes our paper. Technical details appear in the online appendixes.

2 OUR FINANCIAL MARKET MODEL

We begin building our model by defining our complete probability space, \((\Omega, \mathcal{F}_t, \mathbb{Q})\), where \(\mathbb{Q}\) denotes a risk-neutral (martingale probability) measure which is assumed to be equivalent to the real-world (statistical) measure \(\mathbb{P}\). Next we introduce \(H(t)\), which is a default stochastic process that jumps at a random time \(\tau\). Let \(H(t) = 1_{\{\tau < t\}}\), where \(1_{\{\tau < t\}}\) is the indicator function, which is 0 before the default event and 1 otherwise.

We assume that the pension fund’s investment set is a defaultable bond, a stock and a bank account. In the following we describe the dynamics between these securities and wealth.

2.1 The dynamics of financial securities

The promised principal of the defaultable zero-coupon bond is 1. The maturity date is \(T_1\), and the default time is \(\tau\). Following Duffie and Singleton (1999) we define the
valuation formula for a defaultable zero-coupon bond with price \( p(t, T_1) \) as

\[
p(t, T_1) = 1_{\{\tau > t\}} \times \mathbb{E}^Q \left( \exp \left( - \int_t^{T_1} (r + h^Q_s) \, ds \right) \bigg| \mathcal{F}_t \right),
\]

(2.1)

where \( r \) is risk-free rate, \( h^Q \) is default intensity under \( Q \) and \( \iota \) is the loss rate. We assume that all are constant.

Applying Ito’s lemma and \( H(t) = 1_{\{\tau \leq t\}} \) to (2.1), we derive the dynamics of the price process of the defaultable bond:

\[
dp(t, T_1) = p(t, T_1)(r \, dt - (1 - H(t)) \, dM^Q(t)),
\]

(2.2)

where the compensated jump martingale process \( M^Q(t) \) is given by

\[
dM^Q(t) = dH(t) - (1 - H(t)) \, dW^Q(t).
\]

We assume that the stock price and the bank account are given by the following diffusion equation:

\[
dS(t) = S(t)(r \, dt + \sigma_S\, dW^Q(t)),
\]

(2.3)

\[
dB(t) = rB(t) \, dt,
\]

(2.4)

where \( \sigma_S \) is the volatility of the stock and \( W^Q(t) \) is a Brownian motion under \( Q \).

### 2.2 The dynamics of wealth

Contributions payable by both employee and employer are defined under a DC pension plan. In most cases they correspond to a fixed percentage of the salary. Following Zhang et al (2007), we assume that the salary of a pension plan member follows the dynamics described below:

\[
dl(t) = L(t)(\kappa \, dt + \sigma_{L_1}\, dW^Q(t) + \sigma_{L_2}\, dM^Q(t)),
\]

(2.5)

where \( \sigma_{L_1} \) and \( \sigma_{L_2} \) are two volatility scale factors that measure how the risk sources of the stock and the defaultable bond affect the salary, and \( \kappa \) is the expected growth rate of the salary. We assume that \( \sigma_{L_1}, \sigma_{L_2} \) and \( \kappa \) are constants.

We assume that the initial value of the pension account is \( x > 0 \), that the contribution rate is \( c \) and that \( \pi_p(t), \pi_S(t), 1 - \pi_S(t) - \pi_p(t) \) are the proportions of the pension fund invested in the defaultable bond, the stock and the bank account respectively. Thus, the corresponding portfolio process \( X^\pi(t) \) is governed by the following equation:

\[
dX^\pi(t) = X^\pi(t)(\pi_S(t)\, dS(t) + \pi_p(t)\, dp(t, T_1) + (1 - \pi_S(t) - \pi_p(t))\, dB(t))
\]

\[+ cL(t) \, dt. \]

(2.6)
Substituting (2.2)–(2.4) into (2.6), we get
\[ dX^\pi(t) = X^\pi(t)(r\, dt + \pi_S\sigma\, dW^Q(t) - (1 - H(t)))\pi_p\, dM^Q(t)) + cL(t)\, dt. \quad (2.7) \]

Note that contributions are assumed to be invested continuously over time. The expectation of the plan member’s future contribution enables the pension fund to invest the plan member’s contributions optimally. Therefore, we can state the following definitions.

**Definition 2.1** Following Bodie et al (1992) we define the discounted expected future contribution process as
\[ D(t) = \mathbb{E}^Q\left( \int_t^T \frac{B(s)}{B(t)} cL(s) \, ds \mid \mathcal{F}_t \right). \]
We set \( D(0) = d. \)

**Definition 2.2** Following Bodie et al (1992) we define the pension fund wealth process \( V(t) \) as
\[ V(t) = X^\pi(t) + D(t). \quad (2.8) \]

Equation (2.8) should be interpreted as follows: the wealth of the pension fund at time \( t \) is equal to the value of the portfolio \( X^\pi(t) \) plus the discounted expected value of future contributions to the plan.

The process \( D(t) \) is not directly observable. Therefore, it is not certain if the pension fund is likely to base its investment decision upon it. However, an observable alternative is the salary process \( L(t) \). In Lemma 2.3, we show that the processes \( D(t) \) and \( L(t) \) are strongly linked.

**Lemma 2.3** The expected future contributions process \( D(t) \) and the salary process \( L(t) \) are related via the equation
\[ D(t) = \frac{1}{\theta} (\exp(\theta(T - t)) - 1)cL(t), \quad (2.9) \]
where \( \theta = \kappa - r. \)

**Proof** See Appendix A online. \( \square \)

Taking derivatives of (2.9) and using (2.5), we have
\[ dD(t) = D(t)(r\, dt + \sigma_{L_1} dW^Q(t) + \sigma_{L_2} dM^Q(t)) - cL(t)\, dt. \quad (2.10) \]

Taking derivatives of (2.8) and using (2.7) and (2.10), we have
\[ dV(t) = V(t) \left( r\, dt + \frac{(X^\pi(t)\pi_S\sigma_S + D(t)\sigma_{L_1}) dW^Q(t) + (X^\pi(t)((H(t) - 1)\pi_p\tau + D(t)\sigma_{L_2}) dM^Q(t)}{V(t)} \right). \]
Thus, under $Q$, the discounted wealth process is a martingale:

$$
\begin{align*}
\frac{d}{dt} \left( \frac{V(t)}{B(t)} \right) &= \frac{V(t)}{B(t)} \left( \frac{(X^\pi(t)\pi S + D(t)\sigma_{L_1}) dW^Q(t)}{V(t)} 
+ (X^\pi(t)((H(t) - 1)\pi_p + D(t)\sigma_{L_2}) dM^Q(t)} \right).
\end{align*}
$$

Therefore, we write

$$
\mathbb{E}^Q \left( \frac{V(t)}{B(t)} \right) = \frac{V(0)}{B(0)} = v = x + d.
$$

We change the measure from $Q$ to $\mathbb{P}$. This is necessary because the pension fund optimizes utility under the real-world probability measure. Following Girsanov’s theorem we use the Radon–Nikodým density martingale $Z(t) = dQ/d\mathbb{P}$ and

$$
\mathbb{E}^\mathbb{P} \left( Z(t) \frac{V(t)}{B(t)} \right) = \mathbb{E}^Q \left( \frac{V(t)}{B(t)} \right) = v,
$$

where $\mathbb{E}^\mathbb{P}$ is the expectation under $\mathbb{P}$. $Z(t)$ should satisfy

$$
\mathbb{E}^\mathbb{P}(Z(t)) = 1, \quad Z(t) = Z_W(t)Z_H(t),
$$

where

$$
\begin{align*}
Z_W(t) &= \exp \left( -\int_0^t \lambda \, dW(s) - \frac{1}{2} \int_0^t \lambda^2 \, ds \right), \\
Z_H(t) &= \exp \left( \int_0^t \ln \mu \, dH(s) - \int_0^t h^\mathbb{P}(\mu - 1)(1 - H(s)) \, ds \right).
\end{align*}
$$

$\lambda$ is the market price of risk for the stock, $\mu = h^\mathbb{P}/h^Q$ is the jump-risk premium, $h^\mathbb{P}$ is the default intensity under $\mathbb{P}$,

$$
W(t) = W^Q(t) - \int_0^t \lambda \, ds
$$

is a Brownian motion under $\mathbb{P}$, and

$$
M(t) = M^Q(t) - h^\mathbb{P} \int_0^t (\mu - 1)(1 - H(s)) \, ds
$$

is a martingale under $\mathbb{P}$.

### 3 SOLUTION TO THE OPTIMIZATION PROBLEM

The pension fund will maximize the expected utility of terminal wealth. We define our utility function as $U(v)$. We define $J(v)$ as the maximal utility attained by implementing the optimal investment strategy $\pi^*$. We may now define our optimization
problem as

\[ J(v) = \sup_{\pi} \mathbb{E}^\mathbb{P}(U(V(T))) \quad \text{such that} \quad \mathbb{E}^\mathbb{P}\left( Z(t) \frac{V(t)}{B(t)} \right) = v. \]  

(3.1)

Following Cox and Huang (1989) and Duffie (2010), we use the martingale approach to solve our optimization problem.

### 3.1 Maximizing general utility

We assume that the utility function \( U \) is strictly increasing and strictly concave. It therefore follows that the derivative \( U' \) is strictly decreasing and continuous. It also follows that its inverse function, \( I(Y) \), is strictly decreasing and continuous:

\[
I(U'(V)) = V, \quad 0 < V < \infty,
\]

\[
U'(I(Y)) = Y, \quad 0 < Y < U'(0).
\]

The convex dual of \( U \) is the function defined by \( \mathcal{U}(Y) \triangleq \sup_{V \in \mathbb{R}} (U(V) - VY) \). \( \mathcal{U} \) reaches its maximum at \( V = I(Y) \), and

\[
U(V) - VY \leq U(I(Y)) - YI(Y). \quad (3.2)
\]

We interpret \( Y(v) \) as a Lagrange multiplier. We reduce the constrained maximization problem of (3.1) to the unconstrained problem

\[
\mathbb{E}^\mathbb{P}(U(V(T))) + Y(v)(v - \mathbb{E}^\mathbb{P}(N(T)V(T))), \quad (3.3)
\]

where \( N(t) = Z(t)/B(t) \). Applying the result of (3.2), we have

\[
vY(v) + \mathbb{E}^\mathbb{P}(U(V(T)) - Y(v)N(T)V(T)) \leq vY(v) + \mathbb{E}^\mathbb{P}(U(I(Y(v))N(T)) - Y(v)N(T)I(Y(v)N(T))) \quad (3.4)
\]

with equality if and only if

\[
V(T) = I(Y(v)N(T)). \quad (3.5)
\]

Then \( V(T) \) is the optimal terminal wealth.

Using \( V(T) \), the martingale property of \( N(t)V(t) \) and Ito’s lemma, we get an alternative expression for \( d(V(t)/B(t)) \). Comparing this to the corresponding terms in (2.11), we identify the expressions for the optimal portfolio \( \pi^* \). We summarize our results in Proposition 3.1.
Proposition 3.1 Let \( v \) be the endowment for the pension. Let \( V(T) \) be terminal wealth at time \( T \). There exists an optimal investment strategy \( \pi^* \):

\[
\frac{(X \pi^*(t) \pi_S^* \sigma_S + D(t) \sigma_{L_1})}{V(t)} V(t) \frac{B(t)}{E^P(N(T) I(Y(v) N(T)))} = \frac{\phi_1(t) + \eta(t) \lambda}{Z(t)},
\]

\[
\frac{(X \pi^*(t) ((H(t) - 1) \pi_p^* t) + D(t) \sigma_{L_2})}{V(t)} V(t) \frac{B(t)}{E^P(N(T) I(Y(v) N(T)))} = \frac{\phi_2(t) + (1 - \mu) \eta(t)}{\mu Z(t)}.
\]

where

\[
\eta(t) = E^P(N(T) I(Y(v) N(T)) \mid \mathcal{F}_t),
\]

\[
d\eta(t) = \phi_1(t) \, dW(t) + \phi_2(t) \, dM(t).
\]

Proof See Appendix B online.

Proposition 3.1 gives a general solution. In Section 3.2, we assume a constant relative risk aversion (CRRA) utility, which we use to derive the closed-form optimal investment strategy.

### 3.2 Maximizing CRRA utility

Given the pension fund with a CRRA utility function as follows:

\[
U(V) = \frac{V^\gamma}{\gamma}, \quad 0 < \gamma < 1.
\]

The inverse function of \( U' \) is then given by

\[
I(y) = y^{1/(\gamma - 1)}.
\]

Using (2.12), (3.5) and (3.10), we derive

\[
v = E^P(N(T) I(Y(v) N(T))) = Y(v)^{1/(\gamma - 1)} E^P(N(T)^{\gamma/(\gamma - 1)})
\]

and obtain

\[
Y(v) = \left( \frac{v}{E^P(N(T)^{\gamma/(\gamma - 1)})} \right)^{1/(\gamma - 1)}.
\]

Substituting (3.12) into (3.5) and using (3.10), we have the optimal terminal wealth:

\[
V(T) = I(Y(v) N(T)) = \frac{v}{E^P(N(T)^{\gamma/(\gamma - 1)})} N(T)^{1/(\gamma - 1)}.
\]
Substituting (3.13) and $N(T)$ into (3.8) for a CRRA utility yields

$$
\eta(t) = \frac{v}{\mathbb{E}^P(N(T)^{\gamma/(\gamma-1)})} \mathbb{E}^P(N(T)^{\gamma/(\gamma-1)} | \mathcal{F}_t)
$$

$$
= \frac{v N(t)^{\gamma/(\gamma-1)}}{\mathbb{E}^P(N(T)^{\gamma/(\gamma-1)})} \mathbb{E}^P\left( \left( \frac{N(T)}{N(t)} \right)^{\gamma/(\gamma-1)} | \mathcal{F}_t \right). \tag{3.14}
$$

Applying Ito’s lemma to (3.14) we have

$$
\frac{d\eta(t)}{\eta(t-)} = \frac{dN(t)^{\gamma/(\gamma-1)}}{N(t)^{\gamma/(\gamma-1)}} + \frac{d\Omega(t)}{\Omega(t-)}, \tag{3.15}
$$

where

$$
\Omega(t) = \mathbb{E}^P\left( \left( \frac{N(T)}{N(t)} \right)^{\gamma/(\gamma-1)} | \mathcal{F}_t \right). \tag{3.16}
$$

Our last step in determining the optimal portfolio is to find (3.15). Lemmas 2.3 and 3.2 provide the explicit expressions for

$$
\frac{dN(t)^{\gamma/(\gamma-1)}}{N(t)^{\gamma/(\gamma-1)}} \quad \text{and} \quad \frac{d\Omega(t)}{\Omega(t-)},
$$

respectively. Given this, we now have $d\eta(t)/\eta(t-)$.

Comparing this result with (3.9) yields $\phi(t)$. Finally, by substituting the $\phi(t)$ into Proposition 3.1, we obtain the explicit optimal portfolio.

**Lemma 3.2** The derivative of $N(t)^{\gamma/(\gamma-1)}$ is given by

$$
\frac{dN(t)^{\gamma/(\gamma-1)}}{N(t)^{\gamma/(\gamma-1)}} = \frac{\gamma}{(1-\gamma)}(\lambda \, dW(t)) + (\mu^{\gamma/(\gamma-1)} - 1) \, dM(t) + [\] \, dt. \tag{3.17}
$$

**Proof** See Appendix C online.

**Lemma 3.3** Given that $W(t)$ and $H(t)$ are independent of each other, the derivative of $\Omega(t)$ is given by

$$
\frac{d\Omega(t)}{\Omega(t-)} = \left( \mu^{\gamma/(\gamma-1)} \right.

\times \left( \mu^{\gamma/(\gamma-1)}(1 - \exp(P(T-t))) \right.

+ \exp \left( \frac{\gamma}{\gamma-1} (1-\mu)h^P + h^P(T-t) \right) \right)^{-1} - 1 \) \, dM(t) + [\] \, dt. \tag{3.18}
$$

**Proof** See Appendix D online.
Proposition 3.4  When the defaultable bond is not in default, the optimal strategy of the DC pension fund with the CRRA utility function is given by

$$\pi^*_p = \frac{1}{\mu} \left( \mu - \mu^{\gamma/(\gamma-1)} \left( \exp\left( \frac{\gamma}{\gamma-1}(1-\mu)h^p + h^p(T-t) \right) + \mu^{\gamma/(\gamma-1)}(1 - \exp(h^p(T-t))) \right) + 1 \right) \left( 1 + \frac{D(t)}{X^{\pi^*(t)}} \right)$$

$$- \sigma_{L2} \frac{D(t)}{X^{\pi^*(t)}},$$

(3.19)

$$\pi^*_S = \frac{1}{1 - \gamma/\sigma_S} \left( 1 + \frac{D(t)}{X^{\pi^*(t)}} \right) - \frac{\sigma_{L1}}{\sigma_S} \frac{D(t)}{X^{\pi^*(t)}}.$$  

(3.20)

Proof  Using Proposition 3.1, Lemma 3.2 and Lemma 3.3, we directly obtain Proposition 3.4.

From (3.19), it is clear that the optimal investment strategy for the defaultable bond is a decreasing function of the loss rate \(\epsilon\) and the volatility of the salary \(\sigma_{L2}\). The higher the loss rate, the smaller the investor recovery. Holding other factors constant, the pension fund allocates a smaller amount of wealth to the defaultable bond. As the volatility of the salary increases, the pension fund invests smaller amounts of wealth in the defaultable bond. We cannot directly observe the influence of the jump-risk premium \(\mu\), the default intensity \(h^p\) and the time to maturity on the optimal investment strategy for the defaultable bond. Therefore, in the following numerical simulation we analyze the effect of these factors in the pension fund’s decision.

From (3.20) we know that the optimal investment of the stock is similar to that of the defaultable bond. This allows us to proceed to our simulation.

4 SIMULATION

In this section, we develop a numerical simulation that demonstrates the influence of the jump-risk premium \(\mu\), the default intensity \(h^p\) and the time to maturity of the optimal investment strategy for the defaultable bond. The initial values of our parameters are given in Table 1.

Figure 1 provides evidence that there is a positive relationship between optimal investment in the defaultable bond and the jump-risk premium \(\mu\). Holding other factors constant, the pension fund purchases more defaultable bonds when the price of the jump risk is higher. The optimal investment in the defaultable bond increases with a decreasing rate as the jump-risk premium \(\mu\) increases. When the default intensity

Journal of Risk  
www.risk.net/journal

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
TABLE 1 Parameters for our numerical experiment.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
<th>Numerical value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mu)</td>
<td>Jump-risk premium</td>
<td>1–3</td>
</tr>
<tr>
<td>(h_p)</td>
<td>Default intensity</td>
<td>0.01–0.03</td>
</tr>
<tr>
<td>(t)</td>
<td>Loss rate</td>
<td>0.5</td>
</tr>
<tr>
<td>(T - t)</td>
<td>Time to maturity</td>
<td>1–10</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>Risk aversion coefficient</td>
<td>0.5</td>
</tr>
<tr>
<td>(r)</td>
<td>Risk-free rate</td>
<td>0.03</td>
</tr>
<tr>
<td>(\kappa)</td>
<td>Expected growth of salary</td>
<td>0.06</td>
</tr>
<tr>
<td>((\sigma_{L_1}, \sigma_{L_2}))'</td>
<td>Volatility of salary</td>
<td>(0.05,0.01)'</td>
</tr>
<tr>
<td>(c)</td>
<td>Contribution rate</td>
<td>0.15</td>
</tr>
</tbody>
</table>

FIGURE 1 Optimal investment of the defaultable bond versus the jump-risk premium and the default intensity.

\(h_p\) is low, the marginal increase in the optimal investment strategy for the defaultable bond is huge, and vice versa. We interpret our result with the following in mind: the pension fund responds more to the investment return when the default intensity \(h_p\) is low, and responds more to the investment risk when the default intensity \(h_p\) is high. The optimal investment strategy for the defaultable bond is a decreasing function of the default intensity \(h_p\). Greater default intensity \(h_p\) leads to a higher probability of default, thus the pension fund allocates smaller amounts of wealth to the defaultable bond.
We provide evidence in Figure 2 that the optimal investment strategy for the defaultable bond is an increasing function of time to maturity. The pension fund allocates more to defaultable bonds when the investment horizon is longer. When the time to maturity is long, the marginal increase in the optimal investment strategy for the defaultable bond becomes huge, and vice versa. We offer the following explanation: the fund manager permits the choice of a more aggressive investment policy and responds more to the investment return when the time to maturity is long. As the deadline approaches, the fund manager tends toward a more conservative investment policy.

5 CONCLUSIONS

This paper analyzed the problem of optimal investment in the defaultable bond, given an investor with a non-self-financing strategy. We showed how to optimally allocate a DC pension fund’s wealth between a defaultable bond, a stock and a bank account (given a salary is a stochastic process). We used a reduced-form approach to model the defaultable bond, which allowed us to give the dynamics of the defaultable bond. We assumed that the pension fund’s utility was CRRA and used the martingale approach to obtain a closed-form solution to this optimal problem. Our results provide evidence that optimal investment in the defaultable bond is an increasing function of the jump-risk premium and the time to maturity. The optimal investment strategy for
the defaultable bond is decreasing in the default intensity, in the loss rate and in the volatility of salary.

DECLARATION OF INTEREST

This work was supported by the National Natural Science Foundation of China, grant 71301105, 71503154. The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

REFERENCES


Research Paper

A fuzzy data envelopment analysis model for evaluating the efficiency of socially responsible and conventional mutual funds

I. Baeza-Sampere,¹ V. Coll-Serrano,¹ B. M’Zali² and P. Méndez-Rodríguez³

¹Department of Applied Economics, University of Valencia, Avenida Dels Tarongers s/n, 46022 Valencia, Spain; emails: Ismael.Baeza@uv.es, Vicente.Coll@uv.es
²Department of Stratégie, Responsabilité Sociale et Environnementale, Université du Québec à Montréal, Case Postale 8888, Succursale Centre-Ville, Montréal, Canada; email: mzali.bouchra@uqam.ca
³Department of Quantitative Economics, University of Oviedo, Avenida del Cristo s/n, 33006 Oviedo, Spain; email: mpmendez@uniovi.es

(Received March 10, 2015; revised September 28, 2015; accepted September 28, 2015)

ABSTRACT

Although several data envelopment analysis (DEA) models have been proposed in the literature for mutual funds’ performance evaluation, few of them incorporate nonfinancial criteria. In this paper a fuzzy DEA model is used, allowing mutual funds relative performance evaluation in a more realistic and flexible way. We examine the efficiency of forty US large cap equity mutual funds based not only on financial variables but also on nonfinancial ones. To achieve this aim, we extend Basso and Funari’s mutual funds’ ethical level proposing a more reliable fuzzy measure of the social environmental responsibility degree of equity mutual funds. It relies on the corporate social performance of the companies invested in by the mutual funds and on the quality of the management in terms of the transparency and credibility degree of the nonfinancial information provided by the mutual funds. We can conclude that
socially responsible mutual funds show better behavior in terms of efficiency than conventional funds.

**Keywords:** socially responsible investment; equity mutual funds; fuzzy data envelopment analysis (DEA); efficiency; quality management; social environmental responsibility.

## 1 INTRODUCTION

Sustainable and responsible investing (SRI) takes into account not only financial performance but also the investment’s impact on society. The ways in which SRI impacts society are numerous. Some examples are the use of active share ownership and engagement strategies; community investing; or the influence on national and global public policies and the development of global standard-setting organizations as the Carbon Disclosure Project, Ceres, the Council of Institutional Investors or the Global Reporting Initiative.

SRI contributes to the improvement of the Environmental Social and Governance (ESG) practices of companies to generate both positive societal impact and long-term competitive financial returns (Social Investment Forum 2012). Recently, given the causes of the 2008 financial crisis, ESG concerns have become more and more relevant and SRI is gaining adherents even in investment firms that have not historically identified themselves as SRI firms. As an example, the number of signatories of the Principles for Responsible Investment (PRI) is estimated to represent around 20% of the estimated total value of global capital markets.

As reported by the Social Investment Forum in its 2010 report (Social Investment Forum 2010), social investing in the US market enjoyed a growth rate of more than 13% from 2007 to 2010 with 250 socially screened mutual fund products in 2010 with assets of US$316.1 billion. SRI is therefore a growing and important investment field and, consequently, there is an increasing body of academic literature that pays attention to it. Most of the research is focused on the discussion about the financial performance of SRI, in particular of invested funds or mutual funds (see Renneboog et al 2008, for a review of the academic literature on this topic). Very few studies evaluate the performance in terms not only of financial but also nonfinancial criteria. Moreover, few of them can be found analyzing their efficiency.

In this paper, we use data envelopment analysis (DEA) to assess the relative efficiency of a sample of US equity mutual funds. This paper presents two main contributions. First, it extends Pérez-Gladish et al’s social environmental responsibility measurement (Pérez-Gladish et al 2013) proposing a more reliable indicator of the environmental responsibility degree of equity mutual funds which relies on the Corporate Social Performance of the companies invested in by the mutual funds and on the transparency and credibility degree of the nonfinancial information provided by
The efficiency of socially responsible and conventional mutual funds

The main advantage of the proposed measure is that it includes the uncertain and/or imprecise nature of some of the considered nonfinancial inputs and outputs. Secondly, by incorporating the fuzzy measure of social environmental responsibility as an input, it extends Basso and Funari’s DEA model (Basso and Funari 2007, 2010, 2012) giving rise to a fuzzy DEA model, which permits mutual funds relative performance evaluation in a more realistic and flexible way.

The rest of the paper is organized as follows. Section 2 describes the fuzzy DEA model used to measure the efficiency scores of US equity mutual funds. Section 3 specifies the data and variables used in the application. Section 4 presents the main results obtained. Finally, Section 5 shows the main conclusions.

2 METHODOLOGY

Data envelopment analysis (Charnes et al 1994) is a nonparametric technique that allows evaluation of the relative efficiency of a set of decision-making units (DMUs). Conventional DEA models assume data are precisely measured, but this assumption is not always acceptable. In this work we assume the uncertainty and/or imprecision inherent to certain inputs and outputs which we handle through fuzzy numbers.

Let $U$ be a space of points, a fuzzy set in $U$ is $\tilde{A} = \{(x, \mu_\tilde{A}(x)), x \in U\}$, where $\mu_\tilde{A} : U \rightarrow [0, 1]$ is a function. For each $x \in U$, the value $\mu_\tilde{A}(x)$ is called the grade of membership of $x$ in $\tilde{A}$ (Zadeh 1965). A fuzzy number is a convex and normalized ($\sup_{x \in U} \mu_\tilde{A}(x) = 1$) fuzzy set with a piecewise continuous membership function. The most used fuzzy numbers are known as LR-fuzzy numbers (Dubois and Prade 1988). A fuzzy number $\tilde{M}$ is said to be an LR-fuzzy number, $\tilde{M} = (m^L, m^R, \alpha^L, \alpha^R)_{L,R}$, if its membership function has the following form (León et al 2003; Vercher et al 2007):

$$\mu_\tilde{M} = \begin{cases} L\left( \frac{m^L - r}{\alpha^L} \right), & r \leq m^L, \\ 1, & m^L \leq r \leq m^R, \\ R\left( \frac{r - m^R}{\alpha^R} \right), & r \geq m^R, \end{cases}$$

(2.1)

where $L$ and $R$ are reference functions, that is, $L, R : [0, +\infty[ \rightarrow [0, 1]$ are strictly decreasing functions in $\text{supp}(\tilde{M}) = \{r: \mu_\tilde{M}(x) > 0\}$ and upper semi-continuous such that $L(0) = R(0) = 1$. If $L$ and $R$ are linear functions, the fuzzy number is said to be an LR-fuzzy trapezoidal. The triangular fuzzy number, $\tilde{M} = (m, \alpha^L, \alpha^R)_{L,R}$, is a special case of the trapezoidal one, where the upper base degenerates to a point (crisp value).

Fuzzy mathematical programming provides a useful tool to deal with this kind of data (León et al 2003). Hatami-Marbini et al (2011) provides a taxonomy and
review of the fuzzy DEA methods. More recently, several research articles and books covering this topic have been published (Zerafat-Angiz et al. 2013, 2012; Emrouznejad and Tavana 2014; Puri and Yadav 2014; Azadi et al. 2015).

Considering \( n \) DMUs \((j = 1, 2, \ldots, n)\), each using \( m \) fuzzy inputs \((\tilde{x}_{ij}, i = 1, \ldots, m)\) to produce \( s \) fuzzy outputs \((\tilde{y}_{rj}, r = 1, \ldots, s)\), a basic fuzzy Banker, Charnes and Cooper (BCC) input-orientated DEA model can be written as

\[
\min \theta \\
\text{such that} \\
\sum_{j=1}^{n} \lambda_j \tilde{x}_{i,j} \leq \theta \tilde{x}_{i,0}, \quad 1 \leq i \leq m, \\
\sum_{j=1}^{n} \lambda_j \tilde{y}_{r,j} \geq \tilde{y}_{r,0}, \quad 1 \leq r \leq s, \\
\sum_{j=1}^{n} \lambda_j = 1, \\
\lambda_j \geq 0, \quad 1 \leq j \leq n
\]

where \( \tilde{x}_{i0} \) and \( \tilde{y}_{r0} \) denote, respectively, the \( i \)th fuzzy input and \( r \)th fuzzy output of the DMU evaluated, and \( \theta \) its efficiency score.

The problem is that the efficiency of a DMU is given by a fuzzy number with membership function \( \tilde{\mu}_{E_j}, 1 \leq j \leq n \). In the fuzzy DEA literature, several approaches to calculate \( \tilde{\mu}_{E_j} \) can be found. We follow Kao and Liu’s approach (Kao and Liu 2000) (see (2.3)). In this model, for a set of \( n \) DMUs, denote by \( \tilde{x}_{ij} = (x_{ij}^L, x_{ij}^R, \alpha_{ij}^L, \alpha_{ij}^R)_{L_{ij}, R_{ij}} \) \((i = 1, \ldots, m)\) and \( \tilde{y}_{rj} = (y_{rj}^L, y_{rj}^R, \alpha_{rj}^L, \alpha_{rj}^R)_{L_{rj}, R_{rj}} \) \((r = 1, \ldots, s)\), respectively, the fuzzy inputs and fuzzy outputs from the \( j \)th DMU \((j = 1, \ldots, m)\). For a given \( \alpha \in [0, 1] \) the efficiency score \( E_{j0}(\alpha) \) of the \( j \)th DMU is given by means of an interval: \( E_{j0}(\alpha) = [E_{j0}^L(\alpha), E_{j0}^U(\alpha)] \). For calculating the intervals, Kao and Liu propose the following auxiliary models:

\[
E_{j0}^L(\alpha) = \min \theta \\
\text{such that} \\
\sum_{j=1, j \neq j_0}^{n} \lambda_j x_{i,j}(\alpha) + \lambda_{j_0} x_{i,j_0}(\alpha) \leq \theta x_{i,0}, \quad 1 \leq i \leq I, \\
\sum_{j=1, j \neq j_0}^{n} \lambda_j y_{r,j}(\alpha) + \lambda_{j_0} y_{r,j_0}(\alpha) \geq y_{r,0}, \quad 1 \leq r \leq R, \\
\sum_{j=1}^{n} \lambda_j = 1, \\
\lambda_j \geq 0, \quad 1 \leq j \leq n
\]
and

\[ E^U_{j_0}(\alpha) = \min \theta \]

such that

\[ \sum_{j=1, j \neq j_0}^n \lambda_j x_{i,j}^U(\alpha) + \lambda_{j_0} x_{i,j_0}^L(\alpha) \leq \theta x_{i,0}^L, \quad 1 \leq i \leq I, \]

\[ \sum_{j=1, j \neq j_0}^n \lambda_j y_{r,j}^L(\alpha) + \lambda_{j_0} y_{r,j_0}^U(\alpha) \geq y_{r,0}^U, \quad 1 \leq r \leq R, \]

\[ \sum_{j=1}^n \lambda_j = 1, \]

\[ \lambda_j \geq 0, \quad 1 \leq j \leq n, \]

where \( \theta \) is the efficiency of the unit being evaluated \((j_0)\) and \( \lambda_j \) are the weights from which the value \( \theta \) is obtained.

However, the efficiency scores obtained from (2.3) cannot be used to rank units. If a unit shows a lower efficiency ratio, this does not guarantee that another unit should have priority over it (Boscá et al. 2011). In order to solve this problem and to rank units, we propose using Chen and Klein’s index (Chen and Klein 1997), which is given by

\[ CK(\hat{E}_j) = \frac{\sum_{l=0}^N (E_j^U(\alpha_l) - c)}{\sum_{l=0}^N (E_j^U(\alpha_l) - c) - \sum_{l=0}^N (E_j^L(\alpha_l) - d)}, \quad j = 1, \ldots, n, \]

where

\[ \alpha_l = \frac{1}{N}, \quad l = 0, \ldots, N, \]

\[ c_l = \min_{i,j} \{E_{i,j}^L(\alpha_l)\}, \]

\[ d = \max_{i,j} \{E_{i,j}^U(\alpha_l)\}. \]

### 3 DATA AND VARIABLES

DEA has been used for assessing the relative performance of mutual funds. Most of the academic works use a measure of risk (standard deviation of returns, variance, etc) as one of the inputs, and returns (mean return, excess on return, minimum return, skewness, etc) as one of the outputs (see, for example, Murthi et al. (1997); McMullen and Strong (1998); Morey and Morey (1999); Choi and Murthi (2001); Wilkens and Zhu (2001); Galagedera and Silvapulle (2002); Chang (2004); Briec et al. (2004, 2007); Joro and Na (2006); Daraio and Simar (2006); Basso and Funari (2007, 2010, 2012)).

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
Lozano and Gutiérrez (2008) revise and summarize these inputs and outputs together with technology, formulation, orientation, metric and other interesting remarks of the DEA approaches used for mutual funds’ performance evaluation. But as far as we know, only Basso and Funari (2007, 2010, 2012) and more recently Pérez-Gladish et al (2013) include nonfinancial variables in their DEA models for mutual fund performance evaluation; these authors specifically consider mutual funds’ ethical level as an output.

A common characteristic of the above-mentioned studies is that they all use conventional DEA models, and consequently their evaluations are based on the assumption of certain and/or precise data. However, these assumptions could not be realistic, mainly due to (i) the difficulty in measuring certain social, environmental and governance characteristics of the firms and of the social responsibility of mutual funds or (ii) the uncertainty of financial data. In these cases, fuzzy sets theory is a useful tool to reflect this uncertainty and/or imprecision and to incorporate into the model information based not only on historical data but also on the experts’ knowledge.

In this study, 40 US large cap equity mutual funds were analyzed, 23 of them are listed members of the Social Investment Forum (SIF) and the remaining 17 were randomly selected large cap conventional mutual funds. Data comes from the former Kinder, Lydenberg, and Domini Inc (KLD) and Morningstar databases.

We consider two output variables: (1) a fund’s expected return (R), which is the amount one would anticipate receiving on an investment that has various known or expected rates of return; and (2) a social environmental responsibility (SER) index, which is based on the corporate social performance (CSP) of the firms invested in by the equity mutual fund. This paper uses a well-established database from KLD Research & Analytics Inc. that takes into account the multidimensional aspects of CSP (see Pérez-Gladish et al 2012). KLD monitors CSP of US firms according to a set of binary variables, some of which may reflect strengths (S) others concerns (C). In this work, we have transformed concerns into strengths by taking the binary complementaries in order to aggregate them (see (3.1)). If a mutual fund does not present a certain concern, we will consider this fact as a strength:

\[
\text{CSP}_k = \sum_{s=1}^{6} S_{sk} + \sum_{r=1}^{7} C_{rk}, \quad k = 1, \ldots, p, \quad \text{CSP}_k \in \{0, 1, \ldots, 13\}. \tag{3.1}
\]

Once we have obtained a CSP measurement for each of the firms invested in by the considered mutual funds, we can obtain an SER value for each mutual fund:

\[
\text{SER}_j = \sum_{k=1}^{p} \alpha_{kj} \text{CSP}_k, \quad j = 1, 2, \ldots, 40, \tag{3.2}
\]

where \( p \) is the number of firms invested in by mutual fund \( j \), and \( \alpha_{kj} \) is the weighting of firm \( k \) in mutual fund \( j \).
As input variables, we propose to use the following.

(1) Standard deviation of the returns rate (SD), which is a traditional risk measure for the fund investment.

(2) Transaction costs (Murthi et al. 1997; Choi and Murthi 2001). This input is related to the following variables.

(a) Turnover ratio (TR): this is a measure of the portfolio manager’s trading activity, which is computed by taking the lesser of purchases or sales (excluding all securities with maturities of less than 1 year) and dividing by the average monthly net assets.

(b) Gross expense ratio (GER): this represents the total gross expenses (net expenses with waivers added back in) divided by the funds’ average net assets.

(c) Loads: these loads may be charged at the beginning (front load) when the customer invests or they may be deferred till the time the customer withdraws his/her funds (deferred load).

Table 1 shows some basic descriptive statistics for all of the input and output variables defined above.

Thus, values for variables front load and deferred load are considered certain and precise as the charged amount is generally based on the amount of the investment and they remain stable along the investment time.

The values of R, SER, SD, TR and GER will be handled through fuzzy numbers in order to better reflect their uncertainty and to try to incorporate in their construction the expert’s knowledge in addition to the available historical data (data series for the period 1997–2007). Specifically for R, SD and TR, these fuzzy numbers are constructed as fuzzy triangular $LR$ numbers, where $m$ is the median, and $\alpha_L$ and $\alpha_R$ are random percentages of $m$.

In the case of the social, environmental responsibility index (SER) fuzziness will be incorporated based on the quality of the nonfinancial information provided by the mutual fund. The quality management (QM) dimension tries to reflect the transparency and credibility of the information provided by the fund related to the nonfinancial screening and research processes. It includes eleven different criteria ($Q$), ten of which were already considered as criteria in Pérez-Gladish et al. (2012) and an additional criterion: manager’s SRI competence, which refers to SRI education of the fund manager. Measurement of this variable is done through binary variables for each of the criteria. These variables take value 1 if the quality criterion is satisfied and 0 otherwise. The best performance in this dimension implies satisfaction of all criteria.

www.risk.net/journal
<table>
<thead>
<tr>
<th>Variables</th>
<th>Conventional mutual funds (n = 16)</th>
<th>SRI mutual funds (n = 24)</th>
<th>Total (n = 40)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Min.</td>
<td>Max.</td>
<td>Mean</td>
</tr>
<tr>
<td>Outputs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected return (%) (R)</td>
<td>5.03</td>
<td>15.77</td>
<td>9.06</td>
</tr>
<tr>
<td>Social environmental responsibility (SER)</td>
<td>5.62</td>
<td>7.03</td>
<td>6.35</td>
</tr>
<tr>
<td>Inputs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standard deviation (SD)</td>
<td>4.29</td>
<td>10.47</td>
<td>5.58</td>
</tr>
<tr>
<td>Turnover ratio (%) (TR)</td>
<td>4</td>
<td>128</td>
<td>64.54</td>
</tr>
<tr>
<td>Annual report gross expense ratio (%) (GER)</td>
<td>0.51</td>
<td>2.16</td>
<td>1.42</td>
</tr>
<tr>
<td>Front loads</td>
<td>0</td>
<td>5.75</td>
<td>1.72</td>
</tr>
<tr>
<td>Deferred loads</td>
<td>0</td>
<td>5</td>
<td>0.5</td>
</tr>
</tbody>
</table>
The efficiency of socially responsible and conventional mutual funds

FIGURE 1 Social environmental responsibility (SER) fuzzy number.

Source: authors’ own compilation.

Funds not verifying any of the criteria are considered as the worst in terms of QM. (Pérez-Gladish et al 2012, 2013):

\[
QM_j = \sum_{t=1}^{11} Q_{tj}, \quad j = 1, \ldots, 40, \quad QM_j \in \{0, 1, \ldots, 11\}. \tag{3.3}
\]

The SER fuzzy number was computed taking into account the QM, that is to say, transparency and credibility of the information provided by the fund related to the screening and research process. In this sense, the higher the QM the higher our confidence on the SER’s value.

For a fixed crisp SER, Figure 1 shows how to obtain the SER’s triangular LR-fuzzy number for different values of QM. The values \(\alpha^L\) and \(\alpha^R\) represent a percentage of SER which decreases quadratically according to the value of QM. As we can observe, for mutual funds satisfying all the QM criteria (ie, QM = 11), there is no uncertainty related to the transparency and credibility of the information provided by the fund, so the value of the variable measuring the SER of the fund is represented by a crisp or precise number. On the other hand, the lower the number of QM criteria satisfied by the fund, the greater the uncertainty about the SER level of the fund, this leading to a bigger support of the SER fuzzy numbers.

4 EMPIRICAL RESULTS AND DISCUSSION

Since fuzzy numbers R, SER, SD, TR and GER are randomly obtained by considering \(\alpha^L, \alpha^R\) an arbitrary percentage between 0% and 5% of their values, we generated 500 different scenarios in order to obtain more consistent results and to avoid results influenced by a particular situation. For each of these scenarios the models specified in (2.3) were estimated. Then, the average of the funds’ scores was calculated. Table 3 in Appendix A, available online, shows the achieved results.

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
TABLE 2  Chen–Klein Index results and ranking.

<table>
<thead>
<tr>
<th>Fund</th>
<th>Chen–Klein</th>
<th>Rank</th>
<th>Fund</th>
<th>Chen–Klein</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>S5</td>
<td>1</td>
<td>1</td>
<td>S15</td>
<td>0.8744</td>
<td>8</td>
</tr>
<tr>
<td>S9</td>
<td>1</td>
<td>1</td>
<td>S1</td>
<td>0.8729</td>
<td>9</td>
</tr>
<tr>
<td>S16</td>
<td>1</td>
<td>1</td>
<td>S14</td>
<td>0.8713</td>
<td>10</td>
</tr>
<tr>
<td>S17</td>
<td>1</td>
<td>1</td>
<td>S7</td>
<td>0.8586</td>
<td>11</td>
</tr>
<tr>
<td>S18</td>
<td>1</td>
<td>1</td>
<td>S8</td>
<td>0.8586</td>
<td>12</td>
</tr>
<tr>
<td>S19</td>
<td>1</td>
<td>1</td>
<td>C31</td>
<td>0.8509</td>
<td>13</td>
</tr>
<tr>
<td>S22</td>
<td>1</td>
<td>1</td>
<td>C45</td>
<td>0.8447</td>
<td>14</td>
</tr>
<tr>
<td>S24</td>
<td>1</td>
<td>1</td>
<td>S3</td>
<td>0.8248</td>
<td>15</td>
</tr>
<tr>
<td>S25</td>
<td>1</td>
<td>1</td>
<td>S2</td>
<td>0.8192</td>
<td>16</td>
</tr>
<tr>
<td>C27</td>
<td>1</td>
<td>1</td>
<td>C34</td>
<td>0.7883</td>
<td>17</td>
</tr>
<tr>
<td>C38</td>
<td>1</td>
<td>1</td>
<td>C44</td>
<td>0.7846</td>
<td>18</td>
</tr>
<tr>
<td>C41</td>
<td>1</td>
<td>1</td>
<td>C35</td>
<td>0.7533</td>
<td>19</td>
</tr>
<tr>
<td>C46</td>
<td>1</td>
<td>1</td>
<td>C37</td>
<td>0.6996</td>
<td>20</td>
</tr>
<tr>
<td>C28</td>
<td>1</td>
<td>1</td>
<td>C29</td>
<td>0.6895</td>
<td>21</td>
</tr>
<tr>
<td>C33</td>
<td>0.9931</td>
<td>2</td>
<td>C36</td>
<td>0.5922</td>
<td>22</td>
</tr>
<tr>
<td>S11</td>
<td>0.9496</td>
<td>3</td>
<td>C30</td>
<td>0.5708</td>
<td>23</td>
</tr>
<tr>
<td>S10</td>
<td>0.9453</td>
<td>4</td>
<td>S26</td>
<td>0.5395</td>
<td>24</td>
</tr>
<tr>
<td>S12</td>
<td>0.9448</td>
<td>5</td>
<td>C40</td>
<td>0.3029</td>
<td>25</td>
</tr>
<tr>
<td>S13</td>
<td>0.9296</td>
<td>6</td>
<td>S20</td>
<td>0.0210</td>
<td>26</td>
</tr>
<tr>
<td>S6</td>
<td>0.9233</td>
<td>7</td>
<td>S21</td>
<td>0.0014</td>
<td>27</td>
</tr>
</tbody>
</table>

S denotes socially responsible mutual fund. C denotes conventional fund.

A total of 14 funds, representing 35% of the funds analyzed, are efficient in all simulated scenarios and for any \( \alpha \)-cut. Of these, 9 are socially responsible funds (S5, S9, S16, S17, S18, S19, S22, S24 and S25) and 5 conventional funds (C27, C28, C38, C41 and C46).

The funds that are inefficient have, on average, high scores. A total of 12 SRI funds and 6 conventional funds attain efficiency ratios above 90%. Possibly, these results happen due to the fact that virtually all funds invest in similar portfolios, leaving little room for “innovation” and thus limiting the possibility that certain funds significantly stand out above the rest. However, it is interesting to observe the poor ratios attained by socially responsible funds S20 and S21, and to a lesser extent by the conventional mutual fund C40.

In order to rank the analyzed mutual funds we have calculated the Chen–Klein index (see (2.4)). Obtained results are shown in Table 2.

Obviously, for the 14 funds that are efficient in all scenarios, the Chen–Klein index is set to 1 and these funds are ranked in the first position. As shown in Table 2, conventional fund C33 ranks second in the ranking. The next 10 funds in the ranking...
are socially responsible funds, with values for the Chen–Klein index moving between 0.9496 for mutual fund S11 (position 3) and 0.8586 for mutual fund S8 (position 12). Therefore, of the 25 funds best placed in terms of efficiency, 19 are socially responsible funds, which represents 76% of the considered funds.

From the results obtained, it is worth mentioning, undoubtedly, the bad behavior of funds S20 and S21, which are relegated to the last two positions in the ranking. Indeed, mutual fund S20 presents efficiency scores that are between 0.669 and 0.6693 at the lower limit and between 0.6693 and 0.6696 at the upper limit. For mutual fund S21 efficiency ratios are between 0.6624 and 0.6627 at the lower limit and between 0.6627 and 0.663 at the upper limit.

Numerous mutual funds rankings exist and are publicly available in order to assist individual investors in their investment decisions, based only on financial variables. The proposed model in this paper allows ranking mutual funds based on their relative efficiency, which is defined in terms of both financial and nonfinancial criteria.

The efficiency scores obtained can be incorporated by the investor in his/her decision-making model for portfolio selection together with the classical budget and short sales constraints and with any other financial or nonfinancial constraints.

5 CONCLUSIONS

In this study, Kao and Liu’s auxiliary fuzzy DEA models (Kao and Liu 2000) have been applied in order to evaluate 40 US equity mutual funds. With the aim of classifying the mutual funds’ efficiency scores Chein and Klein index has been used.

The main contributions of this work are two: firstly, uncertainty and/or imprecision of the data is incorporated in the model, leading to a more realistic evaluation of the relative performance of the mutual funds. Secondly, a nonfinancial output is considered, the SER, which is related to the transparency and credibility of the nonfinancial information provided by the fund (QM), and it allows us to evaluate the performance of mutual funds in terms of their efficiency.

The results obtained show that 14 mutual funds (35%) are efficient. From these efficient funds, 9 (64.3%) are socially responsible listed funds. Moreover, from the 25 best positioned funds, 19 are nonconventional mutual funds. Thus, all in all, we can conclude that, if the expert knowledge is considered together with the historical information, socially responsible mutual funds show a better performance in terms of efficiency than conventional mutual funds. The investor can incorporate the information provided by the fuzzy DEA analysis in the decision-making about the portfolio selection.
DECLARATION OF INTEREST

This research has been financially supported by the Spanish Ministry of Science and Innovation (project ECO2011-28927). The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

ACKNOWLEDGEMENTS

We would like to thank Blanca Pérez-Gladish for her comments on previous versions of this paper.

REFERENCES


