A Sound Modelling and Backtesting Framework for Forecasting Initial Margin Requirements *†

FABRIZIO ANFUSO‡,
DANIEL AZIZ§, PAUL GILTINAN¶,
KLEARCHOS LOUKOPOULOS‖

January 16, 2016

Abstract

The introduction by regulators of mandatory margining for bilateral OTCs is going to have a major impact on the derivatives market, particularly in light of the additional funding costs and liquidity requirements that large financial institutions will face.

Fabrizio Anfuso, ***Daniel Aziz***, ***Paul Giltinan*** and ***Klearchos Loukopoulos*** propose in the following a simple and consistent framework, equally applicable to non-cleared and cleared portfolios, to develop and backtest forecasting models for Initial Margin.

KEY WORDS: INITIAL MARGIN, BCBS-IOSCO, CCP, OTC, CLEARING, COUNTERPARTY CREDIT RISK, XVA, LIQUIDITY, FUNDING COSTS

1 Introduction

Since the publication of the new BCBS-IOSCO guidance on mandatory margining for non-cleared OTCs [1], there has been a growing interest in the industry for the development of Dynamic Initial Margin models (DIM), see e.g. [2], [3].

The business case is at least two-fold: *i)* The BCBS-IOSCO Initial Margin Requirements (IMR) are very severe and will substantially affect funding

*The views expressed are those of the authors only, no other representation should be attributed

†The authors gratefully acknowledge: *i)* Quaternion Risk Management Ltd. for providing us the QRE software package, used for generating some of the results shown in the present paper. *ii)* Andrew Hudson for the careful review of the manuscript and the highly valuable feedback.

‡Credit Suisse Securities Limited; `dr.fabrizio.anfuso@gmail.com`.

§Credit Suisse Securities Limited; `dnaziz@gmail.com`.

¶Quaternion Risk Management Ltd; `paulgiltinan@outlook.com`.

‖Credit Suisse Securities Limited; `kloukopoulos@googlemail.com`.

costs, XVA and capital. *ii)* The BSBC-IOSCO IMR (B-IMR) set a clear incentive towards clearing; with extensive margining, in the form of Variation and Initial Margins (VM, IM), being the main component of the CCP risk management model (for an extended discussion on the cost of clearing, see e.g. [4]).

Therefore, both for cleared derivatives and non-cleared OTCs in scope for the BCBS-IOSCO IM regulation, current and expected IMR affects significantly the profitability and the risk profile of a given trade.

In the present work, we consider B-IMR as a case study and we show how to include a suitable parsimonious DIM model in the exposure calculation. We propose an end-to-end framework defining also a methodology to backtest the model performance.

The paper is organised as follows: in Sec. 2 the DIM model for B-IMR is presented; in Sec. 3 methodologies for two distinct levels of backtesting analysis are discussed; finally, in Sec. 4, we draw conclusions.

2 How to construct a DIM model?

A DIM model can be used for various purposes. In the computation of Counterparty Credit Risk (CCR) capital exposure or CVA, the DIM model should forecast on a path by path basis the amounts of posted and received IMs at any revaluation point. For this specific application, the key ability of the model is to associate realistic IMR to a simulated market scenario based on a mapping that makes use of a set of characteristics of the path.

The DIM model is *a priori* agnostic to the underlying Risk Factors Evolution (RFE) models used for the generation of the exposure paths (as we will see, dependencies may arise if e.g. the DIM is computed based on the same paths generated for the exposure).

The story becomes different if the goal is to forecast the IMR Distribution (IMRD) at future horizons, either in the real world P or in the market implied Q measures. In this context, the key feature of the model is to associate the right probability weight to a given IMR scenario; hence, the forecasted IMRD becomes also a measure of the accuracy of the RFE models (that ultimately determine the likelihood of the different market scenarios). The distinction between the two cases will become clearer in Sec. 3, where we discuss how to assess model performance.

In the remainder of the paper, we consider BCBS-IOSCO IM as a case study. For B-IMR, the current industry proposal is the ISDA Standard Initial Margin Model (SIMM) [5], a static aggregation methodology to compute IMR based on first order delta-vega trade sensitivities. The exact replication of SIMM in a capital exposure / XVA Monte Carlo framework requires in-simulation portfolio sensitivities to a large set of underlying risk factors, which is unfeasible for most production implementations.

Since the exposure simulation provides portfolio MtM values on the default (time t) and closeout (time $t+\text{MPOR}$, being MPOR the Margin Period Of Risk) grids, Andersen, Phyktin and Sokol [3, 6] have proposed to use this information to infer path-wise the size of of any percentile of the local $\Delta MtM(t, t+\text{MPOR}, \text{path}_i)$ distribution[1] based on a regression on the portfolio $MtM(t)$. This methodology can be further empowered by adding more descriptive variables to the regression, e.g. the values at the default time t of selected risk factors of the portfolio.

With less sophistication, one can drop the path dependence and determine a single $\text{IM}(t)$ for all paths as a percentile of the simulated $\Delta MtM(t, t+\text{MPOR})$ distribution constructed for a given time t by sampling the ΔMtMs over the MPOR from all the simulated paths in the exposure run.

For both approaches, two features are anyway desirable:

1. The DIM model should consume the same paths computed for the exposure so as not to add CPU time burdens.

2. The output of the DIM model (e.g. in a matrix form $\text{IM}(\text{path}_i, t_k)$ being the t_k's the revaluation points on the default grid) should reconcile with the known B-IMR value for $t_k = 0$, i.e. $\text{IM}(\text{path}_i, 0) = \text{IM}_{\text{SIMM}}(0)$ $\forall i$.

Before discussing the modelling further, it is worth to recap few key facts about B-IMR and SIMM [5]:

(a) The MPOR for the IM calculation of a daily margined counterparty is equal to $10d$. This is generally different from the capital exposure calculation where $\text{MPOR} = 20d$ if the number of trades in the portfolio exceed 5000 [7] (the bilateral portfolios of the names in scope for BCBS-IOSCO phase 1 typically fall in this category. The MPOR can even be longer in the case of margin disputes between the counterparties over the last two quarters).

(b) The B-IMR in [1] prescribe to calculate IM segregating trades from different asset classes. This feature is coherently reflected in the SIMM model design[2].

(c) The SIMM methodology consumes trade sensitivities as only input and has a static calibration not sensitive to market volatility (at least until the next model recalibration takes place, i.e. no less than yearly).

These features, together with the requirements 1 and 2 stated above are addressed by our model proposal, as we will see.

[1] The $\Delta MtM(t, t+\text{MPOR}) = MtM(t+\text{MPOR}) - MtM(t)$ distribution is constructed assuming that no cash flows take place in between the default and closeout times.

[2] Two possible levels of bucketing are currently under discussion for SIMM: trade level asset class and risk based [8]. In the present paper, we refer primarily to the former.

For the IM calculation, the starting point is similar to [3, 6], i.e. *i)* to use a regression methodology based on the paths $MtM(t)$ to compute the moments of the local $\Delta MtM(t, t + \mathrm{MPOR}, \mathrm{path}_i)$ distribution and *ii)* to assume that the $\Delta MtM(t, t + \mathrm{MPOR}, \mathrm{path}_i)$ distribution is Normal and is therefore sufficient to determine the drift and volatility to characterise any of its quantiles. Additionally, since the drift is generally immaterial over the MPOR horizon, we do not compute it and set it to 0.

There are multiple regression schemes that can be used to determine the local $\sigma(i, t)$. For the present analysis, we follow the standard American Monte Carlo literature [9] and use a Least Squares Method (LSM) approach with a polynomial basis:

$$\begin{aligned}
\sigma^2(i,t) &= \left\langle (\Delta MtM(i,t))^2 | MtM(i,t) \right\rangle = \sum_0^n a_{\sigma,k} MtM(i,t)^k & (1)\\
\mathrm{IM}^{\mathrm{U}}_{R/P}(i,t) &= \Phi^{-1}(0.99/0.01, \mu = 0, \sigma = \sigma(i,t)), & (2)
\end{aligned}$$

where R/P indicates received and posted respectively and in our implementation the sum in Eq. 1 run up to $n = 2$. We observe that LSM performs well compared to more sophisticated kernel methods (such as e.g. Nadaraya-Watson used in [3, 6]) with the advantage of being parameter free and cheaper from a computational stand point. The local unnormalised posted and received $\mathrm{IM}^{\mathrm{U}}_{R/P}(i, t)$ are calculated analytically in Eqns. 1 and 2 by applying the inverse of the Normal Cumulative Distribution Function $\Phi^{-1}(x, \mu, \sigma)$ for the correspondent quantiles.

As a next step we should account for the $t = 0$ reconciliation and for the mismatch between the SIMM and exposure model calibrations (see respectively items 2, (a) and (c) above). These points can be tackled by scaling $\mathrm{IM}^{\mathrm{U}}_{R/P}(i, t)$ with suitable normalisation functions $\alpha_{R/P}(t)$:

$$\begin{aligned}
\mathrm{IM}_{R/P}(i,t) &= \alpha_{R/P}(t) \times \mathrm{IM}^{\mathrm{U}}_{R/P}(i,t) & (3)\\
\alpha_{R/P}(t) &= (1 - h_{R/P}(t))\sqrt{\frac{10d}{\mathrm{MPOR}}} \cdot \left(1 + (\alpha^0_{R/P} - 1)e^{-\beta_{R/P}(t)t}\right) & (4)\\
\alpha^0_{R/P} &= \sqrt{\frac{\mathrm{MPOR}}{10d}} \cdot \frac{\mathrm{IM}^{\mathrm{SIMM}}_{R/P}(t=0)}{q(0.99/0.01, \Delta MtM(0, \mathrm{MPOR}))}. & (5)
\end{aligned}$$

Where:

i) In Eq. 4, $\beta_{R/P}(t) > 0$ and $h_{R/P}(t) < 1$ with $h_{R/P}(t = 0) = 0$ are four functions to be calibrated (two for received and two for posted IMs). As it will become clearer in Sec. 3, the model calibration generally differs for received and posted DIM models.

ii) In Eqns. 4 and 5, MPOR indicates the MPOR relevant for Basel III exposure. The ratio among MPOR and $10d$ accounts for item (a) and is taken in square root because the underlying RFE models are typically Brownian, at least at short horizons.

iii) In Eq. 5, $\text{IM}^{\text{SIMM}}_{R/P}(t = 0)$ are the $\text{IM}_{R/P}$ computed at $t = 0$ using SIMM, $\Delta MtM(0, \text{MPOR})$ is the distribution of MtM variations over the first MPOR and $q(x, y)$ is a function giving the quantile x for the distribution y.

The normalisation functions $\alpha_{R/P}(t)$ are defined at $t = 0$ such as to reconcile the $\text{IM}_{R/P}(i,t)$ with the starting B-IMR ($\alpha_{R/P}(0) = \alpha^0_{R/P}$). Instead, the functional form of $\alpha_{R/P}(t)$ at $t > 0$ is dictated by what shown in the left panel of Fig. 1: accurate RF models, both in the P and in the Q measure, have either a volatility term-structure or an underlying stochastic volatility process that accounts for the mean-reverting behaviour generally observed from extreme low or high volatility markets to normal market conditions. Since the SIMM calibration is static, the $t = 0$ reconciliation factor is inversely proportional to the current market volatility and not necessarily adequate for the long-term mean level. Hence, $\alpha_{R/P}(t)$ interpolate between the t=0 scaling driven by $\alpha^0_{R/P}$ and the long term scaling driven by $h_{R/P}(t)$.

As we will see, the meaning of $h_{R/P}(t)$ depend on the applications: *i)* For capital and risk models, $h_{R/P}(t)$ are two haircut functions that can be used to reduce the number of backtesting exceptions (see Sec. 3) and ensure that the DIM model is conservatively calibrated. *ii)* For XVA pricing, $h_{R/P}(t)$ can be fine-tuned (together with $\beta_{R/P}(t)$) so as to maximise the accuracy of the forecast based on historical performance.

Notice regarding item (b) above that the $\text{IM}_{R/P}(i,t)$ can be computed on a stand-alone basis for every asset class defined by SIMM (IR/FX, Equity, Qualified and Not Qualified Credit, Commodity) without any additional exposure runs. The $\text{IM}^x_{R/P}(i,t)$ for asset class x is calculated by applying the logic described in Eqns. 1, 2, 3, 4 and 5 to the trades belonging to x (where the scaling functions $\alpha_{R/P}(t)$ are optimally defined and calibrated at asset class level). The total $\text{IM}_{R/P}(i,t)$ is then given by the sum of the $\text{IM}^x_{R/P}(i,t)$'s.

Finally, a comparison between model forecasts and historical realisations for the DIM model defined throughout Eqns. 1, 2, 3, 4 and 5 is shown in the right panel of Fig. 1. As it is evident from the plot, the term structure of $\alpha(t)$ improves the relative accuracy of the DIM model by a significant amount. Notice as well that for this analysis the functions $\beta(t)$ and $h(t)$ have not been specifically calibrated[3]. Accuracy can be further improved

[3] $\beta(t) = 1$ has been chosen based on the heuristics that the range of variation over time for the implied volatility term structures of the FX and IR short rates of the main developed currencies is typically between 50% and 200% of the normal market values and the mean reversion to the long-term mean happens over a time lag of order one year. The haircut function $h(t)$ has been set to 0. See Sec. 3 for a more extended discussion on that.

by an *ad-hoc* calibration targeting a given set of portfolios / asset classes / product types / underlyings.

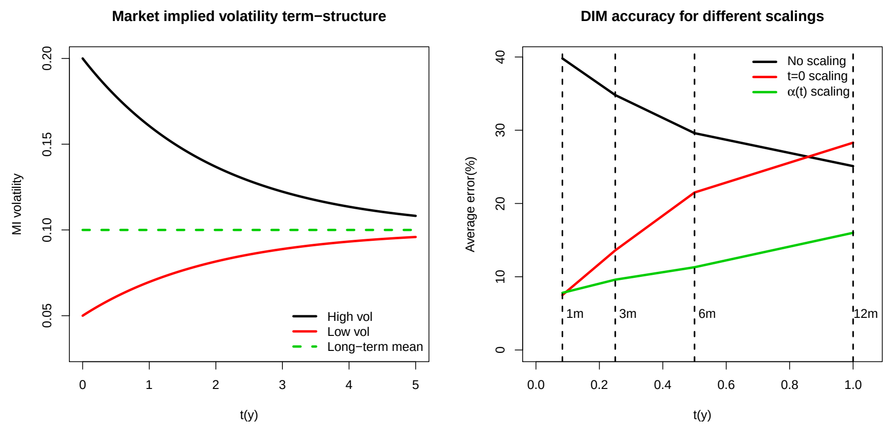

Figure 1: Left panel: Sketch of the term structure of the RF volatility in the case of low (black line) and high (red line) volatility markets (if the RF volatility is a stochastic process itself, the plot can be interpreted as the time dependence of the expected value of the volatility). The dashed green line shows the long term asymptotic behaviour. Right panel: The accuracy of the DIM forecast at different horizons ($h = 1m, 3m, 6m, 12m$) is shown for three choices of scaling (No scaling: $\alpha(t) = 1$; $t = 0$ scaling: as for Eq. 4 with $\beta_{R/P}(t) = 0$ and $h_{R/P}(t) = 0$; $\alpha(t)$ scaling: as for Eq. 4 with $\beta_{R/P}(t) = 1$ and $h_{R/P}(t) = 0$) and for a basket of 102 single-trades portfolios (the products considered, always at the money and of different maturities, include: cross-currency swaps, IR swaps, FX options and FX forwards; where about 75% of the population is made of $\Delta = 1$ trades). The comparison between forecasts and realisations is performed using 7 years of data, monthly sampling and averaging among the posted and received IMs cases. The error metric used is given by $\langle |F_R(t) - G_R(t)| \rangle / \langle G_R(t) \rangle$ (for F_R and G_R, see definitions in Sec.3).

3 How to backtest a DIM model?

So far, we have discussed a DIM model for B-IMR without being too specific on how to assess model performance for the different applications such as CVA and MVA pricing, Liquidity Cover Ratio (LCR) / Net Stable Funding Ratio (NSFR) monitoring [10] and capital exposure. As mentioned above, depending on which application one considers, it may or may not be important to have an accurate assessment of the IMRD. While a perfect model would probably serve well all purposes, it can be the case that the performance of a realistic implementation may differ significantly across applications.

Aim of this section is to introduce two distinct levels of backtesting that can measure the DIM model performance in the two topical cases of: *i)* DIM applications that do not directly depend on the IMRD (such as capital exposure and CVA) and *ii)* DIM applications that do directly depend on the IMRD (such as MVA calculation and LCR / NSFR monitoring). The two corresponding methodologies are presented in subsections 3.1 and 3.2 respectively.

3.1 Backtesting the DIM mapping functions (for capital exposure and CVA)

In a Monte Carlo (MC) simulation framework, the exposure is computed determining the future MtMs of a given portfolio on a large number of forward looking scenarios for the underlying risk factors. On each of these scenarios, the posted and received IMs are provided by the DIM model. If the IMs are segregated and bankruptcy remote: *i)* the received IM can be used to offset the MtM of the portfolio and *ii)* the posted IM does not bear risk. Overall, the exposure is largely reduced (see table in Fig. 2 for a more comprehensive description of the different cases).

To ensure that a DIM model is sound, one should verify that the IM forecast associated to a future scenario is accurate, i.e. we should introduce a suitable historical backtesting framework for DIM.

Let us define generic IMR for a portfolio p (e.g. B-IMR) as:

$$\text{IMR} = g_{R/P}(t = t_\alpha, \Pi = \Pi(p(t_\alpha)), \vec{M}_g = \vec{M}_g(t_\alpha)). \tag{6}$$

Where in Eq. 6:

i) The functions g_R and g_P represents the IMR for received and posted IMs respectively.

ii) $t = t_\alpha$ is the time at which the IMR for the portfolio p are determined.

iii) $\Pi(p(t_\alpha))$ is the trade population of the portfolio p at time t_α.

iv) $\vec{M}_g(t_\alpha)$ is a generic state variable characterising all the $T \leq t_\alpha$ market information required for the computation of the IMR.

Similarly, we define the DIM forecast for the IMR of a portfolio p as:

$$\text{DIM} = f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r}, \Pi = \Pi(p(t_k)), \vec{M}_{\text{DIM}} = \vec{M}_{\text{DIM}}(t_k)). \tag{7}$$

Where in Eq. 7:

i) The functions f_R and f_P represents the DIM forecast for received and posted IMs respectively.

ii) $t_0 = t_k$ is the time at which the DIM forecast is computed.

iii) $t = t_k + h$ is the time for which the IMR are forecasted (over a forecasting horizon $h = t - t_0$).

iv) $\vec{r}$ is a set of market variables whose forecasted values on a given scenario are used by the DIM model as input to infer the IMR. The exact choice of $\vec{r}$ is depending on the DIM model. For the one considered in Sec. 2, $\vec{r}$ is simply given by the MtM of the portfolio.

v) $\vec{M}_{\mathrm{DIM}}(t_k)$ is a generic state variable characterising all the $T \leq t_k$ market information required for the computation of the DIM forecast.

vi) $\Pi()$ is defined as for Eq. 6.

Despite being computed with the use of stochastic RFE models, f_R and f_P are not probability distributions since they do not carry any information regarding the probability weight of a given received / posted IM value. $f_{R/P}$ are instead mapping functions between the set $\vec{r}$ chosen as a representative indicator and the forecasted value for the IM. Since $f_{R/P}$ do not have a probabilistic interpretation, the Probability Integral Transformation (PIT) framework [11] is not suitable for backtesting (while it can be still used in the case of IMRD forecasting, see next section).

Instead, in terms of $g_{R/P}$ and $f_{R/P}$, one can define exception counting tests. The underlying assumption is that the DIM model is calibrated at a given confidence level (CL); therefore it can be tested as a VaR(CL) model. For a portfolio p, a single forecasting day t_k and forecasting horizon h one can proceed as follows:

1. The *forecast functions* $f_{R/P}$ are computed at time t_k as $f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r}, \Pi = \Pi(p(t_k)), \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k))$. Notice that $f_{R/P}$ depends explicitly only on the indicator variable $\vec{r}$ ($\vec{r} = MtM$ for the model considered in Sec. 2).

2. The realised value $\vec{r} = \vec{R}$ of the indicator is determined. For the model considered in Sec. 2, $\vec{R}$ is given by the portfolio value $p(t_k + h)$, where the trade population $\Pi(p(t_k + h)$ at $t_k + h$ differs from t_k only because of trades aging. Aside for aging, no trades are modified, added or removed.

3. The *forecasted values* for the received and posted IMs are computed as $F_{R/P}(t_k + h) = f_{R/P}(t_0 = t_k, t = t_k + h, \vec{r} = \vec{R}, \Pi = \Pi(p(t_k)), \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k))$.

4. The *realised values* for the received and posted IMs are computed as $G_{R/P}(t_k+h) = g_{R/P}(t_0 = t_k+h, \Pi = \Pi(p(t_k+h)), \vec{M} = \vec{M}_g(t_k+h))$.

5. The forecasted and realised values are compared. The received and the posted DIM models are considered independently and a backtesting exception occurs if the F_R (F_P) is above (below) G_R (G_P).

Applying the 1-5 program to multiple sampling points t_k's, one can detect backtesting exceptions for the considered history. The key step is 3, where the dimensionality of the forecast is reduced (from a function to a value) making use of the realised indicator and hence allowing for a comparison with the realised IMR.

The determination of the test p-value requires the additional knowledge of the Test Value Statistics (TVS) that can be derived numerically also if the forecasting horizons are overlapping (see [12]). In the latter situation, it can happen that a single change from a volatility regime to another may trigger multiple correlated exceptions and hence the TVS should adjust the backtesting assessment for the presence of false positives.

An example calculation for a diversified IRS portfolio is shown in Fig. 3, where exceptions tests at $CL = 95\%$ are considered. Over the backtesting history, the DIM model performs generally better for received IM forecasts compared to posted. The functions $h_{R/P}(t)$ are used as haircuts to reduce exceptions and improve the model performance.

Notice in this regard that BCBS-IOSCO netting sets are largely overcollateralised due to the independent requirements for VM and IM. Therefore, the exposure generating scenarios are tail events and hence the impact on Expected Exposure (EE) (and, proportionally, on capital) of a conservative haircut applied to the received IM is rather limited (See the right panel of Fig. 4, where a stylised calculation for different choices of the $\Delta MtM(t, t+MPOR)$ distribution is shown).

3.2 Backtesting the IMRD (for MVA and LCR / NSFR)

The same MC framework can be used in combination with a DIM model to forecast the IMRD at any future horizon (here we implicitly refer to models where the DIM is not always constant across scenarios). The possible business applications of the IMRD are multiple. To mention two that equally apply to the cases of B-IMR and CCP IMR (C-IMR): *i)* Future IM funding costs in the Q measure, i.e. MVA. *ii)* Future IM funding costs in the P measure, e.g. in relation to LCR or NSFR requirements [10].

Our focus is on forecasts in the P measure (tackling the case of the Q measure may require a suitable generalisation of [13]). The main difference with what discussed in Sec. 3.1 is that now the model forecasts to be tested are the numerical IMRDs (one for the received and one for the posted IMs).

	CCR Capital (CVA and Default charges)	Economic CVA	Economic DVA	MVA	LCR / NSFR
BCBS-IOSCO IM posted	Irrelevant: 3rd party custodian, not risk bearing	Irrelevant: 3rd party custodian, not risk bearing	Reduced: IM mitigates counterparty losses in the case of bank own default	Relevant: IM drives funding costs	Relevant: IM drives funding costs
BCBS-IOSCO IM received	Reduced: IM mitigates bank losses in the case of counterparty default	Reduced: IM mitigates bank losses in the case of counterparty default	Irrelevant: 3rd party custodian, not risk bearing	Irrelevant: 3rd party custodian, segregated (*)	Irrelevant: 3rd party custodian, segregated (*)
IM posted to CCP	Reduced: IM is money at risk and should be capitalized with the benevolent CCP regulatory framework (only Default charge in scope)	NA	NA	Relevant: IM drives funding costs	Relevant: IM drives funding costs
IM posted to broker dealer (non clearing member bank clearing via a clearing member bank)	Relevant: IM is money at risk (**)	Relevant: IM is money at risk for the bank in the case of the counterparty default	Reduced: IM mitigates counterparty losses in the case of bank own default	Relevant: IM drives funding costs	Relevant: IM drives funding costs
IM received from client (clearing member bank acting as broker dealer)	Reduced: IM mitigates bank losses in the case of counterparty default	Reduced: IM mitigates bank losses in the case of counterparty default	Relevant: IM is money at risk for the counterparty in the case of the bank own default	Relevant: IM may bring funding benefits	Relevant: IM may bring funding benefits

(*) According to US rules, partial re-hypothecation may be allowed in certain circumstances
(**) The capital rules to be applied depends on the specifics of the client / clearing member contractual agreement. CVA charge may or may not be in scope.

Figure 2: The table shows the main quantities impacted by DIM from regulatory and pricing perspectives for the different flavours of IMR in the context of bilateral and cleared derivatives. the red colour indicates reduced inflow / increased outflow for the bank, while green the vice-versa.

These can be obtained for a given time horizon by associating to every simulated scenario the correspondent IMR computed according to the given DIM model. Using the notation introduced in Sec. 3.1, the numerical representations of the R/P IMRD Cumulative Density Functions (CDFs) of a portfolio p for a given forecasting day t_k and horizon h are given by:

$$\begin{aligned} \mathrm{CDF}_{R/P}(x, t_k, h) &= \#\Big\{ v \in \mathbf{V} \Big| v \leq x \Big\} / N_{\mathbf{V}} \qquad (8) \\ \mathbf{V} &= \Big\{ f_{R/C}(t_0 = t_k, t = t_k + h, \vec{r}_\omega, \Pi = \Pi(p(t_k)), \\ &\quad \vec{M}_{\mathrm{DIM}} = \vec{M}_{\mathrm{DIM}}(t_k)), \forall \vec{r}_\omega \in \Omega \Big\}, \qquad (9) \end{aligned}$$

where: *i)* In Eq. 8, $N_{\mathbf{V}}$ is the total number of scenarios. *ii)* In Eq. 9, $f_{R/P}$ are the functions computed using the DIM model, $\vec{r}_\omega$ are the scenarios for the indicator market variables (the portfolio MtMs in the case discussed in Sec. 2) and Ω is the ensemble of the $\vec{r}_\omega$ spanned by the Monte Carlo simulation.

Received IM, VaR test for CL=95%				
Haircut / Horizon	1m	3m	6m	12m
0	9% (1%)	50% (2%)	14% (3%)	27% (5%)
0.15	100% (n/a)	100% (n/a)	100% (n/a)	100% (n/a)

Posted IM, VaR test for CL=95%				
Haircut / Horizon	1m	3m	6m	12m
0	0% (4%)	0% (10%)	0% (12%)	0% (23%)
-0.15	98% (0.0%)	0% (6%)	0% (10%)	0% (14%)
-0.25	100% (n/a)	65% (6%)	2% (7%)	0% (12%)
-0.3	100% (n/a)	92% (12%)	7% (5%)	0% (11%)

Figure 3: The table shows p-values for exception counting tests at CL=95% for a long-short IRS portfolio as a function of the applied haircut h (see Eq. 4, where $h_{R/P}(t) = h$ for $t > 0$ and $h_{R/P}(t) = 0$ for $t = 0$) and of the backtesting horizon in the cases of received and posted IMs. The number in parenthesis indicates the realised expected shortfall for the given test (i.e. the average size of the observed exceptions). The analysis is performed using 7y of data, monthly sampling frequency and according to the backtesting methodology described in Sec. 3.1. The IRS portfolio constitutes of 12 at the money trades: 3 USD IRS ($T = 5y, 10y, 20y$ and Receiver, Payer, Receiver respectively), 3 EUR IRS ($T = 5y, 10y, 20y$ and Payer, Receiver, Payer respectively), 3 CHF IRS ($T = 5y, 10y, 20y$ and Receiver, Payer, Receiver respectively), 3 JPY IRS ($T = 5y, 10y, 20y$ and Payer, Receiver, Payer respectively).

The IMRD in this form is directly suited for historical backtesting using the PIT framework. Refering to the formalism described in [12] and [14], one can derive the PIT time series $\tau_{R/P}$ of a portfolio p for a given forecasting horizon h and backtesting history $\mathcal{H}_{BT}$ as follows:

$$\tau_{R/P} = \left\{ \mathrm{CDF}(g_{R/P}(t_k + h, \Pi(p(t_k + h)), \vec{M}_g(t_k + h)), t_k, h), \forall t_k \in \mathcal{H}_{BT} \right\}. \tag{10}$$

In Eq. 10, $g_{R/P}$ is the exact IMR function for the IMR methodology we intend to build a forecast for (defined as for Eq. 6) and the t_k's are the sampling points in $\mathcal{H}_{BT}$. Every element in the PIT time series $\tau_{R/P}$ corresponds to the probability of the realised IMR at time $t_k + h$ according to the DIM forecast built at t_k.

As discussed extensively in [12] and [14], one can backtest the $\tau_{R/P}$ using uniformity tests. In particular, in analogy to what is shown in [12] for portfolio backtesting in the context of capital exposure models, one can also use test metrics that do not penalise conservative modelling (i.e. models overstating / understating posted / received IM). In all cases, the correspondent TVS can be derived accordingly using numerical MC simulations.

In this set-up, the performance of a DIM model are not tested in isola-

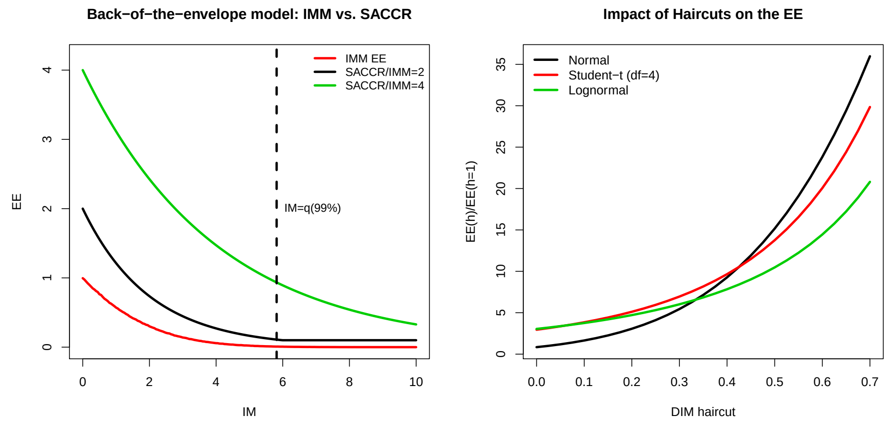

Figure 4: Left panel: The Expected Exposure (EE) for a VM collateralised counterparty is shown as a function of the received IM for: i) a back-of-the-envelope IMM model (continuos red line) and ii) SA-CCR (continuos black and green lines). The IMM exposure is calculated in the case of daily bilateral VM cash margining for a simple model where $\Delta MtM(t, t+\text{MPOR})$ is an iid Normal random variable with zero mean and constant volatility. The SA-CCR exposure is shown for the same VM collateralised example in the cases of SA-CCR add-ons being two (black line) or four (green line) times more conservative than IMM; where the level of conservatism is set for IM=0 by fixing the ratio $\text{PFE}_{\text{SA-CCR}}/\text{EE}_{\text{IMM}}$ to two or four respectively. Right panel: The impact on the EE of a conservative haircut applied to the received IM collateral is shown (where $h = 0$ and $h = 1$ means full IM collateral benefit or no benefit respectively). $\Delta MtM(t, t + \text{MPOR})$ is modelled as an iid random variable drawn by Normal (black line), Student-t (red line) and Lognormal (green line) distributions respectively. The pre-haircut IM is chosen as the 99th percentile of the corresponding distribution.

tion. The backtesting results will be mostly affected by:

1. The choice of $\vec{r}$. As discussed in 3.1, $\vec{r}$ is the indicator used to associate an IMR to a given path / valuation point. If $\vec{r}$ is a poor predictor for the IMR, the DIM forecast will be consequently poor.

2. The mapping $\vec{r} \to$ IMR. Similarly to the previous item, if the the mapping model is not accurate the IMR associated to a given scenario will be inaccurate.

3. The RFE models used for $\vec{r}$. These models ultimately determine the probability of a given IMR scenario. It may so happen that the mapping functions $f_{R/C}$ are accurate or even exact but the probabilities

of the underlying scenarios for $\vec{r}$ are misstated. Hence causing backtesting failures.

Notice that items 1 and 2 above are relevant also for what discussed in Sec. 3.1. Item 3 instead is peculiar of this backtesting variance since it concerns the probability weights of the IMRD.

4 Conclusions

In the previous sections, we have presented a complete framework to develop and backtest IMR forecasting models. Our focus has been on B-IMR and we have shown how to obtain the forward looking IMs from the simulated exposure paths using simple aggregation methods.

The proposed DIM model is suitable both for XVA pricing and capital exposure calculation, where the haircut function $h(t)$ in Eq. 4 can be used either to improve the accuracy (pricing) or to ensure the conservatism of the forecast (capital).

If a financial institution were to compute CCR exposure using the Internal Model Methods (IMM), the usage of a DIM model could reduce CCR capital dramatically, even after the application of a conservative haircut. This should be compared with the regulatory alternative SA-CCR where the benefit from overcollateralisation is largely curbed (see left panel of Fig. 4 and [15]).

As part of the proposed framework, we have also defined a backtesting methodology able to measure performance for the different usages of a DIM model. The approach presented in Sec. 3 is agnostic to the underlying IMR and can be directly applied in other contexts, such as C-IMR.

As a final remark, notice that the DIM model specified in Sec. 2 can be considered for C-IMR as well. The main additional challenge to be tackled is to account for the time dependence of the model calibration underlying the C-IMR (either based on VaR methodologies with enriched historic scenarios such as e.g. PAIRS and HVaR or on rule based approaches such as e.g. SPAN [16]).

References

[1] Basel Committee on Banking Supervision, "Margin requirements for non-centrally cleared derivatives," March 2015.

[2] A. D. Green and C. Kenyon, "MVA by replication and regression," *Risk*, May 2014.

[3] L. Andersen, M. Pykhtin, and A. Sokol, "Modeling Credit Exposure for Margined Counterparties," Risk Minds, 2014.

[4] C. Albanese, “The cost of clearing,” *SSRN pre-print*, 2015.

[5] ISDA, “ISDA SIMM Methodology,” *draft document*, October 2015.

[6] L. Andersen, M. Pykhtin, and A. Sokol, “Modeling Credit Exposure for Margined Counterparties,” WBS, 2015.

[7] Basel Committee on Banking Supervision, “Basel III: A global regulatory framework for more resilient banks and banking systems,” June 2011.

[8] ISDA, “Uncleared Initial Margin Calculations and Processes Related to the ISDA WGMR SIMM Initiative,” *draft document*, June 2015.

[9] F. Longstaff and E. Schwartz, “Valuing american options by simulation: A simple least- squares approach,” *The Review of Financial Studies*, vol. 14, pp. 113–147, 2001.

[10] Basel Committee on Banking Supervision, “Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools,” January 2013.

[11] F. X. Diebold, T. A. Gunther, and A. S. Tay, “Evaluating density forecasts with applications to financial risk management,” *International Economic Review*, vol. 39, no. 4, pp. 863–883, 1998.

[12] F. Anfuso, D. Karyampas, and A. Nawroth, “Credit exposure models backtesting for basel III,” *Risk*, September 2014.

[13] L. Jackson, “Hedge backtesting for model validation,” *Risk*, September 2013.

[14] C. Kenyon and R. Stamm, *Discounting, Libor, CVA and Funding: Interest Rate and Credit Pricing*. Palgrave Macmillan, 2012.

[15] F. Anfuso and D. Karyampas, “Capital flaws,” *Risk*, July 2015.

[16] J. Gregory, *Central Counterparties: Mandatory Central Clearing and Initial Margin Requirements for OTC Derivatives*. Wiley, 2014.