An analytical value-at-risk approach for a credit portfolio with liquidity horizon and portfolio rebalancing
Haohan Huang, Eugene Wang, Huaxiong Huang and Yong Wang

Loss distributions: computational efficiency in an extended framework
Daniel H. Stahl

Default risk of money-market fund portfolios
Matulya Bansal

Are all collections equal? The case of medical debt
Kenneth P. Brevoort and Michelle Kambara
Risk.net in numbers

- **200+** New articles & technical papers (each month)
- **140,000** Users (each month)
- **370,000** Page views (each month)
- **58,000+** articles stretching back 20 years
- **6,500+** on Asset Management
- **6,900+** on Commodities
- **£21,000+** on Derivatives
- **19,400+** on Regulation
- **19,600+** on Risk Management

See what you’re missing

Visit the world’s leading source of exclusive in-depth news & analysis on risk management, derivatives and complex finance now.
The Journal of Credit Risk
EDITORIAL BOARD

Editors-in-Chief
ASHISH DEV Federal Reserve Board
MICHAEL GORDY Federal Reserve Board

Associate Editors
EDWARD ALTMAN NYU Stern School of Business
JENNIE BAI Georgetown University
J. L. BREEDEN Prescient Models LLC
JONATHAN CROOK The University of Edinburgh
DARRELL DUFFIE Stanford University
CHRISTOPHER C. FINGER Federal Reserve Board
KAY GIESECKE Stanford University
GREG GUPTON Federal Reserve Board
JAY HUANG Penn State University
JOHN HULL University of Toronto
ROBERT JARROW Cornell University
AHMET KOCAGIL BlackRock
HOLGER KRAFT Goethe University

ANDRE LUCAS VU Amsterdam
DILIP MADAN University of Maryland
MARCO NALDI Barclays Capital
LORIANA PELIZZON Goethe University
DMITRY PUGACHEVSKY Quantifi
MICHAEL PYKHTIN Federal Reserve Board
DAN ROSEN S&P Capital IQ
PETER RITCHKEN Case Western Reserve University

SUBSCRIPTIONS

The Journal of Credit Risk (Print ISSN 1744-6619 | Online ISSN 1755-9723) is published quarterly 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>£735</td>
<td>€1035</td>
<td>$1215</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 Credit Risk
GENERAL SUBMISSION GUIDELINES

Manuscripts and research papers submitted for consideration must be original work that is not simultaneously under review for publication in another journal or other publication outlets. All articles submitted for consideration should follow strict academic standards in both theoretical content and empirical results. Articles should be of interest to a broad audience of sophisticated practitioners and academics.

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

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

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

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

Before submitting a paper, authors should consult the full author guidelines at: http://www.risk.net/static/risk-journals-submission-guidelines

Queries may also be sent to:
The Journal of Credit 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
With the rewriting of the Basel accords in international banking and their ensuing application, interest in credit risk has never been greater. *The Journal of Credit Risk* is at the forefront in tackling the many issues and challenges posed by the recent financial crisis, focusing on the measurement and management of credit risk, the valuation and hedging of credit products, and the promotion of greater understanding in the area of credit risk theory and practice.

*The Journal of Credit Risk* considers submissions in the form of research papers and technical reports on, but not limited to, the following topics.

- Modeling and management of portfolio credit risk.
- Recent advances in parameterizing credit risk models: default probability estimation, copulas and credit risk correlation, recoveries and loss given default, collateral valuation, loss distributions and extreme events.
- The pricing and hedging of credit derivatives.
- Structured credit products and securitizations, eg, collateralized debt obligations, synthetic securitizations, credit baskets, etc.
- Measuring, managing and hedging counterparty credit risk.
- Credit risk transfer techniques.
- Liquidity risk and extreme credit events.
- Regulatory issues, such as Basel II, internal ratings systems, credit-scoring techniques and credit risk capital adequacy.
CONTENTS

Letter from the Editor-in-Chief vii

RESEARCH PAPERS

An analytical value-at-risk approach for a credit portfolio with liquidity horizon and portfolio rebalancing 1
Haohan Huang, Eugene Wang, Huaxiong Huang and Yong Wang

Loss distributions: computational efficiency in an extended framework 29
Daniel H. Stahl

Default risk of money-market fund portfolios 43
Matulya Bansal

Are all collections equal? The case of medical debt 73
Kenneth P. Brevoort and Michelle Kambara
LETTER FROM THE EDITOR-IN-CHIEF

Ashish Dev
Federal Reserve Board

In this issue we present four research papers.

The first paper is “An analytical value-at-risk approach for a credit portfolio with liquidity horizon and portfolio rebalancing”, by Haohan Huang, Eugene Wang, Huaxiong Huang and Yong Wang. The authors provide a two-period analytical value-at-risk approach for credit portfolios with a liquidity horizon and a constant level of risk. Given any time horizon, a two-period credit portfolio loss model is derived and, at the end of the first period, the portfolio is rebalanced to ensure that it has a constant level of risk as measured by the credit rating. By testing with the Monte Carlo simulation model, it is shown that the accuracy of the analytic model is acceptable over a large range of parameter values.

The issue’s second paper, “Loss distributions: computational efficiency in an extended framework” by Daniel H. Stahl, contributes to the credit modeling literature for mixture models by leveraging an efficient algorithm for computing the density function of the loss distribution and extending the model in two key areas: construction of the systemic variable from a continuous-time process and introduction to semi-endogenous liquidity risk. This generalization allows for time-dependent portfolios, fully accounts for granularity and concentration within the credit portfolio, and does not rely on assumptions that large credit portfolios are asymptotic.

In the third paper, “Default risk of money-market fund portfolios”, Matulya Bansal addresses the problem of quantifying the risk associated with money-market fund (MMF) portfolios. Ever since the Reserve Primary Fund “broke the buck” in 2008, credit risk in MMFs has become an issue of great interest. Different proposals have been proposed to prevent a run on MMFs: for example, the use of capital buffers, liquidity fees and floating net asset values. But very little work has been done on actually measuring the risk. By focusing on default risk – which is the most material driver of portfolio losses – and by using a constant level of risk assumption, the author shows how this problem is similar to that of pricing a collateralized debt obligation (CDO). The author develops a semi-analytical approach to measuring the default risk of MMF portfolios. On using this model to evaluate the portfolios of three of the largest prime MMFs, the author finds that they vary considerably in their default risk.

The fourth and final paper in this issue, “Are all collections equal? The case of medical debt” by Kenneth P. Brevoort and Michelle Kambara, examines the predictive value of medical collections in assessing consumer creditworthiness with credit

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
scoring models. The authors find two main results. First, they find that medical collections are less informative about a consumer’s likelihood of delinquency than are nonmedical collections. Second, they find that medical collections that have been paid in full are less predictive than medical collections that remain unpaid.
An analytical value-at-risk approach for a credit portfolio with liquidity horizon and portfolio rebalancing

Haohan Huang,1,2 Eugene Wang,1 Huaxiong Huang2 and Yong Wang3

1RBC Financial Group, 11th Floor, South Tower, 200 Bay Street, Toronto, ON M5J2J5, Canada; emails: haohan.huang@hotmail.com; eugene.wang@rbc.com
2Department of Mathematics and Statistics, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada; email: hhuang@yorku.ca
3Everbright Securities Company Limited, 20th Floor, 1508 Xinzha Road, Jing'an District, Shanghai 200040, China; email: yong.wang@ebscn.com

(Received June 19, 2014; revised April 13, 2015; accepted June 9, 2015)

ABSTRACT

We provide a two-period analytical value-at-risk (VaR) approach for the credit portfolio with liquidity horizon and the constant level of risk. Given any time horizon, a two-period credit portfolio loss model is derived and, at the end of the first period, the portfolio is rebalanced to ensure a constant level risk of the portfolio measured by the credit rating. The analytical VaR is found by extending the granularity adjustment (GA) approximation. The model is applied to incremental risk charge (IRC), with the liquidity horizon of each asset being six months. By testing with the Monte Carlo simulation model, it is shown that the accuracy of the analytic model is acceptable over a large range of parameters. The model behaves similarly to the standard GA in capturing the concentration risk. We also show the features of liquidity horizon and
constant level of risk are captured adequately in the model. Our analytical approach behaves better than the standard one-period asymptotic single-risk-factor model with and without GA, to achieve a comparable measure to the IRC.

Keywords: analytical value-at-risk; granularity adjustment; incremental risk charge; portfolio rebalancing; liquidity horizon.

1 INTRODUCTION

In response to the financial crisis, the Basel Committee on Banking Supervision (2013b) introduced a capital charge for the incremental risk of unsecured credit products in the trading book. Now generally viewed as part of Basel 2.5, the incremental risk charge (IRC) is measured at a one-year value-at-risk (VaR) at a 99.9th percentile confidence level. The committee allows banks to take a model-based approach and most proposed IRC models are based on a multi-factor multi-period Monte Carlo Simulation model (see Finger 2011; CreditMetrics 2011; Wilkens et al 2013; Yavin et al 2010). In the latest Basel proposal (Basel Committee on Banking Supervision 2013a), IRC would be replaced by incremental default risk (IDR), a two-factor modeling framework is suggested and only default losses are considered.

Several new concepts in credit risk measurement are introduced in the model-based IRC rules. Among them, two are particularly interesting: the liquidity horizon and the constant level of risk over a one-year capital horizon. The liquidity horizon concept allows the bank to model the differences in the underlying liquidity of the trading book position. It represents the time required to sell the position or to hedge all material risk covered by the IRC model in a stressed market. The concept of a constant level of risk over a one-year capital horizon allows a bank to model the portfolio rebalancing of its trading positions in a manner that maintains the initial risk level. In our opinion, proper portfolio rebalancing assumptions are important; they make the IRC model more risk sensitive and relevant to actual trading portfolio behavior. For more liquidity and more highly rated positions, portfolio rebalancing provides a benefit over assuming the same position throughout the capital horizon, but it would be punitive for positions with high default risk. In order to hold the initial risk level, one would have to replace those positions that default within the one-year time horizon, and replaced positions would again carry high default risk.

The liquidity horizon and constant level of risk lead to a high-dimensional portfolio loss function that can be modeled via either convolution or Monte Carlo simulation. In practice, most IRC models rely on the Monte Carlo simulation. Several analytical techniques have been developed for portfolio credit risk modeling. One such technique is the asymptotic single-risk-factor (ASRF) model. The principle of the ASRF model is to model a large credit portfolio via a one-factor model. It assumes that the portfolio is
VaR approach for a credit portfolio with liquidity horizon and rebalancing

infinitely fine grained and driven by one systematic risk factor, with the idiosyncratic risk fully diversified away. It has been the standard capital charge model for the banking book and is also often required as a benchmark by regulators for the trading book. The pros and cons of the ASRF approach have been the topic of extensive research (for a review, see Hibbeln (2010) or Lütkebohmert (2009)).

From a credit portfolio management perspective, the major weakness of the ASRF model is its inability to capture concentration risk. Credit concentrations, including both name concentrations and sector concentrations, are probably the single most important cause of major credit problems, which are behind most major banking disasters, including the most recent financial crisis (see Basel Committee on Banking Supervision 2000, 2006). Within the ASRF approach, the undiversified idiosyncratic risk can be approximated analytically via a granularity adjustment (GA) approximation. GA was first introduced in 2000 by Gordy (see Gordy 2003). The model was improved and put on a more rigorous foundation by Martin et al (2002), Wilde (2001b) and Gouriéroux et al (2000). A survey of these developments and a primer on the mathematical derivation are presented in Gordy (2004), and a rigorous proof of GA has been given recently by Fermanian (2013). In recent years, the concept of GA has also been expanded in both the application to modeling other risks (see Gordy and Marrone 2010) and credit portfolio risk measures other than capital charges (see Düllmann and Masschelein 2006; Pykhtin 2014a,b; Voropaev 2011). The multi-factor extension of the GA technique has been derived by Pykhtin (2004), based on finding a comparable one-factor portfolio with a similar risk profile. The model may also be extended to address the constant level of risk for the credit portfolio.

In this paper, we provide a general framework of a two-period GA VaR model in the context of the IRC modeling framework in which the liquidity horizon and constant level of risk are modeled. At the end of the first period, the portfolio is rebalanced to ensure a constant level of risk as measured by the credit rating. A two-period GA term, which is the second-order Taylor expansion of the two-dimensional portfolio loss function, is calculated directly. The performance of the model is then tested against the two-period IRC Monte Carlo simulation model and also the Monte Carlo model without portfolio rebalancing and the ASRF with and without standard one-period GA to show how concentration risk, liquidity and constant level of risk are captured in the new analytical approach.

The paper is organized as follows. Section 2 introduces the one-period GA. Section 3 presents the general framework of two-period analytical VaR, which is an extension of GA. The behavior of the two-period GA VaR is illustrated in Section 4 by comparing it with ASRF and standard one-period GA, one- and two-period Monte Carlo simulation with and without portfolio rebalancing. Section 5 concludes and discusses possible further applications of the model.
2 THE ASYMPTOTIC SINGLE-RISK-FACTOR MODEL FOR CREDIT PORTFOLIO AND GRANULARITY ADJUSTMENT APPROXIMATION

The general framework of ASRF and its GA are presented in this section. For a credit portfolio, the loss function within a one-factor modeling framework is defined as

\[ L_N = \sum_{i=1}^{N} u_i \mathbf{1}_{\{X_i > U_i\}}. \]  
\[ X_i = \rho_i S + \sqrt{1 - \rho_i^2} \xi_i. \]

where \( L_N \) is the portfolio loss, \( u_i \) is the loss given default (LGD) of the \( i \)th asset, \( \xi_i \) are the idiosyncratic factors, \( \rho_i \) is the positive correlation between negative asset factor \( X_i \) and systematic factor \( S \) and \( U_i \) is the threshold to determine whether the default of the \( i \)th trade will happen. \( S \) and all the \( \xi_i \) are assumed to be independent and identically distributed Gaussian variables \( N(0, 1) \).

Denote by \( \alpha_q(\cdot) \) the \( q \)th percentile value of the random variable, ie,

\[ P(X \leq \alpha_q(X)) = q. \]

Then the \( q \)th percentile VaR of this portfolio, which can be found by Monte Carlo simulation with large time consumption, is denoted by \( \alpha_q(L_N) \). Although it has no direct analytic solution, its approximation can be calculated in ASRF. In ASRF, the most important assumption is that when the portfolio is large enough the individual risk of each trade will be diversified away. With such assumptions and some precise conditions, Gordy (2003) provided a reasonable approximation for a realistic portfolio, as follows:

\[ \alpha_q(L_N) \approx \alpha_q(\mathbb{E}(L_N \mid S)). \]  

If the conditional expectation of loss function \( f(s) = \mathbb{E}[L_N \mid S = s] \) is monotonic, which is the assumption of most models, we have \( \alpha_q[\mathbb{E}(L_N \mid S)] = \mathbb{E}(L_N \mid \alpha_q(S)) \) because of the monotonic property of \( \alpha_q(\cdot) \). If the loss function is defined as in (2.1), we have

\[ \alpha_q(L_N) \approx \alpha_q(\mathbb{E}(L_N \mid S)) \\
= \alpha_q \left[ \sum_{i=1}^{N} u_i \left( 1 - \Phi \left( \frac{U_i - \rho_i S}{\sqrt{1 - \rho_i^2}} \right) \right) \right] \\
= \sum_{i=1}^{N} u_i \left( 1 - \Phi \left( \frac{U_i - \rho_i \alpha_q(S)}{\sqrt{1 - \rho_i^2}} \right) \right). \]
This is how the capital requirement calculation (internal ratings-based approach) is implemented based on the ASRF assumption. In reality, however, an infinitely fine-grained portfolio does not exist, so there is a difference between VaR (ie, $\alpha_q(L_N)$) and $\mathbb{E}(L_N \mid \alpha_q(S))$. The summation of $\mathbb{E}(L_N \mid \alpha_q(S))$ and the difference is considered to be the new VaR. This calculation of the difference is the main procedure of GA.

The key method of GA proposed by Gordy (2003) was the second-order Taylor expansion. We know that

$$
\alpha_q(L_N) = \alpha_q[\mathbb{E}(L_N \mid S) + \varepsilon(L_N - \mathbb{E}(L_N \mid S))]|_{\varepsilon=1}.
$$

Let $z(\varepsilon) = \alpha_q[\mathbb{E}(L_N \mid S) + \varepsilon(L_N - \mathbb{E}(L_N \mid S))]$. Applying the second-order Taylor expansion on $\varepsilon = 0$, we have

$$
z(\varepsilon) \approx z(0) + z'(0)\varepsilon + z''(0)\frac{1}{2}\varepsilon^2. \quad (2.5)
$$

Then

$$
\alpha_q[\mathbb{E}(L_N \mid S) + \varepsilon(L_N - \mathbb{E}(L_N \mid S))]|_{\varepsilon=1} = z(1) \approx z(0) + z'(0) \cdot 1 + z''(0) \cdot \frac{1}{2}
$$

$$
= \alpha_q[\mathbb{E}(L_N \mid S)] + \frac{\partial\alpha_q}{\partial \varepsilon}[\mathbb{E}(L_N \mid S) + \varepsilon(L_N - \mathbb{E}(L_N \mid S))]|_{\varepsilon=0}
$$

$$
+ \frac{1}{2} \frac{\partial^2\alpha_q}{\partial \varepsilon^2}[\mathbb{E}(L_N \mid S) + \varepsilon(L_N - \mathbb{E}(L_N \mid S))]|_{\varepsilon=0}. \quad (2.6)
$$

Then the value of GA, which is the difference between $\alpha_q(L_N)$ and $\alpha_q[\mathbb{E}(L_N \mid S)]$, is approximately the sum of the first and second derivatives.

Thanks to Tasche (2000) and Gouriéroux et al (2000), the following theorem has been proved.

**Theorem 2.1** Consider two random variables $X$ and $Y$ with a joint density function $f(x, y)$. Let $\alpha_q(\cdot)$ be defined as above. Then

$$
\frac{\partial^2\alpha_q(X + \varepsilon Y)}{\partial \varepsilon^2} - \frac{\partial\alpha_q(X + \varepsilon Y)}{\partial \varepsilon} = \mathbb{E}[Y \mid X + \varepsilon Y = \alpha_q(X + \varepsilon Y)], \quad (2.7)
$$

$$
\frac{\partial^2\alpha_q(X + \varepsilon Y)}{\partial \varepsilon^2} = \left[\frac{\partial^2 f(x + \varepsilon y(s))}{\partial s^2} \right]_{s=\alpha_q(x + \varepsilon y)} + \sigma^2(Y \mid X + \varepsilon Y = s) \frac{\partial \ln f_{X+\varepsilon Y}(s)}{\partial s}, \quad (2.8)
$$

where $f_{X+\varepsilon Y}(s)$ is the density function of $X + \varepsilon Y$ and $\sigma^2(Y \mid X + \varepsilon Y = s)$ is the conditional variance of $Y$. 

www.risk.net/journal
Apply Theorem 2.1 by choosing $X = \mathbb{E}(L_N \mid S)$ and $Y = L_N - \mathbb{E}(L_N \mid S)$. Then the first-order term is

$$\frac{\partial \alpha_q}{\partial \varepsilon}[X + \varepsilon Y]_{\varepsilon=0} = \mathbb{E}[Y \mid X = \alpha_q(X)];$$

(2.9)

the second-order term is

$$\left.\frac{1}{2} \frac{\partial^2 \alpha_q}{\partial \varepsilon^2}[X + \varepsilon Y]_{\varepsilon=0}\right|_{s=\alpha_q(X)} = -\frac{1}{2}\left[\frac{\partial \sigma^2(Y \mid X = s)}{\partial s} + \sigma^2(Y \mid X = s) \frac{\partial \ln f_X(s)}{\partial s}\right]_{s=\alpha_q(X)}. \quad (2.10)$$

The first-order term vanishes using the property of $\sigma$-algebra ($S$ can be considered as a $\sigma$-algebra),

$$\mathbb{E}(Y \mid X) = \mathbb{E}(L_N \mid X) - X = \mathbb{E}(\mathbb{E}(L_N \mid S) \mid X) - X = \mathbb{E}(X \mid X) - X = 0,$$

(2.11)

so (2.6) becomes

$$\alpha_q(L_N) = \alpha_q(X) + \text{GA}. \quad (2.12)$$

### 3 ANALYTICAL VALUE-AT-RISK WITH LIQUIDITY HORIZON AND PORTFOLIO REBALANCING

#### 3.1 The two-period credit portfolio VaR measure and its ASRF and GA terms

The two-period credit portfolio valuation and loss model is described in this subsection. In order to model the portfolio rebalancing within a one-year time horizon, we divide the one-year horizon into two half-year periods. In the first half, the credit portfolio follows the standard factor model as outlined in (2.1). Then, at the end of the first period, the portfolio can be rebalanced according to what happens in that period. For example, if one asset defaults in the first period, we choose to replace it with a similar asset of the same LGD and rating (i.e., same default probability and asset correlation). Then we can say that at the end of six months the portfolio is replenished such that it maintains a constant level of risk. For some assets, we can also assume that no action is needed. From the default risk perspective, this has the embedded assumption that this asset has a liquidity horizon of one year. In this paper, we assume that all assets have a six-month liquidity horizon. Therefore, we need to model the losses aggregated in two periods with the losses in the second period conditional on the portfolio rebalancing assumptions.
Let $S_1$, $S_2$ be the realizations of the systematic factor in the end of the first and second period. They are assumed to be independent. Similar to the one-period default model (2.1), the two-step one-factor default model is

$$L_N = \sum_{i=1}^{N} \left[ u_i 1_{\{T_1^{(i)} > U_i\}} + u_i 1_{\{T_2^{(i)} > U_i\}} \right], \quad (3.1)$$

where

$$T_1^{(i)} = \rho_i S_1 + \sqrt{1 - \rho_i^2} \xi_i,$$  \quad (3.2)

$$T_2^{(i)} = \rho_i S_2 + \sqrt{1 - \rho_i^2} \xi'_i,$$  \quad (3.3)

$u_i$ is the LGD of asset $i$; all the $\xi_i$ and $\xi'_i$ are the idiosyncratic factors that are independent of each other and independent across each systematic factor $S_1$, $S_2$; and $\rho_i$ is the positive correlation between the negative asset factor $T_1^{(i)}$ and systematic factor $S_1$. It is the same as the correlation between the negative asset factor $T_2^{(i)}$ and systematic factor $S_2$ since the trade has the same risk profile as the trade in the first period no matter whether it defaults or not. $U_i$ is the threshold used to determine whether the default of the $i$th trade will happen. $S_1$, $S_2$ and all the $\xi_i$ and $\xi'_i$ are assumed to be Gaussian-distributed variables $N(0, 1)$. Although $\sum_{i=1}^{N} u_i$ can be 0 if short positions are allowed, it is highly unlikely to be 0 from the perspective of risk management. So, in the rest of the paper it is assumed to be a nonzero constant.

The second-order Taylor expansion can be applied, similarly to the standard GA. Rewrite $L_N = \alpha_q [\mathbb{E}(L_N \mid S_1, S_2)] + \varepsilon [L_N - \mathbb{E}(L_N \mid S_1, S_2)]_{\varepsilon=1}$. Then use Taylor expansion and proceed as in (2.6):

$$\alpha_q(L_N) = \alpha_q [\mathbb{E}(L_N \mid S_1, S_2)] + \varepsilon [L_N - \mathbb{E}(L_N \mid S_1, S_2)]_{\varepsilon=1}$$
$$\approx \alpha_q [\mathbb{E}(L_N \mid S_1, S_2)]$$
$$+ \frac{\partial \alpha_q}{\partial \varepsilon} [\mathbb{E}(L_N \mid S_1, S_2) + \varepsilon (L_N - \mathbb{E}(L_N \mid S_1, S_2))]_{\varepsilon=0}$$
$$+ \frac{1}{2} \frac{\partial^2 \alpha_q}{\partial \varepsilon^2} [\mathbb{E}(L_N \mid S_1, S_2) + \varepsilon (L_N - \mathbb{E}(L_N \mid S_1, S_2))]_{\varepsilon=0}. \quad (3.4)$$

In (3.4), if we can calculate the value of $\alpha_q [\mathbb{E}(L_N \mid S_1, S_2)]$ and the first and second derivative, we will know the VaR of this portfolio. The sum of the first and second derivatives gives the value of the GA.

Each term in (3.4) will be calculated in the following subsections; a summary of the appropriate cross-references is given in Table 1 on the next page.
### 3.2 The two-period “ASRF” term in (3.4)

Using the formula for $L_N$ in (3.1), $\mathbb{E}(L_N \mid S_1, S_2)$ is calculated as

\[
\mathbb{E}(L_N \mid S_1, S_2) = \sum_{i=1}^{N} \left\{ u_i \left[ 1 - \Phi \left( \frac{U_i - \rho_i S_i}{\sqrt{1 - \rho_i^2}} \right) \right] \right\} + \sum_{i=1}^{N} \left\{ u_i \left[ 1 - \Phi \left( \frac{U_i - \rho_i S_i}{\sqrt{1 - \rho_i^2}} \right) \right] \right\}.
\]

(3.5)

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distributed variable. Define

\[
X = \mathbb{E}(L_N \mid S_1, S_2),
\]

(3.6)

\[
l(s) = \sum_{i=1}^{N} \left\{ u_i \left[ 1 - \Phi \left( \frac{U_i - \rho_i s}{\sqrt{1 - \rho_i^2}} \right) \right] \right\}.
\]

(3.7)

So (3.5) can be written simply as

\[
X = l(S_1) + l(S_2).
\]

(3.8)

Since $\rho_i$ is always positive, it is obvious that $l(s)$ is a strictly monotonically increasing function and

\[
l(-\infty) = 0,
\]

(3.9)

\[
l(+\infty) = \sum_{i=1}^{N} u_i,
\]

(3.10)

where hereafter we use $f(-\infty)$ and $f(+\infty)$ as shorthand for the respective limits.
To calculate \( \alpha_q(X) \), we need to calculate the cumulative distribution function of \( X \), ie,
\[
F_X(t) = P(X \leq t). \tag{3.11}
\]

From (3.9) and (3.10), function \( l(s) \) is bounded between 0 and \( \sum_{i=1}^{N} u_i \). So,
\[
F_X(t) = \begin{cases} 
0 & \text{if } t \leq 0, \\
1 & \text{if } t \geq 2 \sum_{i=1}^{N} u_i.
\end{cases} \tag{3.12}
\]

For any \( 0 < t < 2 \sum_{i=1}^{N} u_i \),
\[
X \leq t \Leftrightarrow l(S_1) + l(S_2) \leq t
\]
\[
\Leftrightarrow \begin{cases} 
\{S_2 \leq l^{-1}(t - l(S_1)), S_1 > l^{-1}(t - \sum_{i=1}^{N} u_i)\} \\
\cup \{S_2 \in \mathbb{R}, S_1 \leq l^{-1}(t - \sum_{i=1}^{N} u_i)\}, & t \geq \sum_{i=1}^{N} u_i, \\
\{S_2 \leq l^{-1}(t - l(S_1)), S_1 < l^{-1}(t)\}, & t < \sum_{i=1}^{N} u_i.
\end{cases} \tag{3.13}
\]

The \( F_X(t) \) in both cases can be calculated, given that the systematic factors \( S_1 \) and \( S_2 \) are normally distributed and uncorrelated. The density function of \( X \) will be calculated here for future use.

(i) In the first case, ie, \( 2 \sum_{i=1}^{N} u_i > t \geq \sum_{i=1}^{N} u_i \),
\[
F_X(t) = \int_{l^{-1}(t - \sum_{i=1}^{N} u_i)}^{+\infty} \phi(s_1)\Phi(l^{-1}(t - l(s_1))) \, ds_1 + \Phi\left(l^{-1}\left(t - \sum_{i=1}^{N} u_i\right)\right), \tag{3.14}
\]
where \( \phi(\cdot) \) is the density function of the standard normal random variable. Note that when \( t = \sum_{i=1}^{N} u_i \),
\[
F_X(t) = \int_{-\infty}^{+\infty} \phi(s_1)\Phi(l^{-1}(t - l(s_1))) \, ds_1.
\]
This does not violate the formula for \( F_X(t) \) in (3.14) since
\[
\Phi(-\infty) = \lim_{x \to -\infty} \Phi(x) = 0.
\]

(ii) In the second case, ie, \( 0 < t < \sum_{i=1}^{N} u_i \),
\[
F_X(t) = \int_{-\infty}^{l^{-1}(t)} \phi(s_1)\Phi(l^{-1}(t - l(s_1))) \, ds_1. \tag{3.15}
\]
Then, using (3.14) and (3.15), we solve for $\alpha_q(X)$ numerically from

$$F_X(\alpha_q(X)) = q.$$  

(3.16)

$f_X(t)$ will be calculated here for future reference. By the rule for change of variables, the density $f_{l(S)}(t)$ of $l(S)$ is given by

$$f_{l(S)}(t) = \frac{\phi(l^{-1}(t))}{l'(l^{-1}(t))} \quad \text{for } t \in \left(0, \sum_{i=1}^{N} u_i\right).$$

Hence,

$$f_X(t) = \int_{y \in (0,1) \cap (t-1,t)} f_{l(S)}(y) f_{l(S)}(t-y) \, dy = \int_{y \in (0,1) \cap (t-1,t)} \frac{\phi(l^{-1}(y)) \phi(l^{-1}(t-y))}{l'(l^{-1}(y)) l'(l^{-1}(t-y))} \, dy.$$  

(3.17)

Reinserting $y = l(s)$ with $dy = l'(s) \, ds$ gives

$$f_X(t) = \int_{l^{-1}(t-\sum_{i=1}^{N} u_i)}^{+\infty} \phi(s) \phi(l^{-1}(t-l(s))) \frac{1}{l'(l^{-1}(t-l(s)))} \, ds.$$  

(3.18)

To simplify the expression for $f_X(t)$, we define a function $g_t(s)$ on

$$\left\{(t,s) \mid (t,s) \in \left[\sum_{i=1}^{N} u_i, 2 \sum_{i=1}^{N} u_i\right] \times \left(l^{-1}\left(\sum_{i=1}^{N} u_i\right) + \infty\right) \cup \left[0, \sum_{i=1}^{N} u_i\right] \times (-\infty, l^{-1}(t))\right\}$$  

(3.19)

as

$$g_t(s) = l^{-1}(t-l(s)).$$  

(3.20)

Since $l(x)$ is a strictly monotonically increasing function, $g_t(s)$ is a strictly monotonically decreasing function with respect to $s$.

Define an interval $\Omega(t)$ as

$$\Omega(t) = \begin{cases} \left(l^{-1}(t-\sum_{i=1}^{N} u_i), +\infty\right), & 2 \sum_{i=1}^{N} u_i > t \geq \sum_{i=1}^{N} u_i, \\ (-\infty, l^{-1}(t)), & \sum_{i=1}^{N} u_i > t > 0. \end{cases}$$  

(3.21)

Then

$$f_X(t) = \begin{cases} \int_{\Omega(t)} \phi(s) \phi(g_t(s)) (l'(g_t(s)))^{-1} \, ds, & \text{if } 2 \sum_{i=1}^{N} u_i > t \geq 0, \\ 0, & \text{otherwise}. \end{cases}$$  

(3.22)
3.3 Calculation of the second and third terms in (3.4)

The second-order term vanishes similarly to (2.11).

Define $Y = L_N - \mathbb{E}(L_N \mid S_1, S_2)$ and use Theorem 2.1:

$$
\frac{1}{2} \frac{\partial^2 \alpha_q}{\partial \varepsilon^2} \left[ \mathbb{E}(L_N \mid S_1, S_2) + \varepsilon(L_N - \mathbb{E}(L_N \mid S_1, S_2)) \right]_{\varepsilon=0} = \frac{1}{2} \frac{\partial^2 \alpha_q}{\partial \varepsilon^2} (X + \varepsilon Y)_{\varepsilon=0} = -\frac{1}{2} \left[ \frac{\partial \sigma^2(Y \mid X = s)}{\partial s} + \sigma^2(Y \mid X = s) \frac{d \ln f_X(s)}{ds} \right]_{s=\alpha_q(X)} \qquad (3.23)
$$

First, the value of $\sigma^2(Y \mid X = s)$ is required:

$$
\sigma^2(Y \mid X = s) = \sigma^2(L_N - X \mid X = s) = \sigma^2(L_N - s \mid X = s) = \sigma^2(L_N \mid X = s) = \mathbb{E}(L_N^2 \mid X = s) - \mathbb{E}^2(L_N \mid X = s). \quad (3.24)
$$

Then $\mathbb{E}(L_N \mid X = s)$ and $\mathbb{E}(L_N^2 \mid X = s)$ are calculated for the value of $\sigma^2(Y \mid X = s)$ in Sections 3.3.1 and 3.3.2.

Second, the derivative of the variance (ie, $\partial \sigma^2(Y \mid X = s)/\partial s$) is calculated in Section 3.3.3.

Finally, $d \ln f_X(s)/ds$ can be calculated based on the formulas for $f_X(s)$ and $f_X'(s)$.

3.3.1 Calculation of $\mathbb{E}(L_N \mid X = s)$

Recall the notation given in Section 3.1.

Then

$$
\mathbb{E}(L_N \mid X = s) = \mathbb{E}\left\{ \frac{\varepsilon}{N} \sum_{i=1}^{N} \left[ u_i 1_{\{T_1(i) > U_i\}} + u_i 1_{\{T_2(i) > U_i\}} \right] \mid X = s \right\} 
= \sum_{i=1}^{N} \left\{ u_i \mathbb{E} \left[ 1_{\{T_1(i) > U_i\}} \mid X = s \right] + u_i \mathbb{E} \left[ 1_{\{T_2(i) > U_i\}} \mid X = s \right] \right\}. \quad (3.25)
$$

By symmetry,

$$
\mathbb{E} \left[ 1_{\{T_1(i) > U_i\}} \mid X = s \right] = \mathbb{E} \left[ 1_{\{T_2(i) > U_i\}} \mid X = s \right] \quad \text{for all } i = 1, 2, \ldots, N. \quad (3.26)
$$

So the formulas for $\mathbb{E}(1_{\{T_1(i) > U_i\}} \mid X = s)$ are sufficient to give the value of $\mathbb{E}(L_N \mid X = s)$. 

www.risk.net/journal
Recall that \( X = l(S_1) + l(S_2) \) is defined in (3.7), and the derivative of the function \( g_s(s_1) \) with respect to \( s \) is calculated as

\[
\frac{\partial g_s(s_1)}{\partial s} = \frac{\partial l^{-1}(s - l(s_1))}{\partial s} = \frac{1}{l'(l^{-1}(s - l(s_1))))} = \frac{1}{l'(g_s(s_1))}.
\] (3.27)

So,

\[
\mathbb{E}[\mathbf{1}_{\{T_1^{(i)} > U_i\}} \mid X = s] = \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \mathbb{P}\left( \xi \geq \frac{U_i - \rho_i S_1}{\sqrt{1 - \rho_i^2}} \wedge S_1 \in [s_1, s_1 + ds_1] \wedge X \in [s, s + ds] \right)
= \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi\left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right]
\times \mathbb{P}(S_1 \in [s_1, s_1 + ds_1] \wedge S_2 \in [g_s(s_1), g_s + ds(s_1)])
= \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi\left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \phi(s_1) ds_1 \phi(g_s(s_1)) \frac{dg_s(s_1)}{ds} ds
= \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi\left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \phi(s_1) \phi'(g_s(s_1))(l'(g_s(s_1)))^{-1} ds_1.
\] (3.28)

Then all \( \mathbb{E}[\mathbf{1}_{\{T_1^{(i)} > U_i\}} \mid X = s] \) can be calculated based on (3.28). Following (3.26), all

\[
\mathbb{E}\left[ \mathbf{1}_{\{T_2^{(i)} > U_i\}} \mid X = s \right]
\]
are known by symmetry. Then \( \mathbb{E}(L_N \mid X = s) \) is calculated by using (3.25).

### 3.3.2 Calculation of \( \mathbb{E}(L_N^2 \mid X = s) \)

Let us assume \( L_N = L_1 + L_2 \), where \( L_i \) for \( i = 1, 2 \) is the loss in the \( i \)th period. Then, using symmetry, we have

\[
\mathbb{E}(L_N^2 \mid X = s) = \mathbb{E}(L_1^2 + L_2^2 + 2L_1 L_2 \mid X = s)
= 2\mathbb{E}(L_1^2 \mid X = s) + 2\mathbb{E}(L_1 L_2 \mid X = s),
\] (3.29)

while, clearly,

\[
\mathbb{E}(L_1^2 \mid X = s) = \sum_{i=1}^{N} u_i^2 \mathbb{E}\left( \mathbf{1}_{\{T_1^{(i)} > U_i\}} \mid X = s \right)
+ \sum_{i,j=1(i \neq j)}^{N} u_i u_j \mathbb{E}\left( \mathbf{1}_{\{T_1^{(i)} > U_i\}} \mathbf{1}_{\{T_1^{(j)} > U_j\}} \mid X = s \right)
\] (3.30)
and
\[ \mathbb{E}(L_1 L_2 \mid X = s) = \sum_{i,j=1}^{N} u_i u_j \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} 1_{\{T^{(j)}_2 > U_j\}} \mid X = s \right). \]  
(3.31)

The value of
\[ \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} \mid X = s \right) \]
was calculated in (3.28). Similarly, following the procedure of (3.28), we can derive
\[ \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} 1_{\{T^{(j)}_2 > U_j\}} \mid X = s \right) \]
\[ = \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \phi\left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \phi\left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \times \phi(s_1) \phi(g(s_1)) (l'(g(s_1)))^{-1} ds_1, \]  
(3.32)

Finally, \( \mathbb{E}(L^2_N \mid X = s) \) is calculated based on (3.28), (3.32) and (3.33). So \( \sigma^2(Y \mid X = s) \) can now be calculated via (3.24).

### 3.3.3 Calculation of \( \partial \sigma^2(Y \mid X = s) / \partial s \)

We have already calculated via (3.24) that
\[ \sigma^2(Y \mid X = s) = \mathbb{E}(L^2_N \mid X = s) - \mathbb{E}^2(L_N \mid X = s). \]  
(3.34)

By adding these values as in (3.25) and (3.29), \( \partial \sigma^2(Y \mid X = s) / \partial s \) can be calculated step by step in a similar procedure.

More specifically, the following should be calculated:
\[ \frac{\partial \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} \mid X = s \right)}{\partial s}, \]  
(3.35)
\[ \frac{\partial \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} 1_{\{T^{(j)}_2 > U_j\}} \mid X = s \right)}{\partial s}, \]  
(3.36)
\[ \frac{\partial \mathbb{E}\left( 1_{\{T^{(i)}_1 > U_i\}} 1_{\{T^{(j)}_2 > U_j\}} \mid X = s \right)}{\partial s}. \]  
(3.37)
Some preparations are required before proceeding with the calculation. First, we define a new function, $\hat{\beta}(\cdot)$, for future derivation:

$$\hat{\beta}(s, s_1) = \frac{\partial \phi(g_s(s_1))}{\partial s} = \phi'(g_s(s_1)) \frac{\partial g_s(s_1)}{\partial s} = \phi'(g_s(s_1)) \frac{1}{l'(g_s(s_1))}. \quad (3.38)$$

Then we find the formula for $f'_X(s)$. Equation (3.22) gives the density function of $f_X(s)$.

However, the values of

$$\lim_{x \to \pm \infty} \frac{l'(x)}{\phi(x)}$$

are required before the analytic formula for $f'_X(s)$ can be derived. Set

$$\lim_{x \to \pm \infty} \frac{l'(x)}{\phi(x)} = \lim_{x \to \pm \infty} \sum_{i=1}^{N} \left\{ u_i \frac{\rho_i}{\sqrt{1 - \rho_i^2}} \exp \left[ -\frac{1}{2} \left( \frac{U_i - \rho_i x}{\sqrt{1 - \rho_i^2}} \right)^2 + \frac{x^2}{2} \right] \right\}$$

$$= \lim_{x \to \pm \infty} \sum_{i=1}^{N} \left\{ u_i \frac{\rho_i}{\sqrt{1 - \rho_i^2}} \exp \left[ -\frac{(2 \rho_i^2 - 1)x^2 - 2 \rho_i U_i x + U_i^2}{2(1 - \rho_i^2)} \right] \right\}$$

$$= 0 \text{ or } \infty. \quad (3.39)$$

In order to continue, we need the following proposition.

**Proposition 3.1** Let

$$F(t) = \int_{b(t)}^{a(t)} f(t, x) \, dx,$$

where $f(t, x) \in C(D)$, $\partial f(t, x)/\partial t \in C(D)$, $D = \{(x, t) \mid x \in [\alpha, \beta], \ t \in [m, n]\}$; $C(D)$ is the set of all continuous functions on $D$. $a'(t)$ and $b'(t)$ exist when $t \in [m, n]$, $\alpha \leq a(t) \leq \beta$ and $\alpha \leq b(t) \leq \beta$. Then

$$F'(t) = f(t, a(t))a'(t) - f(t, b(t))b'(t) + \int_{b(t)}^{a(t)} \frac{\partial f(t, x)}{\partial t} \, dx. \quad (3.40)$$

From (3.39), $\lim_{x \to \pm \infty} l'(x)/\phi(x)$ is $\infty$ or $0$, depending on different sets of $\{\rho_i\}$. If $\lim_{x \to \pm \infty} l'(x)/\phi(x) = 0$, ie, $\lim_{x \to \pm \infty} \phi(x)/l'(x) = \infty$, the continuity in a closed-area condition of Proposition 3.1 is not satisfied. So the formula for $f'_X(s)$ cannot be derived by applying Proposition 3.1 when $s \neq \sum_{i=1}^{N} u_i$. In this case, the $f'_X(s)$ can only be calculated numerically. Then all the first derivatives of the conditional expectations have to be calculated numerically using a simple numerical partial differential equation technique. Clearly, we can find in (3.39) that if each $\rho_i^2 < 0.5$ and $u_i \neq 0$, then $\lim_{x \to \pm \infty} l'(x)/\phi(x)$ is always $\infty$. And the correlation set by risk management practice and regulatory requirements does usually satisfy this inequality. So this restriction can be considered immaterial.
Hereafter, we assume that

$$\lim_{x \to \pm \infty} \frac{\phi(x)}{l'(x)} = 0.$$ 

With this assumption, Proposition 3.1 can be applied as follows.

(i) If

$$2 \sum_{i=1}^{N} u_i > s > \sum_{i=1}^{N} u_i,$$

then, when

$$s = \sum_{i=1}^{N} u_i,$$

$$f_X'(s) = \frac{d}{ds} \int_{\Omega(s)} \phi(s_1) \phi(g_s(s_1))(l'(g_s(s_1)))^{-1} \, ds_1$$

$$= -\phi \left( l^{-1} \left( s - \sum_{i=1}^{N} u_i \right) \right) \phi \left( g_s \left( l^{-1} \left( s - \sum_{i=1}^{N} u_i \right) \right) \right)$$

$$\times \frac{1}{l'(g_s(l^{-1}(s - \sum_{i=1}^{N} u_i))))} \frac{1}{l'(l^{-1}(s - \sum_{i=1}^{N} u_i)))}$$

$$\left[ l'(g_s(s_1)) \frac{\partial \phi(g_s(s_1))}{\partial s} - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] \, ds_1$$

$$= -\phi \left( l^{-1} \left( s - \sum_{i=1}^{N} u_i \right) \right) \left[ \lim_{x \to +\infty} \frac{\phi(x)}{l'(x)} \right] \frac{1}{l'(l^{-1}(s - \sum_{i=1}^{N} u_i)))}$$

$$+ \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2}$$

$$\times \phi(s_1) \left[ l'(g_s(s_1)) \frac{\partial \phi(g_s(s_1))}{\partial s} - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] \, ds_1$$

$$= \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2}$$

$$\times \phi(s_1) \left[ l'(g_s(s_1)) \beta(s, s_1) - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] \, ds_1,$$

(3.41)

when

$$s = \sum_{i=1}^{N} u_i,$$

$$\Omega(s)$$

will be $$(-\infty, +\infty),$$

$$f_X'(s) = \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2} \phi(s_1) \left[ l'(g_s(s_1)) \beta(s, s_1) - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] \, ds_1.$$
(ii) If \( \sum_{i=1}^{N} u_i > s > 0 \),

\[
\begin{align*}
f_X'(s) &= \frac{d}{ds} \int_{\Omega(s)} \phi(s_1) \phi(g_s(s_1)) (l'(g_s(s_1)))^{-1} ds_1 \\
&= \left\{ \phi(l^{-1}(s)) \phi(g_s(l^{-1}(s))) \frac{1}{l'(g_s(l^{-1}(s)))} \\
&\quad + \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2} \\
&\quad \times \phi(s_1) \left[ l'(g_s(s_1)) \frac{\partial \phi(g_s(s_1))}{\partial s} - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] ds_1 \right\} \\
&= \left\{ \phi(l^{-1}(s)) \left[ \lim_{x \to -\infty} \frac{\phi(x)}{l'(x)} \right] \\
&\quad + \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2} \\
&\quad \times \phi(s_1) \left[ l'(g_s(s_1)) \frac{\partial \phi(g_s(s_1))}{\partial s} - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] ds_1 \right\} \\
&= \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2} \\
&\quad \times \phi(s_1) \left[ l'(g_s(s_1)) \beta(s, s_1) - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] ds_1.
\end{align*}
\]

(3.42)

It is interesting to find out that in both cases the formula for \( f_X'(s) \) is the same:

\[
f_X'(s) = \int_{\Omega(s)} \frac{1}{(l'(g_s(s_1)))^2} \phi(s_1) \left[ l'(g_s(s_1)) \beta(s, s_1) - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right] ds_1.
\]

(3.43)

To simplify the rest of the calculations, define a function \( p(s, s_1) \) as

\[
p(s, s_1) = \frac{1}{(l'(g_s(s_1)))^2} \phi(s_1) \left[ l'(g_s(s_1)) \beta(s, s_1) - \phi(g_s(s_1)) \frac{\partial l'(g_s(s_1))}{\partial s} \right].
\]

(3.44)

So \( f_X'(s) \) can be rewritten as

\[
f_X'(s) = \int_{\Omega(s)} p(s, s_1) ds_1.
\]

(3.45)
(3.38) and (3.45), it follows that

$$
\frac{\partial}{\partial s} \mathbb{E}(1_{\{T_1^{(i)} > U_i\}} \mid X = s)
\begin{align*}
&= \frac{\partial}{\partial s} \left\{ \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} \right] ds_1 \right\} \\
&= \frac{1}{f_X^2(s)} \left\{ \frac{\partial}{\partial s} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 \right\} \\
&\quad \times \frac{\partial}{\partial s} \left\{ \frac{\partial}{\partial s} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 \right\} \\
&\quad \times \int_{s_1 \in \Omega(s)} \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 f_X(s) \\
&= \frac{1}{f_X^2(s)} \left\{ f_X(s) \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] p(s, s_1) ds_1 \\
&\quad + \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 \right\} \\
&\quad \times \int_{s_1 \in \Omega(s)} \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 f_X(s) \\
&\quad \times \int_{s_1 \in \Omega(s)} \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 f_X(s) \right\} .
\end{align*}
$$

(3.46)

Similarly, but now using (3.32) instead of (3.28), we obtain

$$
\frac{\partial}{\partial s} \mathbb{E}(1_{\{T_1^{(i)} > U_i\}} 1_{\{T_1^{(j)} > U_j\}} \mid X = s)
\begin{align*}
&= \frac{\partial}{\partial s} \left\{ \frac{1}{f_X(s)} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \Phi \left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \times \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 \right\} \\
&= \frac{1}{f_X^2(s)} \left\{ \frac{\partial}{\partial s} \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \Phi \left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \times \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 \right\} \\
&\quad \times \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 f_X(s) \\
&\quad + \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \Phi \left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \times \phi(s_1) \phi(g(s_1))(l'(g(s_1)))^{-1} ds_1 f_X(s) \right\} .
\end{align*}
$$

www.risk.net/journal

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
Similarly, but now using (3.33) instead of (3.32), we obtain

\[
\frac{\partial}{\partial s} \mathbb{E} \left( \mathbf{1}_{\{T_1^{(i)} > U_i\}} \mathbf{1}_{\{T_2^{(j)} > U_j\}} \mid X = s \right) = \frac{1}{f_X(s)} \left\{ f_X(s) \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \Phi \left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \right. \\
\times \left. \phi(s_1) \phi(g_s(s_1)) (l'(g_s(s_1)))^{-1} \, ds_1 \right\} \\
\times \left\{ \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \beta_i s_1}{\sqrt{1 - \beta_i^2}} \right) \right] ds_1 \right. \\
\left. \times \frac{\rho_j}{\sqrt{1 - \rho_j^2}} \phi(s_1) \phi(g_s(s_1)) (l'(g_s(s_1)))^{-2} \, ds_1 \right\} \\
+ \left. \int_{s_1 \in \Omega(s)} \left[ 1 - \Phi \left( \frac{U_i - \rho_i s_1}{\sqrt{1 - \rho_i^2}} \right) \right] \left[ 1 - \Phi \left( \frac{U_j - \rho_j s_1}{\sqrt{1 - \rho_j^2}} \right) \right] \right. \\
\times \left. \phi(s_1) \phi(g_s(s_1)) (l'(g_s(s_1)))^{-1} \, ds_1 \right\}.
\]

(3.48)
Finally, every component of the second-order derivative in (3.23) is derived, and \( \alpha_q(L_N) \) can be calculated, based on (3.4).

4 NUMERICAL RESULTS

Two sets of numerical results are shown and discussed in this section. First, the accuracy of the analytical VaR model is tested against the Monte Carlo simulation results. The two-period Monte Carlo model with portfolio rebalancing simulates the portfolio loss as outlined in (3.1). If an asset defaults during the first period, the defaulted asset will be replaced with one having the same notional and rating in the second period. It can be viewed as a simplified factor simulation model for IRC.

From a practical application perspective, it will be interesting to see if the model, which is essentially a second-order Taylor expansion, is sufficient to capture risk features in the context of the liquidity horizon and a constant level of risk. We designed several test models: a one-period ASRF model and GA approximation; a one-period Monte Carlo model; and a two-period Monte Carlo model with and without portfolio rebalancing.

The ASRF is calculated based on (2.4). The one-period model is listed in (2.1). The two-period model without portfolio rebalancing is similar to (3.1); however, the loss in the second period is conditional on the default in the first period, ie, the loss function

\[
\tilde{L}_N = \sum_{i=1}^{N} \left[ u_i 1_{\{ T_1^{(i)} > U_i \}} + u_i 1_{\{ T_1^{(i)} \leq U_i \}} 1_{\{ T_2^{(i)} > U_i \}} \right].
\] (4.1)

In all the numerical tests shown below, we assume a credit portfolio in which each asset is modeled as notional = LGD = 1, a specific first-/second-period default probability (PD) and a correlation within the one-factor framework (ie, each asset is correlated to a single common factor). The VaR at any given percentile is expressed as a percentage of the total notional of the portfolio. Due to the consistency of the one- and two-period models, the default probability of the whole period in the one-period model is \( 1 - (1 - PD)^2 \).

4.1 Assessing the accuracy of the model

The results of the two-period GA VaR model are shown in the fourth and fifth columns in Table 2 on the next page and Table 3 on the next page, where we show the 99.9th percentile VaR in different scenarios computed by different models. In Table 2, we assume a portfolio of 100 assets with the uniform first-/second-period PD being 1% and different levels of correlation. The correlation is fixed at 0.5 and the number of assets in the portfolio is changed in Table 3. Compared with the two-period Monte Carlo model results listed in the sixth column of the two tables, the two-period GA VaR
### TABLE 2  
Comparison of asymptotic single-risk-factor VaR, one-period granularity adjustment VaR, two-period granularity adjustment VaR and two-period Monte Carlo VaR with different $\rho$ ($N = 100$, PD = 1%, LGD = 1).

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>ASRF VaR</th>
<th>One-period GA VaR</th>
<th>One-period Monte Carlo VaR</th>
<th>Two-period conditional VaR</th>
<th>Two-period Monte Carlo VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>3.96</td>
<td>10.72</td>
<td>8.00</td>
<td>11.76</td>
<td>8.00</td>
</tr>
<tr>
<td>0.2</td>
<td>7.11</td>
<td>10.72</td>
<td>10.00</td>
<td>10.14</td>
<td>9.00</td>
</tr>
<tr>
<td>0.3</td>
<td>11.84</td>
<td>14.35</td>
<td>14.00</td>
<td>12.32</td>
<td>12.00</td>
</tr>
<tr>
<td>0.4</td>
<td>18.56</td>
<td>20.49</td>
<td>20.60</td>
<td>17.09</td>
<td>17.00</td>
</tr>
<tr>
<td>0.5</td>
<td>27.77</td>
<td>29.32</td>
<td>29.00</td>
<td>24.68</td>
<td>24.10</td>
</tr>
<tr>
<td>0.6</td>
<td>40.05</td>
<td>41.31</td>
<td>41.60</td>
<td>35.76</td>
<td>36.00</td>
</tr>
<tr>
<td>0.7</td>
<td>55.97</td>
<td>56.97</td>
<td>57.00</td>
<td>51.18</td>
<td>50.70</td>
</tr>
<tr>
<td>0.8</td>
<td>75.61</td>
<td>76.35</td>
<td>76.50</td>
<td>71.64</td>
<td>71.40</td>
</tr>
<tr>
<td>0.9</td>
<td>95.20</td>
<td>95.61</td>
<td>95.60</td>
<td>94.16</td>
<td>94.30</td>
</tr>
</tbody>
</table>

All values are given in percent. Here $\rho$ refers to the quantity $\rho_i$ in (2.2), (3.2) and (3.2) and is the square of the corresponding Basel II correlation.

### TABLE 3  
Comparison of asymptotic single-risk-factor VaR, one-period granularity adjustment VaR, two-period granularity adjustment VaR and two-period Monte Carlo VaR with different numbers of assets $N$ ($\rho = 0.5$, PD = 1%, LGD = 1).

<table>
<thead>
<tr>
<th>$N$</th>
<th>ASRF VaR</th>
<th>One-period GA VaR</th>
<th>One-period Monte Carlo VaR</th>
<th>Two-period conditional VaR</th>
<th>Two-period Monte Carlo VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>27.77</td>
<td>35.53</td>
<td>35.00</td>
<td>31.49</td>
<td>30.00</td>
</tr>
<tr>
<td>40</td>
<td>27.77</td>
<td>31.65</td>
<td>32.50</td>
<td>27.23</td>
<td>27.53</td>
</tr>
<tr>
<td>60</td>
<td>27.77</td>
<td>30.36</td>
<td>30.00</td>
<td>25.82</td>
<td>25.17</td>
</tr>
<tr>
<td>80</td>
<td>27.77</td>
<td>29.71</td>
<td>31.25</td>
<td>25.11</td>
<td>25.25</td>
</tr>
<tr>
<td>100</td>
<td>27.77</td>
<td>29.32</td>
<td>29.00</td>
<td>24.68</td>
<td>24.10</td>
</tr>
<tr>
<td>200</td>
<td>27.77</td>
<td>28.55</td>
<td>28.50</td>
<td>23.83</td>
<td>23.95</td>
</tr>
<tr>
<td>400</td>
<td>27.77</td>
<td>28.16</td>
<td>28.50</td>
<td>23.41</td>
<td>23.36</td>
</tr>
<tr>
<td>500</td>
<td>27.77</td>
<td>28.08</td>
<td>27.80</td>
<td>23.32</td>
<td>23.50</td>
</tr>
<tr>
<td>700</td>
<td>27.77</td>
<td>27.99</td>
<td>28.14</td>
<td>23.23</td>
<td>23.42</td>
</tr>
<tr>
<td>1000</td>
<td>27.77</td>
<td>27.92</td>
<td>27.00</td>
<td>23.15</td>
<td>23.36</td>
</tr>
</tbody>
</table>

All values are given in percent.

The model serves as a reasonable approximation. It has the same behavior as the standard GA in the sense that it converges to Monte Carlo in the case of a high correlation and a large portfolio. Since we provide this general two-period GA VaR model for the credit portfolio that accounts for the liquidity horizon and portfolio rebalancing, as
proposed in the IRC model, we also compare our results with Monte Carlo simulated results on a portfolio with long and short positions. The results in Table 4 on page 24 again show that our two-period GA VaR model serves as a reasonable approximation.

Figure 1 shows the ratios of the ASRF, the ASRF plus the standard GA and the two-period GA VaR to the two-period Monte Carlo VaR with respect to different numbers of assets. It can be seen that the two-period GA VaR is a reasonable approximation to the full simulation model.

4.2 Assessing model behavior by comparison with other benchmark models

In order to assess whether the two-period GA VaR can capture the impact of the liquidity horizon and portfolio rebalancing, we first show the differences in VaR via a full Monte Carlo simulation of the one- and two-period models in Table 2 on the facing page and Table 3 on the facing page. The differences are material and become smaller when the correlation is very small and close to 1. In order to understand the
differences, we designed an additional Monte Carlo model in which the portfolio is not rebalanced. The results of three Monte Carlo models at the tail distributions (99th percentile and above) are shown in Figure 2. Initially, given all parameters are the same, the one-period Monte Carlo has the largest VaR at the most percentile points. When we employ the two-period Monte Carlo without portfolio rebalancing, the VaR values at different percentile points reduce. When we switch on the portfolio rebalancing, the VaR values become large but still smaller than those of one-period points. This interesting behavior illustrates two competing factors.

1. The so-called correlation leaking effect within the multi-period factor modeling framework. It is well known that the correlated default scenarios are different in one-period and multi-period simulations, as discussed by Straumann (2009). In our example, a two-period Monte Carlo simulation would cut off the possible joint default events in the first and second periods. This in general leads to fewer joint defaults given the same level of correlation.
FIGURE 3 One-and two-period Monte Carlo path values and two-period conditional VaR, when PD = 7%, N = 200, ρ = 0.5.

(2) The portfolio rebalancing assumption at the end of first period means that, should an asset default in the first period, it will be replaced with a similar one in the second period, and the new asset can also default. This will add more default scenarios than the one-period simulation. This is why, even in the limit of a perfect correlation, the one-period Monte Carlo will be different from the multi-period Monte Carlo.

The results in Table 2 on page 20, Table 3 on page 20 and Table 4 on the next page show that the two-period GA VaR results are close to the full Monte Carlo simulation results and change accordingly when the portfolio rebalancing assumption changes. The two-period GA VaR results at different percentiles are shown in Figure 2 on the facing page and Figure 3. The GA approximation, which is the second-order Taylor expansion, captures the impact of the liquidity horizon and rebalancing, providing a sensible comparable approach to the full Monte Carlo simulation.

The analytical VaR model is further assessed to see how good it is at capturing concentration risk in the presence of portfolio rebalancing. We designed two test cases. In the first case, we assume a portfolio of fifty assets with nonuniform PDs
TABLE 4  Comparison of the two-period granularity adjustment VaR and two-period Monte Carlo VaR with different $\rho$ ($N = 300$, $PD = 1\%$, $LGD = -1$ for the first ten assets and $LGD = 1$ for the rest of the assets).

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>Two-period conditional VaR</th>
<th>Two-period Monte Carlo VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>6.73</td>
<td>5.68</td>
</tr>
<tr>
<td>0.2</td>
<td>7.45</td>
<td>7.18</td>
</tr>
<tr>
<td>0.3</td>
<td>10.53</td>
<td>10.39</td>
</tr>
<tr>
<td>0.4</td>
<td>15.75</td>
<td>15.71</td>
</tr>
<tr>
<td>0.5</td>
<td>23.63</td>
<td>23.50</td>
</tr>
<tr>
<td>0.6</td>
<td>34.91</td>
<td>34.57</td>
</tr>
<tr>
<td>0.7</td>
<td>50.52</td>
<td>50.71</td>
</tr>
<tr>
<td>0.8</td>
<td>71.15</td>
<td>71.29</td>
</tr>
<tr>
<td>0.9</td>
<td>93.89</td>
<td>93.82</td>
</tr>
</tbody>
</table>

All values are given in percent.

FIGURE 4  Comparison of the ratio of ASRF, one-period and two-period analytic VaR to two-period Monte Carlo VaR with respect to different weights of the first asset (1–40%) when $N = 100$, $\rho = 0.5$, $PD = 0.1\%$. 
TABLE 5 Comparison of asymptotic single-risk-factor VaR, one- and two-period granularity adjustment VaR and two-period Monte Carlo VaR with respect to mixed default probabilities ($\rho = 0.5$, $N = 50$, LGD = 1).

<table>
<thead>
<tr>
<th>PD (%)</th>
<th>ASRF VaR</th>
<th>One-period GA VaR</th>
<th>Two-period conditional VaR</th>
<th>Monte Carlo VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed</td>
<td>31.17</td>
<td>34.31</td>
<td>31.70</td>
<td>31.90</td>
</tr>
<tr>
<td>5</td>
<td>61.32</td>
<td>64.70</td>
<td>62.44</td>
<td>62.47</td>
</tr>
<tr>
<td>1</td>
<td>27.77</td>
<td>30.89</td>
<td>26.38</td>
<td>26.39</td>
</tr>
<tr>
<td>0.10</td>
<td>6.18</td>
<td>8.60</td>
<td>7.42</td>
<td>7.80</td>
</tr>
</tbody>
</table>

All values are given in percent.

TABLE 6 Comparison of asymptotic single-risk-factor VaR, one- and two-period granularity adjustment VaR and two-period Monte Carlo VaR with respect to different notional weights of the first asset (the other assets are equally weighted, PD = 1%, $N = 100$, $\rho = 0.5$, LGD = 1).

<table>
<thead>
<tr>
<th>Weight (%)</th>
<th>ASRF VaR</th>
<th>One-period GA VaR</th>
<th>Two-period conditional VaR</th>
<th>Monte Carlo VaR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>27.77</td>
<td>29.32</td>
<td>24.68</td>
<td>24.10</td>
</tr>
<tr>
<td>11</td>
<td>27.77</td>
<td>30.89</td>
<td>26.40</td>
<td>26.28</td>
</tr>
<tr>
<td>21</td>
<td>27.77</td>
<td>35.60</td>
<td>31.55</td>
<td>33.77</td>
</tr>
<tr>
<td>31</td>
<td>27.77</td>
<td>43.44</td>
<td>40.14</td>
<td>42.85</td>
</tr>
<tr>
<td>41</td>
<td>27.77</td>
<td>54.42</td>
<td>52.17</td>
<td>51.13</td>
</tr>
</tbody>
</table>

All values are given in percent.

with the results shown in Table 5. Table 6 shows the case with nonuniform notional of the asset in the portfolio by change the weight of one asset. We can see that in both cases the GA VaR model behaves reasonably well, since our approach is more consistently closer to the true (i.e., two-period Monte Carlo) results. More results with different notional weights are plotted in Figure 4 on the facing page. We can clearly see that it is necessary to make a granularity adjustment in the two-period model to capture concentration risk, and a two-period GA does a better job than the one-period GA.

Note also that, due to the multiple defaults, the 100th percentile for two-period VaR is always higher than that for the one-period VaR since the entire portfolio can default multiple times. It is difficult to show in Figure 2 on page 22, but if we increase the PD to 7%, we can clearly see that two models cross at the 99.9th percentile. With a higher PD, we have more chances that one asset defaults in the first period and the replacement in the second period asset also defaults. Figure 2 on page 22 and Figure 3 on page 23 show that the tail distributions in the presence of the liquidity
horizon and portfolio rebalancing are different from standard one-period models such as ASRF. Depending on the credit quality of the portfolio, standard ASRF can be both conservative and aggressive.

5 DISCUSSION

This paper provides a general two-period GA VaR model for the credit portfolio that accounts for liquidity horizon and portfolio rebalancing. The portfolio is rebalanced at the end of the first period so that the constant level of risk can be maintained. The profit and loss (P&L) and VaR contributions from the second period are conditional on the portfolio rebalancing assumptions. The methodology is an extension of the GA model.

We examined the accuracy of the model by comparing it against a two-period Monte Carlo model with portfolio rebalancing. As expected, our analytical model has very similar behavior to the standard GA in capturing concentration risk of the credit portfolio. The model’s behavior is also assessed by comparing it against a one-period Monte Carlo model, a two-period Monte Carlo with and without portfolio rebalancing, a standard ASRF and a standard (one-period) GA. Our main conclusions can be summarized as follows.

- As expected, when compared with two-period Monte Carlo with portfolio rebalancing, our analytical model serves as a reasonable approximation and has very similar behavior to the standard GA in capturing concentration risk of the credit portfolio.

- More importantly, the two-period GA VaR model captures the impact of the liquidity horizon and portfolio rebalancing as confirmed by Monte Carlo simulation. Our method takes a similar approach to the standard Monte Carlo-based IRC model, with much higher computational efficiency.

- We also show that the standard one-period ASRF (with and without standard GA) is not enough to achieve a risk measure comparable with the liquidity horizon and portfolio rebalancing. The tail distributions with and without portfolio rebalancing are different due to two competing factors: the default correlation and its relationship with asset correlations are different for different time windows; portfolio rebalancing allows multiple defaults. This addresses the fact that the defaulted asset will be replaced with another asset, which can default again. Because of this nonzero multiple default feature, the standard ASRF model will always be aggressive at the 100th percentile but can be both conservative and aggressive for other percentiles, depending on the credit quality of the portfolio.
In this paper we consider only the one-factor case. It can be expanded to the multi-factor case by proposed by Pykhtin (2004). In the actual credit portfolio, different trades are assigned different liquidity horizons, which can be modeled readily in the current approach. For example, the longer liquidity horizon can be modeled by assuming no portfolio rebalancing at the end of the first period. Although the rating is taken as the constant level of risk measure, we do not consider the rating migration P&L. Extending the current approach to include rating P&L is straightforward, but makes the analytical solution too complicated.

The model can also be extended to other types of risk factors, by modeling the portfolio rebalancing assumption via some discretized values such as rating ranks. The portfolio rebalancing assumption can be either exposure based, as is the case for the credit portfolio discussed in this paper, or risk based, which is defined as the sensitivity to the risk factors. In our opinion, this direction of research is important in order to address liquidity modeling, as discussed in the trading book fundamental review in Basel Committee on Banking Supervision (2013a).

DECLARATION OF INTEREST

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

REFERENCES


Research Paper

Loss distributions: computational efficiency in an extended framework

Daniel H. Stahl

BB&T, 602 Lockland Ave, Winston Salem, NC 27103; email: dstahl@bbandt.com

(Received January 20, 2015; revised May 27, 2015; accepted June 8, 2015)

ABSTRACT

Credit risk models are difficult to implement in an actionable form. While near real-time results are required for pricing credits, making origination decisions and optimizing portfolio allocation, the complexity of credit models (which are often employed on portfolios spanning millions of exposures) typically requires either expensive Monte Carlo simulations or the imposition of inflexible assumptions to compute the portfolio loss distribution. The contribution of this paper to the credit risk literature is twofold: Credit Suisse’s CreditRisk⁺ framework is significantly generalized, and an algorithm from the option pricing literature is introduced to retain precision and speed for the computation of the loss distribution even for very large portfolios. This generalization allows for time-dependent portfolios, fully accounts for granularity and concentration within the credit portfolio and does not rely on assumptions of asymptotically large credit portfolios. The algorithm allows even a standard laptop to precisely compute an entire bank’s loss distribution. The core results of this paper are as follows: the model uses a stochastic process as the state variable yet retains tractability; the model integrates liquidity risk into the credit loss distribution; the algorithm from the option pricing literature is leveraged for efficient computation.

Keywords: CreditRisk⁺; mixture models; credit portfolio; Fourier inversion; loss distribution.
1 INTRODUCTION

It is a common credit risk modeling technique to introduce random variables that jointly affect the probability of default of distinct creditors. This class of models is termed “mixture” models due to the “mixing” of the default probability with other random variables. Categories of mixture models differ in the specification of the random variables that effect the probability of default. Merton (1974) created one of the first mixture models for estimating the probability of default for a single firm by defining default as the realization of negative equity at the some fixed time horizon. The firm’s assets are modeled by a diffusion process that serves as the “mixing” random variable. Since default is determined by an underlying economic factor, models for a firm’s assets are dubbed “factor” or “structural” models. Black and Cox (1976) extend Merton (1974) by defining default as the first occurrence of negative equity at any time prior to the time horizon. Both of these models feature endogenous default. “Intensity” or “reduced-form” models are another strand of research that treats the cause of default as exogenous. In these models the default intensity (which, for small probabilities of default, is roughly the probability of default) follows a latent stochastic process. Jarrow and Turnbull (1995) pioneered this approach; Duffie (2005) provides a detailed overview.

Merton, Black and Cox, and Jarrow and Turnbull focus on modeling individual defaults. For many banks the entire loss distribution of credit portfolios is of more concern than pricing individual defaults. Early research in modeling the loss distribution extended Merton’s model. Vasicek (1987, 1997) uses the Merton framework, with each firm’s asset returns following correlated diffusion processes. He then channels the Capital Asset Pricing Model by reducing each firm’s asset returns into an idiosyncratic component and a systemic component. The systemic component drives correlation between defaults. In the idealized portfolio of homogeneous and atomic assets, the Vasicek model has an analytical solution for the portfolio density. Extensions to Vasicek include multi-dimensional variations of systemic factors (Fok et al 2014). A common constraint on these models is that \( \lim_{n \to \infty} \sum_{i}^{n} w_{i}^{2} \to 0 \), where \( w_{i} \) is the relative exposure of asset \( i \). For large banks, \( n \) is typically rather large and the assets are relatively diverse, satisfying the constraint.

Several industry models, eg, JP Morgan’s CreditMetrics (JP Morgan 1997), use a more sophisticated version of Merton’s model and compute the loss distribution via Monte Carlo techniques. These models capture granularity and exposure concentrations, but are computationally expensive. Credit Suisse First Boston’s CreditRisk+ (Credit Suisse 2001) took a different approach to mixture models. Instead of modeling underlying factors, the probability of default is a linear combination of the mixture variables. The mixtures in these models lose the economic interpretation of a firm’s assets, but gain considerable computational advantage. However, the computational
advantage is obtained by rounding exposures to integer values. For a very diverse portfolio, this rounding leads to inaccurate modeling. As default is exogenous, these models share conceptual similarities to reduced-form models.

A common trade-off with credit models is that either restrictive assumptions are introduced (e.g., an “infinitely” large portfolio in Vasicek’s model or the rounding of exposures in the CreditRisk+ model), or the computation of the loss distribution is prohibitively time intensive. The trade-off between complexity and accuracy is ubiquitous in any computational field, but is especially onerous when modeling credit portfolios where millions of data points must be analyzed while still providing timely results for business decisions. Indeed, computational convenience is essential for providing actionable information for a variety of banking decisions. Loan committees need to make real-time decisions about the acceptability and pricing of the risk in the loans they approve. Risk and capital committees make decisions for optimal portfolio allocations among loan asset classes, requiring efficient algorithms for the computation of marginal risks of individual assets and asset classes within the loan portfolio. The contribution of this paper to the credit risk literature is twofold: the CreditRisk+ framework is significantly generalized to include liquidity risk, stochastic exposures and stochastic processes as mixture variables; and an algorithm from the option pricing literature is introduced to retain precision and speed for the computation of the loss distribution even for very large portfolios. This framework allows for time-dependent portfolios, fully accounts for granularity and concentration within the portfolio, does not rely on asymptotically large portfolios and can compute an entire bank’s loss distribution on a standard laptop.

2 MODEL DESCRIPTION

2.1 Assumptions

There exists a portfolio of defaultable assets. The random variable describing this portfolio resides in a probability space \((\Omega, \mathcal{G}, \mathbb{P})\) with a filtration \(\mathcal{G}_t\) and a filtration \(\mathcal{F}_t \subset \mathcal{G}_t\) generated by \(W_t\), a one-dimensional Brownian motion. It is natural to consider the time period \(0 \leq t \leq T < \infty\). Let \(\tau_j, 0 < j \leq n\), be stopping times with respect to \(\mathcal{G}_t\) and let \(X_{ij}\) be functions of \(\tau_j\) such that

\[
X_{ij} = \begin{cases} 
  l_j, & \tau_j \in [t, T], \\
  0, & \tau_j \notin [t, T],
\end{cases}
\]

(2.1)

where \(l_j\) are mutually independent random variables. Fixing \(T\), \(X_{ij}\) can be considered functions of \(\tau_j\) and hence are also random variables with the following distribution:

\[
\begin{cases} 
  l_j, & \mathbb{P}(\tau_j \in [t, T] | \mathcal{G}_t), \\
  0, & 1 - \mathbb{P}(\tau_j \in [t, T] | \mathcal{G}_t).
\end{cases}
\]

(2.2)
$X_t^j$ can be interpreted as a return on a defaultable asset. Each $X_t^j$ has a random exposure of $l_j$ and a potential deterministic gain $R_j$ at some time $T$ if default has not occurred in $[0, T)$. For simplicity, it is possible to consider the exposure $l_j$ and a zero gain: once the loss distribution for the portfolio is computed, the entire distribution can be shifted by the sum of the expected returns to recover the profit and loss distribution. The random variable describing the portfolio loss is denoted $X_t = \sum_j X_t^j$. From the definition of $X_t^j$, $X_0 = 0$.

The stopping times are distributed as follows:

1. $\mathbb{P}(\tau_j \in [0, T] \mid \mathcal{F}_T) = p_j Y_T$,
2. $\tau_j$ and $\tau_k$, $j \neq k$, are independent conditioned on $\mathcal{F}_T$,

where each $p_j$ is a constant and $Y_t$ is a random process adapted to $\mathcal{F}_T$. $\mathcal{F}_T$ represents the information available only from the Brownian motion. Knowledge of $\mathcal{F}_T$ does not imply knowledge of the number of defaults in $[0, t]$. However, knowledge of $\mathcal{G}_t$ implies knowledge of all defaults in $[0, t]$. $\mathcal{F}_t$ is useful since, conditioned on $\mathcal{F}_T$, the defaults are mutually independent. While the stopping times are independent conditioned on $\mathcal{F}_T$, they are unconditionally dependent. The unconditional dependence of the stopping times drives correlation between $X_t^j$. The rationale for using both $\mathcal{F}_t$ and $\mathcal{G}_t$ is perhaps best exemplified by considering a possible Monte Carlo estimation of the portfolio loss. Generating a single path of $Y_t$ (that is adapted to $\mathcal{F}_t$) yields information on the probability of default, but does not actually provide any information on the actual default event. Once a single path of $Y_t$ is generated, the probability of default for each asset can be determined. The process $X_t$ can then be simulated using the probability of default generated by $Y_t$. The process $X_t$ is adapted to $\mathcal{G}_t$, containing all the information in $\mathcal{F}_t$ and the information for the actual defaults. For a detailed discussion on the subject of filtrations in the context of defaults, see Duffie (2001, Appendix I).

Finally, $Y_t$ is defined as

$$Y_t = \int_0^t Z_s \, ds,$$

and $Z_s$ solves the following stochastic differential equation (SDE):

$$dZ_s = \alpha(1 - Z_s) \, dt + \sigma \sqrt{Z_s} \, dW_t. \tag{2.4}$$

$Z_s$ is mean-reverting, nonnegative and has a long-run expected value of 1 (see online Appendix A.1). Taking the differential of $p_j Y_t$, $d(p_j Y_t) = p_j Z_t \, dt$. The long-run instantaneous default rate is thus $p_j \, dt$, permitting the interpretation of $p_j$ as the long-run default rate for $X_t^j$ per unit time. $Y_t$ can be interpreted as the effect the state of the economy has on the probability of default.
Technically, $Y_t$ should be constrained to $[0, 1/p_k]$, where $p_k$ is the large $p_j$. Without this constraint the probability of default can be greater than 1. However, relaxing this constraint allows the model to have a tractable solution. In most practical applications the probability that any $p_jY_t$ is greater than 1 is too small to be consequential.

In summary, the model contains a portfolio of $n$ assets with the following features:

1. they have correlated defaults;
2. the default correlation is driven by a latent $\mathcal{F}_t$-measurable variable $Y_t$;
3. each asset has a random but uncorrelated exposure $l_j$.

### 2.2 Characteristic function

The characteristic function of a random variable is the Fourier transform of its density. By the Fourier inversion theorem, taking the inverse Fourier transform of the Fourier transform recovers the density (Hewitt and Stromberg 1965). The aim of this section is to derive an analytic solution to the characteristic function of the portfolio loss distribution. The characteristic function of $X_t$ is

$$\phi_X(u, t) = \mathbb{E}[e^{uiX_T} \mid \mathcal{G}_t]$$

$$= \mathbb{E}[\mathbb{E}[e^{uiX_T} \mid \mathcal{F}_T] \mid \mathcal{G}_t]$$

$$= \mathbb{E}\left[\prod_j (p_j(Y_T - Y_t)e^{uil_j} + 1 - p_j(Y_T - Y_t)) \mid \mathcal{G}_t\right]$$

$$= \mathbb{E}\left[\prod_j (p_j(Y_T - Y_t)\mathbb{E}[e^{uil_j}] + 1 - p_j(Y_T - Y_t)) \mid \mathcal{G}_t\right]$$

$$= \mathbb{E}\left[\prod_j (p_j(Y_T - Y_t)\phi_{l_j} + 1 - p_j(Y_T - Y_t)) \mid \mathcal{G}_t\right],$$

where $\phi_{l_j}$ is the characteristic function of $l_j$. Note that, if an asset has defaulted in $[0, t]$, then at time $t$ it has value 0 with probability 1. The loss has already been realized and does not contribute to the loss distribution conditioned on $\mathcal{G}_t$.

For portfolios of reasonable size, the idiosyncratic nature of $l_j$ diversifies away the risk of loss for each $l_j$, causing the distribution to converge to the distribution where exposures are constant. As shown in online Appendix A.5, the contribution of the variance of $l_j$ to the portfolio is proportional to $n$, which is quickly dwarfed by a term that is proportional to $n^2$. However, there is no computational penalty for including a stochastic $l_j$, and the advanced approach in Basel II encourages stochastic exposures (Basel Committee on Banking Supervision 2006).
If $p_j(Y_T - Y_t)$ is small (as is typical of default models), then the following approximation holds by Taylor’s theorem:

$$\phi_X(u, t) \approx \mathbb{E} \left[ \exp \left( \sum_j p_j(Y_T - Y_t)(\phi_{l_j}(u) - 1) \right) \mid \mathcal{G}_t \right]. \quad (2.5)$$

If the exposures are deterministic, the expression within the expectation is the characteristic function of a Poisson random variable and the Taylor approximation is equivalent to $X^j_t$ following a Poisson distribution conditional on $\mathcal{F}_T$ instead of a Bernoulli distribution. As a Poisson random variable has sample space over the integers, the assumption that $X^j_t$ has a Poisson distribution implies that assets can have multiple “defaults”, though the probability of more than one default is typically very small.

Substituting $p_j(Y_T - Y_t) = p_j \int_t^T Z_s \, ds$ into the approximate characteristic function yields the following expression for $\phi_X$:

$$\phi_X(u, t) \approx \mathbb{E} \left[ \exp \left( \int_t^T Z_s \, ds \sum_j p_j(\phi_{l_j}(u) - 1) \right) \mid \mathcal{G}_t \right]$$

$$\approx \mathbb{E} \left[ \exp \left( v \int_t^T Z_s \, ds \right) \mid \mathcal{G}_t \right],$$

where $v = \sum_j p_j(\phi_{l_j}(u) - 1)$. Since $v$ is deterministic, $\mathbb{E}[\exp(v \int_t^T Z_s \, ds) \mid \mathcal{G}_t]$ is the moment-generating function of $\int_t^T Z_s \, ds$.

The Feynman–Kac theorem states that, for a suitable function $h$ and diffusion process $dB_t = \alpha(B_t, t) \, dt + \sigma(B_t, t) \, dW_t$,

$$g(B_t, t) = \mathbb{E} \left[ \exp \left( \int_t^T r(B_s) \, ds \right) h(B_T) \mid \mathcal{F}_t \right]$$

satisfies the following partial differential equation (PDE) (Øksendal 2007):

$$\frac{\partial g}{\partial t} + \frac{\partial g}{\partial b} \alpha(b, t) + \frac{\partial^2 g}{2 \partial b^2} \sigma^2(b, t) + rg = 0,$$

$$g(b, T) = h(b). \quad (2.6)$$

Letting

$$\phi(u, t) = \eta(z, t; v) = \mathbb{E} \left[ \exp \left( v \int_t^T Z_s \, ds \right) \mid \mathcal{G}_t \right]$$

and leveraging the Feynman–Kac theorem, $\eta(z, t; v)$ satisfies

$$\frac{\partial \eta}{\partial t} + \frac{\partial \eta}{\partial z} \alpha(1 - z) + \frac{1}{2} \frac{\partial^2 \eta}{\partial z^2} \sigma^2 z + vz = 0,$$

$$\eta(z, T; v) = 1. \quad (2.7)$$
Since this PDE is affine in \( z \), the solution has the form \( \eta(z, t; v) = \exp(A(t, T) + C(t, T)z) \) (Duffie and Kan 1996). The set of SDEs that satisfy this affine form includes the Vasicek (1977) and Cox et al (1985) interest rate models in one dimension and the Heston (1993) model in two dimensions. These “affine” models are computationally tractable even in higher dimensions, since they only require solving a system of ordinary differential equations (ODEs) (Duffie and Kan 1996). In the case where \( z \) satisfies (2.4), the moment-generating function has an analytical solution (Dufresne 2001)

\[
\eta(z, t; v) = \left( \frac{e^{\alpha(T-t)/2}}{\beta} \right)^{2\alpha/\sigma^2} \exp\left( \frac{v Z_t}{\gamma} \frac{2 \sinh(y(T-t)/2)}{\beta} \right),
\]

where

\[
\beta = \cosh\left( \frac{\gamma(T-t)}{2} \right) + \alpha \sinh\left( \frac{\gamma(T-t)}{2} \right) \quad \text{and} \quad \gamma = \sqrt{\alpha^2 - 2v\sigma^2}.
\]

Substituting \( v = \sum_j p_j (\phi_lj(u) - 1) \) yields the analytical expression for \( \phi(u, t) \):

\[
\phi(u, t) = \left( \frac{e^{\alpha(T-t)/2}}{\beta} \right)^{2\alpha/\sigma^2} \exp\left( \frac{2Z_t}{\gamma\beta} \sum_j p_j (\phi_lj(u) - 1) \sinh(y(T-t)/2) \right).
\]

Appendix A online gives the analytical expressions for the moments of the loss distribution.

### 2.3 Comparison to intensity models

Intensity models often use option pricing theory to model the default of a bond (Duffie 2005). In particular, the probability of survival in \([0, t]\) is modeled as follows:

\[
1 - \bar{p}_t = \mathbb{E}\left[ \exp\left( -\int_0^t \lambda_s \, ds \right) \right].
\]

The expectation is under the risk-neutral measure and \( \lambda \) satisfies the following SDE:

\[
d\lambda_t = \alpha(b - \lambda_t) \, dt + \sigma \sqrt{\lambda} \, d\bar{W}_t.
\]

This is a similar process to \( Z_t \) in (2.4). Recall that the probability of a default not occurring in the mixture model is, by the Poisson approximation, \( \mathbb{E}[e^{-p_j Y_t}] \), where

\[
p_j Y_t = \int_0^t p_j Z_s \, ds.
\]

Letting \( \lambda_s = p_j Z_s \), it is clear that the two models coincide on the probability of a default not occurring. Thus, the model proposed in this paper can be considered an extension of the intensity model in much the same way that Vasicek extended Merton’s model to compute the distribution of the entire portfolio.
3 LIQUIDITY RISK

One of the major contributions of this paper is the introduction of semi-endogenous liquidity risk to credit portfolio modeling. When credit losses start increasing, debt holders may become concerned and start pulling funding. To meet these funding requirements, additional assets must be liquidated, likely at fire-sale prices. While the probability of a liquidity crisis is nonzero at any level of credit loss, the probability of such a crisis increases as credit losses mount. To retain an analytic expression for the characteristic function, the frequency of a liquidity crisis is modeled as a linear function of credit losses, and the losses due to the liquidity crisis are assumed to be deterministic. In addition, the liquidity crisis can only occur at $T$, when the credit losses in $[0, T]$ have been made public. Let $C(t, T)$ be a random variable on $\mathbb{R}^+$ that represents the potential US dollar loss to the portfolio caused by fire-selling assets within the portfolio. The distribution of $C(t, T)$ is

$$C(t, T) = \begin{cases} 
\lambda(t, T) & \text{with probability } q(X_T - X_t) + qX_t = qX_T, \\
0 & \text{with probability } 1 - q(X_T - X_t) - qX_t = 1 - qX_T.
\end{cases} \quad (3.1)$$

Here $\lambda$ is the US dollar loss in the event of a liquidity crisis and is a deterministic function of $t$ and $T$, while $q$ is a nonnegative constant that scales the US dollar loss $X_T$ so that the probability remains feasible. For large loan portfolios, $q$ should be rather small. Letting $t = 0$ for simplicity, the total characteristic function is

$$\mathbb{E}[e^{ui(X_T + C_T)}] = \mathbb{E}[\mathbb{E}[e^{uiX_T} e^{uiC_T} \mid \mathcal{F}_T]]$$

$$\approx \mathbb{E}[\mathbb{E}[e^{uiX} \exp(qX(e^{ui\lambda} - 1)) \mid \mathcal{F}_T]]$$

$$= \mathbb{E}[\mathbb{E}[\exp((ui + q(e^{ui\lambda} - 1))X) \mid \mathcal{F}_T]]$$

$$\approx \phi(u - iq(e^{ui\lambda} - 1), 0).$$

For $\lambda = 0$ or $q = 0$, the characteristic function reduces to (2.8); as it should. Figure 1 on the facing page shows the distributions of credit portfolios with various parameters.

4 NUMERICAL INVERSION

4.1 Algorithms for inversion

With an analytic expression available for the characteristic function of $X_T$, it is possible to invert the function to recover the loss distribution. Letting $f(x)$ be the density of $X_T$, by the Fourier inversion theorem (Hewitt and Stromberg 1965),

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \phi(u, t) \, du. \quad (4.1)$$
To solve this integral numerically, $u$ and $x$ must be discretized. The loss distribution has support on $[0, x_{\text{max}}]$, where $x_{\text{max}}$ is the sum of the total possible exposure on each asset. $x_{\text{max}}$ is an extreme upper bound on the possible losses in the portfolio, and it is often possible to truncate this range to provide superior accuracy for the same number of discrete intervals. To obtain the numeric solution at each discrete point in $x$, (4.1) must be numerically integrated for each discrete $x$. A naive implementation of this numeric integration has complexity $O(mh)$, where $m$ is the number of discrete steps in the $u$ domain and $h$ is the number of discrete steps in the $x$ domain. Using the fast Fourier transform (FFT) reduces this complexity to $O(m \log_2(m))$ (Fang and Oosterlee 2008). However, the FFT requires the same number of discrete steps in both the $x$ and $u$ domains. Fang and Oosterlee (2008) use a cosine series expansion (COS) to invert the characteristic function instead of the FFT. This expansion separates the discretization of $u$ and $x$ and provides exponential convergence for suitable functions. For credit portfolio losses, the computationally difficult part is discretization of $u$: for each discrete $u$ the entire characteristic function must be recomputed. In a portfolio of $n$ loans, the computation time for each discrete $u$ is $O(n)$. Hence, the COS algorithm is perfectly suited for inverting characteristic functions, since a fine mesh in $x$ can be achieved while still using relatively few calls to the characteristic function. The pseudocode for the COS algorithm is presented in Appendix B online.
TABLE 1  Convergence (in log relative error) for $h = 1024, n = 10000, \alpha = 0.3, t = 1, Z_0 = 1.1, p = 0.03, n = 2^b, \sigma = 0.5.$

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT $M_1$</td>
<td>0.77</td>
<td>1.83</td>
<td>0.25</td>
<td>-2.33</td>
<td>-9.84</td>
<td>-16.84</td>
<td>-26.90</td>
<td>-33.77</td>
</tr>
<tr>
<td>COS $M_1$</td>
<td>-3.73</td>
<td>-2.32</td>
<td>-7.93</td>
<td>-8.25</td>
<td>-11.07</td>
<td>-19.02</td>
<td>-28.14</td>
<td>-33.08</td>
</tr>
<tr>
<td>FFT $M_2$</td>
<td>4.76</td>
<td>4.48</td>
<td>2.38</td>
<td>0.77</td>
<td>-7.34</td>
<td>-13.49</td>
<td>-23.46</td>
<td>-30.42</td>
</tr>
<tr>
<td>COS $M_2$</td>
<td>0.33</td>
<td>-2.01</td>
<td>-2.64</td>
<td>-5.47</td>
<td>-10.49</td>
<td>-18.43</td>
<td>-27.95</td>
<td>-29.89</td>
</tr>
</tbody>
</table>

TABLE 2  Convergence (in log relative error) for $h = 1024, n = 10000, \alpha = 0.3, t = 1, Z_0 = 1.1, p = 0.03, n = 2^b, \sigma = 1.$

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT $M_1$</td>
<td>0.11</td>
<td>1.49</td>
<td>-2.86</td>
<td>-1.59</td>
<td>-5.17</td>
<td>-6.61</td>
<td>-9.23</td>
<td>-8.37</td>
</tr>
<tr>
<td>FFT $M_2$</td>
<td>4.289</td>
<td>4.29</td>
<td>1.73</td>
<td>1.39</td>
<td>-1.61</td>
<td>-3.70</td>
<td>-6.18</td>
<td>-5.33</td>
</tr>
<tr>
<td>COS $M_2$</td>
<td>0.11</td>
<td>-2.94</td>
<td>-3.50</td>
<td>-7.17</td>
<td>-10.42</td>
<td>-13.57</td>
<td>-15.61</td>
<td>-15.73</td>
</tr>
</tbody>
</table>

Let $m_f$ be the number of discrete steps in the FFT algorithm, let $m_c$ be the number of steps in $u$ for the COS algorithm and let $h$ be the number of steps in $x$ for the COS algorithm. The FFT algorithm for inverting the characteristic function has complexity $O(m_f(n + \log_2(m_f)))$, while the COS algorithm has complexity $O(m_c(n + h))$. As $n$ grows, the COS algorithm converges to complexity $O(n m_c)$, while the FFT algorithm, for fixed $m_f$, has complexity $O(n m_f)$. Thus, the two algorithms have similar complexity for $m_f = m_c$. The similarity is confirmed in numerical tests in Table 5 on the facing page and Table 6 on page 40.

4.2 Algorithm results

The moments of the distribution are known analytically and the first two are given in appendix A online. Using these moments is it possible to test the accuracy of each algorithm. For the first and second moments, the expectation is numerically integrated over the approximate density from each of the two algorithms for a variety of parameters. Letting $M_i$ represent the $i$th moment of the distribution, the relative error $|M_i - \hat{M}_i|/M_i$ is then computed and compared. The results are displayed in Table 1, Table 2, Table 3 on the facing page and Table 4 on the facing page. Overall, the COS algorithm performs better than the FFT algorithm. The COS algorithm
TABLE 3  Convergence (in log relative error) for $h = 1024$, $n = 10000$, $\alpha = 0.3$, $t = 1$, $Z_0 = 1.1$, $p = 0.0005$, $n = 2^b$, $\sigma = 0.5$. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT $M_1$</td>
<td>0.62</td>
<td>2.01</td>
<td>-0.68</td>
<td>-2.71</td>
<td>-2.17</td>
<td>-2.26</td>
<td>-1.81</td>
<td>-1.93</td>
</tr>
<tr>
<td>COS $M_1$</td>
<td>-1.93</td>
<td>-2.95</td>
<td>-3.95</td>
<td>-7.88</td>
<td>-8.34</td>
<td>-8.38</td>
<td>-9.24</td>
<td>-16.51</td>
</tr>
<tr>
<td>FFT $M_2$</td>
<td>4.13</td>
<td>5.20</td>
<td>1.70</td>
<td>0.45</td>
<td>1.08</td>
<td>0.14</td>
<td>1.10</td>
<td>0.92</td>
</tr>
<tr>
<td>COS $M_2$</td>
<td>0.86</td>
<td>-2.23</td>
<td>-2.85</td>
<td>-6.43</td>
<td>-11.21</td>
<td>-12.93</td>
<td>-11.59</td>
<td>-12.41</td>
</tr>
</tbody>
</table>

TABLE 4  Convergence (in log relative error) for $h = 1024$, $n = 10000$, $\alpha = 0.3$, $t = 1$, $Z_0 = 1.1$, $p = 0.0005$, $n = 2^b$, $\sigma = 1$. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT $M_1$</td>
<td>0.87</td>
<td>1.57</td>
<td>0.49</td>
<td>-1.30</td>
<td>-0.61</td>
<td>-0.82</td>
<td>-0.60</td>
<td>-0.66</td>
</tr>
<tr>
<td>COS $M_1$</td>
<td>-1.71</td>
<td>-4.28</td>
<td>-4.67</td>
<td>-8.19</td>
<td>-6.89</td>
<td>-7.12</td>
<td>-8.01</td>
<td>-14.87</td>
</tr>
<tr>
<td>FFT $M_2$</td>
<td>3.57</td>
<td>4.75</td>
<td>3.54</td>
<td>1.93</td>
<td>2.42</td>
<td>1.75</td>
<td>2.17</td>
<td>2.06</td>
</tr>
<tr>
<td>COS $M_2$</td>
<td>0.69</td>
<td>-2.11</td>
<td>-3.88</td>
<td>-6.51</td>
<td>-10.40</td>
<td>-10.29</td>
<td>-10.55</td>
<td>-11.32</td>
</tr>
</tbody>
</table>

TABLE 5  Time (in seconds) for $n = 10000$, $h = 1024$, $\alpha = 0.3$, $\sigma = 1$, $t = 1$, $Z_0 = 1$ using one core of an Intel Core i5-2520M 2.5 GHz. 

<table>
<thead>
<tr>
<th>$m$</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT</td>
<td>0.031</td>
<td>0.032</td>
<td>0.047</td>
<td>0.124</td>
<td>0.234</td>
<td>0.468</td>
<td>0.968</td>
<td></td>
</tr>
<tr>
<td>COS</td>
<td>0.031</td>
<td>0.016</td>
<td>0.062</td>
<td>0.062</td>
<td>0.125</td>
<td>0.250</td>
<td>0.499</td>
<td>1.029</td>
</tr>
</tbody>
</table>

retains acceptable accuracy over a range of parameter values, while the FFT algorithm becomes unstable. This is especially evident when the volatility of $Z_t$ is large or when the $p_j$ are small, as shown in Tables 2–4. As most loan portfolios tend to be composed of relatively safe assets with a small probability of default, the accuracy of the COS algorithm in such a scenario is vital.

The algorithm shows remarkable speed even on minimal hardware and with large portfolios. Table 5 shows that on a single core of an Intel Core i5 computer the algorithm can accurately compute the entire distribution of 10 000 assets in under one-tenth of a second. Even with $n = 10 000 000$ the density can be computed within
TABLE 6 Time (in seconds) for $n = 10,000,000$, $h = 1024$, $\alpha = 0.3$, $\sigma = 1$, $t = 1$, $Z_0 = 1$ using one core of an Intel Core i5-2520M 2.5 GHz.

<table>
<thead>
<tr>
<th>$m$</th>
<th>8</th>
<th>16</th>
<th>32</th>
<th>64</th>
<th>128</th>
<th>256</th>
<th>512</th>
<th>1024</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFT</td>
<td>6.88</td>
<td>14.40</td>
<td>28.85</td>
<td>58.09</td>
<td>125.81</td>
<td>270.15</td>
<td>568.96</td>
<td>1088.43</td>
</tr>
</tbody>
</table>

30 seconds, demonstrating the efficiency of this method for the computation of even very large portfolios (see Table 6).

5 CONCLUSION

This paper contributes to the credit modeling literature for mixture models by leveraging an efficient algorithm for computing the density function of the loss distribution and extending the model in two key areas: constructing the systemic variable from a continuous-time process and introducing semi-endogenous liquidity risk. The model remains tractable yet with sufficient flexibility to provide at least a first approximation to the actual portfolio loss distribution. Having the ability to quickly compute a loss distribution opens up a broad range of possibilities for financial institutions. The speed of the algorithm makes computing, pricing and allocating risk a relatively simple matter. A loan committee can efficiently price individual credits by computing incremental or marginal risk. Banks can assess the risk contribution of arbitrary sections of the loan portfolio. As a simple example, the bank could assess an individual loan officer’s portfolio to determine bonuses or uncover underwriting issues. Boards can stress the loan portfolio by shocking the mixture processes and seeing the results in real time.

While this paper generalizes and extends the CreditRisk$^+$ model, there are a number of avenues for further work. Since affine processes are computationally convenient even in higher dimensions, a multidimensional version of the model should be entirely feasible. Each dimension could represent a concentration within the portfolio: for instance, loans within a region experiencing slower economic growth are more risky for the same probability of default and exposure as loans in an economically healthy region. In order to incorporate the multidimensional process, the characteristic function would have to be modified so that $v = p_j \sum_k w_k Y_{t,k}$, where $\sum_k w_k = 1$. Each $w$ would represent the weight that each concentration has on $X^j$. For example, if every asset in the portfolio has systemic risk with the overall state of the economy, then every asset will have a positive weight for that systemic sector.
Multidimensional models add some numerical complexity. The additional computation is due to the more complicated characteristic function. For example, independent processes linearly increase the number of calls to the one-dimensional characteristic function, since each process has the same characteristic function (though with different parameters). Exploring the trade-off between flexibility and computational time is another area of future research.

An additional path for future research is to explore other affine processes. In particular, Gaussian processes are more analytically tractable and retain an affine structure in multidimensional models with instantaneous correlation. However, Gaussian processes can become negative. The trade-off between the instantaneous correlation available to Gaussian processes and the potential for negative values should be explored to determine which affine processes are the most flexible for credit modeling purposes.

DECLARATION OF INTEREST

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

REFERENCES


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

Default risk of money-market fund portfolios

Matulya Bansal

343 Gold Street, Apartment 2114, Brooklyn, NY 11201, USA; email: mbansal11@gsb.columbia.edu

(Received December 9, 2014; revised April 13, 2015; accepted June 25, 2015)

ABSTRACT

We address the problem of quantifying the risk associated with money-market fund (MMF) portfolios. As MMFs are intended to be perpetual investment vehicles but hold short-maturity instruments, this problem presents some unique modeling challenges. By focusing on default risk, the most material driver of portfolio losses, and by using a constant level of risk assumption, we show how this problem is similar to that of pricing a collateralized debt obligation. Drawing upon this insight, we present a novel semianalytic approach to compute the default risk of MMF portfolios. Upon using this model to evaluate the portfolios of three of the largest prime MMFs, we find that they vary considerably in their default risk, which in our stylized model increases with the risk horizon. This suggests that to be effective and yet not punitive, any regulatory proposal (eg, establishing capital buffers, imposing liquidity fees, etc) to address the systemic risk posed by MMFs should be based on the riskiness of the MMF portfolios as well as their liquidation strategy.

Keywords: money-market fund (MMF); financial regulation; default risk; portfolio replenishment; collateralized debt obligation (CDO) pricing.

1 INTRODUCTION

In the aftermath of the 2008 credit crisis, money-market funds (MMFs) continue to be a significant part of the global fixed-income markets. Even with yields near zero,
Moody’s (2013) reports that, as of August 2013, investors held more than US$2.5 trillion and US$1 trillion in MMFs in the United States and Europe, respectively: a testament to their enduring appeal.

The reason for this continued popularity of MMFs seems to arise from the perceived simplicity and attractiveness of the product: the ability to invest and redeem at par (US$1.00) and yet earn in excess of short-dated government debt (eg, T-bills) and bank deposits. However, as evidenced during the crisis (in particular, from the failure of the Reserve Primary Fund), the redemption at par is far from guaranteed. As MMF portfolios can be concentrated, a single default can cause the MMF net asset value (NAV) to fall below par and even to break the buck (ie, cause the MMF NAV to fall below US$0.995), and thereby trigger a run on the fund. Given the interconnectedness between funds (see, for example, Office of Financial Research 2013, Figure 6), such a run can quickly cascade into a systemic stress. Given their large size, addressing the systemic risk posed by MMFs continues to be an active area of interest. Recent proposals toward this (see, for example, European Commission 2013; European Systemic Risk Board 2012; Financial Stability Oversight Council 2013; Securities and Exchange Commission 2014; Squam Lake 2011) can be classified into four categories: floating the NAV, enforcing liquidity fees, enabling redemption gates and establishing capital buffers.

The proposal to float the NAV of MMFs seeks to directly address the perception of stability associated with MMFs. Under the current regulations, MMFs are not required to report the mark-to-market (MTM, also referred to as the shadow NAV) of their portfolio. Rather, MMFs are allowed to follow amortized cost accounting, under which, unless the shadow NAV of their portfolio falls below US$0.995, the fund NAV can be reported to be par (hence the name stable NAV MMFs). Under this proposal, the shadow NAV of each MMF would float, and thus any concerns about the credit or liquidity of the fund would be promptly reflected in its pricing (hence the name floating NAV MMFs). Hence, in the event of an obligor default, there would not be a need to defend the NAV of the fund from breaking the buck. Despite making the ownership of risk by the investors more transparent, it is not evident how effective floating NAV would be in preventing redemptions as long as the future losses on the fund are perceived to exceed its floating NAV. Moreover, as it introduces the possibility of incurring a loss on an MMF investment even under normal market environments, this proposal is strongly opposed by some in the industry (see, for example, Investment Company Institute 2013). Nevertheless, this proposal became a rule in the United States (to be implemented in October 2016) when the US Securities and Exchange Commission (SEC) adopted it for institutional prime and tax-exempt MMFs, while allowing treasury, government and retail prime and tax-exempt MMFs to continue to report the fund NAV at par (see Securities and Exchange Commission 2014).
The proposals on establishing liquidity fees and redemption gates seek to prevent a run on any given MMF (and hence the industry) by stemming the outflow of funds in the event of a stress. Under the proposals, this would be achieved by enabling the MMF under stress to impose liquidity fees on withdrawals and to suspend redemptions. While clearly disincentivizing withdrawals, it is unclear how effective the liquidity fees would be if the potential loss is perceived to be greater than the fee imposed. Similarly, it is unclear how effective halting redemptions temporarily would be if the stress event lasts longer. Further, it is likely that the market would try to front-run the enforcement of such fees and gates by accelerating redemption requests in anticipation of such a move, thereby at least partly thwarting their purpose (see, for example, International Cash Distributors 2012b). In any event, these proposals became a rule in the United States (to be implemented in October 2016) when the SEC adopted it for both retail as well as institutional, prime and tax-exempt MMFs, wherein under periods of illiquidity, the fund managers for these MMFs can choose to impose liquidity fees of up to 2% as well as suspend redemption requests for up to ten days in a ninety-day period (see Securities and Exchange Commission 2014).

The final proposal seeks to address the systemic risk posed by MMFs by requiring fund sponsors to establish capital buffers, much akin to capital held by the banks for their loan portfolios. Such buffers, similar to an equity tranche, are intended to bear the first loss on the fund, assure the investors of the safety of their investment and prevent a run on the fund (and thus the broader MMF industry). As with liquidity fees, for a given MMF the effectiveness of the capital buffer would depend on the buffer size relative to the perceived loss. While a large buffer would reduce the probability of a run, the associated costs may render the MMF product unattractive. At present, a 3% buffer has been proposed in European Systemic Risk Board (2012) and European Commission (2013), and this has led to questions on the viability of MMFs (see, for example, International Cash Distributors 2012a).

These different proposals to address the systemic risk of MMFs have naturally led to much discussion on the pros and cons of each approach, as well as that of one approach over the others. However, there has been little work on measuring the risk of the MMF portfolios themselves. This is important, as discussed above, to calibrate the size of any capital buffer or liquidity fees, as well as to enable the fund investors to assess the magnitude of the risk they undertake as a part of their MMF investments.

The aim of this paper is to propose an approach to quantifying the risk associated with an MMF portfolio, thereby allowing the fund investors, the sponsor and the regulators to assess its risk profile and, as a result, institutionalize as well as calibrate the mechanisms for handling any potential losses. Toward this, we develop a novel semianalytic approach that accounts for the short-dated nature of MMF holdings, the fact that the risk horizon can be larger than the maturity of the fund holdings and the limited granularity and heterogeneity of MMF portfolios. Using publicly
available data, we use the proposed approach to estimate the portfolio default risk and probability of breaking the buck for three of the largest prime MMFs. We hope that our results will serve as a basis for an informed discussion regarding the risk associated with MMF portfolios.

The remainder of this paper is organized as follows. This section concludes with a brief literature review. In Section 2, we present the problem of computing the tail risk of an MMF portfolio. This includes a discussion of the relevant risk factors and the replenishment assumptions, as well as a mathematical formulation of the problem. In Section 3, we present the modeling assumptions and computational techniques employed to solve this problem, and discuss some variations. In Section 4, we present the results of applying our approach to three prime MMF portfolios. We conclude in Section 5.

**Literature review**

While we are not aware of any study that attempts to quantify the default risk associated with MMF portfolios, our work builds upon previous research in several areas. Our paper builds upon the collateralized debt obligation (CDO) pricing literature that focuses on using semianalytic approaches to computing the portfolio loss distribution of obligors underlying a tranche. Two of the earliest and most significant papers in this area are Hull and White (2004) and Andersen et al (2003), who model default codependence via a copula and construct the conditional portfolio loss distribution in a bottom-up fashion by using the conditional loss distribution for each of the reference obligors. A survey of such factor-based CDO pricing approaches can be found in Andersen and Sidenius (2005). The approach proposed in this paper shares the factor-based, bottom-up, conditional portfolio loss distribution construction approaches discussed in these papers. However, unlike a CDO that references a basket of unique obligors and has a defined contractual maturity, our approach seeks to build this loss distribution for MMF portfolios, which are intended to be perpetual investment vehicles, and which typically hold multiple securities referencing the same obligor but having different maturities.

Our work is also related to the literature on portfolio risk modeling for traditional credit portfolios (for example, Basel Committee on Banking Supervision 2005; Gordy and Lutkebohmert 2013), where analytic approaches are proposed to estimate tail risk. The former employs the large homogeneous pool assumption wherein it assumes that any idiosyncratic risk is diversified away and only the systemic risk needs to be accounted for. The latter seeks to improve upon this assumption by accounting for the marginal impact of granularity due to each obligor. The main attraction of these approaches is that they are analytic and, hence, easy to implement and quick to run. Their main drawback is that portfolio loss aggregation is based on approximations,
so it may underestimate or overestimate the portfolio risk. Our approach employs a bottom-up approach to build the loss distribution of the portfolio, thereby ensuring that the portfolio loss aggregation is exact. This, however, also makes our approach semianalytic, wherein numerical integration is required. The other aspect differentiating the problem we address from both these problems is the need to model portfolio replenishment.

Our paper contributes to the growing literature on the risks of MMFs. Several researchers have analyzed the risks in MMFs from different perspectives. For example, Baba et al (2009) provides a fascinating narrative on the role of MMFs in the credit crisis of 2008. Similarly, Kacperczyk and Schnabl (2013) examine the behavior of MMFs during the credit crisis of 2008 with an emphasis on understanding whether MMFs loaded on to risk during the crisis. Proposals from Squam Lake (2011), the Financial Stability Oversight Council (2013), the European Commission (2013), the European Systemic Risk Board (2012) and the Securities and Exchange Commission (2014) outline approaches to address the systemic risk posed by MMFs. Our work is complementary in that instead of providing a historical perspective, analyzing MMF behavior or proposing a mechanism for reducing the systemic risk posed by MMFs, we propose an approach to quantifying the default risk associated with a given MMF portfolio. To our knowledge, our work is the first such attempt.

Our work is also related to the evolving literature on measuring portfolio risk in the presence of portfolio replenishment strategies. A prominent example is Glasserman (2012), which studies the impact of rebalancing frequency on the portfolio profit and loss distribution under the assumption that rebalancing targets a constant volatility. As our work focuses on the default risk associated with MMF portfolios, the rebalancing approach proposed in this paper instead assumes that the default exposure to obligors remains constant. This rebalancing approach, while taking into account the ongoing nature of MMFs and the short maturities of the constituent securities, allows us to use a semianalytic calculation to quantify the default risk of MMF portfolios.

2 PROBLEM CONTEXT AND FORMULATION

2.1 MMF portfolios

MMFs are a category of open-ended mutual funds that invest in short-dated fixed-income securities. Unlike most other mutual funds, however, under the current rules MMFs are allowed to use amortized cost accounting to compute their NAV. Unless the difference becomes significant, they do not need to disclose the MTM of their portfolio. For this reason, their portfolio holdings are subject to additional regulations. For example, in the United States, MMFs are required to comply with Rule 2a7 of the SEC’s Investment Act of 1940, which restricts the credit quality, maturity and liquidity
of their holdings. As a result, MMFs can only invest in first- or second-tier issuers, with exposure to any single issuer not exceeding 5% and 3%, respectively (some securities, eg, government debt, are exempt). The maximum maturity of any holding is limited to 397 days, and the portfolio weighted average maturity is limited to sixty days. MMFs are also required to maintain a daily and weekly portfolio liquidity of 10% and 30%, respectively. (See Securities and Exchange Commission (2010) for details. Note that, while the new rule (Securities and Exchange Commission 2014) clarifies the rules governing MMF portfolio holdings, it does not stipulate any changes to them.)

While MMFs are a popular product in Europe as well as in the United States, in the remainder of this paper we will use the case of MMFs within the latter (the portfolio holdings of which are regulated under the SEC Rule 2a7) for discussion.

MMFs can be categorized by the types of securities they hold, eg, into government, agency, municipal and prime MMFs. In this paper, we focus on prime MMFs, which constitute some of the biggest MMFs and are characterized by high risk and high yields within the MMF world. Unlike government, agency and municipal MMFs, each of which target a subset of the securities issued or backed by the federal, state and local governments, and government-backed agencies, the portfolios of prime MMFs may comprise government, agency and municipal securities, and repurchase agreements (repos) as well as certificates of deposit (CDs), commercial paper (CP) and asset-backed CP (ABCP) issued by financial institutions.

### 2.2 Risks in MMF portfolios

While Rule 2a7 limits the risks in MMF portfolios, being fixed-income portfolios, they are exposed to interest rate risk, credit spread risk, liquidity risk and default risk. Interest rate risk and credit spread risk for an MMF arise from the possibility that interest rates and credit spreads rise, respectively, thereby leading to an MTM loss for the fund. Given the short duration of MMF portfolio holdings, an increase in interest rates or credit spreads is unlikely to lead to material losses for an MMF. In particular, Securities and Exchange Commission (2010) notes that an MMF with a weighted average maturity of sixty days would be able to withstand a jump in yield of 300 basis points (bps) without causing the buck to break. Further, as stable NAV MMFs do not have to follow MTM accounting, such MMFs can hold securities that have incurred an MTM loss to maturity; in this case, unless there are defaults, no realized loss results. Even for floating NAV MMFs, minus excessive redemptions and in the absence of defaults, the NAV would return to par as the affected holdings (typically short-dated due to Rule 2a7) mature. Hence, interest rate risk or credit spread risk by themselves are unlikely to be material drivers of portfolio loss for MMFs.
Liquidity risk for an MMF is the risk that one or more of the portfolio holdings become illiquid or incur an MTM loss, so that selling them would result in a realized loss, thereby restricting the MMF’s ability to service redemptions by selling such securities. Given the liquid investments that MMFs hold to comply with Rule 2a7, this is usually not an issue, as any redemption requests that arise as a part of business-as-usual can be serviced using the liquid holdings. The illiquid holdings (usually having a short maturity, given Rule 2a7 requirements) can then, upon maturing, be used for servicing any redemptions or invested in liquid securities, thereby addressing this temporary illiquidity. Hence, under normal circumstances, liquidity risk by itself is also unlikely to cause an MMF to break the buck or to lead to material losses.

Default risk is the risk that one or more obligors in the MMF portfolio default, thereby causing losses in the MMF portfolio. While Rule 2a7 limits MMF exposure in any individual first-tier and second-tier issuer to be 5% and 3%, respectively, if the recovery is low, a default of an obligor with such exposure can cause the stable NAV MMFs to break the buck and lead to material losses for floating NAV MMFs. The now classic example is the Reserve Primary Fund, which broke the buck and led to an industry-wide crisis as a result of the Lehman default in September 2008 (see Baba et al. (2009) for details). Accounting for default risk is therefore a key aspect of MMF portfolio risk modeling.

In the discussion so far, while we considered interest rate risk, credit spread risk, liquidity risk and default risk individually, we did not consider the possibility that a combination of these risks surface together. In particular, if the shadow NAV of a stable NAV MMF falls below par (e.g., due to credit spread widening), redemption requests can result in a run on the MMF. In this case, as worried investors seek to redeem their investments at par, the MMF is forced to sell high-quality liquid securities and is left with less liquid securities that have incurred an MTM loss. This, in turn, causes a further reduction in the shadow NAV of the fund, causing a further increase in investor redemption requests. Unless the fund sponsor steps in, redemptions are halted or some sort of backstop is provided, such a scenario can quickly morph into a run on the fund. Similarly, for a floating NAV MMF, a run can result if the future losses are perceived to be higher than those reflected in the current NAV. Finally, such a run can occur even in the presence of a capital buffer or liquidity fees, if future losses are perceived to be larger than the buffer size and fees, respectively.

In such a situation, given the fall in market price of MMF holdings that the affected MMF needed to liquidate in order to satisfy redemption requests, the similarity of holdings across MMFs and a general increase in risk premiums in the market, the shadow NAVs of the other stable MMFs as well as the NAVs of the other floating MMFs are also very likely to fall. If the investors also perceive these other MMFs as being risky, this may in turn trigger a run on the broader MMF industry. Given the size and importance of the MMF industry, a run on the industry has the potential to
threaten the functioning and stability of the broader financial market (the collapse of
the Reserve Primary Fund is illustrative, see Baba et al (2009)).

This paper does not directly deal with the run risk or systemic risks posed by MMFs. Instead, the scope is limited to modeling and quantifying the risk of a given MMF portfolio. More specifically, given that in the absence of a systemic stress or a run, interest rate risk, credit spread risk and liquidity risk are unlikely to be significant drivers of MMF portfolio loss, we focus on quantifying the MMF portfolio risk due to defaults. Despite the limitation of not modeling MMF runs, our analysis helps us to better understand the risk profile of MMF portfolios. This knowledge in itself can help reduce the possibility of a run. For example, if, following an understanding of MMF portfolio risks, a capital buffer is established or liquidity fees are imposed that exceed the size of the losses perceived during a stress, a run on the fund and thus the broader MMF industry can be averted.

2.3 Risk horizon

To estimate the risk of a portfolio, one needs to choose a risk horizon. For MMFs, the choice of risk horizon is far from obvious. On the one hand, MMFs are intended to be perpetual vehicles, much like the loan or trading businesses of banks. So it is reasonable to assume that, despite occasional setbacks, and unless something fundamental changes and offering the product becomes uneconomical, the fund sponsor would seek to continue to administer this product. This reasoning is supported by the several capital injections or support measures that various sponsors have historically provided to their sponsored MMFs (see, for example, Moody’s 2010). Hence, it is sensible to choose a relatively long risk horizon, eg, one year (which is comparable to the maximum maturity of any MMF holding in the United States and is commonly used to estimate the capital requirements of bank loan portfolios).

On the other hand, as the maturities of MMF holdings are quite short, if the MMF sponsor desires it, even without selling the portfolio holdings, the portfolio can be unwound in a few months by not reinvesting and allowing the portfolio holdings to mature. For this reason, a shorter risk horizon, eg, three months (which is comparable to the limit on the portfolio weighted average maturity of any MMF in the United States), is also reasonable, as in an adverse scenario, minus a run, the fund can be unwound using this approach. This, however, assumes that in the event of a stress, the fund sponsor would choose to unwind the fund over supporting it as an ongoing investment vehicle.

As it depends on the individual MMF’s strategy to support (or not support) the fund in the event of a loss, in this paper, we do not have an outright view on the right risk horizon to use for estimating the risk of MMF portfolios. Rather, driven by the above considerations, we use two different risk horizons: three months and one year.
2.4 Replenishment strategy

Given a risk horizon, one needs to consider the need to model portfolio replenishment. For MMFs, where the risk horizon can be significantly larger than the maturity of any of the portfolio holdings, not modeling portfolio replenishment would lead to an underestimation of risk as any maturing positions would convert to cash (which we assume to be risk free). Moreover, it would not be reflective of the operation of the fund in practice, wherein minus any redemption requests and inflows, and ex-yield distribution and fees, cash obtained from maturing positions is typically invested into new positions. Hence, it is imperative that portfolio replenishment be modeled for estimating the risk of MMF portfolios.

Toward modeling portfolio replenishment, one needs to understand the nature of MMF investment strategies. Portfolio investment strategy, including both asset allocation and reinvestment strategy, is the primary dimension that differentiates one prime MMF from another. While the initial asset allocation of an MMF can be captured using the publicly disclosed MMF holdings, modeling its reinvestment strategy, which depends on factors such as the portfolio managers’ desire to enhance yield, their tolerance of portfolio risk and their reading of the market conditions and investor sentiment, is difficult, and beyond the scope of this paper. Instead, we restrict ourselves to the following simple reinvestment strategy: we assume that the portfolio manager seeks to maintain the same unconditional default risk profile, i.e., a notional, unconditional probability of default (PD) and loss given default (LGD), over the risk horizon. In more detail, this corresponds to the following course of action on behalf of the portfolio manager.

Upon the maturity of an instrument, and given that the obligor it references has not defaulted or been downgraded, we assume that the portfolio manager reinvests the amount received (minus any yield and fees) into another security issued by the same obligor and of the same type, so that the notional, PD and LGD for this obligor remains unchanged.

In the event of an obligor default, we assume the portfolio manager makes up for any losses and invests an amount equal to the original notional invested in the defaulted obligor into a new obligor (one that does not exist in the portfolio). This new obligor has the same rating as the initial rating of the defaulted obligor it is replacing, and in the corresponding instruments, so that the notional, PD and LGD associated with the new obligor is the same as that initially for the defaulted obligor. Note that this implicitly assumes that obligors with the same rating have the same PD, and that LGD for any given security type does not vary across obligors.

In the event that an obligor is downgraded, we assume the portfolio manager chooses to hold the corresponding positions to maturity or default, whichever occurs
first. In case the positions corresponding to the downgraded obligor mature before defaulting, we assume that the proceeds are reinvested into a new obligor with the same rating as the initial rating of the downgraded obligor, and into the same security types. In case the downgraded obligor defaults first, we assume that the portfolio manager makes up for any losses and invests an amount equal to the original investment in this defaulted obligor into a new obligor that has the same rating as its initial rating, as well as into the corresponding security types.

In the case where all securities referencing the downgraded obligor have the same maturity, this replenishment strategy ensures that, at any given time, either the downgraded or the new obligor (but not both) exists in the portfolio. In case the securities referencing the downgraded obligor have different maturities, we assume that the portfolio manager adopts the following rollover strategy. Each maturing security referencing the downgraded obligor, except the one with the longest maturity at the time of the downgrade, is rolled over so that it continues to reference the downgraded obligor. Upon the maturity of the security that had the longest maturity at the time of the downgrade, all securities corresponding to this downgraded obligor are sold and reinvested into corresponding securities referencing the new obligor. We assume that there is a one-to-one correspondence between idiosyncratic credit spread change and rating change, and so, unless the downgraded obligor is downgraded again, in which case the choice of the security with the longest maturity at the time of the downgrade is again reset, the above does not result in any realized loss due to the downgrade. We also assume that any potential fall/rise in the MTM of securities referencing the downgraded obligor due to interest rate movements or nonidiosyncratic credit spread changes is offset by the corresponding fall/rise in the MTM of the securities referencing the new obligor. Overall, this implies that, unless the downgraded obligor defaults before it is replaced, the portfolio does not incur any realized loss.

We do not model any inflows or redemptions, and we assume that the portfolio yield income is positive and distributed between the investors and the fund sponsor (in the form of yield and fee, respectively). This implies that the portfolio notional also remains constant over the risk horizon.

The proposed replenishment strategy corresponds to a constant level of risk assumption (see Basel Committee on Banking Supervision 2009) in that it maintains the same level of default risk over the risk horizon. While simplistic, as noted in Basel Committee on Banking Supervision (2009), it is consistent with the Basel II framework, and the choice of a risk horizon together with a constant level of risk assumption reflects the view that an MMF is an ongoing investment vehicle that would continue to operate over the risk horizon, and whose default risk profile would not be dramatically altered in the event of any losses experienced over the risk horizon.
2.5 Problem formulation

Consider an MMF portfolio that has positions referencing \( n \) obligors, \( i = 1, 2, \ldots, n \). We refer to these obligors as obligor 1, obligor 2 and so on. Let us denote the risk horizon as \( T \), the PD associated with obligor \( i \) over \( T \) as \( \text{PD}_i \) and the default time associated with obligor \( i \) as \( t_i \). We assume that there exists a deterministic mapping from an obligor rating to its PD, so that any two obligors with the rating have the same PD. For each obligor \( i \), the portfolio investments may be in any of \( n_s \) different security types, eg, CP, CDs or repos. Let us denote the LGD associated with security type \( k \in \{1, 2, \ldots, n_s\} \) as \( \text{LGD}_k \). For each obligor \( i \) and security type \( k \), the investments may be spread across \( n_{i,k} \) different maturities, eg, CP referencing a given obligor but staggered across maturities. Let us denote the notional associated with obligor \( i \), security type \( k \) and maturity \( r \), \( r \in \{1, 2, \ldots, n_{i,k}\} \), as \( N_{i,k,r} \); the notional associated with obligor \( i \) and security type \( k \) as \( N_{i,k} \); the notional associated with obligor \( i \) as \( N_i \); and the total portfolio notional as \( N \). Then, by definition,

\[
N_{i,k} = \sum_{r=1}^{n_{i,k}} N_{i,k,r}, \quad N_i = \sum_{k=1}^{n_s} N_{i,k}, \quad N = \sum_{i=1}^{n} N_i.
\]

Before formulating the problem associated with the replenishment strategy of Section 2.4, let us consider the problems corresponding to two other strategies: the no-replenishment (or buy-and-hold) strategy and the replenish-unless-default strategy. Under the former, any downgraded positions are held to maturity, and any matured or defaulted positions are converted into cash. In this case, the portfolio loss \( L_{\text{no replenishment}}^{\text{pfl}} \) is given by

\[
L_{\text{no replenishment}}^{\text{pfl}} = \sum_{i=1}^{n} \sum_{k=1}^{n_s} \sum_{r=1}^{n_{i,k}} 1\{t_i < t_{i,k,r}\} \cap (t_i < T) \} N_{i,k,r} \cdot \text{LGD}_k. \quad (2.1)
\]

Next, consider the replenish-unless-default strategy, under which any downgraded positions are held to maturity, and any matured or defaulted positions are converted into cash. In this case, the portfolio loss \( L_{\text{replenish unless default}}^{\text{pfl}} \) can be written as

\[
L_{\text{replenish unless default}}^{\text{pfl}} = \sum_{i=1}^{n} 1\{t_i < T\} \sum_{k=1}^{n_s} N_{i,k} \cdot \text{LGD}_k. \quad (2.2)
\]

Finally, let us consider the replenishment strategy discussed in Section 2.4, wherein positions referencing defaulted obligors are also replenished. As with (2.2), upon the default of an obligor, the individual maturities of securities do not matter. In this
case, however, we need to consider the possibility that an obligor that replaces a defaulted obligor may itself default. Toward this, note that the replenishment strategy in Section 2.4 implies that the number of obligors in the MMF portfolio remains constant over the risk horizon. In case obligor \(i\) defaults (or is downgraded) and is replaced, we continue to label the new obligor as obligor \(i\). The replenishment strategy in Section 2.4 ensures that the notional, PD and LGD associated with this new obligor remain the same as initially associated with the obligor it replaced. This considerably simplifies the problem, as we do not need to track the obligors in the portfolio on an individual basis. Instead, it suffices to track the number of defaults associated with each of the obligor indexes. Denoting by \(N_i^T\) the number of defaults corresponding to obligor index \(i\) that occur over \(T\), the portfolio loss, \(L_{\text{full replenishment}}\), is given by

\[
L_{\text{full replenishment}} = \sum_{i=1}^{n} \sum_{w=1}^{N_i^T} \sum_{k=1}^{n_s} N_{i,k} \text{LGD}_k^w = \sum_{i=1}^{n} N_i^T \sum_{k=1}^{n_s} N_{i,k} \text{LGD}_k^w.
\]

(2.3)

Under the assumption that at most one default per obligor index occurs over the risk horizon, (2.3) reduces to (2.2).

### 2.6 Connection with CDO pricing

The problem of measuring the default risk of an MMF portfolio seems to be, on the surface, very different from that of pricing a CDO. In particular, MMFs are ongoing investment vehicles with no defined maturity, wherein the portfolio may contain multiple securities referencing a given obligor, with each security having a potentially different maturity. On the other hand, a CDO has a contractually specified maturity, and the underlying portfolio references a single bond or credit default swap (CDS) for a given obligor. Using the modeling assumptions in Sections 2.2–2.4, we have been able to establish a close connection between the two problems. In particular, upon identifying the term \(\sum_{k=1}^{n_s} N_{i,k} \text{LGD}_k^w\) with the loss associated with the default of obligor \(i\), we note that (2.2) is also what we would use when computing the loss distribution for a CDO, approaches for the solving of which are discussed, for example, in Hull and White (2004) and Andersen et al (2003). Problem (2.1) is similar to the problem of pricing an \(n\)th-to-default CDO, and under certain assumptions on default arrival times, the loss distribution can be solved, for example, using Monte Carlo simulation, as discussed in Andersen et al (2003). While (2.3) is slightly different in that we also need to model the number of defaults that can occur over the risk horizon for any given obligor index, in Section 3, we show how this can be achieved by extending the framework adopted in Andersen et al (2003) and Hull and White (2004).
3 MODELING AND COMPUTATION

Toward solving (2.3), we note from Sections 2.4 and 2.5 that it suffices to model defaults on a per-obligor basis, and that, for any given obligor index \( i \), the notional associated with any security \( k \), \( N_{i,k} \), is fixed and known in advance. Together, this implies that we can assign to each obligor index \( i \) an index-specific LGD, \( \text{LGD}_i \), given by \( N_i \text{LGD}_i = \sum_{k=1}^{n_s} N_{i,k} \text{LGD}_k \). Then, abbreviating \( L_{\text{full replenishment}} \) to \( L \), and denoting the loss corresponding to obligor \( i \),

\[
L_i = \sum_{w=1}^{N_i} N_i \text{LGD}_w,
\]
as \( L_i \), (2.3) can be rewritten as

\[
L := L_{\text{full replenishment}} = \sum_{i=1}^{n} \sum_{w=1}^{N_i} N_i \text{LGD}_w = \sum_{i=1}^{n} L_i.
\]

(3.1)

Toward solving (3.1), we need to model the following: the LGD corresponding to each obligor index, the default codependence structure between obligors and the number of defaults that can occur over the risk horizon for any given obligor index. To simplify this problem, in what follows, we assume that the LGD associated with any obligor index is deterministic (we show how this assumption can be relaxed in Section 3.2.2).

To capture the default codependence between obligors, we use the factor copula framework discussed in Hull and White (2004) and Andersen et al (2003). In particular, given the absence of market-implied information, we use the Gaussian and the \( t \) copulas. The Gaussian copula is the standard copula employed in the CDO market, so it is a natural choice. The \( t \) copula implies greater tail dependence and allows us to quantify model uncertainty. Note that our formulation does not depend on this assumption; if desirable, another copula could be used. Given the copula choice, we use a factor specification to capture the default correlation between obligors. Specifically, corresponding to each obligor \( i \), we define a random variable \( X_i \) given by

\[
X_i = \beta_i^T M + \sqrt{1 - \beta_i^T \beta_i} \epsilon_i,
\]

(3.2)

where \( \beta_i \) is the vector of loadings of obligor \( i \) onto the systemic factors \( M \), and \( \epsilon_i \) is a standard normal random variable. The systemic factors are assumed to be independent standard normal random variables. Given the replenishment strategy in Section 2.4,
(3.2) assumes that any new obligor corresponding to index $i$ continues to follow this specification.

Under the copula framework, the realization of $X_i$ is mapped to the default of the obligor $i$ using a percentile-to-percentile transformation, i.e., $P$(obligor $i$ defaults) = $P(X_i < Q_i^{-1}(PD_i))$, where $Q_i^{-1}(\cdot)$ represents the inverse cumulative distribution function (CDF) of $X_i$. Given (3.2) and a realization $m$ of the systemic factors $M$, this implies that for the Gaussian copula the conditional PD associated with obligor $i$ can be written as

$$
P(D_i \mid m) = \Phi\left(\frac{\Phi^{-1}(PD_i) - \beta_i^T m}{\sqrt{1 - \beta_i^T \beta_i}}\right),
$$

(3.3)

where $\Phi(\cdot)$ and $\Phi^{-1}(\cdot)$ represent, respectively, the normal CDF and the normal inverse CDF. For the $t$ copula, we use the identity that if $Y$ is an $n$-dimensional standard normal vector with correlation matrix $\Sigma$, and $g$ is a chi-squared variable with $v$ degrees of freedom, then $Z = (v/g)Y$ is an $n$-dimensional $t$ distributed vector with $v$ degrees of freedom and correlation matrix $\Sigma$. As a result, given the systemic factors’ realization $m$ and the chi-squared realization $\omega$, the conditional PD associated with obligor $i$ can be written as

$$
P(D_i \mid m, \omega) = \Phi\left(\frac{\sqrt{\omega/v} t_v^{-1}(PD_i) - \beta_i^T m}{\sqrt{1 - \beta_i^T \beta_i}}\right),
$$

(3.4)

where $t_v^{-1}(\cdot)$ denotes the inverse CDF of a $t$ distribution with $v$ degrees of freedom.

In what follows, we abbreviate $(m, \omega)$ to $m$ for the $t$ copula.

Given $m$, and having computed $PD_i \mid m$ toward computing the conditional loss distribution associated with obligor index $i$, $L_i \mid m$, we assume that defaults arrive as a Poisson process, with the arrival intensity calibrated such that the probability of no default arriving over the risk horizon matches the conditional no-default probability obtained in (3.3) or (3.4) for a Gaussian copula and a $t$ copula, respectively. Hence, we set

$$
e^{-\lambda_i^m T} = 1 - P(D_i \mid m),
$$

(3.5)

where $\lambda_i^m$ denotes, conditional on $m$, the Poisson arrival intensity associated with obligor index $i$. While this assumption admits the possibility that any given number of defaults occurs over the risk horizon with a positive probability, for economic as well as computational reasons, for each obligor index we would cap the number of defaults following which a reinvestment would be made over the risk horizon. Let us denote this cap by $n_d$ and note that it can be set as high as is desirable. Then, $L_i \mid m$
can be written as follows:

\[
L_i | m = \begin{cases} 
0 & \text{with probability } e^{-\lambda_i^m T}, \\
\text{sLGD}_i N_i & \text{with probability } e^{-\lambda_i^m T} \left( -\frac{\lambda_i^m T}{s!} \right)^s, \quad 1 \leq s \leq n_d - 1, \\
n_d \text{sLGD}_i N_i & \text{with probability } 1 - e^{-\lambda_i^m T} - \sum_{s=1}^{n_d-1} e^{-\lambda_i^m T} \left( -\frac{\lambda_i^m T}{s!} \right)^s. 
\end{cases}
\] (3.6)

Note that \( L_i | m \) is a discrete distribution in which the probability that the number of defaults exceeds \( n_d \) is assigned to the event that \( n_d \) defaults occur.

Following (3.1), the conditional loss distribution for the portfolio, \( L | m \), can be written as \( L | m = \sum_{i=1}^{n} L_i | m \). In particular, to obtain \( L | m \), we need to convolve \( L_i | m, i = 1, 2, \ldots, n \). To achieve this, we use the computational approach discussed in Hull and White (2004), wherein individual losses are scaled into a multiple of a common loss unit; subsequently, an iterative approach is used to build the conditional portfolio loss distribution. Toward this, let us denote by \( u \) the loss unit size used to scale losses (this choice is discussed in Section 3.1.4). Then, the scaled conditional loss distribution corresponding to obligor index \( i \), \( L^u_i | m \), can be written as

\[
L^u_i | m = \begin{cases} 
0 & \text{with probability } e^{-\lambda_i^m T}, \\
\text{floor} \left( \frac{\text{sLGD}_i N_i}{u} + 0.5 \right) & \text{with probability } e^{-\lambda_i^m T} \left( -\frac{\lambda_i^m T}{s!} \right)^s, \quad 1 \leq s \leq n_d - 1, \\
\text{floor} \left( \frac{n_d \text{sLGD}_i N_i}{u} + 0.5 \right) & \text{with probability } 1 - e^{-\lambda_i^m T} - \sum_{s=1}^{n_d-1} e^{-\lambda_i^m T} \left( -\frac{\lambda_i^m T}{s!} \right)^s. 
\end{cases}
\] (3.7)

Equation (3.7) is the same as (3.6) except that loss sizes have been scaled and rounded off. In what follows, we would denote the \( j \)th largest loss, \( j \in \{0, 1, \ldots, n_d\} \), in the support of \( L^u_i | m \) as \( l_i^j \). As a result, the maximum possible portfolio loss is given by \( n_l = \sum_{i=1}^{n} l_i^{n_d} \) loss units.

Toward obtaining the scaled conditional portfolio loss distribution, \( L^u | m \), define \( L^{(i),m,u} \), the scaled conditional portfolio loss distribution when losses due to the obligor indexes \( 1, 2, \ldots, i \) have been accounted for. Note that the support of \( L^{(i),m,u} \), \( i = 0, 1, \ldots, n \) is \( \{0, 1, \ldots, n_l\} \), and by definition \( L^u | m = L^{(n),m,u} \). Further, \( P(L^{(0),m,u} = 0) = 1 \), ie, having accounted for zero obligors corresponds to zero loss
with probability 1. Next, we perform the following computation for each of the obligor indexes, \( i = 1, 2, \ldots, n \), and for each possible portfolio loss size \( x \in \{0, 1, \ldots, n_l\} \):

\[
P(L(i,m,u) = x) = \sum_{0 \leq j \leq n_d, j|l_i| \leq x} P(L(i-1,m,u) = x - l_i) P(L_i^u = l_i | m).
\]  

Equation (3.8) states that, given \( m \), the event of a loss of size \( x \) being obtained upon adding up losses due to obligor indexes 1, 2, \ldots, \( i \) can occur only if a loss of size \( l_i \leq x \) occurs for obligor \( i \), and the sum of losses due to obligor indexes 1, 2, \ldots, \( i - 1 \) is \( x - l_i \). Equation (3.8) provides an iterative way to construct \( L^u | m \). Given \( L^u | m \), \( L^u \) is obtained by integrating over \( m \):

\[
P(L^u = x) = \int P(L^u = x | m) dP(m).
\]  

If no rounding error was introduced in (3.7), multiplying \( L^u \) by \( u \) gives \( L \). Otherwise, we obtain an approximation to \( L \).

### 3.1 Discussion

#### 3.1.1 Static versus dynamic model

The model proposed in (3.2)–(3.9) is static, in that the realization of systematic factors \( m \) does not evolve over time. In the absence of sufficient marked-implied information, specifying a dynamic model for the evolution of the systemic factors over the risk horizon is challenging; hence, for simplicity, we work in a static setting. As can be seen above, a significant advantage of this approach is that it allows us to use a semianalytic approach to compute the portfolio loss distribution; this is opposed to having to resort to Monte Carlo simulation, which can be computationally expensive.

#### 3.1.2 Poisson arrival of defaults

Our assumption that defaults arrive as a time-homogeneous Poisson process is equivalent to assuming that, for any given obligor index, default arrival times are independent and have an exponential distribution with a rate given by (3.5). The usage of (3.5) over the entire risk horizon itself implies that any obligor corresponding to obligor index \( i \) continues to follow the default codependence specification in (3.2). In the absence of market-implied information about default arrival times, and the potential default codependence of any new obligor, these assumptions introduce model-specification risk. To mitigate it, we supplement our results in Section 4 with sensitivity analysis with respect to the conditional loss distribution. We note that it can be easily replaced in our setting.
3.1.3 Computational complexity

Given \( u \) and \( n_l \), and conditional on systemic factor realizations \( m \), the steps outlined in (3.8) require \( O(n n_d n_l) \) work. Assuming that each systemic factor is discretized into \( n_m \) points, and that the dimensionality (number of systemic factors in (3.2), plus one for the chi-squared random variable in (3.4) if using a \( t \) copula) of the problem is \( d \), the computational complexity of the problem is given by \( O(n n_d n_l n_m^d). \)

3.1.4 Loss unit size

From the discussion above, we note that the choice of loss unit size \( u \) leads to a trade-off between computational time and accuracy. In our implementation, we set the loss unit size to be \( u = \min\{\min_{i=1,2,...,n}\{\text{LGD}_i N_i\}, \theta \sum_{i=1}^n N_i\} \), where \( \theta \) is a parameter that can be tuned to improve accuracy. To ensure that the error introduced by the choice of \( \theta \) is acceptable, the solution thus obtained could be compared with a smaller value of \( \theta \) or, alternatively, with the case where \( u \) is set to be the greatest common divisor of \( \text{LGD}_i N_i \) for all \( i = 1, \ldots, n \), so that the rounding error in (3.7) is zero. For an alternative approach to setting \( u \), see Andersen et al (2003).

3.1.5 Truncating integration of right tail

The computation in (3.2)–(3.9) is different from that discussed in the CDO pricing literature (for example, Andersen et al 2003; Hull and White 2004) in that, instead of computing an expected value, we need to compute the tail of a distribution. In particular, assuming that the highest loss quantile that we want to compute corresponds to a loss of \( f \% \) of the portfolio notional, the calculations outlined in (3.8)–(3.9) can be reduced to

\[
n_l f = \min\left\{n_l, \lceil \frac{f}{u} \sum_{i=1}^n N_i \rceil \right\}
\]

loss sizes ranging from 0 to \( n_l f \), thereby reducing the computation required by a factor \( \alpha = n_l f / n_l \). The choice of \( f \) could be informed by an understanding of the portfolio, or successively refined.

3.1.6 Sensitivities

As highlighted in Andersen et al (2003), sensitivities are easily computed in the framework discussed in Section 3 and, unlike in a Monte Carlo setting, they do not pose any additional numerical challenges. For example, to estimate portfolio risk upon removing a given obligor, one can skip the corresponding obligor index from
the computation in (3.8) or, alternatively, set the PD or LGD associated with this obligor to be zero.

3.2 Variations and extensions

3.2.1 Capping maximum loss

In (3.6), we have capped the number of defaults across obligor indexes at $n_d$. If the loss upon default, $\text{LGD}_i N_i$, varies significantly across obligor indexes, then it may instead be more natural to cap the total loss due to a given obligor index $i$ or, alternatively, the total portfolio level loss. The latter is particularly relevant, as it can be viewed as a modeling abstraction of the decision to liquidate the MMF portfolio when total losses hit a tolerance threshold. Both these variations are conveniently handled in our formulation. Toward capping the loss due to each obligor index at a given loss size $l_{\text{max}}^i$, define $n_{i,d}^{\text{max}}$, the maximum number of defaults allowed for obligor index $i$ given the loss cap $l_{\text{max}}$, as

$$n_{i,d}^{\text{max}} = \max_s \{s \mid s \text{LGD}_i N_i \leq l_{\text{max}}\}.$$ 

Then, replacing $n_d$ by $n_{i,d}^{\text{max}}$ in (3.6) results in the desired outcome. Toward capping the portfolio loss distribution at a given limit $l_{\text{pfl}}^{\text{max}}$, define the capped conditional portfolio loss distribution, $L_{\text{capped},u} \mid m$, as

$$L_{\text{capped},u} \mid m = \begin{cases} x & \text{with probability } P(L^u = x \mid m), \ x < l_{\text{pfl}}^{\text{max}}, \\ l_{\text{pfl}}^{\text{max}} & \text{with probability } \sum_{s=l_{\text{pfl}}^{\text{max}}}^{n_i} P(L^u = s \mid m), \ x \geq l_{\text{pfl}}^{\text{max}}, \end{cases}$$

where $L^u \mid m$ is given by (3.8). The unconditional capped portfolio loss distribution can then be obtained by integrating over $L_{\text{capped},u} \mid m$. Note that the formulation in (3.10) assumes that the portfolio manager would be able to liquidate the fund precisely when the portfolio loss hits the target threshold $l_{\text{pfl}}^{\text{max}}$. In practice, it is likely that there will be a slippage between the time this decision is made and that by which the portfolio liquidation is complete.

3.2.2 Stochastic LGD

Under the assumption that the LGD distribution has a discrete support, stochastic LGDs can be easily incorporated in our framework. Since any arbitrary distribution can be discretized to the desired degree of accuracy, this requirement does not lead to any loss of generality. Toward incorporating stochastic LGDs, denote the discrete LGD distribution associated with obligor index $i$ as $\text{SLGD}_i$. Then, (3.6) can be rewritten...
as follows:

\[
L_i^{\text{SLGD}} \mid m = \begin{cases} 
0 & \text{with probability } e^{-\lambda_i^m T}, \\
\sum_{j=1}^s \text{SLGD}_i N_i & \text{with probability } e^{-\lambda_i^m T} \left(\frac{(-\lambda_i^m T)^s}{s!}\right), 1 \leq s \leq n_d - 1, \\
\sum_{j=1}^{n_d} \text{SLGD}_i N_i & \text{with probability } 1 - e^{-\lambda_i^m T} - \sum_{s=1}^{n_d-1} e^{-\lambda_i^m T} \left(\frac{(-\lambda_i^m T)^s}{s}\right).
\end{cases}
\]

While an additional step of convolving the loss distributions is required, the conditional loss distribution \(L_i^{\text{SLGD}} \mid m\) continues to have a discrete support, so that the convolution approach described in Section 3 can again be used to obtain the conditional portfolio loss distribution. In addition to better reflecting reality, stochastic LGDs, by distributing the probability mass over a range of loss sizes, smooth the portfolio loss distribution.

4 EXPERIMENTS

4.1 MMF portfolios

We used our model to quantify the default risk of three prime MMFs. With total assets under management across these three MMFs exceeding US$150 billion, these were among the largest prime MMFs available, and were therefore selected. In what follows, for the purpose of anonymity, we will label these MMFs as Fund 1, Fund 2 and Fund 3, respectively. The fund holdings were sourced from publicly disclosed information as of March 31, 2013.

4.2 Numerical parameters and implementation

The numerical parameters that we need to specify for the implementation of our model are \(\theta, n_d, n_m\) and \(f\). The parameter \(\theta\), which determines the loss unit size, was adaptively set to be \(1 \times 10^{-4}\), as reducing it further did not have a material impact on the results. The number of discretizations of each of the systemic factors, \(n_m\), was adaptively set to a value within the range \(2^5\)–\(2^{12}\) to ensure that the convergence error due to \(n_m\) was within \(1 \times 10^{-5}\) (we used the trapezoidal rule for numerical integration). For parameter \(n_d\), which determines the maximum number of defaults that can occur for a given obligor index over the risk horizon, we used the values 1, 2, 5 and 10. The parameter \(f\), which determines the computation needed in (3.8) and (3.9), was set to be 5%. For the portfolios we studied, this was sufficient to estimate the quantiles of interest. To smooth the portfolio loss distribution, we assumed that \(\text{LGD}_i N_i, i = 1, 2, \ldots, n\), follows a uniform distribution over the support \([\text{LGD}_i N_i - \eta u, \text{LGD}_i N_i + \eta u]\).
\[ \eta \mu \], where the parameter \( \eta \) was experimentally set to be 5. We implemented the model in C++.

### 4.3 Model parameters

The model parameters that we need to specify for our implementation are the obligor PDs, security LGDs, the correlations between obligors and the degrees of freedom for the \( t \) copula. As the securities held by MMFs typically reference the various central, state and local governments, banks and other large financial companies, the majority of which are very highly rated, scant historical default data is available to estimate these parameters. Similarly, while instruments such as CDSs trade for some of these obligors, it is hard to strip out the risk premiums embedded in the prices so that they can be used for computing risk under the real-world measure. For this reason, our estimates are necessarily subjective. To address the resulting parameter uncertainty, we have supplemented our results in Section 4.5 with sensitivity analysis.

#### 4.3.1 PD

For each obligor, we mapped the median of publicly available ratings from Standard & Poor’s (S&P), Moody’s and Fitch into a one-year PD. Based on Standard & Poor’s (2013), this mapping is (AAA,0), (AA+,0), (AA,1), (AA-,2), (A+,5), (A,6), (A-,7), (BBB+,15) and (BBB,24), where the first entry of each pair denotes the S&P rating, and the second entry denotes the corresponding PD in bps. Ratings below BBB were not needed for the portfolios we considered. One limitation associated with the usage of these PD values is that, being one-year PDs, they may not adequately account for the possibility that under the replenishment strategy of Section 2.4, a downgraded obligor may be replaced in the MMF portfolio before it defaults. Given the lack of empirical data, and as PD estimation is not the focus of this paper, we did not investigate this further.

#### 4.3.2 LGD

Our choice of LGD for the different security types held by MMFs is based on subjective judgment and estimates available in the literature. We discuss these choices in turn. As US treasuries and agencies are backed by the credit of the US government, we set their LGD to 0%. For the same reason, and as repos are overcollateralized, we set the LGD for treasury and agency repos to 0%, and the LGD for repos backed by other assets to 5%. For ABCP, in line with Acharya et al (2013), we set the LGD to 10%. For bank CDs, based on FitchRatings (2012), we set the LGD to 20%. For municipal bonds, based on Moody’s (2012), we set the LGD to 35%. While some variable rate demand notes have a backstop letter of credit, because we are unable to identify the ones that do, we set their LGD to 35% as well. Based on Mattin (2013),
which notes that the recovery on Lehman-issued CP was 56%, we set the LGD for financial CP to 45%. As the details of the other CP category are not known to us, we set its LGD to also be 45%. For all other securities, we assume that the LGD is 60%, which is consistent with the 40% recovery assumption made in the corporate CDS market.

4.3.3 Correlation

We estimate correlations using two widely used approaches. First, we use the Basel II formula from Basel Committee on Banking Supervision (2005); second, we proxy asset correlation using equity correlation estimated from historical returns. While neither of these approaches is perfect (e.g., the former was intended for wholesale credit portfolios, while the latter is thought to overestimate asset correlation, see Dullmann et al. (2008)), we use the estimates obtained from these approaches as bounds for our analysis.

For the PD values chosen in Section 4.3.1, the Basel correlation formula (Basel Committee on Banking Supervision 2005, p. 17) results in correlations ranging from 22% to 24%. The factor loading of each obligor ($\beta_i$ in (3.2)) was set to be the square root of this correlation.

To estimate equity correlations, we used historical, currency-adjusted, weekly, equity log returns for three major financial companies from each of the United States, Canada, Europe and Australia, over the period March 2008–March 2013, to construct a $12 \times 12$ correlation matrix. The choice of the financial companies (JP Morgan, Wells Fargo, State Street, Bank of Montreal, Bank of Nova Scotia, Royal Bank of Canada, Westpac, Commonwealth Bank of Australia, Australia and New Zealand Banking Group, Barclays, BNP Paribas and Deutsche Bank) was subjective, and influenced by their representativeness as well as their occurrence as a holding in the chosen MMF portfolios. The estimated correlations ranged from 55% to 90% within regions, and from 35% to 65% across regions (for brevity, we have not presented the pairwise correlations). A one-factor and a two-factor principal component analysis (PCA) decomposition of this matrix was used to estimate factor loadings, and each obligor in the MMF portfolio was randomly mapped to one of the three financial companies in the region to which it belonged.

4.3.4 $t$ copula: degrees of freedom

The parameter $v$ allows us to explore a range of codependence structures. Using the tail-dependence criteria discussed in Demarta and McNeil (2005) as a guide toward ensuring that for the range of correlations considered in Section 4.3.3, the resulting tail dependence is neither extreme nor trivial, we chose two values for $v$: 5 and 15.
4.4 Metrics

We use two metrics to measure the risk of MMF portfolios:

(i) the quantiles of the loss distribution;

(ii) the probability that loss exceeds a given threshold.

The former captures the tail of the distribution, while the latter helps reflect it into a more interpretable measure, eg, the probability that loss exceeds 50bps would reflect that the buck has broken. The measures are complementary, eg, higher default correlation (keeping everything else fixed) would cause quantiles to become higher as we move into the tail, while the probability of breaking the buck may (or may not) become smaller.

4.5 Numerical results

Table 1 on the facing page, Table 2 on page 66 and Table 3 on page 67 show the results of using our approach to quantify the default risk for Fund 1, Fund 2 and Fund 3, respectively, for the two risk horizons of three months and one year, and under various choices of copulas, correlation levels, PDs and LGDs. The first column in each row lists the choice of the risk horizon $T$ (three months (3M) or one year (1Y)), the copula (Gaussian, $t$ with degrees of freedom $v = 15$ ($t_{15}$) or $t$ with degrees of freedom $v = 5$ ($t_5$)), the correlation level (one-factor PCA decomposition of the empirical correlation matrix (1f) or Basel II ASRF correlations (ASRF)), the maximum number of defaults allowed per obligor index $n_d$ (1, 2, 5 or 10) and optionally whether PD or LGD were tweaked from their base-case values, specified in Sections 4.3.1 and 4.3.2, respectively (by 1bp and 5%, respectively). The remaining columns show the corresponding tail loss quantiles (99th, 99.9th and 99.97th) and the probability of loss exceeding a threshold (25bps, 50bps, 75bps, 100bps and 250bps). For brevity, we have limited the presented results to the most illustrative combinations of parameter choices.

From our stylized model, we observe that for each prime MMF, the choice of the risk horizon has the largest impact on both the tail quantiles and the probability of breaking the buck, with both increasing with the risk horizon. For example, for Fund 1, under the Gaussian copula, 1f correlations and $n_d = 1$, the 99.9th quantile and the probability of breaking the buck for the 3M and the 1Y risk horizons are 14bps and 74bps, and 0.04% and 0.14%, respectively. While the magnitude of these metrics is driven by the various model assumptions, the increase with respect to the risk horizon simply reflects that, under a stress scenario, funds would experience higher losses if they were to continue to operate as ongoing concerns in the face of defaults. This suggests that any capital requirement for MMFs should be closely linked to their liquidation strategy, in particular, whether the fund sponsor would seek to support or liquidate the fund in the event of a stress.
### TABLE 1  Tail quantiles and probabilities of loss exceeding given thresholds for Fund 1.

<table>
<thead>
<tr>
<th>Horizon, copula, correlation, $n_d$, tweak</th>
<th>Loss quantiles</th>
<th>Probability loss exceeds threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>99th</td>
<td>99.9th</td>
</tr>
<tr>
<td>3M, Gaussian, 1f, 1 default</td>
<td>0</td>
<td>14</td>
</tr>
<tr>
<td>3M, $r_{15}$, 1f, 1 default</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>3M, $r_5$, 1f, 1 default</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>3M, $r_5$, 1f, 10 defaults</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td>1Y, $r_{15}$, 1f, 1 default</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td>1Y, $r_5$, 1f, 1 default</td>
<td>0</td>
<td>64</td>
</tr>
<tr>
<td>1Y, Gaussian, ASRF, 1 default</td>
<td>1</td>
<td>73</td>
</tr>
<tr>
<td>1Y, $r_{15}$, 1f, 2 defaults</td>
<td>0</td>
<td>74</td>
</tr>
<tr>
<td>1Y, $r_5$, 1f, 2 defaults</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>1Y, $r_5$, 1f, 10 defaults</td>
<td>0</td>
<td>69</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default, LGD + 5%</td>
<td>1</td>
<td>82</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default, PD + 1bp</td>
<td>1</td>
<td>101</td>
</tr>
</tbody>
</table>

We observe that in our stylized model, risk horizon has the largest impact on both the tail quantiles and the probabilities that loss exceeds given thresholds. A heavier tailed copula shifts the loss distribution into the tail and leads to a higher probability of extreme losses. Increasing $n_d$ has a small impact at $T = 3$M, and it is the most noticeable for the $r_5$ copula, at high quantiles and for $T = 1$Y.
We observe that in our stylized model, risk horizon has the largest impact on both the tail quantiles and the probabilities that loss exceeds given thresholds. A heavier tailed copula shifts the loss distribution into the tail and leads to a higher probability of extreme losses. Increasing $n_d$ has a small impact at $T = 3M$, and it is the most noticeable for the $t_5$ copula, at high quantiles and for $T = 1Y$. 

<table>
<thead>
<tr>
<th>Horizon, copula, correlation, $n_d$, tweak</th>
<th>Loss quantiles</th>
<th>Probability loss exceeds threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>99th</td>
<td>99.9th</td>
</tr>
<tr>
<td>3M, Gaussian, 1f, 1 default</td>
<td>0</td>
<td>41</td>
</tr>
<tr>
<td>3M, $t_{15}$, 1f, 1 default</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>3M, $t_5$, 1f, 1 default</td>
<td>0</td>
<td>21</td>
</tr>
<tr>
<td>3M, $t_5$, 1f, 10 defaults</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default</td>
<td>0</td>
<td>119</td>
</tr>
<tr>
<td>1Y, $t_{15}$, 1f, 1 default</td>
<td>0</td>
<td>125</td>
</tr>
<tr>
<td>1Y, $t_5$, 1f, 1 default</td>
<td>0</td>
<td>119</td>
</tr>
<tr>
<td>Gaussian, ASRF, 1 default</td>
<td>2</td>
<td>104</td>
</tr>
<tr>
<td>1Y, $t_{15}$, 1f, 2 defaults</td>
<td>0</td>
<td>129</td>
</tr>
<tr>
<td>1Y, $t_5$, 1f, 2 defaults</td>
<td>0</td>
<td>125</td>
</tr>
<tr>
<td>1Y, $t_5$, 1f, 10 defaults</td>
<td>0</td>
<td>125</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default, LGD + 5%</td>
<td>0</td>
<td>139</td>
</tr>
<tr>
<td>1Y, Gaussian, 1f, 1 default, PD + 1bp</td>
<td>0</td>
<td>103</td>
</tr>
</tbody>
</table>
We observe that in our stylized model, risk horizon has the largest impact on both the tail quantiles and the probabilities that loss exceeds given thresholds. A heavier-tailed copula shifts the loss distribution into the tail and leads to a higher probability of extreme losses. Increasing \( n_d \) has a small impact at \( T = 3M \), and it is the most noticeable for the \( t_5 \) copula, at high quantiles and for \( T = 1Y \).
Comparing across funds, we find that Fund 1 has the lowest tail loss, followed by Fund 2 and Fund 3, respectively. For example, under the setting of a Gaussian copula, $1\text{f}$ correlations, $n_d = 1$ and $T = 1\text{Y}$, the 99.9th and the 99.97th quantiles for Fund 3 exceed the corresponding quantiles for Fund 1 by about 185bps and 313bps, respectively. Similarly, under the same setting, compared with a 0.14% probability of breaking the buck for Fund 1, the corresponding probability for Fund 3 is 0.55%, thereby suggesting that Fund 1 is less risky compared with Fund 3. To understand the cause of these differences, we computed the notional weighted PD and LGD, and the average expected loss as a fraction of notional for each portfolio. These metrics were 0.73bps, 11.7% and 0.24bps for Fund 1; 1.67bps, 16.0% and 0.45bps for Fund 2; and 3.54bps, 22.6% and 1.09bps for Fund 3, respectively. This is directionally consistent with the results in Table 1 on page 65, Table 2 on page 66 and Table 3 on the preceding page. Together, these results show that MMFs vary considerably in the riskiness of their portfolios. They also suggest that any capital requirements or liquidity fees for MMFs should account for such differences.

Regarding the need to model replenishment upon default, we observe that increasing $n_d$ has a very small impact for $T = 3\text{M}$. For $T = 1\text{Y}$, the effect is more noticeable, particularly for the 99.97th quantile, the $t_5$ copula and for Fund 3. This is expected, as the likelihood of multiple defaults per obligor index increases with PD (which, as noted above, is the highest for Fund 3), and the probability that multiple obligor indexes incur defaults simultaneously is higher for the heavier tailed $t_5$ copula. These results, in addition to providing sensitivity to our assumptions regarding the conditional loss distributions of obligor indexes, demonstrate the importance of modeling portfolio replenishment upon default for MMF portfolios when measuring risk at high quantiles and for long risk horizons.

In addition to the above, Table 1 on page 65, Table 2 on page 66 and Table 3 on the preceding page shows that a heavier tailed copula as well as a higher correlation level shifts the loss distribution into the tail (to see this, compare, eg, the loss probabilities upon switching from ASRF to $1\text{f}$ correlation levels), provide sensitivity with respect to the various model parameters (eg, PD) as well as confirm the observation in Section 3.1.1 that estimating the default losses for these high-quality MMF portfolios using a Monte Carlo simulation would have been computationally expensive.

In Table 4 on the facing page, we present the tail quantiles of the loss distribution for each of the three prime MMFs upon using the two analytic approaches mentioned in Section 1: the Basel II ASRF formula without and with granularity adjustments (labeled as ASRF and GA, respectively). For the Basel II ASRF approach, we used the contractual maturity of the underlying positions, together with the Basel maturity adjustment (see Basel Committee on Banking Supervision (2005) for details), to estimate the tail quantiles for $T = 1\text{Y}$. For granularity adjustment, we used the simplified granularity adjustment approach, together with parameters specified in
Gordy and Lutkebohmert (2013). Upon comparing the results from these approaches with a setting employing the same parameters in our framework (namely, a Gaussian copula, ASRF correlation, $n_d = 1$ and $T = 1Y$), we find that the Basel II ASRF approach underestimates the risk of MMF portfolios at high quantiles (99.9th and 99.97th) while overestimating it at lower quantiles (99th). This is expected, as MMF portfolios are far from homogeneous and fine grained, which is what they are assumed to be in the Basel II ASRF formula. However, applying granularity adjustment leads to an overestimation of risk across quantiles, thereby demonstrating the need for an accurate loss aggregation approach, such as the one used in this paper, to estimate the default risk of MMF portfolios.

5 CONCLUSION

We proposed a semianalytic approach to quantify the default risk of MMF portfolios. The approach takes into account the short-dated nature of MMF holdings and the need to model portfolio replenishment, the existence of portfolio holdings with different maturities and concentration risk. Upon applying this approach to three of the largest prime MMFs, we found that MMFs vary considerably in their default risk. This suggests that enforcing fixed fees, such as the liquidity-dependent but not risk-dependent 1% or 2% liquidity fees proposed in the United States (see Securities and Exchange Commission 2014), or a static capital buffer across funds, such as the 3% buffer proposed in Europe (see European Commission 2013), would fail to account for the heterogeneity in MMF portfolios. As such, these fees are likely to either be punitive or unable to address the systemic risk associated with MMFs. The results of our stylized model also show that default risk increases with the risk horizon, which suggests that any requirement to establish liquidity fees or capital buffers should be closely linked to the portfolio liquidation strategies established by MMFs. Overall, this implies that while the proposals of the Securities and Exchange Commission (2014) and European Commission (2013) are steps in the right direction, unless the liquidity fees and capital buffer size are calibrated to the riskiness of the MMF portfolio, during a crisis, a run on MMFs could still result.
In addition to providing a way to quantify the default risk of MMF portfolios, the proposed model can be used to answer interesting questions related to MMFs (for example, the variation in risk over time due to changing portfolio characteristics and economic conditions, or the impact of changes in diversification rules). The approach is general and could also be used to compute the default risk of other portfolios where replenishment needs to be modeled.

DECLARATION OF INTEREST

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

ACKNOWLEDGEMENTS

The author would like to thank Anupam Sahay, Ashish Dev, Chris Turner, Jim Sexton, Jimmie Irby, Jonathan Wendin, Ke Wan and Nan Zhou for several insightful discussions. Any mistakes are the author’s own.

REFERENCES


Moody’s (2010). Sponsor support key to money market funds. Report, August.


Research Paper

Are all collections equal? The case of medical debt

Kenneth P. Brevoort and Michelle Kambara

Office of Research, Consumer Financial Protection Bureau, 1275 N First Street NE, Washington, DC 20002, USA; emails: kenneth.brevoort@cfpb.gov, michelle.kambara@cfpb.gov

(Received March 27, 2015; revised July 10, 2015; accepted August 12, 2015)

ABSTRACT

Bills for unreimbursed medical care may be reported to national credit reporting agencies by third-party debt collectors. The use of this information in credit scoring models, which have not traditionally distinguished collection accounts for medical bills from other collection accounts, has been controversial because of the unique characteristics of medical debt. This paper explores the predictive value of medical collections in the context of a credit scoring model. We find that medical collections are less predictive of future credit performance than nonmedical collections. We also find that medical collections that have been paid in full are less predictive than those that remain unpaid. These results suggest that the practice of treating all collections the same over-penalizes the credit scores of consumers with medical collections and reduces the predictiveness of credit scoring models.

Keywords: credit scoring; medical debt; collection accounts.

1 INTRODUCTION

Adverse health events are often cited as “trigger” events associated with financial hardship and credit defaults (see, for example, Elmer and Seelig 1999; Avery et al
Such events can cause income disruptions and increase household expenditure, particularly when they involve prolonged or permanent changes in a consumer’s health status. The result can be an increase in consumer indebtedness (Kim et al 2012; Babiarz et al 2013; Mohanan 2013).

The amount of medical debt carried by US households has grown over time (Doty et al 2008), and this debt poses challenges for households. Several studies have found a relationship between medical debt and consumer bankruptcy (Domowitz and Sartain 1999; Himmelstein et al 2005; Jacoby and Holman 2010; Zhu 2011). Consumers with medical debt are also more likely to forgo medical care (Kalousova and Burgard 2013).

One of the potential challenges that medical debt poses for households is its effect on credit access. If medical debts are reported to national credit reporting agencies (NCRAs), the existence of these debts (whether eventually paid off or not) can reduce a consumer’s credit score, making credit access both more difficult and more expensive. Concerns about the effect medical debt has on credit scores have led some, such as Rukavina (2014), to suggest banning the use of this information in credit scoring models. These concerns have also drawn the attention of policymakers (Menendez 2012).

Those who have advocated limiting the use of medical debt in credit scoring models have done so, at least in part, in the belief that this information is not informative about a consumer’s creditworthiness. These arguments are based on the unique aspects of the accounts that are reported to the NCRAs as medical debt.

Only bills for unreimbursed medical care are reported to the NCRAs as medical debt. These are bills for medical services that the consumer either has not paid or has left unpaid for a time. Debts can be incurred when paying for medical services, such as when one uses a credit card. Debts can also be incurred indirectly, such as when a consumer pays for the services with cash but is left with insufficient funds to cover other expenses and, as a result, has to take on more debt. Both of these can justifiably be considered medical debt, but they are not reported to the NCRAs as such. Instead, debts accumulated in reimbursing health-care providers appear as credit card or other debts, indistinguishable from debts accumulated for nonmedical purposes.

Almost all of the bills for unreimbursed medical care that are reported to NCRAs are reported by third-party collection agencies. We refer to these as “medical collections” to distinguish this type of debt from broader definitions of medical debt, which could include debt incurred in reimbursing providers for medical care. There are two

---

1 While not addressing medical debt directly, Gross and Souleles (2002) show that credit delinquencies are related to a lack of health insurance, and Gross and Notowidigdo (2011) show that expansions of Medicaid eligibility decrease consumer bankruptcies.
potential differences between medical collections and debts for reimbursed medical care that are important for the issues we examine.

The first is that consumers may not be aware of debts for unreimbursed care. Paying for medical services with cash, checks or credit cards requires consumer participation, so consumers should be aware of the amount they owe. This is not necessarily true for unreimbursed care. A consumer with medical insurance may make the co-payment required by their insurance and consider the bill paid; however, if the insurer does not cover the expected amount, the consumer may be liable for the balance. Such debts can be reported as medical collections without informing the consumer.

The second is that consumers may not view debts from unreimbursed care as legitimate. In the previous example, if the consumer believes their insurance should cover the entire balance, they might not acknowledge the debt as their own. Or, if a long time has passed, the consumer may no longer recognize the charges. In contrast, consumers who acquire debt as a result of paying for medical care have implicitly acknowledged the legitimacy of the debt in doing so.2

These unique characteristics of medical collections have led some to suggest that such debts are not informative about creditworthiness. This belief is reinforced by the notion that medical expenses are more likely than the various forms of nonmedical debt to be incurred by necessity. If medical debt essentially reflects the bad luck associated with illness, then the presence of an unpaid bill for medical care may say less about a consumer’s creditworthiness than other unpaid bills.

If medical collections are not informative about creditworthiness, then credit scoring models that incorporate this information may penalize these consumers, despite the debts not indicating a higher likelihood of future delinquency. On the face of it, this concern may appear to be without merit. In an empirically derived credit scoring model, a variable that is not predictive will receive no weight, even if it is included as a predictive factor. However, when it comes to accounts reported by third-party collection agencies, which account for almost all of the medical debt reported to NCRAs, credit scoring models generally do not differentiate according to the source of the debt. This means that the score penalty for an unpaid medical bill will be the same as unpaid cell phone bills, or other nonmedical collection. As long as nonmedical collections predict future performance, it is possible that consumers are being over-penalized for their medical debts.

In addition to not differentiating medical from nonmedical collections, credit scoring models have not traditionally distinguished collections that have been paid in full from those that remain unpaid. Paid collections are debts that were delinquent at some

2 Clearly, there are some exceptions to this implicit acknowledgment, such as when the debts are fraudulent or when the debts were for goods or services that were not provided as agreed.
point in time, reported to NCRAs and later paid by the consumer. In such cases, the collection remains on the consumer’s credit record for up to seven years, though it is recorded as having been paid in full.

While one might expect paid collections to provide a less negative signal about a consumer’s creditworthiness than an unpaid collection, credit scoring models have not traditionally distinguished between paid and unpaid collections. This equal treatment has generated concerns about the appropriateness of this treatment, and the US Congress has considered legislation in response. The Medical Debt Responsibility Act of 2013 would require NCRAs to remove paid medical collections from a consumer’s credit record forty-five days after they are paid.

Despite the concerns about medical collections, very little is known about their value in predicting credit performance. This paper seeks to fill this gap, and to inform the debate on the use of medical collections in credit scoring models, by empirically evaluating the predictive value of medical collections in the context of generic credit history scoring models.

We use a new source of data, acquired by the Consumer Financial Protection Bureau (CFPB) from one of the NCRAs. This data set, the Consumer Credit Panel (CCP), contains the same information used by industry to build and validate credit scoring models. We use this data to relate the credit scores of a random sample of over four million consumers in September 2011 to their performance on credit obligations during the next two years. We assess credit performance using two measures commonly used when validating scoring models.

We use this data to conduct two analyses. First, we evaluate whether medical and nonmedical collections are equally predictive of credit performance. We do this by comparing the performance of consumers with medical collections with the performance of consumers who have the same credit scores and the same number of total collections (medical plus nonmedical) but fewer medical collections. If medical and nonmedical collections are equally informative about creditworthiness, which is the implicit assumption of credit scoring models that treat all third-party collections equally, then the performance of these two groups should be the same. In contrast, if medical collections provide a less-negative signal about a consumer’s creditworthiness, then we would expect to observe lower delinquency rates among consumers with more medical collections.

3 Collections that are paid in full before being reported to NCRAs likely are never reported. Debt collectors often use the fact that the debt has not yet been reported to induce consumers to pay the collection so as to avoid the debt being reported.

4 The most recent version of VantageScore (version 3.0) excludes collections that have been paid in full. Previous versions of VantageScore did not distinguish paid from unpaid collections.
Second, we examine the predictive value of paid medical collections using a similar approach. Comparing people with the same credit scores and the same number of medical collections, we evaluate whether there is a relationship between the share of medical collections that remain unpaid and credit performance. If paid and unpaid medical collections are equally predictive, which is also an implicit assumption of many credit scoring models, then we should observe comparable performance, regardless of the mix of paid and unpaid medical collections. If paid medical collections are less informative than unpaid medical collections, then we would expect lower delinquency rates for consumers with more paid medical collections.

Our results suggest that the practice of affording equal treatment to all collections, regardless of type, is likely suboptimal. We find that consumers with medical collections have lower delinquency rates during the performance period than consumers who have fewer medical collections (holding constant their credit scores and number of total collections). We also find that consumers with more paid medical collections have lower delinquency rates than consumers with fewer paid medical collections. These results suggest that the equal treatment of all collections may be over-penalizing the scores of consumers who have medical collections, particularly those with paid medical collections, and potentially reducing the predictiveness of the credit scoring models that use this equal treatment.

The remainder of the paper details our analyses. Section 2 documents the data used in our analyses. Sections 3 and 4 present the results of our analyses, exploring the relative predictiveness of medical and nonmedical collections and of paid and unpaid medical collections, respectively. Section 5 examines the robustness of our results, and Section 6 discusses the conclusions we draw from these analyses.

2 DATA

The data used in this study comes from the CFPB’s CCP. The CCP is a longitudinal, nationally representative sample of de-identified credit records from one of the NCRAs. The full data set covers the end of each calendar quarter from June 2004 through December 2012, with monthly updates thereafter.5 At each time period, the entire credit record, less any information that could directly identify the consumers or creditors in the sample, is supplied.6 This data provides account-level information about each consumer’s debts and includes a commercially available credit score calculated at the end of each quarter.

---

5 The CCP also includes credit records from each September from 2001 through 2003.
6 For more information about the contents of credit records, see Avery et al (2003).
TABLE 1  Summary statistics.

(a) Medical collections

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence</td>
<td>4,627,075</td>
<td>0.195</td>
<td>0.396</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>900,131</td>
<td>3.803</td>
<td>5.673</td>
<td>2</td>
</tr>
<tr>
<td>Number paid</td>
<td>900,131</td>
<td>0.447</td>
<td>1.296</td>
<td>0</td>
</tr>
<tr>
<td>Number ≥ US$100</td>
<td>900,131</td>
<td>2.765</td>
<td>4.456</td>
<td>1</td>
</tr>
<tr>
<td>Original amount (thousands)</td>
<td>900,131</td>
<td>2.114</td>
<td>7.160</td>
<td>0.620</td>
</tr>
<tr>
<td>Current balance (thousands)</td>
<td>900,131</td>
<td>1.978</td>
<td>7.177</td>
<td>0.491</td>
</tr>
</tbody>
</table>

(b) Nonmedical collections

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence</td>
<td>4,627,075</td>
<td>0.254</td>
<td>0.436</td>
<td>0</td>
</tr>
<tr>
<td>Number</td>
<td>1,177,440</td>
<td>2.902</td>
<td>2.827</td>
<td>2</td>
</tr>
<tr>
<td>Number Paid</td>
<td>1,177,440</td>
<td>0.357</td>
<td>0.838</td>
<td>0</td>
</tr>
<tr>
<td>Number ≥ US$100</td>
<td>1,177,440</td>
<td>2.359</td>
<td>2.269</td>
<td>2</td>
</tr>
<tr>
<td>Original Amount (000s)</td>
<td>1,177,440</td>
<td>2.693</td>
<td>7.881</td>
<td>1.012</td>
</tr>
<tr>
<td>Current Balance (000s)</td>
<td>1,177,440</td>
<td>2.659</td>
<td>7.189</td>
<td>0.828</td>
</tr>
</tbody>
</table>

(c) Performance metrics

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Any account</td>
<td>3,705,062</td>
<td>0.124</td>
<td>0.330</td>
<td>0</td>
</tr>
<tr>
<td>New</td>
<td>745,978</td>
<td>0.110</td>
<td>0.313</td>
<td>0</td>
</tr>
<tr>
<td>Credit score</td>
<td>4,181,081</td>
<td>693.2</td>
<td>107.2</td>
<td>716</td>
</tr>
</tbody>
</table>

SD denotes standard deviation. The incidences for medical and nonmedical collections are calculated using the entire sample population. Other variables about medical and nonmedical collections are calculated conditional on having a medical or nonmedical collection, respectively. Performance measures are calculated for those consumers with observable performance during the October 2011–September 2013 performance period. Credit scores are calculated for those consumers with a sufficient credit history to be considered scorable by the credit scoring model.

Following conventional practice in evaluating credit scoring models, we use data from two points in time, September 2011 and September 2013 (Board of Governors of the Federal Reserve System 2007; Avery et al 2012). Using two time periods allows consumers’ credit characteristics, including their credit scores, at the earlier point (September 2011) to be related to performance on credit obligations during the ensuing two-year “performance period” (October 2011 to September 2013). We use September 2011 as the start of the performance period because it is the earliest date...
for which medical collections can be identified in the CCP. We use September 2013 as the end of the performance period to provide a twenty-four-month performance period, a standard length of time used to evaluate credit scoring models (Board of Governors of the Federal Reserve System 2007).

Summary statistics from the credit record data are provided in Table 1 on the facing page, which shows information about the incidence of medical and nonmedical collections. This data shows that medical collections appear on about one in five credit records, compared with about one in four for nonmedical collections (or just under one in three for either type of collection). Consumers with medical collections had an average of 3.8 collections each (the median is 2), of which 0.4 had been paid in full.

The summary statistics in Table 1 on the facing page also include the two measures of credit performance. The first is the “any account” performance measure, which describes performance on new and existing accounts. A “new account” is one that was opened during the first three months of the performance period (October to December 2011). An “existing account” is one that was open and in good standing at the start of the performance period. We use the any account performance measure because it approximates the type of performance that is used to construct generic credit history scoring models, such as FICO or VantageScore, and because we observe performance on new and existing accounts for most consumers in our sample.

The second performance measure, the “new account” measure, is limited to performance on new accounts. In some ways, performance on new accounts is a superior measure for evaluating credit scores, because none of the information about these accounts was used to calculate the score. The drawback of evaluating performance on new accounts is that only a small portion of our sample, about 16%, opened a new account on which performance could be assessed.

Both performance measures are indicator variables that take a value of 1 if the consumer is ninety days past due or worse on one of their accounts during the performance period. Otherwise, the indicator variable takes on a value of 0 as long as the

---

7 One of the variables necessary to determine if a third-party collection was medical or nonmedical was not available for earlier dates in the archives from which the CCP was drawn. As a result, we cannot evaluate the predictive value of medical collections for performance periods starting before September 2011. While we have no reason to believe this performance period will generate unusual results, because we can only evaluate a single performance period, we cannot rule out that other performance periods might produce different results.

8 The same cannot be said for existing accounts, which is why we only evaluate existing accounts that were in good standing at the start of the performance period. Accounts that were open and delinquent at the start of the performance period would both cause a lower credit score and almost guarantee a delinquency during the performance period (unless the account cured during the first month of the performance period).
consumer has at least one account with payments due during the performance period. For consumers without new or existing accounts that have observable performance, the performance measure is not applicable and we exclude these observations from our analyses.

These performance measures are highly correlated with the credit scores calculated from just before the start of the performance period, as shown in Figure 1. In both panels, the blue circles show the delinquency rate (the mean of the performance measure) at each score. (The red line shows the fitted values for a logit estimation that includes a cubic spline of the credit score.) The negative monotonic relationships between credit scores and delinquency rates suggest that, for the population as a whole, the credit score effectively rank orders consumers according to the credit risk they pose.

3 MEDICAL AND NONMEDICAL COLLECTIONS

Credit scores are empirically derived signals about a consumer’s likelihood of default that are based on credit history information. Most types of credit accounts, such as credit cards, can either positively or negatively affect a consumer’s credit score. For example, a credit card with a long history or low utilization rate can increase a consumer’s score, while a credit card with a history of missed payments could decrease it. Collections, whether medical or nonmedical, are different in that they should provide a
consistently negative signal. This unidirectional effect makes analyzing the predictive value of collections easier than it is for other types of credit.

Credit scores provide an ordinal ranking of consumers in terms of the credit risk they pose. By themselves, credit scores do not predict the (cardinal) delinquency rate. Delinquency rates are affected by macroeconomic conditions that change over time and are not factored into scoring models. As a result, the delinquency rate associated with any score level changes over time. Evaluating the predictiveness of credit scores, therefore, involves testing whether the credit scores rank order consumers according to credit risk. The downward relationship between credit scores and the two measures of performance, shown in Figure 1 on the facing page, suggest that the credit scores effectively rank order consumers overall during the performance period.

To illustrate our empirical approach, consider two groups of identical consumers, $M$ and $N$. All of the consumers in both groups have identical credit records and no collections. As a result, all of the consumers have the same credit scores, $\tilde{s}$, and we would expect both groups to exhibit similar delinquency rates.

Now assume that we add a medical collection to the credit records of consumers in group $M$ and a nonmedical collection to the records for group $N$. In a credit scoring model that penalizes medical and nonmedical collections equally, the credit scores for both groups would decrease by the same amount, say $\tau$, to $\tilde{s} - \tau$.

If the scoring model’s treatment of collections correctly reflects the information these accounts provide about creditworthiness, then both groups should exhibit the same delinquency rates. Moreover, we would expect delinquency rates for both groups to be consistent with those of consumers without collections who also have scores of $\tilde{s} - \tau$ (and higher than those of other consumers with scores of $\tilde{s}$).

In contrast, if medical collections are completely uninformative about a consumer’s credit risk, but the credit scoring model penalizes them anyway, then we would expect these two groups to exhibit different delinquency rates. Consumers in group $M$ would continue to exhibit delinquency rates consistent with other consumers with scores of $\tilde{s}$. In contrast, consumers in group $N$ should exhibit higher delinquency rates than the consumers in group $M$, despite the fact that both groups have identical scores. We might also expect delinquency rates for consumers in group $N$ to be higher than those of consumers without collections who also have scores equal to $\tilde{s} - \tau$, since the equal score penalty likely understates the signal provided by nonmedical collections.

In between these two extremes is a middle ground in which medical collections are informative about a consumer’s creditworthiness, but less so than nonmedical collections. In this case, we would expect consumers in both groups to have higher

---

9 Some, such as Bellotti and Crook (2009) and Crook and Bellotti (2010), have proposed incorporating macroeconomic conditions into scoring models, though these approaches have not yet been widely adopted in the United States.
delinquency rates than other consumers with scores of $\tilde{s}$. However, we would expect higher delinquency rates for group $N$ than for group $M$.

This illustrative example is simplified in that it only considers consumers who have one medical collection, one nonmedical collection or neither. Consumers do not have both or multiples of either. Yet, in practice, most consumers with collections tend to have both types and more than one of each. Nevertheless, the framework just described can be easily extended to cover this situation.

To handle these cases, we focus on comparing consumers who have the same credit scores and the same number of total collections (medical plus nonmedical), but different mixes of medical and nonmedical collections. Because these consumers have the same credit scores (the equivalent of $\tilde{s} - \tau$ in the above example), we would expect both consumers to have been similarly penalized for their collections and, consequently, the comparable characteristics for the noncollections portion of their credit records. If medical collections are less informative than nonmedical collections, delinquency rates should be lower for consumers with more medical collections.

We investigate this possibility by estimating logit models for our two performance measures. Suppressing individual subscripts, we use $y^k$ to denote the value of performance measure $k$ (where $k \in \{\text{any, new}\}$) for consumer $i$. We assume the existence of an unobserved latent variable, $y^{k*}$, which determines whether a consumer becomes delinquent. We observe delinquency $y^k = 1$ if $y^{k*} > 0$, and $y^k = 0$ otherwise. The equation for this latent variable is specified as

$$y^{k*} = f(s) + g(m, n) + \epsilon,$$  

(3.1)

where $s$ is consumer $i$’s credit score, $m$ and $n$ are the numbers of medical and nonmedical collections on the credit record of consumer $i$, respectively, and $\epsilon$ is an independent and identically distributed error term from a type I extreme value distribution. We estimate the function $f(x)$ using a cubic spline with knots at each quintile, which provides a flexible functional form that captures the relationship between credit scores. As shown in Figure 1 on page 80, the cubic spline representation (shown by the red line) fits the patterns observed in the data very well. Finally, the function $g(m, n)$ is defined as

$$g(m, n) = \sum_{a=0}^{6} \sum_{b=0}^{6-a} I(a = m)I(b = n)\beta_{m,n} + I(m + n > 6)\zeta,$$  

(3.2)

where $I(c)$ is an indicator function that takes on a value of 1 if the condition $c$ is true (or 0 otherwise), and $\beta_{m,n}$ and $\zeta$ are coefficients to be estimated. To identify the coefficients, one category has to be excluded, so we normalize $\beta_{0,0} = 0$. Because samples sizes are decreasing in $t$, and because the samples are further diluted by an
increasing number of possible combinations of \( m \) and \( n \) for each \( t \), we focus our analysis on consumers who have six or fewer collections.

While collections should provide a negative signal about a consumer’s creditworthiness, this information is already incorporated in the credit scores that we observe. If the score penalty imposed on consumers with \( m \) medical and \( n \) nonmedical collections accurately reflects the signal that the collections provide, then we would expect \( \beta_{m,n} = 0 \). When \( \beta_{m,n} < 0 \), these consumers have lower delinquency rates than consumers with the same credit scores but no collections. We refer to this as “overperformance”. When \( \beta_{m,n} > 0 \) we say consumers “underperformed” their scores, meaning that they had higher delinquency rates than other consumers with the same scores.

Our analysis focuses on comparing consumers with the same number of total collections. If the predictive value of medical collections is comparable with nonmedical collections, then performance should be unrelated to the mix of the two. That is, for a given number of total collections, \( t = m + n \), we would expect \( \beta_{0,t} = \beta_{1,t-1} = \cdots = \beta_{t,0} \). Moreover, we would expect this relationship to hold for all six levels of \( t \) included in our analysis. We use this as our null hypothesis and test whether this equality holds jointly for all levels of \( t \) and for each \( t \) individually.

If the null hypothesis is rejected, this suggests that medical and nonmedical collections are not equally informative. By itself, however, it is insufficient to establish that medical collections are less predictive. If they were, we would expect the estimated coefficients to be monotonically decreasing in the number of medical collections; that is, \( \beta_{0,t} > \beta_{1,t-1} > \cdots > \beta_{t,0} \) for all \( t \).

To make comparing the estimated coefficients easier, we present them graphically. The coefficients from the estimation using the any account performance measure as the dependent variable are provided in Figure 2 on the next page. Coefficients are displayed with a 95% confidence interval and are grouped by the total number of collections. For comparison, we also show, in gray, the coefficients from an estimation that uses a single effect for each value of \( t \), which imposes the null hypothesis.

Using a likelihood ratio test, we are able to reject the null hypothesis that medical and nonmedical collections are equally informative.\(^{10}\) Likelihood ratio tests also reject the individual null hypotheses at the 0.1% level for each \( t \).

The estimated coefficients, however, do not have the strict monotonic relationship that we would expect if medical collections are less informative than nonmedical collections. Instead, for \( t \in \{2, 3, 4, 5\} \), consumers with one or two medical collections and multiple nonmedical collections have higher delinquency rates than consumers with the same number of total collections, but all of one type or the other. This suggests that medical and nonmedical collections may provide complementary

\(^{10}\) The chi-squared statistic for this test is 272.63, with twenty-one degrees of freedom. The reported \( p \)-value is \( 2.2 \times 10^{-16} \).
FIGURE 2 Coefficient estimates for medical and nonmedical collections ($\beta_{m,n}$) using any account performance measure.

information about a consumer’s creditworthiness. If such a complementarity exists, then not differentiating between these collections in the scoring model would miss this information. The result would be the pattern observed in the data, which would not necessarily be inconsistent with medical collections being over-penalized.

Instead, if we compare the performance of consumers who only have medical collections with the performance of consumers with only nonmedical collections but
the same \( t \), consumers with medical collections have significantly lower delinquency rates. This suggests that medical collections are likely less informative and, hence, they are being over-penalized by the scoring model.

The results of our logit estimations using the new account performance measure are shown in Figure 3. Unlike the results using the any account performance measure, these results indicate that consumers with collections generally underperformed their
scores; that is, they had higher delinquency rates than other consumers with the same scores and no collections. This is shown by the gray areas, which provide the estimates from an estimation that imposes the null hypothesis.

This underperformance appears to derive entirely from underperformance by consumers with nonmedical collections as indicated by three facts. First, a likelihood ratio test again rejects the null hypothesis that medical and nonmedical collections are equally informative. Similarly, the individual null hypotheses are also rejected at the 0.1% level for each \( t \). This indicates that there were statistically significant differences in the predictive value of medical and nonmedical collections. Second, for each \( t \), the estimated coefficients exhibit the monotonic pattern that would be expected if medical collections were less informative than nonmedical collections. Third, consumers with only medical collections overperformed their scores by a statistically significant amount for all \( t \).

These three facts suggest strongly that medical collections are less informative than nonmedical collections about consumer creditworthiness. Credit scoring models that treat both types the same, by imposing equal penalties, are likely over-penalizing consumers for their medical collections and reducing the predictiveness of the scoring models.

While medical collections appear less informative than nonmedical collections, they are not necessarily uninformative about consumer creditworthiness. As stated at the outset, if medical collections are uninformative, then the delinquency rates of consumers with medical collections should be comparable with those of consumers whose scores are higher by the amount of the score penalty (\( \tau \) in our earlier example). Since we do not know what the scores of these consumers would have been had they not had collections, we cannot accurately identify either \( \tau \) or \( s \).

A rough approximation, however, can be made based on estimates provided in news reports (Johnson 2012). These results suggest that the score penalty for someone with a 680 credit score will be between forty-five and sixty-five points. Using the midpoint of this range as our estimate, this implies that someone with a 625 score likely was penalized fifty-five points for having a collection.

To approximate the amount of over-penalization implied by our estimates, we calculate the fitted value for someone with a 625 score and a single medical collection as \( \hat{f}(625) + \hat{\beta}_{1.0} \). We then solve for the score level, \( s' \), at which consumers without

---

11 The chi-squared statistic from this test equals 847.36 with twenty-one degrees of freedom. The reported \( p \)-value was \( 2.2 \times 10^{-16} \).
12 The same source also reports that someone with a 780 score would be penalized between 105 and 125 points. Since most of the consumers in our sample have much lower credit scores than this would imply, we focus on the estimated score penalties for someone with a 680 score.
collections have the same delinquency rate. This could be interpreted as a “correct” score level for the level of performance we observed. We then use \( s' = 625 \) as an estimate of the amount of over-penalization.

Our estimates of over-penalization work out to 11.2 points for the any account performance measure and 8.1 points for the new account performance measure. These estimates are both well below the published score penalty for having a collection account. This suggests that while medical collections appear to be over-penalized, they likely contain significant predictive value for consumer creditworthiness.

4 PAID AND UNPAID MEDICAL COLLECTIONS

To test the possibility that paid medical collections are less informative about creditworthiness than medical collections that remain unpaid, we use an empirical approach similar to the one in Section 3. Specifically, we compare the performance of consumers with the same credit scores and the same number of medical collections, but with different mixes of paid and unpaid medical collections. If paid medical collections are less informative than unpaid medical collections, we would expect to see delinquency rates that decrease in the proportion of paid medical collections.

Using \( p \) and \( u \) to denote the number of paid and unpaid medical collections on consumer \( i \)’s credit record, respectively, we specify the equation for the latent performance variable as

\[
y^{k*} = f(s) + h(p, u) + \epsilon, \tag{4.1}
\]

where the terms are as defined in (3.1) and \( h(p, u) \) is given by

\[
h(p, u) \equiv \sum_{a=0}^{6} \sum_{b=0}^{6-a} I(a = p)I(b = u)\gamma_{p,u} + I(p + u > 6)\eta. \tag{4.2}
\]

As with the previous analysis, if paid and unpaid medical collections are equally informative about a consumer’s creditworthiness, then we would expect for a given number of medical collections, \( m \), that \( \gamma_{0,m} = \gamma_{1,m-1} = \cdots = \gamma_{m,0} \). As our null hypothesis, we test whether this equality holds for \( m \in \{1, \ldots, 6\} \). We also test whether this equality holds for each \( m \) individually. In contrast, if paid collections are less informative than unpaid collections, and are thus being over-penalized by scoring models that treat paid and unpaid collections equally, then we would expect a monotonic relationship among the coefficients, such that \( \gamma_{0,m} < \gamma_{1,m-1} < \cdots < \gamma_{m,0} \).

\[\text{13 In other words, we find the value of } s' \text{ that solves } \hat{f}(s') = \hat{f}(625) + \hat{\beta}_{1,0}.\]
The estimated $\gamma_{p,u}$ coefficients are shown, along with a 95% confidence interval, from the estimation using the any account performance measure in Figure 4 and from the estimation using the new account performance measure in Figure 5 on the facing page.
FIGURE 5 Coefficient estimates for unpaid and paid collections ($\gamma_{p,u}$) using new account performance measure.

<table>
<thead>
<tr>
<th>One account</th>
<th>Two accounts</th>
<th>Three accounts</th>
<th>Four accounts</th>
<th>Five accounts</th>
<th>Six accounts</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{0,1}$</td>
<td>$\gamma_{0,2}$</td>
<td>$\gamma_{0,3}$</td>
<td>$\gamma_{0,4}$</td>
<td>$\gamma_{0,5}$</td>
<td>$\gamma_{0,6}$</td>
</tr>
<tr>
<td>$\gamma_{1,0}$</td>
<td>$\gamma_{1,1}$</td>
<td>$\gamma_{1,2}$</td>
<td>$\gamma_{1,3}$</td>
<td>$\gamma_{1,4}$</td>
<td>$\gamma_{1,5}$</td>
</tr>
<tr>
<td>$\gamma_{2,0}$</td>
<td>$\gamma_{2,1}$</td>
<td>$\gamma_{2,2}$</td>
<td>$\gamma_{2,3}$</td>
<td>$\gamma_{2,4}$</td>
<td>$\gamma_{2,5}$</td>
</tr>
<tr>
<td>$\gamma_{3,0}$</td>
<td>$\gamma_{3,1}$</td>
<td>$\gamma_{3,2}$</td>
<td>$\gamma_{3,3}$</td>
<td>$\gamma_{3,4}$</td>
<td>$\gamma_{3,5}$</td>
</tr>
<tr>
<td>$\gamma_{4,0}$</td>
<td>$\gamma_{4,1}$</td>
<td>$\gamma_{4,2}$</td>
<td>$\gamma_{4,3}$</td>
<td>$\gamma_{4,4}$</td>
<td>$\gamma_{4,5}$</td>
</tr>
<tr>
<td>$\gamma_{5,0}$</td>
<td>$\gamma_{5,1}$</td>
<td>$\gamma_{5,2}$</td>
<td>$\gamma_{5,3}$</td>
<td>$\gamma_{5,4}$</td>
<td>$\gamma_{5,5}$</td>
</tr>
<tr>
<td>$\gamma_{6,0}$</td>
<td>$\gamma_{6,1}$</td>
<td>$\gamma_{6,2}$</td>
<td>$\gamma_{6,3}$</td>
<td>$\gamma_{6,4}$</td>
<td>$\gamma_{6,5}$</td>
</tr>
</tbody>
</table>

Likelihood ratio tests allow us to reject the null hypotheses for both the any account and new account performance measures at the 0.1% level.\(^{14}\) We are also able to reject

\(^{14}\) The chi-squared statistic for the test involving the any account performance measure is 369, and for the new account performance measure it is 370. Both statistics have twenty-one degrees of freedom and reported \(p\)-values of \(2.2 \times 10^{-16}\).
the individual null hypotheses for both performance measures. The hypotheses are rejected at the 0.1% level in all cases, except $m = 4$ using any account performance ($p$-value = 0.02) and $m = 5$ using new account performance ($p$-value = 0.002).

The estimated coefficients for both the any account and new account performance measures suggest that consumers with paid medical collections have lower delinquency rates during the performance period than consumers with the same number of medical collections but fewer of them paid. The coefficients do not exhibit the strict monotonic relationship that was expected, but this may reflect the small sample sizes available for these estimations.

Nevertheless, the coefficients for people who had all paid medical collections, $\gamma_{m,0}$, are less than the coefficients for consumers with all unpaid medical collections, $\gamma_{0,m}$, for all six levels of $m$. Moreover, for the any account performance measure, we can reject the hypothesis that $\gamma_{0,m} = \gamma_{m,0}$ at the 0.1% level for $m \in \{1, 2, 3\}$ and at the 5% level for $m \in \{4, 5, 6\}$. For the new account performance measure, we can reject this hypothesis at the 0.1% level for $m \leq 4$ and at the 5% level for $m = 6$. Only the hypothesis for $m = 5$ for the new account performance measure appears statistically insignificant ($p$-value = 0.12).

To further test whether there is a relationship between the mix of paid and unpaid medical collections (for a fixed number of total collections), we also estimate logit

### TABLE 2 Coefficient estimates for paid and unpaid collections with linear effects.

<table>
<thead>
<tr>
<th>Number of medical collections</th>
<th>Any account</th>
<th>New account</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept $\alpha_{p+u}$</td>
<td>Slope $\beta_{p+u}^L$</td>
</tr>
<tr>
<td>1</td>
<td>$-0.020^{**}$</td>
<td>$-0.172^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>2</td>
<td>$-0.025^{**}$</td>
<td>$-0.094^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.009)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>3</td>
<td>$-0.044^{***}$</td>
<td>$-0.071^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>4</td>
<td>$-0.037^{*}$</td>
<td>$-0.033^{**}$</td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>5</td>
<td>$-0.049^{**}$</td>
<td>$-0.056^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.013)</td>
</tr>
<tr>
<td>6</td>
<td>$-0.043$</td>
<td>$-0.058^{***}$</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

Statistical significance shown by *,**, and *** for the 5%, 1% and 0.1% levels, respectively.
models that impose a linear relationship. Specifically, we replace $h(p, u)$ in (4.1) with

$$h'(p, u) = \sum_{a=1}^{6} I(p + u = a)(\alpha_{p+u} + \gamma_{p+u}^L p) + I(p + u > 6)\eta^M. \quad (4.3)$$

Instead of the individual effects for each combination of $p$ and $u$ (for $p + u < 7$) that are shown in Figure 4 on page 88 and Figure 5 on page 89, this formulation estimates a linear function with its own intercept, $\alpha_{p+u}$, and slope, $\gamma_{p+u}^L$, for each level of $m = p + u$. The results of these estimations for our two performance measures are shown in Table 2 on the facing page.

The linear effects indicate that there is a statistically significant relationship between the mix of paid and unpaid medical collections. Each of the slope coefficients, in both estimations, are statistically significant at the 1% level (in all but one case, they are also significant at the 0.1% level). Using a likelihood ratio test, we are unable to reject the hypothesis that the coefficients in Figure 4 on page 88 and Figure 5 on page 89 are linear for either performance measure at the 1% level.15 This is consistent with the existence of a monotonic relationship between paid medical collections and overperformance.

While the evidence in favor of a monotonic relationship across the coefficients in our estimations comparing paid and unpaid medical collections is somewhat weaker than it is in our estimations comparing medical and nonmedical collections (which we believe is primarily a result of the former being based on fewer observations), the magnitude of the effects here is notably larger than those observed in the results from Section 3. This suggests that paid medical collections are being over-penalized by a much larger amount than medical collections in general.

To gauge the extent of over-penalization relative to the reported score penalty from having a collection, we use the same approach from the previous section. Again, we solve $\hat{f}(s') = \hat{f}(625) + \hat{\beta}_{1,0}$ for $s'$ and use $s' - 625$ as an estimate of the amount of over-penalization. The results of this calculation suggest over-penalization of fifteen points using the any account performance measure and twenty-two points using the new account performance measure. While the over-penalization of paid medical collections appears larger than for medical collections generally, these estimates remain well below the reported score penalty of around fifty-five points. This suggests that paid medical collections also retain significant predictive value.

---

15 For the any account measure, the chi-squared test statistic is 20.552 and the $p$-value is 0.1518. For the new account measure, the chi-squared statistic is 26.137 and the $p$-value is 0.03661. Both tests had fifteen degrees of freedom.
FIGURE 6  Credit scores by number and amount of debts in collection.

5 ROBUSTNESS

So far, we have compared the performance of consumers with the same credit scores and the same number of collection accounts. One potential explanation for the differences we find is that, as shown in Table 1 on page 78, medical collections tend to be smaller than nonmedical collections, and paid medical collections tend to be smaller than unpaid. If smaller collections are a less negative signal about creditworthiness than larger collections, then our results may reflect differences in the size of these debts, rather than differences between medical and nonmedical collections.

Our focus on the number of accounts in collection is driven by anecdotal evidence gained from conversations with industry modelers who have told us that, in credit scoring models like the one used in this study, the number of collections is more important than the amount of collections.16 The patterns we observe in the data are consistent with the number of collections being a more important factor in credit scoring models than the size of debts in collection.

For example, Figure 6 shows how the average credit score varies with the number of collection accounts and with the total amount of those collections (in US dollars).

---

16 One aspect of the size of a collection that is very important is whether the collection is above the minimum size threshold, below which the collection account may be ignored by the credit scoring model. In this section, we refer to the number and amount of collections above the US$100 threshold for the credit score used in this study.
Because the distributions of the number and amounts of collections are both highly skewed, the average credit score reported for consumers with twenty-one collection accounts reflects the average across consumers with twenty-one or more accounts. Similarly, the average credit score for people with US$21 000 in collections reflects the average across consumers with at least US$21 000 in collections. As shown, the relationship between credit score and the number of collections is much stronger than the relationship between score and total collection amounts. Indeed, for larger amounts, there appears to be little relationship between credit score and collection amount. In fact, the correlation between credit score and the number of collection accounts (0.51) is almost double the correlation between credit score and the total amount of collections (0.26). This is consistent with the number of accounts being a more important predictive factor in the credit scoring model than the amount of debt in collection.

Nevertheless, to determine if our results are driven by differences in the amount of debt in collection, we modify the equations used earlier to control for collection amounts. In place of (3.1), we specify our equation for the latent delinquency variable as

\[ y^k = f(s) + g(m,n) + a_{m+n} \tau_{m+n} + \epsilon; \]  

in place of (4.1), we specify an equation of the form

\[ y^k = f(s) + h(p, u) + a_m \tau_m + \epsilon, \]  

where \( a_{m+n} \) is the total amount of collections accounts (medical plus nonmedical) and \( a_m \) is the total amount of medical collections. All other terms are as described earlier for (3.1) and (4.1), respectively.

The results of these estimations are provided in Table 3 on the next page. There are two relevant takeaways from these results. First, the estimated coefficients on the total amount of collections, \( \tau_{m+n} \) and \( \tau_m \), are small. The largest effect is observed in the estimation of (5.1) using the any account performance measure (column 1), which indicates that a US$1000 increase in the collection amount decreases the log odds of delinquency by 0.007. To put this in context, the expected delinquency rate for someone with a credit score of 625, one nonmedical collection and a total collection amount of US$400 is 22.5%. Increasing the total amount of collections to US$1400 reduces the delinquency rate by 0.1 percentage points to 22.4%. The effects from the other three estimations are smaller in magnitude and sometimes positive. We believe this is consistent with the view that the amount of debts in collection has little marginal value in predicting performance.

---

17 Less than 1% of consumers with collections have twenty-one or more accounts, and about 4% have amounts of US$21 000 or more.
TABLE 3  Coefficient estimates with controls for collection amounts.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>Any account</th>
<th>New account</th>
<th>Coefficient</th>
<th>Any account</th>
<th>New account</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amount, $\tau_m+n$</td>
<td>-0.007***</td>
<td>0.002*</td>
<td>Amount, $\tau_m$</td>
<td>-0.001*</td>
<td>0.006***</td>
</tr>
<tr>
<td>One account</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0,1}$</td>
<td>-0.102***</td>
<td>0.151***</td>
<td>$\gamma_{0,1}$</td>
<td>-0.019**</td>
<td>0.051***</td>
</tr>
<tr>
<td>$\beta_{1,0}$</td>
<td>-0.147***</td>
<td>-0.125***</td>
<td>$\gamma_{1,0}$</td>
<td>-0.191***</td>
<td>-0.372***</td>
</tr>
<tr>
<td>Two accounts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0,2}$</td>
<td>-0.108***</td>
<td>0.242***</td>
<td>$\gamma_{0,2}$</td>
<td>-0.027**</td>
<td>0.031</td>
</tr>
<tr>
<td>$\beta_{1,1}$</td>
<td>-0.088***</td>
<td>0.120***</td>
<td>$\gamma_{1,1}$</td>
<td>-0.089***</td>
<td>-0.262***</td>
</tr>
<tr>
<td>$\beta_{2,0}$</td>
<td>-0.184***</td>
<td>-0.131***</td>
<td>$\gamma_{2,0}$</td>
<td>-0.235***</td>
<td>-0.347***</td>
</tr>
<tr>
<td>Three accounts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0,3}$</td>
<td>-0.098***</td>
<td>0.398***</td>
<td>$\gamma_{0,3}$</td>
<td>-0.038**</td>
<td>-0.002</td>
</tr>
<tr>
<td>$\beta_{1,2}$</td>
<td>-0.086***</td>
<td>0.178***</td>
<td>$\gamma_{1,2}$</td>
<td>-0.142***</td>
<td>-0.310***</td>
</tr>
<tr>
<td>$\beta_{2,1}$</td>
<td>-0.140***</td>
<td>0.083*</td>
<td>$\gamma_{2,1}$</td>
<td>-0.178***</td>
<td>-0.266**</td>
</tr>
<tr>
<td>$\beta_{3,0}$</td>
<td>-0.215***</td>
<td>-0.206***</td>
<td>$\gamma_{3,0}$</td>
<td>-0.239***</td>
<td>-0.491***</td>
</tr>
<tr>
<td>Four accounts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0,4}$</td>
<td>-0.071***</td>
<td>0.532***</td>
<td>$\gamma_{0,4}$</td>
<td>-0.031</td>
<td>0.034</td>
</tr>
<tr>
<td>$\beta_{1,3}$</td>
<td>-0.045*</td>
<td>0.343***</td>
<td>$\gamma_{1,3}$</td>
<td>-0.109**</td>
<td>-0.105</td>
</tr>
<tr>
<td>$\beta_{2,2}$</td>
<td>-0.050*</td>
<td>0.136**</td>
<td>$\gamma_{2,2}$</td>
<td>-0.035</td>
<td>-0.110</td>
</tr>
<tr>
<td>$\beta_{3,1}$</td>
<td>-0.168***</td>
<td>-0.089</td>
<td>$\gamma_{3,1}$</td>
<td>-0.184**</td>
<td>-0.577***</td>
</tr>
<tr>
<td>$\beta_{4,0}$</td>
<td>-0.215***</td>
<td>-0.269***</td>
<td>$\gamma_{4,0}$</td>
<td>-0.153*</td>
<td>-0.499***</td>
</tr>
<tr>
<td>Five accounts</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta_{0,5}$</td>
<td>-0.029</td>
<td>0.688***</td>
<td>$\gamma_{0,5}$</td>
<td>-0.043*</td>
<td>0.031</td>
</tr>
<tr>
<td>$\beta_{1,4}$</td>
<td>0.009</td>
<td>0.499***</td>
<td>$\gamma_{1,4}$</td>
<td>-0.081</td>
<td>-0.125</td>
</tr>
<tr>
<td>$\beta_{2,3}$</td>
<td>0.000</td>
<td>0.322***</td>
<td>$\gamma_{2,3}$</td>
<td>-0.227***</td>
<td>-0.394*</td>
</tr>
<tr>
<td>$\beta_{3,2}$</td>
<td>-0.125***</td>
<td>0.161**</td>
<td>$\gamma_{3,2}$</td>
<td>-0.204*</td>
<td>-0.265</td>
</tr>
<tr>
<td>$\beta_{4,1}$</td>
<td>-0.169***</td>
<td>0.034</td>
<td>$\gamma_{4,1}$</td>
<td>-0.428***</td>
<td>-0.873**</td>
</tr>
<tr>
<td>$\beta_{5,0}$</td>
<td>-0.232***</td>
<td>-0.223*</td>
<td>$\gamma_{5,0}$</td>
<td>-0.215*</td>
<td>-0.251</td>
</tr>
</tbody>
</table>

Statistical significance indicated by *, ** and *** for the 5%, 1% and 0.1% levels, respectively.

The second takeaway is that, controlling for the total amount of debt in collections does not alter the main results of our analysis. Consumers with more medical collections have lower delinquency rates than other consumers with the same credit scores, the same total number of collections and the same amount of debt in collection. Similarly, paid medical collection accounts also appear to be associated with
lower delinquency rates than unpaid medical collections. Our results, therefore, do not appear to be notably affected by differences in the sizes of debts in collection.\textsuperscript{18}

\section*{6 CONCLUSIONS}

This paper examines the predictive value of medical collections in assessing consumer creditworthiness with credit scoring models. We find two main results. First, we find that medical collections are less informative than nonmedical collections about a consumer’s likelihood of delinquency. Second, we find that medical collections that have been paid in full are less predictive than medical collections that remain unpaid.

These results suggest that consumers with medical collections are being overly penalized for these accounts; that is, their credit scores appear to be between eight and eleven points lower than is justified by their subsequent credit performance. Similarly, we find that paid medical collections are being over-penalized by fifteen to twenty-two points. This over-penalization is likely material for many consumers with collections, whose credit scores tend to be below average. While the relationship between credit scores and loan pricing differs across lenders and time periods, data published by FICO provides some insight into the relationship between credit scores and interest rates.\textsuperscript{19}

The FICO data indicates that interest rates on thirty-year fixed-rate mortgages for consumers with scores of between 620 and 639 are around 5.3\%. Between 39 and 53\% of consumers with a medical collection and scores in this range would see their scores increase into the next-highest score range (640–659) if their scores were no longer penalized by eight to eleven points. The percentages for similar consumers with a paid medical collection would be even higher. By moving up this one credit tier, the interest rate charged to a consumer would decline to 4.8\%. On a US$200 000 mortgage, this would save such a consumer over US$100 per month, or about US$24 000 over the entire thirty-year term of the loan.

This over-penalization suggests that credit scoring models can be made more predictive by no longer treating all collection accounts equally. Allowing medical and nonmedical collections, as well as paid and unpaid medical collections, to make different contributions in calculating a credit score will generate more accurate predictions about the likely default rates of consumers with medical collections. This will enhance model predictiveness and increase the credit scores of consumers with medical collections, thereby expanding their access to credit.

\textsuperscript{18}In these estimations, collection amounts enter as linear functions. The results shown here are robust to using the more flexible cubic spline functional form used for the credit score.

\textsuperscript{19}Interest rates are published by credit score band in FICO’s Loan Savings Calculator, available online at www.myfico.com/myfico/creditcentral/loanrates.aspx. The rates here were published on July 7, 2015.
While this study has focused entirely on differences between paid and unpaid collections for medical debt, the large effect that we find may represent a broader tendency of paid collections to be a less negative signal about creditworthiness than unpaid collections. Credit scoring models might also be improved by differentiating between paid and unpaid collections more broadly, though establishing a similar relationship for nonmedical collections is left as a topic for future research.

Because of data limitations, we are unable to establish the underlying reasons why these differences exist. At the beginning of this paper, we discussed how issues related to medical billing could result in consumers being unaware of their outstanding medical bills, or in consumers viewing the debts as illegitimate (if, for example, the consumer believes that the debt should be covered by their insurance). While our results certainly do not prove that these factors are causing the difference in predictiveness of medical collections, they are consistent with such theories. As such, the differences may reflect issues related to medical billing practices. If true, efforts to improve medical billing and to improve how unpaid medical debts are reported to credit reporting agencies may reduce some of the differences between medical and nonmedical collections. Like allowing medical and nonmedical collections to have different effects on scores, such efforts could enhance the predictiveness of credit scoring models and better align the credit scores of consumers with medical collections with their subsequent performance on credit obligations.

DECLARATION OF INTEREST

The views expressed are those of the authors and do not necessarily reflect those of the Consumer Financial Protection Bureau or the United States.

ACKNOWLEDGEMENTS

We thank Ron Borzekowski, Brian Bucks, Jonathan Crook, Debb Gordon, Dan Grodzicki, Corey Stone and participants of the Credit Scoring and Credit Control XIII conference at the University of Edinburgh for helpful comments.

REFERENCES


The Journal of Credit Risk
INDEX OF PAPERS

To order back issues, please contact:
Head Office: Incisive Media (c/o CDS Global), Tower House, Sovereign Park, Market Harborough, Leicestershire, LE16 9EF, UK; Tel: +44 (0)870 787 6822; fax: +44 (0)870 607 0106; e-mail: incisivehv@subscription.co.uk or subscribe online at www.risk.net/static/risk-journals-subscribe

Volume 11/Number 1

RESEARCH PAPERS

Sovereign risk and the pricing of corporate credit default swaps
Matthias Haerri, Stefan Morkoetter and Simone Westerfeld

Forecasting credit card portfolio losses in the Great Recession: a study in model risk
José J. Canals-Cerdá and Sougata Kerr

How banks’ capital ratio and size affect the stability of the banking system: a simulation-based study
Mitja Steinbacher and Matjaz Steinbacher

The relationship between counterparty default and interest rate volatility and its impact on the credit risk of interest rate derivatives
Geoffrey R. Harris, Tao L. Wu and Jiarui Yang

Volume 11/Number 2

RESEARCH PAPERS

Systematic risk and yield premiums in the bond market
Liang Fu, Austin Murphy and Terry Benzschawel

Default predictors in credit scoring: evidence from France’s retail banking institution
Ha-Thu Nguyen

TECHNICAL REPORT

The robustness of estimators in structural credit loss distributions
Enrique Batiz-Zuk, George Christodoulakis and Ser-Huang Poon

To subscribe to a Risk Journal visit Risk.net/subscribe or email corpsubs@risk.net
Volume 11/Number 3

RESEARCH PAPERS

Hermite approximations in credit portfolio modeling with probability of default–loss given default correlation 1–20
Anthony Owen, James Bryers and François Buet-Golfouse

Time series models for credit default swap premiums 21–44
Márton Eifert

Counting processes for retail default modeling 45–72
Nicholas M. Kiefer and C. Erik Larson

Credit risk: taking fluctuating asset correlations into account 73–90
Thilo A. Schmitt, Rudi Schäfer and Thomas Guhr

Volume 11/Number 4

RESEARCH PAPERS

An analytical value-at-risk approach for a credit portfolio with liquidity horizon and portfolio rebalancing 1–28
Haohan Huang, Eugene Wang, Huaxiong Huang and Yong Wang

Loss distributions: computational efficiency in an extended framework 29–42
Daniel H. Stahl

Default risk of money-market fund portfolios 43–71
Matulya Bansal

Are all collections equal? The case of medical debt 73–97
Kenneth P. Brevoort and Michelle Kambara

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