How confident are we of margin model procyclicality measurements?

Pedro Gurrola-Perez[(1)]

April 6, 2023

Abstract:

Measuring the responsiveness of a market risk model is relevant whenever the focus is on evaluating if a model is over- or under-reacting to changes in market conditions. Such is the case, for example, in the discussion about procyclical effects of the initial margin models used both in the centrally cleared and non-centrally cleared worlds to estimate the potential future exposure of portfolios. By definition, these models are sensitive to changes in market risk and, as a consequence, when market risk increases, initial margin requirements will tend to increase. To mitigate procyclical effects, CCPs have put in place different procyclicality mitigation tools across their risk management arrangements. However, after the Covid-19 stress, there have been renewed discussions about further monitoring, measuring, and mitigation of models' procyclicality. This paper contributes to the discussion by bringing the attention to the fact that the standard measures of model responsiveness are random variables and, as such, are subject to uncertainty. Therefore, any decision or policy making that is based on these measures requires, to be robust, to take into consideration the impact that uncertainty will have on expected outcomes. To estimate such impact, the analysis examines the case of some typical margin models, both empirically and within a Monte Carlo simulation setting. The results show there is a significant amount of uncertainty when measuring responsiveness, which raises questions about the effectiveness of model-focused, hard-rule approaches to procyclicality.

Keywords: Central counterparties (CCPs), procyclicality, initial margin (IM), margin models, filtered historical simulation.

JEL classification: G17, C60, G23, G01.

[(1)] World Federation of Exchanges. Email: pgurrola@world-exchanges.org

Acknowledgements: I would like to thank Fernando Cerezetti, Alex Comninos, Eddie Gerba, Richard Haynes, Emily Hendrix, Petros Katsoulis, David Murphy, Charlie Ryder, Paul Spang, members of the WFE CCP Working Group and participants at a Bank of England seminar and at the WFE Clearing and Derivatives Conference 2023 (WFE Clear) for helpful conversations and comments. Any errors are solely mine.

WFE Working Papers: WFE Working Papers present research in progress by the author(s) and are published to share early insights, stimulate debate, and elicit feedback. Comments on Working Papers are welcomed, and may be sent to Head of Research, World Federation of Exchanges, research@world-exchanges.org.

.

1 Introduction

Quantile-based market risk models, like value-at-risk (VaR), aim at measuring potential future exposure by forecasting a distribution of returns to estimate the worse potential outcomes with a given confidence level. While the performance of these models is linked to meeting their target coverage, it may be desirable to have a model that satisfies some additional properties. For example, if there is a set of models, all of them achieving similar target coverage but showing different levels of responsiveness to the same sequence of events, a risk modeller may prefer to choose those models that tend to over- or under-react as little as possible. He will therefore need to distinguish between models with different levels of responsiveness.

In particular, in the case of the initial margin (IM) models used in bilaterally and centrally cleared trades, model responsiveness became a key area of focus for regulators after the Great Financial Crisis of 2007-2008 (CGFS, 2010; CPMI, 2012). The concern was that, in a situation of financial stress, a large increase in margin requirements would add liquidity pressure to the participants at a time when they are already struggling; a situation that could potentially amplify the stress and affect financial stability. To reduce the effects of such procyclical behaviour with respect to the market, regulators and industry have put several measures in place. However, questions from a subset of participants and policymakers about whether initial margin calls for centrally cleared trades are still too procyclical seem to come back every time there is a shock in the market and margin calls increase, as has happened during the Covid-19 crisis in March 2020, or after the Russian invasion of Ukraine.[1]

A fundamental element in the discussion and in the implementation of any procyclicality mitigation tool is our ability to measure the responsiveness of a model. Several measures have been suggested by regulators and in the academic literature, including measuring the range of margin called over a certain period, or the largest cumulative margin increase in n-days over a given period. Although these measures only capture some aspect of the model responsiveness and they can only be used as a tool for comparison between models, they are sometimes intended to be used for decision making: the risk modeller is expected to consider a set of selected historical periods, estimate the corresponding margins produced by a set of acceptable models and, using these measures, select the model that exhibits lower levels of responsiveness across the sample periods. In some cases, there may also be requirements to set procyclicality "targets" or "thresholds" against which the risk modeller will validate the choice of model; see, for example, ESMA (2018, 2022).

However, to the best of my knowledge, little consideration has been given to the fact that, for a given model, these measures are real functions on a space of random events (the P&L scenarios). That is, they are random variables and, as such, they are characterized by their probability distribution function. Therefore, a robust decision making process based on these measures requires acknowledging the presence of uncertainty in the measurements, for example, by estimating their sensitivity to the choice of scenario. From a policy perspective, while one cannot legislate across all states of nature, without quantifying the uncertainty surrounding procyclicality measurements there is a risk of prescribing rules which have a significant chance of resulting ineffective. It will also be difficult to judge whether a "too procyclical" behaviour in the future is a failure of the model's anti-procyclicality credentials, or the likely consequence of the uncertainty in our measurements.[2]

The importance of measuring the sensitivity to the choice of scenario can be illustrated with a simple hypothetical example.

[1] In fact, at the time of writing, there are several regulatory initiatives re-examining the question of IM procyclicality. See, for example, the consultations by the European Securities and Markets Authority (ESMA, 2022) and by the BIS-CPMI-IOSCO (CPMI, 2022).

[2] While the expression "too procyclical" is often used to characterise a situation of margin flagrantly over- or -under reacting, it is worth noting that there is no intrinsic notion of "correct" procyclicality.

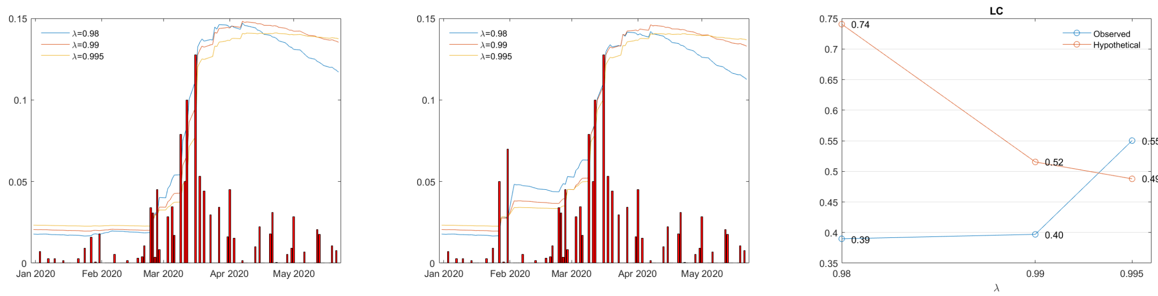

Figure 1: The graphs illustrate the impact that a small change in the historical series of returns can have on the procyclicality measurements. Here, the baseline model is an FHS VaR with a 1-year lookback period and decay factors of 0.98, 0.99 and 0.995. The left panel shows the losses observed in the S&P500 (SPX) from January to May 2020 (100 observations). The centre panel shows the model response in a hypothetical scenario which coincides with the observed one except on three days of January, where large losses have been introduced. The right hand panel shows the procyclicality estimates for the real and hypothetical scenarios using the 5-day cumulative increases $LC(5)$. The models achieve similar coverage across scenarios: six breaches for $\lambda = 0.99$ and seven breaches for $\lambda = 0.995$.

Consider the S&P500 in the first quarter of 2020, when volatility spiked across markets, and consider a filtered historical simulation (FHS) VaR margin model with the decay factor λ calibrated at 0.99 and 0.995 (section 3.1 provides a detailed description of how these models work, and what is the impact of the parameter λ but, for the time being, it is sufficient to know that higher λ's produce a less responsive output). The left panel in Figure 1 shows the margin generated under each one of these models.

Now, consider a hypothetical scenario which is identical to the observed one except that on January 23 and 27 losses were larger than what was observed (center panel of Figure 1).

Suppose that the choice of the less procyclical model is based on estimating the largest cumulative 5-day increase during the period, denoted $LC(5)$ (see Section 3, for details).

It turns out that even with such a small change in the history of events, the choice of the less procyclical model would be completely different, as the third panel in Figure 1 shows. Since the two models provide similar coverage, a risk modeller concerned about the short-term increases in margin and using the 5-day cumulative measure will choose, under the hypothetical scenario, the model with $\lambda = 0.995$ as the less procyclical ($LC(5) = 0.49$), just to find out, if the real scenario crystallizes, that this instead is the most procyclical! Moreover, contrary to intuition, when considering the real scenario, models with lower responsiveness (that is, with higher λ) are not those attracting lower procyclicality values, which raises doubts about the power of these measures to discriminate between models with different levels of responsiveness.

It can be argued that the above example is a singular instance of an infinite set of potential scenarios and that the consistency of the measurements can only be proved or disproved in the long run, in the same way that, if the model is fit for purpose, backtesting will show convergence between expected and observed outcomes as the sample size increases.[3] It can also be argued that "mechanically" creating hypothetical scenarios without any consideration of the underlying dynamics (e.g. autocorrelations) is a flawed approach. Both observations are true, and that is why the probability distribution of procyclicality measurements needs to be estimated and it needs to be done in a way that is consistent with the underlying dynamics.

Given the random nature of measures of model responsiveness, two related questions arise: given a

[3]Backtesting refers to the process of comparing realized outcomes against expected outcomes. In the case of VaR models with a x% coverage, it is expected that, on average, observed losses will not be larger than the VaR value more than x% of the time. To test the significance of backtesting results different statistical tests can be applied.

set of scenarios and a set of acceptable margin models, how sensitive are the measures to the choice of scenarios? And, how powerful are these measures to discriminate between models with different levels of responsiveness?

The purpose of this paper is to provide an answer to these questions. This has relevant consequences for policy making since the greater the uncertainty in responsiveness measurements, or the more sensitivity to the underlying scenario, the less certainty we can place on approaches to mitigate future responsiveness that are based on one-fits-all, deterministic rules.

The analysis will be performed using filtered historical simulation (FHS) VaR models, with conditional volatility estimates generated with an exponentially weighted moving average (EWMA) process. Not only are these FHS-EWMA models (or some variations of them) often used by central counterparties (CCPs) in their IM modelling, but they are designed to adjust their *responsiveness* through a parameter λ (the decay factor). This allows to know in advance what is the true responsiveness of different models and test the power of the measures to capture such behaviour.

First, some empirical evidence is gathered by considering the historical returns of three products: the S&P 500, Brent Oil and Brazilian Real. For each of these series, a set of samples (scenarios) will be selected using a 4-year rolling window. For each scenario, the corresponding margin is then estimated and the standard procyclicality measures are applied. One can then test the consistency of the resulting measurements across the scenarios.

But, while empirical testing is fundamental to support the analysis, it limits the universe of possible outcomes to some combination of what has already occurred in the past. To provide a forward-looking view which can account for previously unseen scenarios and to better capture the uncertainty around the measurements, the analysis will also estimate the probability distribution of these measurements within a Monte Carlo simulation setting, sampling scenarios generated by a GARCH process which has been fitted to the S&P500.

The present work is related to the strand of research on procyclicality that looks at the IM model in isolation and which examines, either theoretically or empirically, the relation between the model's inputs and outputs (Abruzzo and Park, 2014; Murphy, Vasios and Vause, 2014; Gurrola-Perez and Murphy, 2015; Murphy, Vasios and Vause, 2016; Maruyama and Cerezetti, 2019; Murphy and Vause, 2021; Gurrola-Perez, 2022; Kahros and Weissler, 2021). Previous analyses have often considered historical samples and, when doing simulations, comparisons are done focusing on mean values. Our contribution highlights the importance of quantifying uncertainty in procyclicality measurements.

The rest of the paper is organised as follows. Section 2 discusses the main measures of model responsiveness and their condition as random variables. In Section 3 I establish the notation and formalize the definitions. Section 4 presents the empirical analysis. In Section 5, I present and discuss the results of the simulation framework. Section 6 discusses how the results affect the assessment of procyclicality mitigation tools. Section 7 concludes.

2 Definition and measurement of procyclicality

According to the Principles for Financial Market Infrastructures (PFMIs), the global principles set by CPMI and IOSCO, procyclicality "typically refers to the changes in risk-management requirements or practices that are positively correlated with market, business or credit cycle fluctuations and that may cause or exacerbate financial instability" (CPMI (2012), 3.5.6 and Annex H).

The main concern was that, in a situation of financial stress, a large increase in margin requirements would add liquidity pressure to the members at a time when they are already struggling, which, under particular states of the system, could potentially amplify the stress and affect financial stability.

While the principle is clear and is useful to guide risk management decisions, its operational implementation (that is, its implementation in terms of specific tests and measurements that can be used to identify, quantify, and forecast procyclical behaviours) is not straightforward. There are at least three reasons for this. First, in terms of the wider system, it would require to identify the thresholds beyond which a margin requirement may exacerbate a financial stress. But the core characteristics of a complex system (non-linear interactions, emergence, feedback loops, etc.) are in fundamental contradiction with such approach. Second, one would need to identify *all* the channels through which the IM interacts with the wider system and measure those interactions. Clearing members' balance sheets, discussed in Faruqui et al. (2018) are not the only channel; there are also interconnections through the repo markets (Benos et al., 2022; Gerba and Katsoulis, 2021), and through the members' network of clients: in CPMI (2022) it is reported that, during the Covid-19 events, it was generally not the clearing members, but their clients, who varied in their liquidity preparedness. Third, in terms of the IM models, such implementation is limited by the fact that reducing model responsiveness is constrained by the need of preserving its risk sensitivity and of keeping the costs of clearing economically efficient (Gurrola-Perez, 2022; Murphy and Vause, 2021).

An additional problem, which is the focus of this paper, arises from the fact that there is uncertainty in measurements of responsiveness and of margin costs.

2.1 Measuring model responsiveness

While the notion of procyclicality is often applied to the IM model itself, referring instead to model *responsiveness*, as in CPMI (2022), or to model *reactiveness* as in Houllier and Murphy (2017), is more adequate and brings clarity to the discussion. The term procyclicality implies some correlation with a "cycle" in the system, which may only have an indirect connection with the model's input or parameters and which often is more a shock than a cycle (e.g., the "dash for cash" in 2020); it is a property of the system, "the reinforcing effect of collateralized trades" as expressed in Kahros and Weissler (2021), and not of the IM model in isolation.

In contrast, responsiveness is a property that can be made operational by describing the model's output under a given input. In other disciplines, like macroeconomics or signal processing, the responsiveness of a system or model is often synthesized in the form of a function (the *impulse response function* or IRF) which describes the evolution of the system along a time horizon when a brief signal or shock (the *impulse*) is applied. By using a basic impulse, IRFs allow to compare models on an equal footing and reduce the space of potential inputs to be tested. In the context of IM models and considering volatility as the main input, Murphy, Vasios and Vause (2014) already hinted at the idea of an IRF for margin models when they examined how different models respond to pre-specified short-term and medium-term volatility bumps.

In the case of IM models, the measures of model responsiveness that have been suggested in the literature, either by academics or by regulators, and which have become more or less standard, are basically measures of the size and/or speed of the variation in the IM model output (see Figure 2):

Standard deviation (SD) (ESMA, 2022).

Peak-to-trough (PT) ratios (Murphy, Vasios and Vause, 2014; Murphy and Vause, 2021). This is the ratio of the maximum (or nearly) initial margin required for a static portfolio to the minimum (or nearly) margin required during a fixed observation period I, for a given underlying process. For example, Murphy and Vause (2021), consider the ratio between the 99.7th percentile and the 0.3rd percentile of IM requirements:

$$PT = \frac{m_T^{99.7}}{m_T^{0.3}} \tag{1}$$

It is mainly a measure of the *amplitude* of the response to a given scenario.

n-day largest cumulative increase (LC(n)) (Murphy and Vause, 2021) This captures the largest (or nearly) cumulative increase in margin over an n-day period assuming a constant portfolio over a

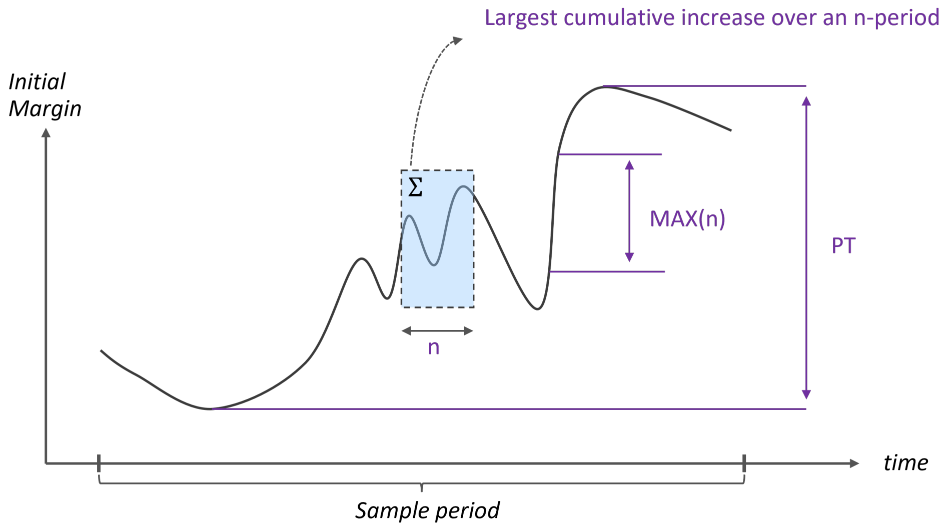

Figure 2: The figure shows a stylized representation of the aspects captured by the standard measures of model responsiveness. MAX(n) is the maximum change between the first and last days across n-day periods. Peak-to-through ratios (PT) capture the range of margin values observed in a given period. All the measures are a function of the scenario chosen (sample period).

fixed observation period. It is a measure of the *speed* of the incremental response to a sequence of events. The margin call could cumulate between day 1 and day n of the window. That is,

$$LC_t = \max_{w,W}(m_{t-w} - m_{t-W}), \quad \forall\{0 \leq W \leq ndays, 1 \leq w < W\} \tag{2}$$

and the LC measure up to t is the X percentile of $\{LC_t | t < T\}$.

n-day largest increase, MAX(n) The largest increase between the first and the last day in a n-day period.

While the idea behind these measures is intuitively clear, they have some limitations:

A1) They are not a function of the model only. Instead, they are a function of the models and the scenarios: a change in scenario can lead to different measurements for the same model. And, similarly, given a fixed scenario, different models can produce different measurements.[4]

A2) Since there is no "correct" responsiveness value associated to the model which can be used as reference of benchmark, the measurements by themselves do not contain any significant information.[5] For example, knowing that a model has a PT value of, say, 3.2, is irrelevant, unless information about the underlying scenarios and the PT produced by other models in identical scenarios is provided. One can only use these measures as a comparison/ranking tool: given a set of admissible models (for example, models with the correct coverage), and a set of scenarios, choose those that over- or under-react as little as possible under the given set of scenarios.

A3) There is no reference to the liquidity or market risk cycle against which procyclicality is predicated. Therefore, depending on the state of the system, the same level of model responsiveness can lead to

[4] They are not functions of the volatility of the underlying scenario either, as two scenarios with the same volatility can lead to different measures of procyclicality.

[5] Compare with backtesting, where we know, *by definition*, what is the expected coverage of a model (e.g. 99% VaR) and we can *statistically* test whether the breaches observed are consistent with the target coverage to decide whether a model is fit for purpose.

completely different systemic outcomes. For example, had market participants been better prepared on their liquidity arrangements during March 2020, the "dash-for-cash" would presumably have been less likely to happen, even with the same level of initial margin calls.

A4) They measure different aspects of model responsiveness (e.g., PT is mostly about amplitude, MAX(n) mainly about speed, and SD is dispersion around the mean). There is no guarantee they will be consistent.

2.2 Measures of IM model responsiveness are random variables

A key observation is that, for market risk models based on empirical distributions, *standard measures of model responsiveness are random variables defined over a set of potential P&L scenarios*. In other words, they are functions that, given a margin model, associate a real number to a random sample selected from the universe of all potential P&L scenarios. This is a consequence of (A1) above and the fact that the measures make no reference to any particular underlying process.[6] In fact, any function that assigns a real value to the output of a margin model and which also depends on the underlying scenario (for example, a function that assigns a cost to the margin or a function that counts the number of breaches), is similarly a random variable on the space of P&L scenarios.

Although the above observation has attracted little attention, it has important consequences. First, to characterise a random variable one needs information about its probability distribution. Therefore, to be robust, an analysis based on these measures should inform *at least* of the mean and the dispersion around the mean. Without quantifying the uncertainty in the measurements (for example, by estimating their sensitivity to the choice of scenario or, even better, their whole probability distribution) it is not known how reliable these measurements are.

In summary, because measures of responsiveness are random variables, a robust assessment of model responsiveness requires to *estimate the probability that, for a given set of scenarios, and for a given set of acceptable models, the given model produces the less responsive outcome, under a set of measures.*

In the next section the above discussion will be formalized.

3 Notation and definitions

Let

- Ω be a set of scenarios ω generated by the same data generating process (dgp).
- $\mathcal{M}(\lambda)$ a set of margin models parametrized by λ, $0 \leq \lambda \leq 1$.
- $\Gamma : \mathcal{M} \times \Omega \to \mathbb{R}$, a function that assigns to each model and each scenario a real value $\Gamma(\lambda, \omega)$. For example, Γ could represent
  - A measure of model responsiveness
  - A cost function: the cost of margin produced by a margin model.
  - A backtesting function: a function that assigns the number of breaches produced by a margin model.

For a fixed λ, the function $\Gamma_\lambda : \Omega \to \mathbb{R}$, with $\Gamma_\lambda(\omega) = \Gamma(\lambda, \omega)$ defines a random variable. This is true whether Γ refers to margin costs, the sequence of backtesting breaches, or any of the standard measures of responsiveness.

[6] Compare with the case of IRFs, which use a specific input to characterize the response that is associated to the model.

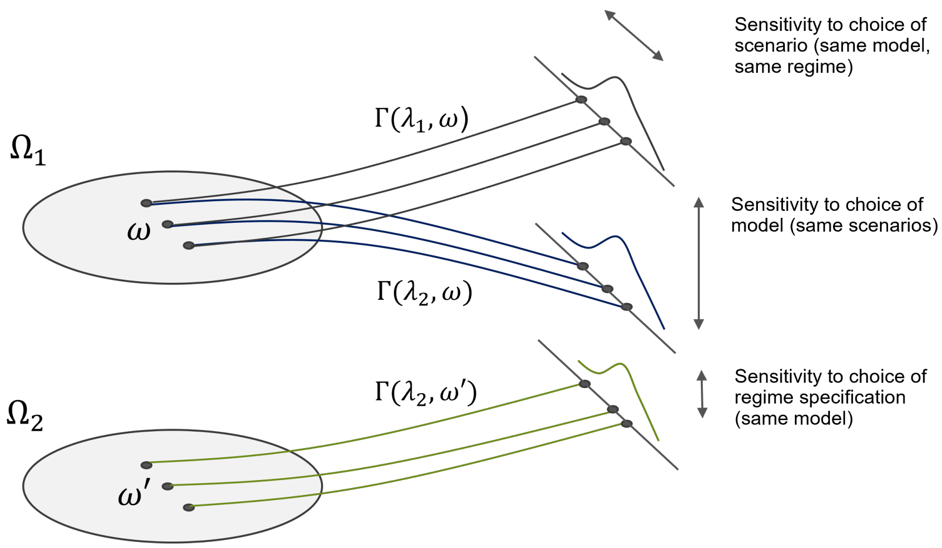

Figure 3: The figure illustrates three sources of uncertainty in procyclicality measurement: sensitivity to 1) changes in scenario ω, 2) changes in the model (parametrized in this case by λ), and 3) changes in the specification Ω defining the set of all possible scenarios. Adapted from the 'Hawkmoth effect' diagram in Thompson and Smith (2019)

With regards to the specification of the measures to be tested, we will consider the 5-day cumulative increase, $LC(5)$; the maximum one-day increase, $MAX(1)$; and the PT ratios and standard deviation, SD, both estimated over the observation period.

From the discussion in section 2.2, it is clear there are three sources of uncertainty: uncertainty regarding the choice of dynamic process Ω, the choice of scenario ω, and the choice of model in $\mathcal{M}$ (Figure 3). The research question can therefore be formulated as follows:

> Suppose we make a procyclicality assessment based on a measure Γ. How sensitive is our assessment to 1) changes in scenario ω, 2) changes in the model parameter λ, and 3) changes in the specification Ω defining the set of possible scenarios.

3.1 FHS models

To estimate initial margins (i.e., potential future exposure), it is common to estimate the distribution of returns and use quantile-based risk measures like Value-at-Risk (VaR) or Expected Shortfall (ES).[7] These estimates are used directly to determine the IM, or are sometimes used as an input to a scenario-based model. If the forecast distribution is based on historical returns then, to account for the presence of conditional volatility and to improve the responsiveness of the model, one can estimate VaR using a Filtered Historical Simulation (FHS) process where the historical sample is rescaled using the most recent conditional volatility estimate. These models (or some variants of them) are common across CCPs. When conditional volatility is estimated using an EWMA or GARCH process, we have the further advantage that we know, by design, which models are more or less responsive to the arrival of new information. In the case of using an EWMA parametrization, the models can be linearly parametrized by the decay factor λ, $0 \leq \lambda \leq 1$ which determines the responsiveness of the model to new information.

Unless otherwise stated, we will estimate percentage VaR and we will follow the convention of quoting VaR (and portfolio losses) as positive numbers.

[7] VaR is the maximum expected loss of a portfolio over a given time horizon and with a given probability and ES is the expected loss beyond a given VaR. See Dowd (2005) for a general reference.

The FHS-EWMA models are defined as follows. For each day T, let $\mathcal{D}_T$ be the distribution of returns at time T. In the case of historical simulation models, $\mathcal{D}_T$ is the empirical distribution of the observations $y_T = \{r_{T-N}, ..., r_{T-1}\}$, where N is the length of the look-back period.

In the filtering approach proposed by Hull and White (1998) the empirical distribution $\mathcal{D}_T$ is rescaled so that its volatility matches the most recent volatility estimate. Returns are first standardized to obtain residuals which are assumed to be approximately stationary. This residual sample is then rescaled using a conditional volatility estimate that will be derived from an EWMA volatility updating scheme. In other words, on each day T, instead of directly calculating a percentile from y_T, each one of the historical returns r_t in y_T is divided by the volatility estimate σ_t for day t. The resulting standardized returns are then multiplied by the conditional volatility σ_T to obtain the rescaled returns

$$R_t = r_t \times \frac{\sigma_T}{\sigma_t} \tag{3}$$

In the case of using EWMA volatility estimates (as is often the case with CCPs), the volatility forecast at time t for time $t+1$ can be calculated using the recursive formula

$$\sigma_{t+1}^2 = \lambda\sigma_t^2 + (1-\lambda)r_t^2 \tag{4}$$

The decay factor λ determines how the volatility estimates incorporate past and new information. The sensitivity to new information increases as λ gets close to 0, while the persistence of the past will increase as λ is closer to 1. The FHS model is therefore more reactive as the parameter λ decreases. CCPs tend to apply decay factors in the higher range (> 0.97). Our example will use $\lambda = 0.96, 0.97, 0.98, 0.99$ and 0.995.

4 Sampling from historical series

We can start the analysis by sampling from historical series of returns. We will consider the P&L of holding a long position in, respectively, the S&P 500, the Brent Oil and the Brazilian Real. For each one of these products, we estimate around 15 years of returns (Figure 4), from which we generate 1000-day rolling samples of returns and we estimate the margin required in the last 250 days of each sample under each of the models to be tested.[8] We then measure the procyclicality of each model over each one of these samples using the standard measures. Finally, we rank the models according to each measure and we estimate the frequency of consistent rankings across simulations.

Table 1, shows the ratio of success obtained under the different measures and for the five different models. Here, success is interpreted as correctly ranking the models according to their λ. For example, a 0.76 ratio indicates that in 76% of scenarios the pair of models were correctly ranked (i.e., lower procyclicality measurement coincides with higher λ) and 24% showed the ranking reversed. Without any a priori knowledge of the model, this can be interpreted as having 24% chance of getting the initial selection wrong.

One perhaps could improve the chances of getting greater consistency by choosing a stressed period as initial testing sample. However, reversals also occur within samples chosen from stress periods. The example in Figure 5 illustrates this. When looking at $MAX(1)$, if tested at the period of high volatility observed at the end of 2011, the model with $\lambda = 0.98$ would have been (wrongly) chosen as the less responsive, compared to the model with $\lambda = 0.995$). But this choice would have proved wrong throughout 2017 and 2018, and at the start of 2020. Both models provide similar coverage, with around 1.6% and 1.4% of breaches.

[8]Clearly, the rolling samples overlap. However, because margins are path dependent, even if two samples overlap except for the first and the last values, the margins can be completely different. The Monte Carlo simulation in the next section will allow to generate completely independent samples.

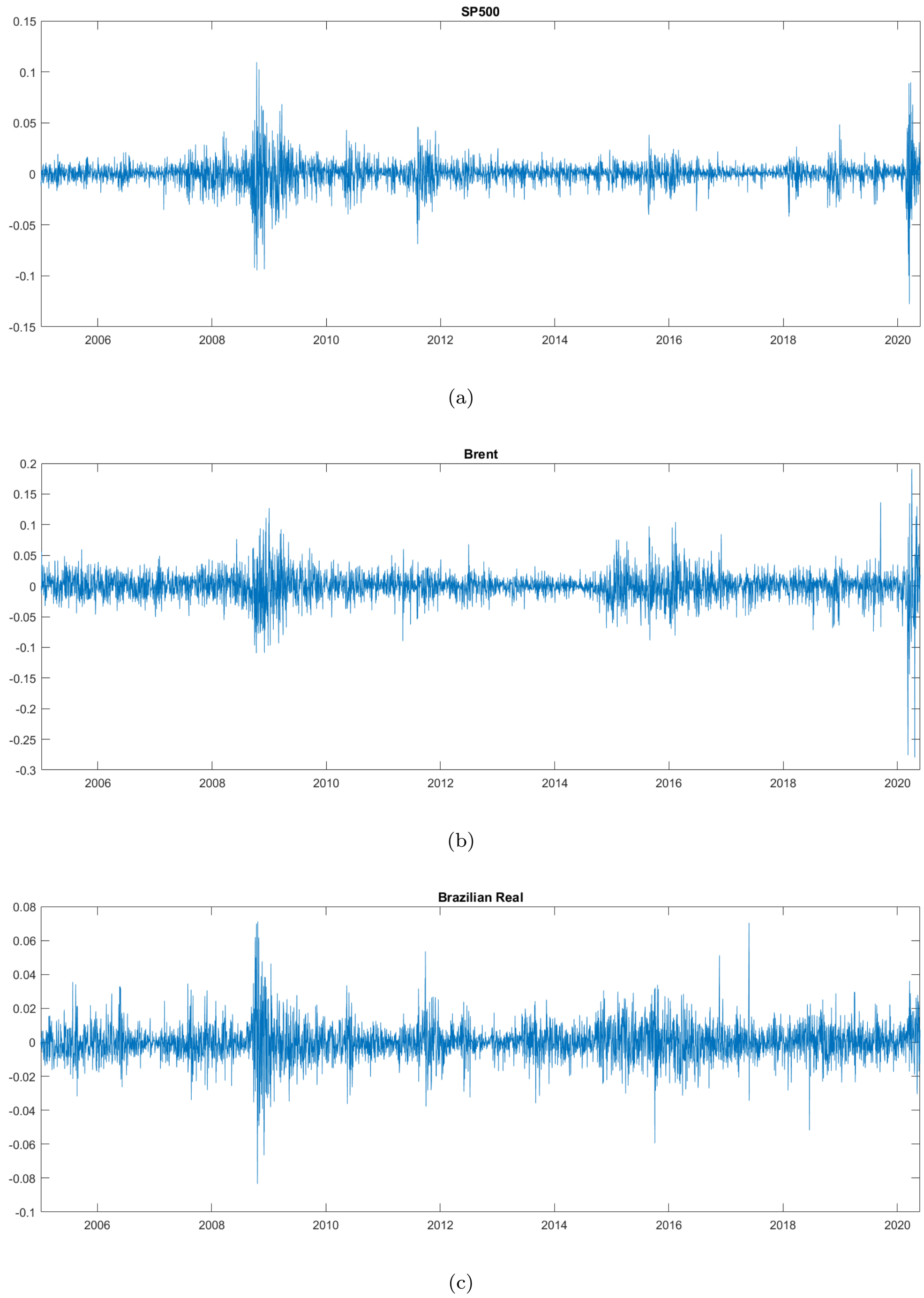

Figure 4: The figures show the return series of the S& P500, Brent Oil and the Brazilian Real. The datasets comprise from 4 January 2005 to 26 May 2020 (3940 observations). Data source: Refinitiv.

Table 1: The table shows the success ratios for standard responsiveness measures when comparing betwen pairs of models with different λ. $\Gamma(\lambda)$ represents the outcome of applying the responsiveness measure Γ to the set of scenarios. For each simulation, success is interpreted as having $\Gamma(\lambda_i) > \Gamma(\lambda_j)$ when $\lambda_i < \lambda_j$. The table also shows the ratio of coincidence between the different measures.

	$\Gamma()$	$\Gamma(96) > \Gamma(97)$	$\Gamma(97) > \Gamma(98)$	$\Gamma(98) > \Gamma(99)$	$\Gamma(99) > \Gamma(995)$	$\Gamma(96) > \Gamma(995)$
S&P500	PT	0.83	0.91	0.84	0.78	0.90
	SD	0.74	0.73	0.73	0.69	0.75
	MAX(1)	0.89	0.76	0.85	0.73	0.85
	LC(5)	0.91	0.87	0.83	0.76	0.91
	Coincidence	0.66	0.54	0.56	0.40	0.56
Brent	PT	0.94	0.92	0.80	0.75	0.90
	SD	0.81	0.78	0.68	0.61	0.72
	MAX(1)	0.95	0.83	0.75	0.65	0.93
	LC(5)	0.92	0.93	0.91	0.87	0.99
	Coincidence	0.75	0.69	0.49	0.39	0.68
Brazilian Real	PT	0.97	0.94	0.89	0.85	0.93
	SD	0.80	0.76	0.78	0.68	0.86
	MAX(1)	0.84	0.90	0.69	0.74	0.87
	LC(5)	0.93	0.99	0.88	0.81	0.98
	Coincidence	0.70	0.67	0.51	0.45	0.73

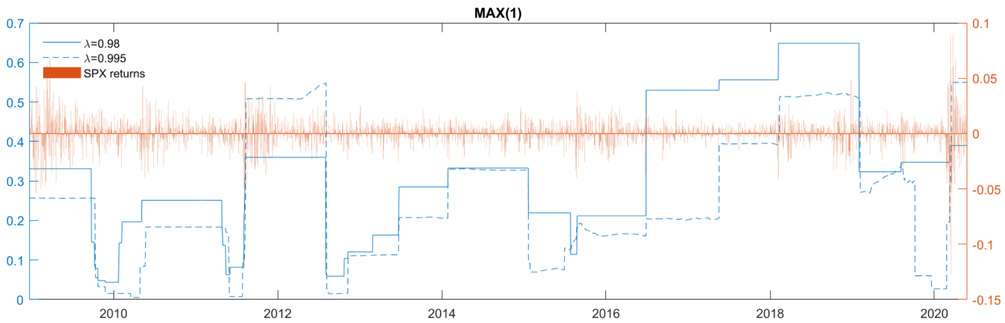

Figure 5: The figures show the MAX(1) values obtained when measuring model responsiveness over rolling samples of the S&P500 time series. The FHS models have $\lambda = 0.98$ and 0.995. As is typical of rolling window estimations, there are sharp jumps (ghost effects), produced by events falling out of the window.

5 Monte Carlo simulation

Since historical sampling limits the universe of possible outcomes to combinations of past scenarios, to obtain forward looking estimates and quantify uncertainty around the forecasts one can approach the question using Monte Carlo simulations. Broadly speaking, the process will consist of the following steps:

1. Estimate a GARCH specification fitted to a real series of returns (e.g., the S&P500), choosing two regimes: "normal" (N) and "stressed" (S).
2. For each regime, simulate 10,000 paths, each with 1,000 observations, as innovations from the corresponding GARCH process.
3. For each simulated path,
   - Estimate initial margin according to five margin models which, by design, have different levels of responsiveness.
   - Estimate model responsiveness using standard measures: SD, PT, MAX(1), and LC(5).
   - Rank models.

4. Estimate the frequency of consistent rankings across simulations.

5.1 Data generating process (dgp)

For the data generating process, consider a GARCH(1,1) process with normal innovations,

$$X_t = \sigma_t z_t, \qquad z_t \sim \mathcal{N}(0,1) \tag{5}$$

$$\sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{6}$$

where the innovations z_t are independent and identically distributed (iid), and $\omega > 0$, $\alpha > 0$, and $\beta \geq 0$ are constants.

The sum $\alpha + \beta$ controls the persistence of volatility and α controls its *responsiveness*: larger α means that large price moves in one period generate large price movements in the next period.

Note that the additional condition $\alpha + \beta < 1$ ensures that (X_t, σ_t) admits a stationary ergodic distribution (the unconditional distribution). Glasserman and Wu (2016) also introduced the parameter κ, which is defined as the strictly positive solution to the equation

$$\mathbb{E}[(\alpha z^2 + \beta)^{\kappa/2}] = 1 \tag{7}$$

The parameter κ describes the tail of the unconditional distribution. Smaller values of κ corresponding to the heavier tails and to a wider gap between the average conditional volatility and the unconditional one Glasserman and Wu (2016).

For the purposes of identifying two different regimes, the GARCH specification is fitted to two subsets of the S&P 500 data. The first, labelled "normal (N)" will cover the period until October 2019. The second one, labelled "stressed (S)' will cover from December 2019 to May 2022, to capture the Covid-19 stress. Table 2 presents the estimation of the GARCH coefficients when fitted to the subsets of S&P 500 data.

Table 2: The table shows the parameters obtained from fitting a GARCH(1,1) specification to two subsets of the S&P 500 data. The first, labelled "normal' (N), covers the period until October 2019. The second one, labelled "stressed' (S), covers from December 2019 to May 2022.

	Normal (N)	p-value	Stress (S)	p-value
α	0.1207	0.0000	0.2628	0.0000
β	0.8566	0.0000	0.7159	0.0000
$\alpha + \beta$	0.9774		0.9786	
κ	4.85		2.60	
Unconditional vol	0.0102		0.0208	

To illustrate the testing design, Figure 6 shows an example of simulated path (1,000 observations), with two FHS-EWMA VaR margin estimations over the last 250 days of the path, using a 250-day lookback period.

For the last part of the analysis, for each simulated path and for the different margin models, the "procyclicality" of margins is estimated using standard responsiveness measures.

The models are then ranked according using these measures and the model showing lower responsiveness is selected.

Finally, the frequency of success in the selection (i.e., the number of times the procyclicality measure selected the less procyclical model) is estimated. The ratio of success relative to the number of simulations is a measure of the power of the test; the closer the ratio is to 1, the more powerful the test.

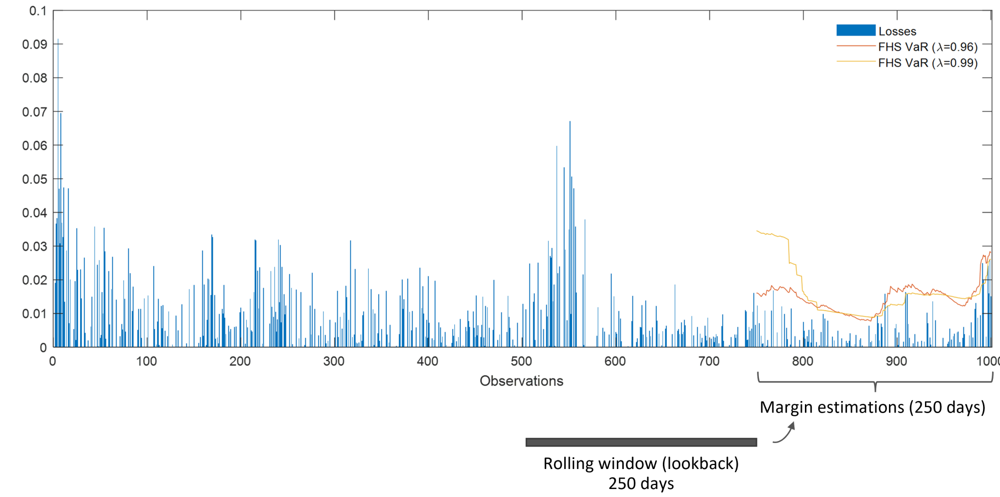

Figure 6: The figure illustrates the generation of margin for each simulated scenario. Given a 10000-day simulated path of returns, generated by a given GARCH specification, the FHS VaR for the last 250 days is estimated using 250-day rolling windows.

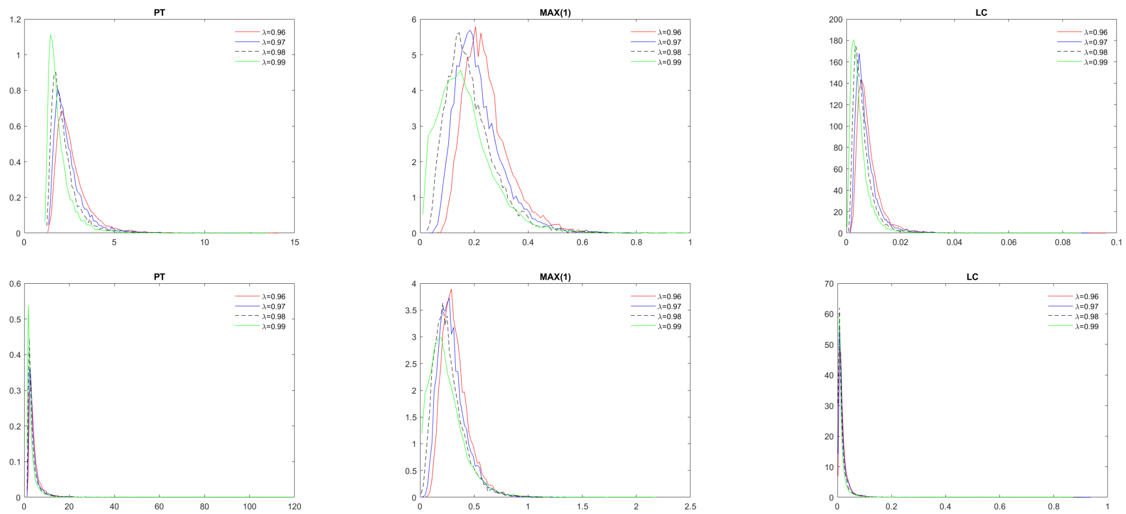

Figure 7: The figures show the distribution of values for each measure of responsiveness: the PT-ratio, MAX(1) and LC(5), and for each λ. The top (resp. bottom) row corresponds to the normal (resp. stressed) regime. Sampling was done considering 10,000 simulations of a GARCH(1,1) process.

5.2 Simulation results

Figure 7 shows the distributions of the measures Γ obtained for the different λ. In all cases there is significant dispersion around the mean, but this becomes more notable with $MAX(1)$. In all cases, there are large, extreme positive values. Finally, although the stressed period (S) appears to attract less disperse measurements, this is only a question of scale; they are in fact significantly more fat-tailed.

Table 3, shows the ratios of success, where success is equivalent of providing a ranking that is consistent with the responsiveness of the model, as defined by the parameter λ. In other words, if we compare to models, say one with $\lambda = 0.97$ and the other with $\lambda = 0.98$ we would expect that, when measuring procyclicality under measure Γ we would obtain $\Gamma(0.98) < \Gamma(0.97)$. When this is the case in a given simulation, we consider the ranking as "success". If this doesn't happen it means that the ranking provide by the measure Γ is reversed. The higher the success ratio, the more consistent is our measure in picking up the correct model.[9]

Table 3: The table shows the success ratios when comparing pairs of models with different λ. Success is defined as providing a ranking that is consistent with the responsiveness of the model, as defined by the parameter λ.

Regime	$\Gamma()$	$\Gamma96 > \Gamma97$	$\Gamma97 > \Gamma98$	$\Gamma98 > \Gamma99$	$\Gamma99 > \Gamma995$	$\Gamma96 > \Gamma995$
N	PT	0.92	0.89	0.85	0.83	0.95
	SD	0.76	0.76	0.76	0.82	0.87
	MAX(1)	0.79	0.75	0.65	0.46	0.74
	LC(5)	0.90	0.87	0.81	0.64	0.93
	Coincidence	0.60	0.56	0.48	0.35	0.64
S	PT	0.89	0.85	0.82	0.84	0.91
	SD	0.69	0.70	0.71	0.78	0.79
	MAX(1)	0.75	0.70	0.62	0.54	0.72
	LC(5)	0.88	0.85	0.83	0.75	0.94
	Coincidence	0.52	0.49	0.42	0.38	0.56

The results in Table 3 suggest that the resulting rankings are not very consistent. For example, PT consistently provides higher success ratios, but this is probably the less useful measure as it doesn't capture short-term changes. At the other extreme, it is clear that MAX(1) is the least reliable with success ratios as low as 46%.

When comparing across different λ, success ratios tend to decrease as λ closer to 1; that is, as models become less responsive, the discriminatory power of the measures decreases. On the other hand, coincidence between measures is very low (between 35% and 64%), which may not be surprising as they are measuring different things.

Finally, it is also interesting to note that, while the lack of consistency is more or less the same between regime N or S, the procyclicality measures tend to be slightly larger for the less volatile period (regime S). This may seem surprising, but only because there is a misconception that testing the model under stressed conditions would somehow increase the chances of the model performing well (procyclicality speaking) under all conditions. This is true when we are testing the model actual performance (its ability to target a coverage level) because there exists a link between events at the tail of the forecast distribution and the frequency and size of breaches. But such link is inexistent when talking about procyclicality. In other words, as the example shows, testing responsiveness under stressed conditions does not guarantee capturing the most extreme responsiveness values.

One could think that the differences leading to ranking reversals could be very small, in which case such reversals could be more a consequence of minor differences or noise. To test this, one can estimate the differences $\Gamma(\lambda_1)$-$\Gamma(\lambda_2)$ for pairs of models with $\lambda_1 < \lambda_2$. The distribution of these differences show that negative values (i.e., wrong rankings) are frequent and can be significant, as depicted in Figure 8.

These results highlight the importance of quantifying uncertainty in any procyclicality assessment.[10] A full statement of model procyclicality should reference the sensitivity to the different choices involved in our measurement. In other words, what we can only assert that "under a procyclicality measure Γ, and for a given type of scenarios Ω, we expect model M_1 to be less procyclical than model M_2, $X\%$ of the time".

[9]One could argue that a ratio close to zero can also be useful, as it shows consistency in the choice, even if the choice does not coincide with the decay factor. Indeed, from a decision-making perspective, the situation deteriorates as the ratio moves away from 1 or 0.

[10]As robustness test, one can verify the sensitivity of the ranking success ratios to the number of simulations. The results suggest stability after approximately 2000 simulations.

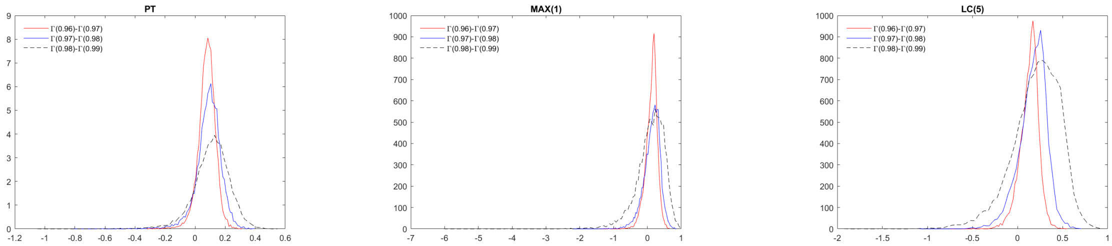

Figure 8: The figures illustrate the distribution of the differences between pairs of measurements

6 Uncertainty and procyclicality mitigation strategies

Intuitively, one can think of three broad approaches to moderate model responsiveness: recalibration of the model, introducing a drag or friction force, or overlaying a countercyclical component. Each will affect the model's responsiveness in different ways and to a different degree.

1. **Model (re)calibration** This could be the case if the model includes parameters that can affect responsiveness. For example, longer lookback periods in historical VaR tend to produce less responsive output. In FHS-EWMA VaR models, larger decay factors correspond to less responsive models too.

2. **Adding a drag or a friction force.** For example, averaging the model output with a less responsive component (the drag force) to slow down overall responsiveness. An example could be assigning a 25% weight to a component estimated using stressed observations, as in ESMA (2013). Other option is to create a friction force where the weight changes dynamically, so that higher weights are assigned in low-volatility conditions (and vice versa), as proposed in Murphy and Vause (2021). An approach where the drag force is based on implicit volatility, scaling up (resp. down) the key risk drivers during quiet (resp. stressed) periods, has been proposed by Wong and Zhang (2021).

3. **Adding countercyclical overlays**

   - **Exogenous to the model:** For example, a buffer that could be depleted when markets become stressed and replenished as markets return to normal. To be effective or even beneficial, the overlay needs to be correctly linked to the relevant "stressed/normal" cycle. [11] The "buffer" tool required by ESMA (ESMA (2013, 2022)), would be an example.

   - **Endogenous to the model:** For example, floors based on (some) model response to past stressed market conditions. For example, ESMA requires a floor defined by a fixed 10-year historical period (ESMA (2013, 2022)).

There may also be tools that combine two or more of these approaches. For example, the dynamic cap and floor of (Goldman and Shen, 2020) is based on minimizing a loss function with two competing objectives: risk sensitivity (model accuracy) and mitigation of procyclicality.

Different authors have already pointed out the limitations of the "buffer", the "stress", and the "floor", as defined in ESMA (2013); see, for example, Glasserman and Wu (2016); Maruyama and Cerezetti (2019); Murphy, Vasios and Vause (2016) or Gurrola-Perez (2022). But, from the perspective of the

[11] Countercyclical buffers used in the banking sector can be a related example. But in that case the countercyclical buffer is typically set up by the authority (e.g.,the ECB in the case of the EU) and its calibration is determined by a set of clearly defined economic indicators. Two problems with applying this mechanism to IM are that the financial indicators are unclear and the time-frame is much more shorter.

present analysis, it is worth noting that any of these "antiprocyclicality" (APC) tools is subject to parameter estimation risk, which will add up to any measurement errors in the procyclicality assessment. This is clearly the case, for example, for the buffer, where the choice of the times for consumption and replenishment can determine whether the buffer is beneficial, irrelevant, of even detrimental (Murphy, Vasios and Vause, 2016).

Moreover, the analysis of the impact of APC tools only makes sense under constraints on (maximum) cost and on (minimum) risk sensitivity Gurrola-Perez (2022). But any measure of cost which is a function of the margin levels, like the average margin required in a given sample period (Murphy and Vause, 2021; ESMA, 2022), presents the same challenges as measures of procyclicality: it is a random variable that depends on the dynamics of underlying process, and therefore there is some degree of uncertainty in its estimation.

So given 1) the uncertainty surrounding measurements Γ of cost and responsiveness, 2) the different parameter choices involved in calibrating an APC tool, and 3) the cost-benefit constraints imposed on the problem, how reliable an assessment of the effectiveness of these APC tools can be?

To illustrate this, consider the impact of adding a static floor, which is the simpler of the APC tools (it involves only one parameter, the floor level) and therefore involves less parameter estimation risk. As in Murphy and Vause (2021), margin cost, C, can be measured as the average margin required in a given sample period. To illustrate the impact of uncertainty when estimating the trade-off between model responsiveness and margin costs, the top panels in Figure 9 show the mean values together with the first and third quantiles of the distributions of $LC(5)$, $MAX(1)$ and C, for different floor levels.

The two top graphs show, for example, that an 10% average increase in IM costs corresponds to a 6.25% reduction in average LC(5) and a 9% reduction in average MAX(1). But there is substantial dispersion around the average values and, in the case of costs, the estimation error increases as the floor increases.

Although the above statistics already give an indication of the effects of uncertainty when measuring the impact of adding a floor, the whole picture only emerges if we look at the full distribution of outcomes. To illustrate this, the bottom graphs in Figure 9 show the points obtained for each individual simulation in the particular case when the floor level is set at $f = 0.015$. This floor level is the average, across all simulations, of the minimal margin value observed in each simulation. The graphs also indicate (with a black asterisk) the point corresponding to the average reduction in procyclicality. In both cases, the dispersion around the average values is substantial, confirming the limitations of relying on average procyclicality values to assess the impact of an APC measure without quantifying the uncertainty in those measurements. In fact, there are many instances in which the introduction of the floor increases costs by 20% or 25% without delivering any reduction of procyclicality at all.[12]

Finally, since there is no absolute reference or target, it is not known what an optimal trade-off (if any) could be.

7 Conclusions

By bringing the attention to the fact that standard responsiveness (procyclicality) measures are random variables on the space of P&L scenarios, the analysis has shown that the assessment of model responsiveness using these measures can be significantly sensitive to the choice of scenario and of underlying process. In some cases, forecasting which model among a set of acceptable models will be the less responsive in a future scenario is almost equivalent to throwing a coin. Moreover, the assessment is also sensitive to the measure chosen.

[12] In a few cases, procyclicality is reduced to zero, corresponding to situations in which the floor was high enough to make the margin irrelevant.

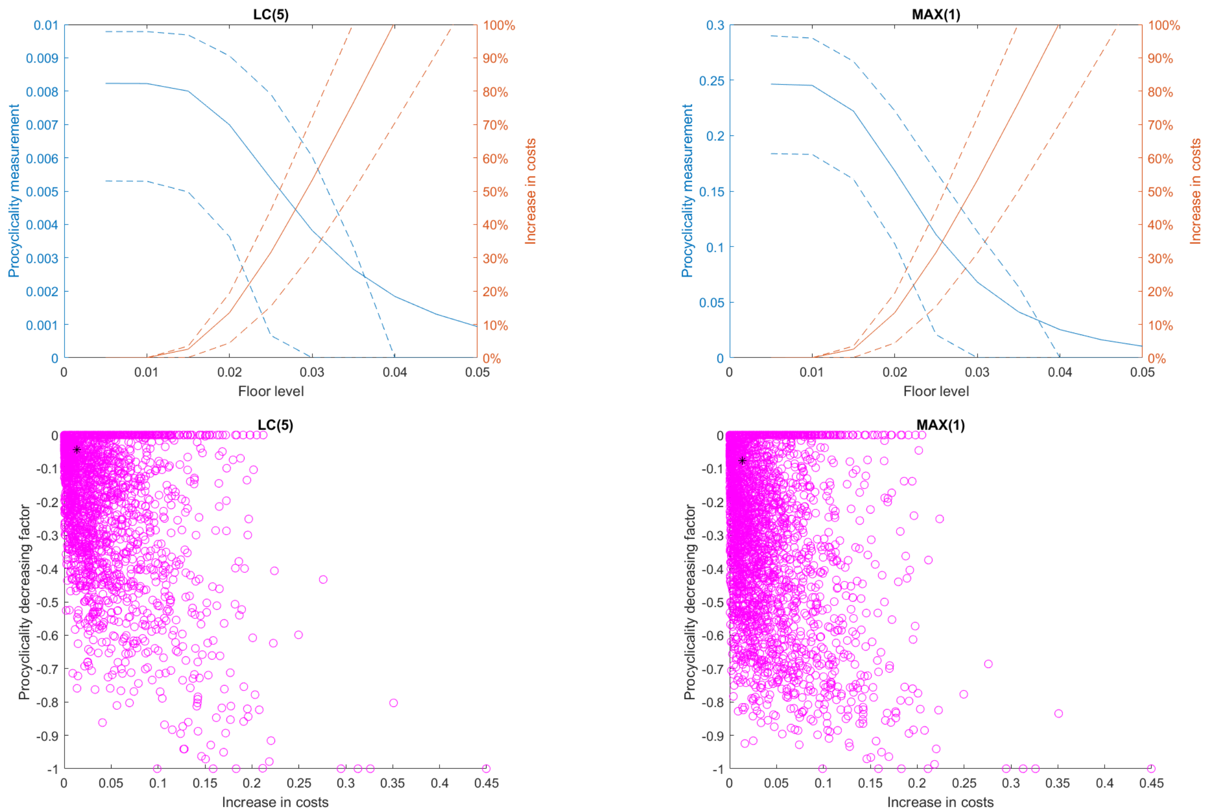

Figure 9: The figures at the top show the procyclicality values of LC(5) (left panel) and MAX(1) (right panel), and the estimated margin costs for different floor levels. The continuous line is the mean value and dotted lines represent 1st and 3rd quartiles. The model is an FHS VaR with $\lambda = 0.98$ *The bottom figures show the individual reduction in procyclicality and increase in cost points obtained for each simulation for a floor level of 0.015. The left panel shows the case of* $LC(5)$ *and the right panel shows the case of* $MAX(1)$*. The black asterisk shows the point corresponding to the average change in cost and in procyclicality: (0.0138,-0.0429) for* $LC(5)$ *and (0.0138,-0.0762) for* $MAX(1)$*.*

The significant uncertainty in measurements implies that, when estimating the trade-off between costs, procyclicality, and risk sensitivity, there is a considerable probability of getting the conclusions wrong for a large set of scenarios. This is even more true when we move out of model-land and into the real world, where events deviate from historical patterns and do not necessarily accommodate to our hypothetical scenarios: we can have a very good model (procyclicality speaking) and still have a probability of obtaining a more "procyclical" outcome under a future *real* scenario. And this will not be a failure of the anti-procyclicality credentials of the model, but only a consequence of the uncertainty in our predictions.

All the above highlights the importance that quantifying uncertainty has for a robust decision making process.

The results have also implications for policymaking.

For example, in addition to the fact that there is no intrinsic responsiveness benchmark associated to a model, large estimation errors in measurements of model responsiveness imply that requirements to define hard thresholds or targets and to implement specific tools in case of breaching those thresholds could be inherently ineffective.

Uncertainty around the costs and procyclicality measurements, together with the additional model risk brought by APC tools, also means that there is some non-negligible probability that, in some future scenarios, we will find these tools to be miscalibrated, inefficient, or even detrimental. This, together with the empirical evidence that their effectiveness can be very sensitive to the calibration and to the type of portfolios on which they are applied (Kahros and Weissler, 2021), should advise against prescribing them as one-size-fits-all solutions.

The above does not mean that CCPs should stop adopting margin arrangements that, “to the extent practical and prudent, limit the need for destabilizing, procyclical changes” (PFMIs). But the presence of uncertainty in model procyclicality forecasts draws a limit to what can be achieved through model-focused, hard-rule approaches, and underlines the importance of expert judgement to address model responsiveness on a case-by-case basis.

References

Abruzzo, Nicole and Yang-Ho Park (2014). An Empirical Analysis of Futures Margin Changes: Determinants and Policy Implications. Finance and Economics Discussion Series 2014-86, Board of Governors of the Federal Reserve System (U.S.), revised 02 Sep 2014.

Bank of England (2020). What role did margin play during the COVID-19 shock? Bank Overground, 10 June 2020. https://www.bankofengland.co.uk/bank-overground/2020/what-role-did-margin-play-during-the-covid-19-shock

Benos, Evangelos, Gerardo Ferrara, and Angelo Ranaldo (2022). Collateral cycles. Bank of England Staff Working Paper 966. March 2022

CFTC: Commodity Futures Trading Commission (2013). Enhanced risk management standards for systemically important derivatives clearing organizations, Final Rule 17 CFR Part 39, August. www.cftc.gov/ucm/groups/public/@lrfederalregister/documents/file/2013-19791a.pdf.

Cominetta, Matteo, Michael Grill, and Audrius Jukonis (2019). Investigating initial margin procyclicality and corrective tools using EMIR. European Central Bank, Macroprudential Bulletin, 29 October 2019. https://www.ecb.europa.eu/pub/financial-stability/macroprudential-bulletin/html/ecb.mpbu201910_5~6c579ba94e.en.html#toc1

CGFS: Committee on the Global Financial System (2010) The role of margin requirements and haircuts in procyclicality. https://www.bis.org/publ/cgfs36.pdf.

CPMI: Committee on Payment and Settlement Systems (2012) Principles for Financial Market Infrastructures. http://www.bis.org/cpmi/publ/d101a.pdf.

CPMI: Committee on Payment and Settlement Systems (2022) Review of Margining Practices. September 2022. https://www.bis.org/bcbs/publ/d537.pdf.

Dowd, Kevin (2005). Measuring market risk, 2nd Edition, Wiley & Sons.

ESMA: European Securities and Markets Authority (2013). European Markets Infrastructure Regulation (EMIR), Regulatory technical standards on requirements for central counterparties. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:052:0041:0074:EN:PDF.

ESMA: European Securities and Markets Authority (2018). Guidelines on EMIR Anti-Procyclicality Margin Measures for Central Counterparties. https://www.esma.europa.eu/sites/default/files/library/esma70-151-1496_guidelines_on_ccp_apc_margin_measures.pdf.

ESMA: European Securities and Markets Authority (2022). Consultation paper- Review of RTS No 153/2013 with respect to procyclicality of margin. https://www.esma.europa.eu/sites/file/122618.

ESRB: European Systemic Risk Board (2020). Mitigating the procyclicality of margins and haircuts in derivatives markets and securities financing transactions, ESRB Report, January 2020.

Faruqui, Umar, Wenqian Huang, and Elöd Takáts (2018). Clearing risks in OTC derivatives markets: the CCP-bank nexus. BIS Quarterly Review, December 2018.

FSB: Financial Stability Board (2020). Holistic review of the March market turmoil. Financial Stability Board, 2020. https://www.fsb.org/wp-content/uploads/P171120-2.pdf

Gerba, Eddie and Katsoulis, Petros, The Repo Market Under Basel III (December 17, 2021). Bank of England Working Paper No. 954, Available at SSRN: https://ssrn.com/abstract=4012925 or http://dx.doi.org/10.2139/ssrn.4012925

Glasserman, Paul and Qi Wu (2016). Persistence and Procyclicality in Margin Requirements. Office of Financial Research, Working Paper 17-01.

Goldman, E., and Shen, X. (2020). Procyclicality mitigation for initial margin models with asymmetric volatility. The Journal of Risk 22(5), 1–41 (https://doi.org/10.21314/JOR.2020.435).

Gurrola-Perez, Pedro and David Murphy (2015). Filtered historical simulation Value-at-Risk models and their competitors. *Bank of England Staff Working Paper* No. 525. March 2015.

Gurrola-Perez, Pedro (2018). The validation of filtered historical Value-at-risk models. Journal of Risk Model Validation 12 (1), 85-112.

Gurrola-Perez, Pedro (2022). The procyclicality of CCPs initial margin models: systemic problems need systemic solutions. *Journal of Financial Market Infrastructures* (2022).

Houllier, Melanie and David Murphy (2017). Borderline: judging the adequacy of return distribution estimation techniques in initial margin models. Bank of England Staff Working Paper No. 673. https://www.bankofengland.co.uk/-/media/boe/files/working-paper/2017/borderlinejudgingtheadequacyofreturndistributionestimationtechniquesininitialmarginmodels.pdf

Houllier, Melanie and David Murphy (2017). Initial margin model sensitivity analysis and volatility estimation. *Journal of Financial Market Infrastructures* 5(4), 77–103. DOI: 10.21314/JFMI.2017.078

Hull, John and Alan White (1998). Incorporating volatility updating into the historical simulation method for Value-at-Risk. *Journal of Risk*, Volume 1, 1998.

Kahros, A. and Marco Weissler (2021). Mitigating margin procyclicality – effectiveness of APC measures during the COVID-19 stress. Unpublished manuscript. Presented at the WFE Clearing and Derivative Conference 2021.

King, Thomas, Travis D. Nesmith, Anna Paulson, and Todd Prono (2020). *Central Clearing and Systemic Liquidity Risk*, Finance and Economics Discussion Series 2020-009. Washington: Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2020.009.

Maruyama, Atsushi and Fernando Cerezetti (2019). Central counterparty anti-procyclicality tools: a closer assessment. *Journal of Financial Market Infrastructures*, 7 (4), June 2019, 1-25. DOI: 10.21314/JFMI.2018.110 https://www.risk.net/journal-of-financial-market-infrastructures/6785891/central-counterparty-anti-procyclicality-tools-a-closer-assessment.

Murphy, David, Michalis Vasios, and Nicholas Vause (2014). An investigation into the procyclicality of risk-based initial margin models. *Bank of England Financial Stability Paper* No. 29.

Murphy, David, Michalis Vasios, and Nicholas Vause (2016). A comparative analysis of tools to limit the procyclicality of initial margin requirements. Bank of England Staff Working Paper No. 597.

Murphy, David and Nicholas Vause (2021).A CBA of APC: analysing approaches to procyclicality reduction in CCP initial margin models. Bank of England Staff Working Paper No. 950

O'Neill, Cyan and Nicholas Vause (2018). Macroprudential margins: a new countercyclical tool? Bank of England Staff Working Paper no. 765.

Thompson, E. and L. A. Smith(2019). Escape from model-land. *Economics*, Discussion paper no 2019-23, March 08, 2019. http://www.economics-ejournal.org/economics/discussionpapers/2019-23

Wong, Lauren W. and Yang Zhang (2021). Procyclicality control in risk-based margin models. *Journal of Risk* 23(5), 1-24. DOI: 10.21314/JOR.2021.010