WFE RESEARCH WORKING PAPER NO. 6

THE IMPULSIVE APPROACH TO PROCYCLICALITY

Measuring the reactiveness of risk-based initial margin models to changes in market conditions using impulse response functions

PEDRO GURROLA-PEREZ, DAVID MURPHY

APRIL 2024

WORLD FEDERATION OF EXCHANGES

The Impulsive Approach to Procyclicality

Measuring the reactiveness of risk-based initial margin models to changes in market conditions using impulse response functions

Pedro Gurrola Perez* David Murphy†

This version: 15th February, 2024

Abstract

In recent years, many derivatives market participants received large margin calls in episodes of elevated market volatility such as the onset of the Covid-19 global pandemic and the illegal Russian invasion of Ukraine. The lack of some market participants' preparedness to meet margin increases resulted in liquidity stress and reinvigorated the policy debate about how reactive margin should be to changes in market conditions. This debate has been hampered by the lack of a generally accepted way of measuring the reactiveness of the models used to calculate initial margin. The first contribution of this paper is to provide such a measure. We consider a step function in volatility, and examine the responses of various initial margin models to paths of risk factor returns consistent with this impulse, introducing the impulse response function as a convenient means of presenting this reaction.

The results presented demonstrate that a model's impulse response is a robust and useful measure of its reactiveness. This approach could be used both to measure initial margin model reactiveness, or procyclicality as it is often termed, and to capture the uncertainty in this measurement. It also provides significant, novel insights into the behaviour of some economically important margin models. In particular, the tendency of some filtered historical simulation value at risk models to over-react to sharp stepwise increases in volatility is demonstrated and the reasons for it are explored. The behaviour of two widely-used anti-procyclicality tools, the buffer and the use of a stressed period, are also analysed: the latter is found to be more successful at mitigating procyclicality than the former. The paper concludes with a discussion of the policy implications of the results presented.

Keywords: Anti-procyclicality, impulse response function, initial margin model, margin model response, procyclicality, volatility estimation

JEL codes: G13, C52, C12

Acknowledgements: The authors would like to thank Fernando Cerezetti, Viktoria Hackenberg, Richard Haynes, Rohit Koovarjee, Rama Kumanduri, Travis Nesmith, Lauren Wong, Michael Wood, the two anonymous referees from the WFEClear conference, and participants at the CFTC seminar on clearing risk for helpful comments on earlier versions of the paper.

Declaration of Interest: Pedro Gurrola-Perez is Head of Research at the World Federation of Exchanges, a trade association of exchanges and CCPs. David Murphy reports no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Disclaimer: The views expressed here are the authors' and do not necessarily represent those of the WFE or its members, or the London School of Economics and Political Science.

*Head of Research, World Federation of Exchanges. *Email*: pgurrola@world-exchanges.org

†Visiting Professor in Practice, Law School, London School of Economics and Political Science, Houghton Street, London, UK. *Email*: D.Murphy3@lse.ac.uk.

1 Introduction

The early 2020s have been characterised by a number of episodes of elevated volatility in financial markets. These include the onset of the Covid-19 global pandemic in March 2020 and the illegal Russian invasion of Ukraine in February 2022. In these episodes, there was a large, broad-based increase in margin requirements across the financial system, with commodity markets particularly affected in the latter episode, as discussed in [FSB 2023]. Some market participants appeared to be unprepared for this, and hence suffered liquidity stress. The question therefore naturally arose whether the margin calls were justified. This partly turns on the reactiveness of initial margin to changes in market conditions, the topic of this paper.

Margin calls in periods of increased market volatility are an important matter as most derivatives (and many other types of transactions) are margined, so changes in margin requirements can have a broad impact.[1] The use of central counterparties ('CCPs') is widespread in modern derivatives markets, in part due to regulatory requirements,[2] so the behaviour of CCP margin models is particularly important here. When market prices move substantially, as they do in periods of market stress, there are both initial and variation margin calls: the former because initial margin (or 'IM') is calculated using risk sensitive models, and these models, especially the models typically used by large CCPs, often react to increases in volatility; the latter because variation margin (or 'VM') settles or collateralises the value of affected portfolios, and these values change substantially.

A recent authority report into the events of the early Covid period[3] confirms this account, saying:

> *The aggregate changes in stocks and flows of margin differed in size across markets and CCPs. Daily CCP VM calls were large, and significantly higher than the average flows observed between January and February 2020, increasing from around $25 billion to a peak of $140 billion... The total IM requirement across CCPs increased by roughly $300 billion over March 2020... Market volatility and model reactions to volatility were responsible for the majority of the peak increase in IM requirements, with changes in volumes and risk positions playing a smaller role.*

Broadly, this is approximately a 40% increase in IM. To put this into context, the VIX volatility index increased by about 400% in the same period and, according to [BCBS et al. 2022], margin calls consumed less than 2.5% of the liquid resources of large intermediaries. Nevertheless, meeting margin calls during these episodes created substantial funding stress on some market participants and, in a subset of cases, prompted authority intervention.[4]

The issue is particularly pertinent because it has been known for many years that initial margin models can sometimes over-react to increases in market volatility. This 'excessive procyclicality' is undesirable, as concerns about margin calls, or the calls themselves, can force market participants to adjust positions just when market conditions are inhospitable or even, in extremis, force them to default.[5] Thus, the key post-crisis international standard for CCPs, the *Principles for Financial Market Infrastructures*, or PFMIs, requires that:[6]

> *A CCP should appropriately address procyclicality in its margin arrangements... in a period of rising price volatility or credit risk of participants, a CCP may require additional initial margin for a given portfolio... This could exacerbate market volatility further, resulting in additional margin requirements.*

However, what it means to 'appropriately address procyclicality' is disputed, as we discuss next.

---

[1] Policy changes after the 2008 global financial crisis are a major factor in the ubiquity of margin. As [BCBS and IOSCO 2013] details, significant bilateral OTC derivatives market participants are required to post margin to each other, and many derivatives must now be cleared, either because they are exchange-traded, and clearing is ubiquitous for exchange-traded derivatives, or because they are subject to a clearing mandate. Cleared transactions are typically subject to margin requirements.

[2] In addition to the clearing mandate, discussed the previous footnote, there are significant incentives to clear in many cases: see [FSB 2018] for more details.

[3] See [BCBS et al. 2022].

[4] See, for instance, the UK Treasury Energy Markets Finance Scheme, 2022, which was introduced to address "the extraordinary liquidity requirements faced by energy firms operating in UK wholesale gas and electricity markets as a result of margin calls".

[5] For a further discussion of the risks created by margin-induced funding stress, see [Bakoush et al. 2019], [King et al. 2020] and [Brunnermeier and Pedersen 2009].

[6] See [CPMI and IOSCO 2012] at 3.6.10.

1.1 Issues with measuring initial margin model reactivity

There is general agreement that initial margin should not be excessively procyclical. The difficulty is understanding what is an over-reaction to market conditions, and what is merely (desirable) risk sensitivity. This is a sufficiently important problem that the authorities recently proposed further work "to understand the degree and nature of CCP margin models' responsiveness to volatility and other market stresses and to explore appropriate ways to analyse, compare and set baseline expectations as to [initial margin model] procyclicality in various settings": see [BCBS et al. 2022].

Understanding the degree and nature of CCP margin models' responsiveness is difficult because initial margin models take a representation of a portfolio and time series of risk factor returns and produce a margin estimate for that portfolio. We can directly observe how a given model reacts to a given set of risk factor returns – such as those in March 2020 – but this only allows us to compare different models' reaction to the same episode, not to assess which reaction, if any, is *excessively* procyclical (and, by extension, which is not reactive enough). Moreover, a new stressed episode will undoubtedly have a new set of risk factor returns, and it might well produce a different procyclicality-ranking of models than the one suggested by the original episode.

The underlying issue here is that risk factor returns are random variables taken from some distribution. Even if the time series of the conditional variance of this distribution is known, there are many paths of risk factor returns consistent with this time series, and different paths usually lead to a different set of margin calls. If one return path samples the tails of the distribution a little more than another, it will typically lead to higher margin calls, as [Gurrola Perez 2023] investigates. A means of comparing the average degree of margin model reactivity to a change in market conditions, and the variability of this reactivity depending on the particular path of returns, is needed.

1.2 Policy context and prior work on initial margin model procyclicality analysis and mitigation

The 'excessive procyclicality' question relates to another important policy issue in derivatives markets, that of the need for 'anti-procyclicality' or 'APC' tools. European regulation[7] requires that CCPs incorporated in the EU use one of three prescribed tools to mitigate margin model procyclicality,[8] while CCP regulation in the United States has no such prescriptive requirement and instead follows a principles-based approach. Two major jurisdictions therefore differ on the need for CCPs to apply specific APC tools: EU CCPs must use one of three APC tools to mitigate their model's procyclicality no matter how large or small it is, while systemically important US derivatives CCPs[9] must simply demonstrate that they meet the PFMI requirement, using whatever approach they determine to be most appropriate. These differences further highlight the complexity of the procyclicality question.

These issues have been extensively studied. The variability of margin requirements, and the implications of these swings, are considered in [Cominetta et al. 2019] (using real data) and [Glasserman and Wu 2018] (using a theoretical model which emphasises the difference between through-the-cycle or unconditional margin and point-in-time or conditional margin). Most authors, including [Gurrola Perez 2021] and [Maruyama and Cerezetti 2019], as well as regulators, consider that this variability may give rise to systemic risks, although [Lewandowska and Glaser 2017] questions this.

The properties of the three APC tools required in Europe, various competitors to them, and their systemic context, have been extensively studied: see, for instance, [Gurrola Perez 2021], [Kahros and Weissler 2022], [Maruyama and Cerezetti 2019], [Murphy et al. 2014], [Murphy et al. 2016], [Murphy and Vause 2021] and [Wong and Zhang 2021].

For our purposes, there are four main threads in this literature:

- *The dimensions of margin procyclicality.* Different aspects of margin procyclicality are potentially of policy relevance. These include the degree of variation of margin over the long term; the size of short term margin calls during high volatility periods; and the types of market participants

---

[7] See [EU 2012] and, for a further discussion, [ESMA 2018].

[8] The EU position is currently under review, although there seems little prospect of the requirement to use one of the three tools being lifted: see [ESMA 2022].

[9] Only systemically important derivatives clearing organisations are required to be PFMI compliant in the United States.

affected (and hence their access to funding liquidity): see [Murphy et al. 2014] and [Murphy and Vause 2021].

- *Model variability.* Different, otherwise acceptable[10] initial margin models have significantly different reactions to changes in market conditions. Thus, it may be that some models require procyclicality mitigation while some do not.
- *The difficulty of designing good, general-purpose APC tools.* A good tool should be effective at mitigating procyclical margin calls for a wide range of portfolios in different market conditions at reasonable cost, while ensuring that the margin required for a portfolio is always prudent, given its risks. There are questions as to whether any of the APC tools currently proposed unambiguously pass this test, as [Kahros and Weissler 2022], [Maruyama and Cerezetti 2019] and [Murphy and Vause 2021] discuss.
- *The question of mutualization.* One means of reducing margin model procyclicality is simply to reduce the amount of risk covered by margin in a crisis and increasing the amount covered by the CCP's default fund. The (controversial) argument is that increasing mutualization in a crisis can be a good solution if it reduces the risk of funding-stress-created defaults.[11]

All of this suggests that the APC debate would be illuminated by a better measure of the reactiveness of an initial margin model, and hence the ability to require mitigation by the CCP only where a model over-reacts, and where that over-reaction creates funding liquidity risk for the affected parties.

1.3 Our contribution

In this paper, impulse response functions ('IRFs') are proposed as a tool for analysing margin model reactiveness.[12] This approach does not rely on particular risk factor paths, as many paths consistent with a given time series of conditional volatility are considered. The average response across all simulations is a good measure of a model's typical reaction, while low (5%) and high (95%) percentiles of the distribution of reactions illustrate the potential variation in response for different risk factor paths consistent with the time series of volatility. This approach therefore gives substantial insights into model reactiveness. Section 2 describes the IRF idea, illustrating the reaction of several well-known volatility estimators to a simple scenario (or 'impulse') where volatility increases.

Section 3 introduces a number of standard initial margin models and illustrates their reaction to stressed episodes by presenting their impulse response functions for a step function increase in volatility and thereby illustrating the different degrees of reactiveness of different models. One of the classes of model introduced, filtered historical simulation ('FHS'), is widely used by large CCPs.[13] Interestingly, we uncover a situation where the margin estimates produced by FHS models which use an EWMA volatility filtering scheme are biased, further substantiating the utility of the IRF technique.

APC tools are then considered. Section 4 presents the impulse response functions of all of the models considered with two of the EU APC tools – the buffer and the stressed period, both discussed below

[10] 'Acceptable' in the sense of passing tests for risk sensitivity, such as backtesting.

[11] One could argue that a margin model need only cover its target confidence level *unconditionally*, so that a model targeting, say, 99%, need only ensure that margin is sufficient to absorb 99% of the unconditional return distribution of a portfolio, not 99% of conditional returns. This would require other resources (such as the default fund in a CCP or equity capital in a bilateral situation) to support the risk in excess of the 99th percentile of unconditional returns. This can result in vastly reduced margin procyclicality, and is broadly the approach adopted in the industry standard bilateral OTC derivatives margin model (albeit with a target confidence interval much in excess of 99%): see [ISDA 21] for a discussion of the bilateral situation, and [Goldman and Shen 2020], [Raykov 2018] and [Wang et al. 2022] for a discussion of an approach to cleared initial margin procyclicality mitigation which mutualise more risk in higher volatility periods.

[12] This reactiveness is an important (but not the only) contributor to margin procyclicality. Three effects are often identified: variation margin, as a result of price changes; the fact that initial margin models typically produce a percentage risk estimate, so if this is applied to a higher price, initial margin increases; and changes in the risk estimate itself. We capture the second and third of these effects.

[13] For the purposes of our analysis, we test some specific, relatively simple, types of FHS models. The FHS models used by large CCPs are not necessarily of this type, nor are they always calibrated in the same way: see [Barone-Adesi et al. 1998], [Boudoukh et al. 1998] and [Hull and White 1998] for a further discussion of FHS model types. Moreover, even where CCPs do use the type of model we discuss, this is sometimes not the only approach used in the determination of initial margin.

– applied. The IRFs of the mitigated models provide useful insights into the behaviour of these tools. Measures of procyclicality introduced in earlier literature and a measure of model reactiveness are also presented for the models with and without the use of an APC tool. As before, the average across all simulations of the measures provides insights into the typical procyclicality and reactiveness, while high and low percentiles illustrate the variability of response. In particular, we find a significant, and previously undocumented over-response of FHS models.

Stressed episodes in markets are not just characterised by an increase in volatility. The tails of the return distribution may fatten too. In order to illustrate the effect of this phenomenon, section 5 presents IRFs and procyclicality measures for an impulse with increased volatility and fatter tails. Section 6 concludes with a discussion of the policy implications of our analysis.

2 The Impulse Response Methodology

This section introduces the primary tool used to visualise the response of initial margin models in this paper, the impulse response function ('IRF'). The IRF of several common volatility estimators are presented: this illustrates the approach in a simple context.

2.1 Impulse response functions

The IRF of a dynamical system is its output when presented with a change in input, known as the impulse. It illustrates the reaction of the system as a function of time. IRFs are commonly used to analyse the response of components in signal processing applications such as audio processing or optics.[14]

Risk estimators can be thought of as dynamical system in that they typically take as input the prices of risk factors and some representation of a portfolio of financial instruments whose value is sensitive to these risk factors, and output a risk estimate. If we hold the portfolio fixed, the input can be seen as the impulse, and the change in risk estimate as the response.

For the sake of clarity, it may be helpful to discuss the proposed IRF methodology further. It is a tool that allows the reactiveness of market risk models to be measured (and, hence, compared) in a statistically robust way. That is, it provides procyclicality measurements which are a function of volatility and which are independent of any particular path. It provides the full probability distribution of measurements and therefore captures the uncertainty surrounding those measurements. This distribution is, in some sense, the 'signature' of a model, in that it captures exactly how the model's margin estimates can change in response to a change in risk factor dynamics. In contrast, because of their path dependency, procyclicality measurements based on individual historical or theoretical paths are not statistically robust and give very little sense of the range of responses possible.

2.2 The IRF for an unweighted volatility estimator

This idea can be illustrated with volatility estimation. Consider a single risk factor, and suppose that its volatility is constant for some period, then increases, and stays at the new higher level for a period: a 'step function' increase. A very simple form for the distribution of returns $r(t)$ on day t will be assumed:

$$r_t \sim \mathbf{N}(0, \sigma(t))$$

where $\sigma(t)$ is the chosen step function and $\mathbf{N}(\mu, \sigma)$ is the normal distribution with mean μ and standard deviation σ.

Volatility estimators are based on risk factor returns, so in order to examine the range of possible responses of a given estimator, we need to:

- Simulate a path of returns consistent with the intended volatility impulse;
- Use the volatility estimator to calculate an estimate of volatility along this path;
- Repeat the process many times to obtain a distribution of volatility estimates at each point in time.

[14] [Najim 2006] gives an introductory account of the use of IRFs in signal processing.

A perfect estimator would accurately track the impulse: in reality, all estimators are error prone and lag, as evidence about changes in volatility conditions accumulates slowly.

To make this concrete, we will take a step function where the volatility of daily returns is constant at 1% for 500 days, then jumps to 3% and stays there for another 500 days: other choices of volatility dynamics are of course possible. A very simple volatility estimator is the standard deviation of returns in some window: 250 days, say. Figure 1 illustrates the IRF of this volatility estimator using 200,000 simulations. True volatility is in grey and average estimated volatility is in purple. We also capture the 5th and 95th percentiles of the estimates of volatility in light blue: this illustrates the range of estimates possible along different paths consistent with the chosen impulse.

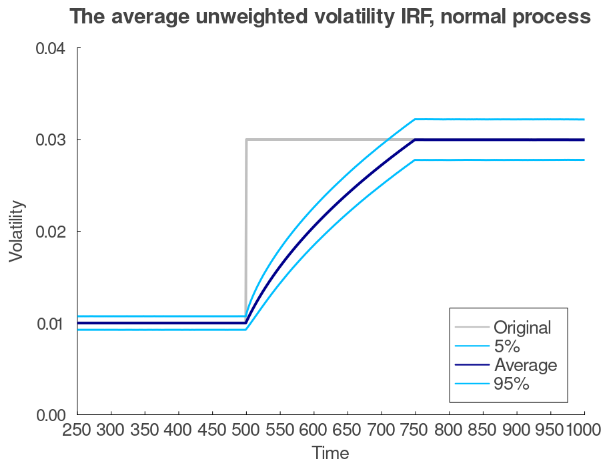

Figure 1: The response of an unweighted volatility estimator to a step function increase.

The unweighted volatility estimator converges close-to-linearly to the correct value, hitting this value once the returns with the lower volatility are no longer used in the window that determines the estimate of volatility. There is some spread between the percentiles, corresponding to the fact that some paths preferentially sample to tails of the distribution of returns, leading to higher estimates, while some oversample the centre, leading to lower estimates.

2.3 The IRFs for exponentially weighted moving average volatility estimators

It is well-known that unweighted volatility estimators converge slowly to the true value: as a result, exponentially weighted moving average volatility ('EWMA') estimators are often used. These estimate the volatility for day t, $\widetilde{\sigma}_t$, as:

$$\widetilde{\sigma}_t^2 = \lambda \widetilde{\sigma}_{t-1}^2 + (1 - \lambda) r_{t-1}^2$$

Here $0 \leq \lambda \leq 1$ is known as the decay parameter: lower λ EWMA estimators forget old returns faster. This can be seen in the impulse response functions: Figure 2 shows the IRFs for $\lambda = 0.97$ and $\lambda = 0.99$ EWMA volatility estimators. Both estimators have convex convergence, and the lower λ estimator converges more quickly than the higher λ one. Note too that, because the lower λ EWMA volatility estimators weights more recent data more highly, variation in those returns has a higher impact on the estimate. Therefore, the confidence bounds on the estimators are wider: this is the price of faster convergence.

3 Results for Margin Models without Procyclicality Mitigation

The IRF analysis technique will be applied to a selection of standard margin models in this section. Specifically, a portfolio which is long one unit of a risk factor will be considered, and six different methods of calculating margin will be used. We begin by setting out the models, then present their IRFs.

3.1 Margin models

All of the models will target the 99th percentile of returns. This percentile will be estimated using the following approaches:

1. Historical simulation ('HS') value at risk ('VAR'), with a 250 day window;

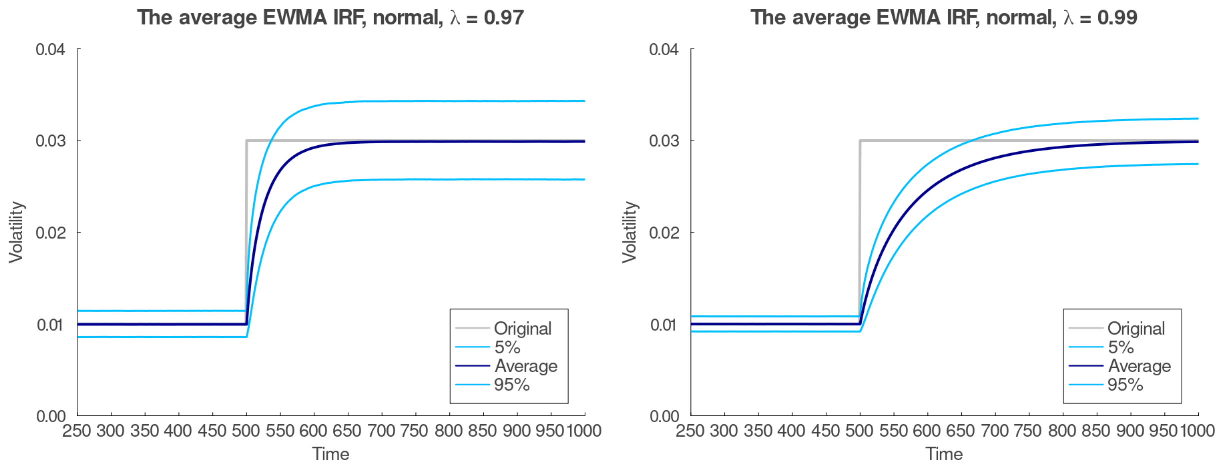

Figure 2: The response of $\lambda = 0.97$ and $\lambda = 0.99$ EWMA volatility estimators to a step function increase.

2. Parametric VAR, using a normal distribution assumption and estimating volatility using an unweighted historical estimate (i.e. the first of our volatility estimators above);
3. Parametric VAR, using a normal distribution assumption and estimating volatility using an EWMA estimate with $\lambda = 0.97$ (i.e. the second of our volatility estimators above);
4. Parametric VAR, using a normal distribution assumption and calculating volatility using an EWMA estimate with $\lambda = 0.99$;
5. Filtered historical simulation ('FHS') VAR, using an EWMA estimate with $\lambda = 0.97$;
6. Filtered historical simulation ('FHS') VAR, using an EWMA estimate with $\lambda = 0.99$.

The first four approaches are very well-known. FHS is slightly less so, although it is widely used by CCPs for margin calculation. The central idea is that instead of using the returns in an N-day window $[r_{t-N}, r_{t-N+1}, \ldots, r_{t-1}]$ to calculate VAR at day t, as in historical simulation, filtered returns are used instead. These are

$$\left[\frac{r_{t-N}}{\widetilde{\sigma}_{t-N}}, \frac{r_{t-N+1}}{\widetilde{\sigma}_{t-N+1}}, \ldots, \frac{r_{t-1}}{\widetilde{\sigma}_{t-1}}\right]$$

where $\widetilde{\sigma_T}$ is volatility estimated using EWMA, at time T as above. For a given filtered return $\frac{r_T}{\widetilde{\sigma}_T}$, the denominator is known as the 'devol' or 'devolatilise' volatility.

The FHS VAR at α confidence for day t is then:

$$\widetilde{\sigma_t} \times Q_\alpha \left[\frac{r_{t-N}}{\widetilde{\sigma}_{t-N}}, \frac{r_{t-N+1}}{\widetilde{\sigma}_{t-N+1}}, \ldots, \frac{r_{t-1}}{\widetilde{\sigma}_{t-1}}\right]$$

where $\widetilde{\sigma_T}$, which rescales the standardised (or devolatilised) returns, is the forecast or 'revol' volatility for the day risk is being estimated on, and Q_α is the α-quantile of the (finite) distribution of devolatilised returns. The accurate estimation of quantiles for VAR calculation is a delicate matter: see Appendix 1 for a discussion of this issue, and [Gurrola Perez and Murphy 2015] and [Gurrola Perez 2018] for further motivation and analysis of FHS VAR models.

3.2 IRFs for the unmitigated models: historical simulation VAR

Figure 3 shows the IRF for the first margin model, historical simulation VAR. Here the average margin across the simulations is in dark green, and the 5th and 95th percentiles are in sea green: the correct level of margin based on the true volatility of returns is shown in gray. Margin calculated using historical simulation is unbiased. It is imprudently low immediately after the step-up, and converges to the true VAR as the lower volatility returns disappear from the estimation window. The error bounds for margin increase with volatility.

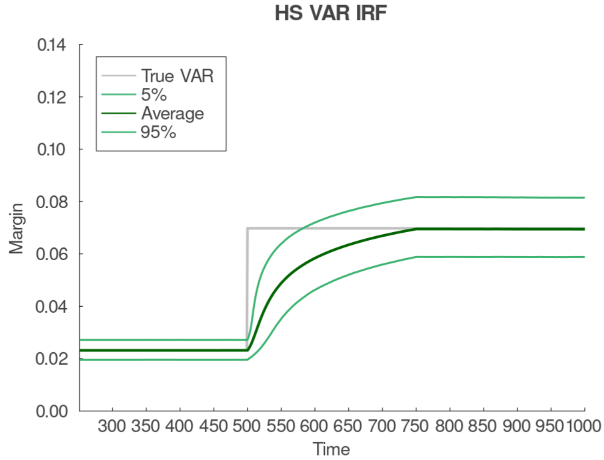

Figure 3: The response of an unweighted historical simulation VAR model to a step function increase.

3.3 IRFs for the unmitigated models: parametric VAR

Parametric VAR with an unweighted volatility estimator converges nearly linearly to the true VAR, as Figure 4 illustrates. It too has an initial period where margins are below the true VAR after the step-up.

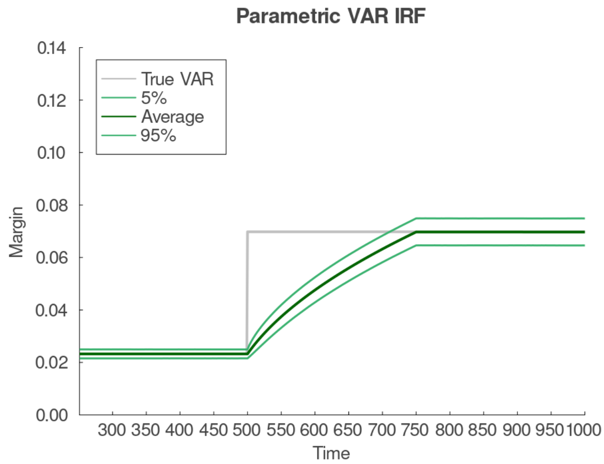

Figure 4: The response of an unweighted parametric VAR model to a step function increase.

If the model instead uses an EWMA volatility estimator, convergence is faster, with a speed determined by λ. Figure 5 illustrates the IRFs.

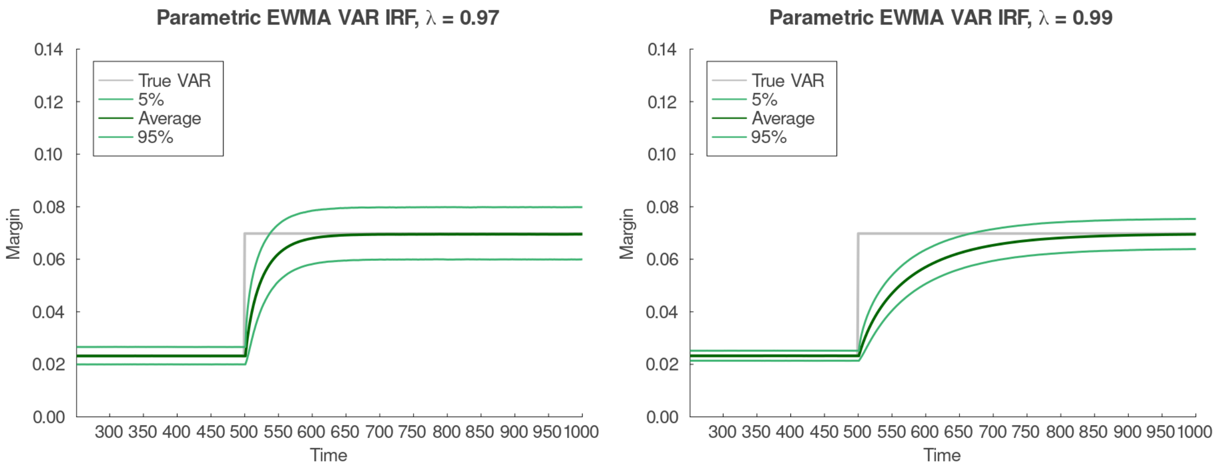

Figure 5: The response of $\lambda = 0.97$ and $\lambda = 0.99$ EWMA parametric VAR models to a step function increase.

The confidence bounds are tighter on the higher λ margin estimates, reflecting the longer effective history of data used in them. For the more reactive model in particular, the spread between margin estimates on different paths can be significant: the difference between the 5th percentile of margin and the 95th in the high volatility period is roughly three quarters of the true margin in the starting period. On a relatively unlucky path, a market participant could find their margin requirement was substantially higher than on a more fortunate one, even though both paths have the same structure of conditional volatility. This nicely illustrates the trade-offs involved in risk model design: the price of faster reactivity is more dispersion in the risk estimate.

3.4 IRFs for the unmitigated models: FHS VAR

The IRFs for the FHS VAR models analysed here are very interesting. They are presented in Figure 6. The models initially produce a margin below the true VAR, but then over-react. This is because the devol volatilities are too low immediately after the step, so the filtered returns are too big. Once these overly-large filtered returns have fallen out of the window, FHS converges to close to the true value. However, this only happens slowly because the filtered returns immediately after the step have very low devols, compared to the correct value, so they are much too large, and likely determine the VAR estimate until they fall out of the window 250 days later. The over-estimate of margin thus typically lasts for nearly the entire window.

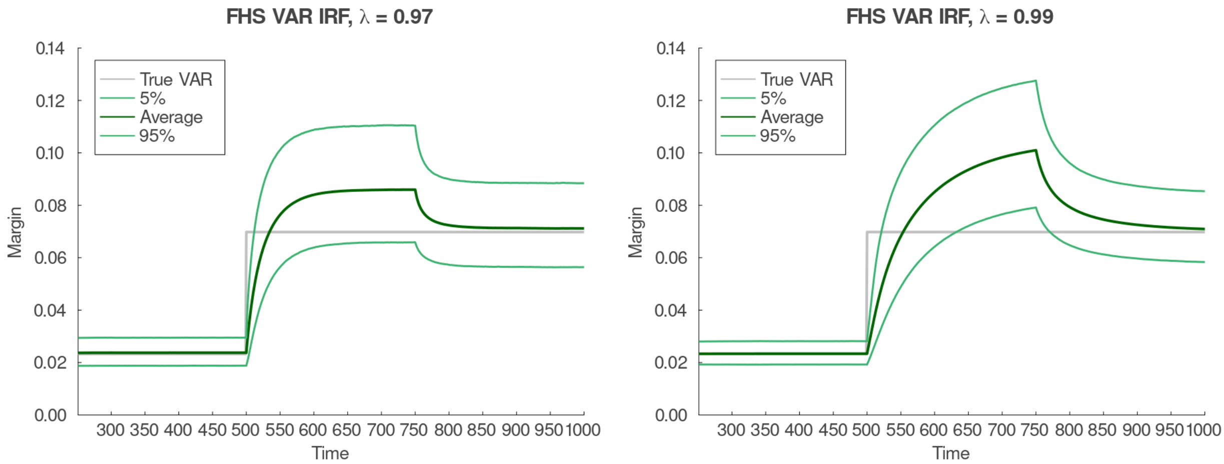

Figure 6: The response of $\lambda = 0.97$ and $\lambda = 0.99$ FHS VAR models to a step function increase.

The spread of estimates is wider for the higher λ model and estimates from this model over-react more. The latter is because the higher λ devol volatility is more wrong, as it takes longer to converge to the true value. The error in revol partially compensates for this, but revol, being at the end of the window, is closer to being correct than devol, so it does not fully compensate for the error in devol. This also suggests that, in situations where the build-up of volatility is more gradual or when using a conditional volatility estimator that provides a better fit, the error in devol volatility will be smaller and the tendency to over-react will decrease. The influence of window length is important here too: because the over-reaction in FHS only falls out of the filtered returns once they have left the window, the over-reaction lasts longer in models with longer windows.

3.5 Bias in FHS models

Even once an FHS model has converged to a new steady state, a small amount of bias is present. To see this, consider the returns used in historical simulation VAR estimation: $[r_{t-N}, r_{t-N+1}, \ldots, r_{t-1}]$. Clearly if the r_T are IID samples from zero-drift normal process $\mathbf{N}(0, \sigma)$, the standard deviation of these returns is an unbiased estimator of σ, and $Q_{0.99}$ of these returns converges to the correct value, $\Phi^{-1}_{0.99}(0, \sigma)$, where $\Phi^{-1}_{\alpha}(0, \sigma)$ is the inverse of the cumulative normal distribution with mean zero and standard deviation σ evaluated at α.

Now consider the filtered returns. These are:

$$\left[\frac{r_{t-N}}{\widetilde{\sigma}_{t-N}}, \frac{r_{t-N+1}}{\widetilde{\sigma}_{t-N+1}}, \dots, \frac{r_{t-1}}{\widetilde{\sigma}_{t-1}}\right]$$

Note that if r_T is normally distributed, then r_T^2 is χ-square distributed with one DF. The weighted sum of two χ-square variables is described by a scaled infinite sum of gamma distributions, but all that matters for our purposes is that it has non-zero variance. This applies to $\widetilde{\sigma_T}^2 = (1-\lambda)(r_{T-1}^2 + \lambda r_{T-2}^2 + \lambda^2 r_{T-3}^2 + \dots)$, and thus variance in EWMA volatility estimates widens the distribution of filtered returns compared to the unfiltered case. Thus, the variance of the distribution of the filtered returns is more than 1 in expectation. The consequence of this is that FHS does not converge to the true volatility for a process with normal returns: for $\lambda = 0.97$ with a 250 day window, the result is about 2% too high.

4 Results for Margin Models with APC Tools

European regulation requires central counterparties to use one of three anti-procyclicality ('APC') tools, as noted in the Introduction. In brief, these are:

1. A buffer of 25% of margin, released when conditions become more volatile.
2. Using a stressed period in margin estimation, so that margin is based 75% on current conditions and 25% on stressed ones;
3. A floor based on the average ten year unweighted volatility.

These tools are discussed in detail in the literature: see for instance [Murphy et al. 2014], [Murphy et al. 2016], [Maruyama and Cerezetti 2019], [Murphy and Vause 2021] and [Gurrola Perez 2021]. For our purposes, we will consider the first two tools, as the ten year floor has no effect once margin is above the floor level, which is typically the case in stressed episodes.

4.1 IRFs for margin models incorporating the buffer

A key issue in determining the effect of the buffer is when it is released. Here an omniscient risk manager who releases the buffer the day after volatility increases is assumed. Figure 7 shows the resulting IRF for the historical simulation and unweighted parametric VAR models.

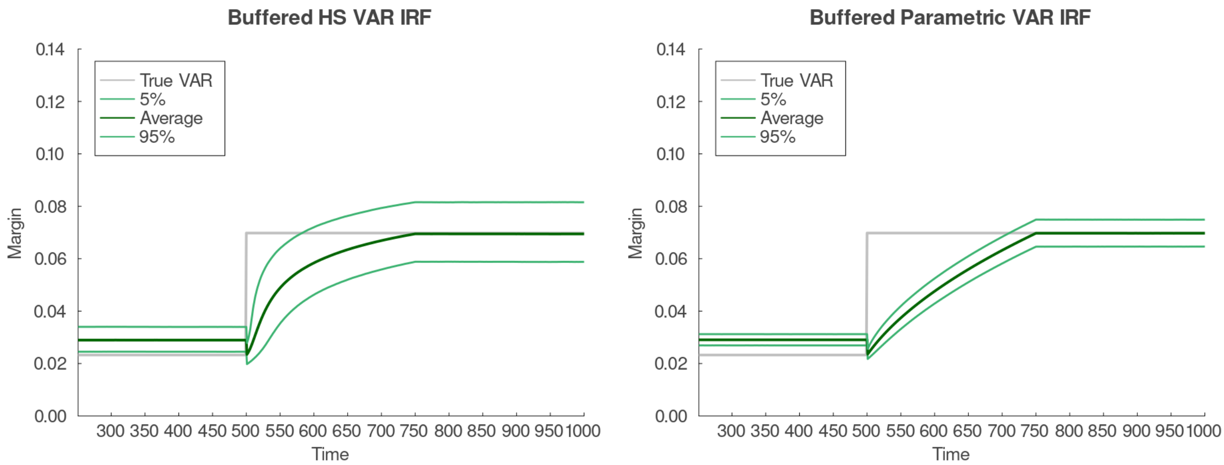

Figure 7: The IRFs of the buffered unweighted historical simulation and parametric VAR models.

The buffer keeps margin higher in the low volatility period. However, under the conditions of the current experiment, it has a really undesirable effect: it reduces margin after the stressed episode begins, but before the model has a chance to react to it, meaning that margin is just as imprudently low as it was with the unmitigated model early in the stressed episode. The cost of excess margin in the low volatility period does not lead to any benefit after the buffer has been released: the margin calls are exactly the same size as they would have been without APC. All the buffer does is provide market participants with extra liquidity at the start of the stressed episode, but this is only sufficient to fund less than 10% of the margin they will eventually be called for. Moreover, a similar pattern is observed for the EWMA parametric VAR models: see Figure 8.

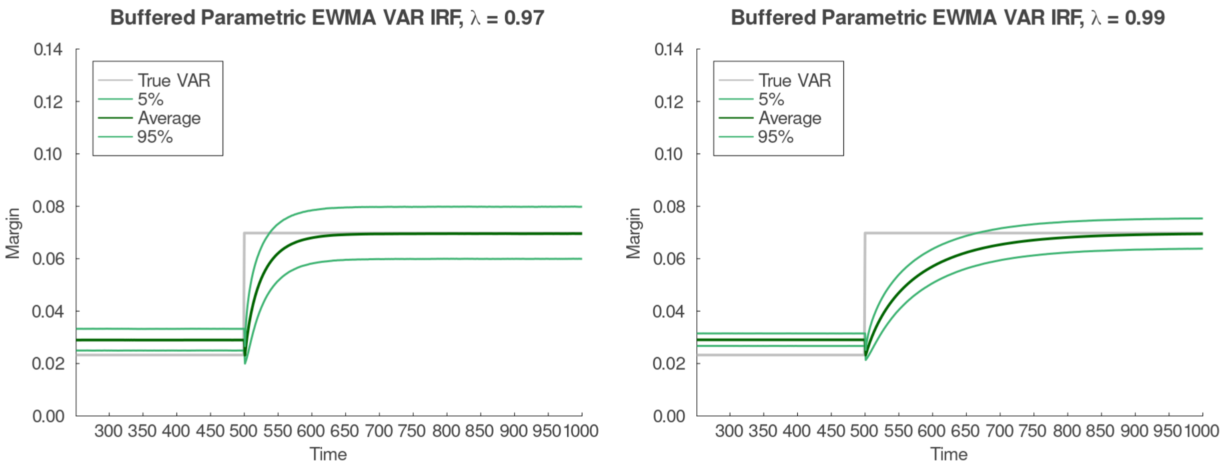

Figure 8: The IRF of $\lambda = 0.97$ and $\lambda = 0.99$ buffered EWMA parametric VAR models.

Buffered FHS VAR models demonstrate an unfortunate trifecta in the situation analysed. They demand margin well above the true VAR in the low volatility period, due to the presence of the buffer; they are imprudent immediately after episode begins, as the buffer is released but the model has not yet reacted; and they display a continuing over-reaction to the episode, as before. Figure 9 illustrates this.

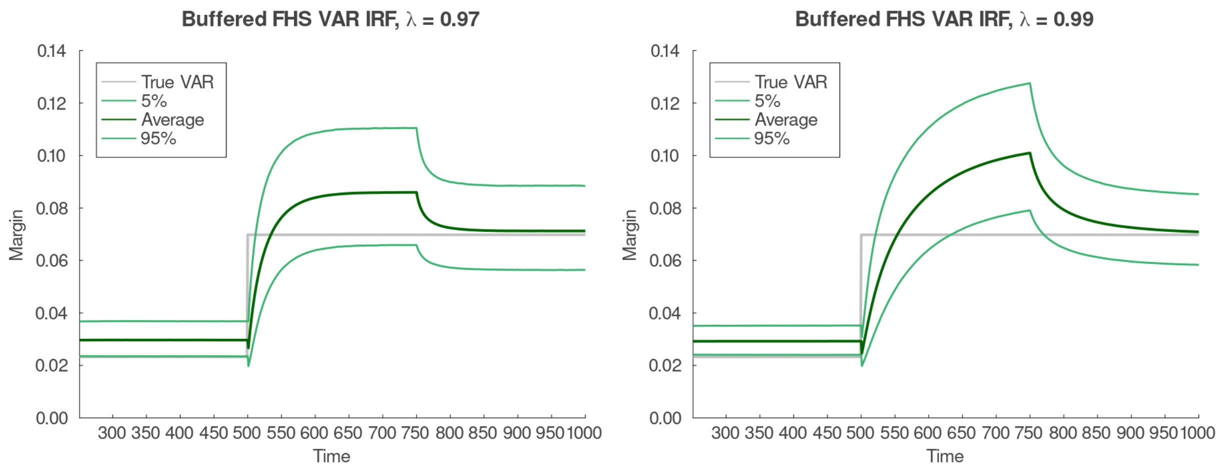

Figure 9: The response of $\lambda = 0.97$ and $\lambda = 0.99$ buffered FHS VAR models.

The results presented here are a caution against the use of a margin buffer without a deep understanding of the reactiveness of the model it is being applied to. The buffer release must be appropriately timed not to the stressed episode, but to a particular model's reaction to it. At least here, a later or slower buffer release would have been more effective, and a different buffer release rule is analysed in Appendix 2.

4.2 IRFs for margin models incorporating a stressed period

In keeping with the previous assumption of an omniscient risk manager, it is assumed that the stressed period used for APC purposes is exactly the episode which occurs, i.e. a daily volatility of returns of 3%. The IRFs for the first two margin models using this APC tool are shown in Figure 10. The effect of the stressed period APC tool in increasing margin in the low volatility period is evident. The spread of margin estimates between the 5th and 95th percentiles decreases due to the use of the fixed stressed period. Otherwise, the shape of margin reactiveness is similar to that for the unmitigated models. The parametric VAR models with EWMA volatility are somewhat quicker to react, as before.

The IRFs for these models with the stressed period APC tool applied are presented in Figure 11: the parametric VAR with the lower λ performs best, converging fast to the correct value without over-reacting, but showing a wider spread of margin estimates than the unweighted and higher λ EWMA parametric VARs. As before, the picture is not encouraging for the FHS VAR models analysed here. The use of

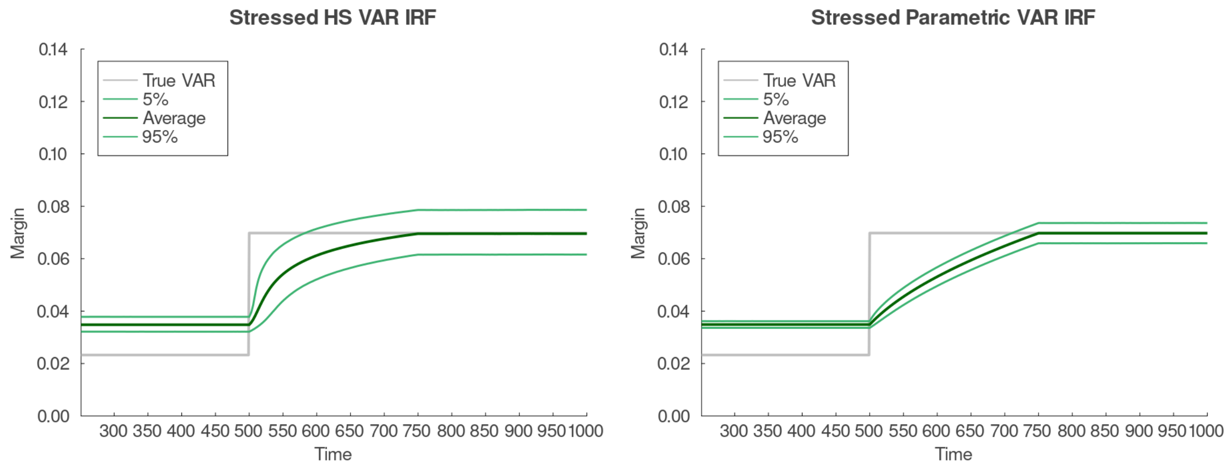

Figure 10: The IRFs of the historical simulation and parametric VAR models with the stressed period APC.

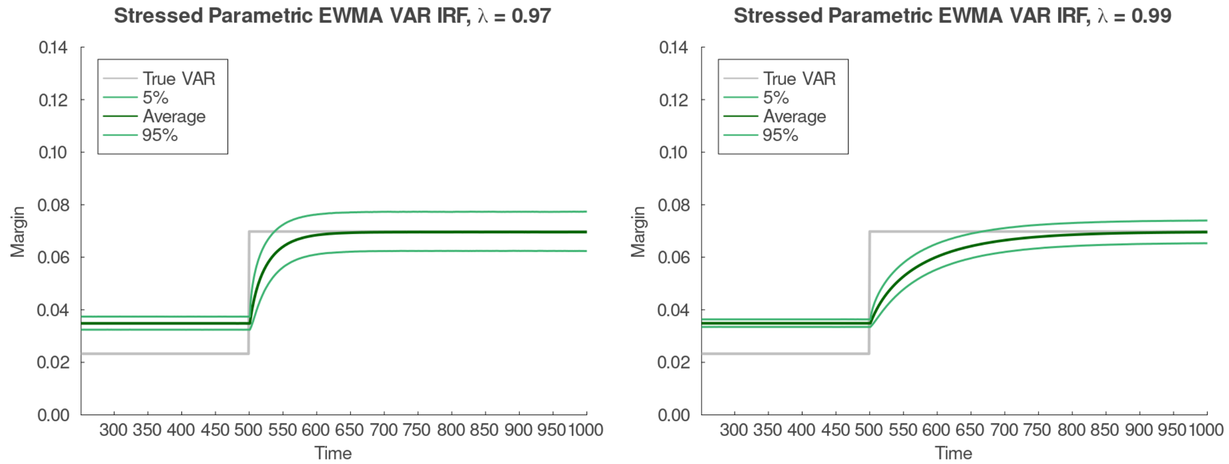

Figure 11: The IRFs of $\lambda = 0.97$ and $\lambda = 0.99$ EWMA parametric VAR models with the stressed period APC.

the stressed period reduces the models' over-reaction and spread, but the phenomena of a period where margin estimates are below the true VAR followed by over-reaction, with a fairly wide interquantile range of margin estimates, are still evident in Figure 12.

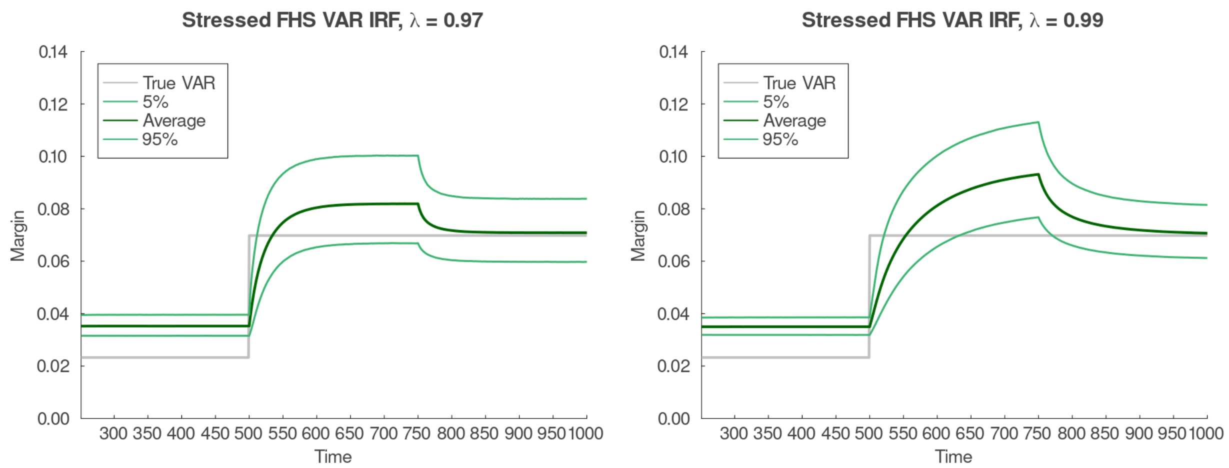

Figure 12: The IRFs of $\lambda = 0.97$ and $\lambda = 0.99$ FHS VAR models with the stressed period APC.

4.3 Procyclicality measurement: the peak-to-trough ratio and delay measure

The question naturally arises as to how well the APC tools reduce procyclicality and what their effect on margin reactiveness is. In order to investigate this, we consider four measures. First, note that, in the setting analysed, a perfect margin model would have margin after the step-up in volatility three times bigger than before. One obvious measure is therefore the *relative peak-to-trough* ratio, defined for a path of returns as:

$$\left(\frac{\text{Peak margin on the path}}{\text{Trough margin on the path}}\right) \Big/ \left(\frac{\text{True peak margin}}{\text{True trough margin}}\right)$$

The 5th, average and 95th percentile of the relative[15] peak-to-trough ratio are calculated. If the relative P/T ratio is less than one, then procyclicality has been mitigated, while if it is over one, it has been amplified. For the percentiles of this measure, note that the β (= 5% or 95%) percentile of the ratio across all simulations is given, *not* the ratio at the β percentile of the distribution of margin estimates.

The other aspect of interest is model reactiveness. Convergence to the correct level of margin is often slow, so we measure how long the model takes to raise margin to 90% of the correct level after stressed episode begins. The number of days needed for this is termed the model's delay. If margin never reaches this level, the delay is taken as 500 days (i.e. the end of the simulation).

	Relative P/T			**Delay (in days)**		
Model	5%	Average	95%	5%	Average	95%
Historical simulation	0.98	1.20	1.45	46	169	379
Unweighted parametric VAR	1.01	1.10	1.20	162	198	236
$\lambda = 0.97$ EWMA parametric VAR	1.27	1.43	1.63	22	50	93
$\lambda = 0.99$ EWMA parametric VAR	1.04	1.14	1.25	91	152	235
$\lambda = 0.97$ FHS VAR	1.41	1.84	2.38	9	29	61
$\lambda = 0.99$ FHS VAR	1.33	1.75	2.29	17	47	92

Figure 13: The relative P/T and delay measures for the margin models without APC tools.

4.4 The relative peak-to-trough ratio and delay measure for the models without an APC tool

Figure 13 presents the relative P/T ratio and delay measures for each of the six margin models without APC tools. Of the models analysed, the FHS ones have highest P/T ratios, with average relative P/T ratios of 1.84 and 1.75. Against this, they get to our target margin level much faster. Unweighted parametric VAR has a relative P/T ratio scarcely over 1, but it takes on average 200 days to get to 90% of the true margin, while $\lambda = 0.97$ FHS does it on average in less than 30 days.

The delay measure is important because it captures the period during which margin is imprudently low. The length of the average delays reported in Figure 13 suggests that there is a least a month after a large step change in volatility when most margin models are exposed to this vulnerability, and much longer for some types of model. An elevated number of exceptions – days when losses are larger than margin – are to be expected in this period.

4.5 The relative peak-to-trough ratio and delay measure for the buffered models

The buffer with full release after the step-up does very little for procyclicality. Its only effect is to remove a small amount of volatility in the trough of margin, as this is now always on the day after episode begins, rather than at a random point in the low volatility period depending on the path. Thus, the relative P/T measures are slightly smaller than before, and the delays are very similar. The results are reported in Figure 14.

4.6 The relative peak-to-trough ratio and delay measure for the models with a stressed period

The stressed period APC tool performs better: it produces a more substantial lowering of relative P/T ratios, and it reduces the delay before the model reaches our target level, albeit not substantially in many cases: see Figure 15.

[15] The absolute peak-to-trough ('P/T') ratio was first introduced in [Murphy et al. 2014].

	Relative P/T			**Delay (in days)**		
Model	5%	Average	95	5%	Average	95%
Historical simulation	0.88	1.09	1.33	46	168	376
Unweighted parametric VAR	0.93	1.03	1.14	162	199	236
$\lambda = 0.97$ EWMA parametric VAR	1.07	1.24	1.43	22	50	93
$\lambda = 0.99$ EWMA parametric VAR	0.96	1.07	1.18	91	152	234
$\lambda = 0.97$ FHS VAR	1.17	1.52	1.98	9	29	62
$\lambda = 0.99$ FHS VAR	1.13	1.51	1.99	17	47	91

Figure 14: Relative P/T and delay measures for the margin models with the buffer.

	Relative P/T			**Delay (in days)**		
Model	5%	Average	95%	5%	Average	95%
Historical simulation	0.65	0.74	0.84	38	138	291
Unweighted parametric VAR	0.67	0.71	0.75	147	182	218
$\lambda = 0.97$ EWMA parametric VAR	0.77	0.83	0.90	19	43	78
$\lambda = 0.99$ EWMA parametric VAR	0.68	0.72	0.77	78	127	192
$\lambda = 0.97$ FHS VAR	0.84	1.00	1.20	8	26	54
$\lambda = 0.99$ FHS VAR	0.82	0.99	1.21	15	43	82

Figure 15: The relative P/T and delay measures for the models with the stressed period APC tool.

The trade-off between model reactiveness and procyclicality is evident in these results. Smaller delays tend to be associated with higher average relative P/T ratios. The $\lambda = 0.97$ EWMA parametric VAR model with the stressed period APC tool represents a good trade-off: it has an average relative P/T of 0.83, but an average delay of only 43 days. Of course, the choices of a model which exactly matches the underlying process and of a perfectly calibrated stressed period flatters the results here: this issue is examined further below. Nevertheless, the relatively good performance of the simpler models is interesting.

4.7 Procyclicality measurement: relative n day calls

The long term behaviour of margin is not the only phenomenon of interest. Large, short term margin calls can impose significant funding liquidity stress on market participants. For this reason, [Murphy et al. 2014] introduced 5- and 30-day call measures: these record the maximum increase in margin demanded over the relevant number of days on a particular risk factor path. We will report n day calls as a fraction of the true pre-step-up margin (i.e. 2.33%), so an average relative 5-day margin of 0.7 means that the worst 5-day increase in margin will, on the average path, be called for 0.7× the starting level of margin. As before, we record the 5th percentile, average, and 95th percentile of these short and medium term funding stress measures over the simulated paths.

Figure 16 sets out the results for our 5- and 30-day call measures for the margin models without the use of APC tools. The difference in the average between the models tested is interesting: unweighted parametric only imposes relatively small liquidity demands over the short and medium term, while the FHS VAR models generate much larger calls, especially the smaller λ one. On one hand, this is the flipside of their responsiveness: a more responsive model will demand more margin than a less responsive one when volatility increases. However, as we have seen, the FHS VAR models over-react, so some of these demands are unnecessary given the change in risk.

The confidence intervals are also wide: the 95th percentile path demands over 75% more liquidity over 5 days than the average path for the FHS models. Clearly, even for the same volatility increase, some paths demand much more funding liquidity than others.

Model	Relative 5-day calls 5%	Average	95%	Relative 30-day calls 5%	Average	95%
Historical simulation	0.39	0.69	1.14	0.66	1.09	1.66
Unweighted parametric VAR	0.11	0.16	0.23	0.35	0.46	0.60
$\lambda = 0.97$ EWMA parametric VAR	0.48	0.69	1.01	1.08	1.50	1.99
$\lambda = 0.99$ EWMA parametric VAR	0.21	0.31	0.46	0.58	0.81	1.07
$\lambda = 0.97$ FHS VAR	0.65	1.08	1.76	1.33	2.07	3.08
$\lambda = 0.99$ FHS VAR	0.42	0.81	1.44	0.95	1.62	2.56

Figure 16: The relative 5- and 30-day call measures for the models without APC tools.

4.8 The 5- and 30-day call measures for the buffered models

As might be expected from the relative P/T ratio results, in the situation analysed, the buffer does little to reduce procyclicality as measured by short and medium term funding stress, and the spread of results continues to be large.

Model	Relative 5-day calls 5%	Average	95%	Relative 30-day calls 5%	Average	95%
Historical simulation	0.39	0.69	1.14	0.65	1.08	1.65
Unweighted parametric VAR	0.11	0.15	0.22	0.35	0.46	0.59
$\lambda = 0.97$ EWMA parametric VAR	0.48	0.69	1.01	1.07	1.49	1.98
$\lambda = 0.99$ EWMA parametric VAR	0.21	0.31	0.46	0.58	0.80	1.07
$\lambda = 0.97$ FHS VAR	0.64	1.07	1.75	1.28	2.00	3.01
$\lambda = 0.99$ FHS VAR	0.41	0.81	1.43	0.93	1.59	2.52

Figure 17: The relative 5- and 30-day call measures for the buffered models.

Figure 17 sets out the results supporting this. Given the significant cost of the buffer in demanding more margin in the low volatility period, this lack of mitigation is disappointing.

4.9 The 5- and 30-day call measures for the models incorporating the stressed period

The stressed period tools has better performance. It consistently reduces funding demands by 25%, just as one would expect from the definition of the tool. Consistent with this, the spread of results is also decreased in all cases. It remains large for FHS VAR, though: the 95^{th} percentile call is over $2.6\times$ that of the 5^{th} percentile. The results are detailed in Figure 18. Even after APC mitigation, the short and medium term liquidity demands made by the FHS models, particularly on the unlucky paths, are large: the 30-day margin call is roughly twice the pre-crisis margin at the 95^{th} percentile. Of course, this is without the liquidity impact of variation margin, which might make the situation substantially better or worse.

Model	Relative 5-day calls 5%	Average	95%	Relative 30-day calls 5%	Average	95%
Historical simulation	0.29	0.52	0.85	0.49	0.82	1.24
Unweighted parametric VAR	0.08	0.12	0.17	0.26	0.34	0.45
$\lambda = 0.97$ EWMA parametric VAR	0.36	0.52	0.76	0.81	1.13	1.49
$\lambda = 0.99$ EWMA parametric VAR	0.16	0.23	0.34	0.44	0.60	0.80
$\lambda = 0.97$ FHS VAR	0.49	0.81	1.32	1.00	1.55	2.32
$\lambda = 0.99$ FHS VAR	0.31	0.61	1.07	0.71	1.21	1.91

Figure 18: The relative 5- and 30-day call measures for the models using a stressed period.

5 Results for an Episode with Fatter Tails

Increases in volatility are not the only features of risk factor returns which can change in a crisis. The tails of the distribution can fatten too.[16] In order to explore the impact of this, this section considers a situation where the distribution of returns after the step-up is modified. Instead of using normal returns, we assume that a student-t distribution with three degrees of freedom. The Student-t distributed returns are scaled to produce the same 99th percentile as before, i.e. the path of true margin is identical, but the tails of the distribution of returns are fatter and hence, because the total probability mass of the distribution remains the same, the standard deviation of returns is smaller.

5.1 IRFs for EWMA volatility estimators with a Student-t stressed episode

The IRFs for EWMA volatility estimators using $\lambda = 0.97$ and $\lambda = 0.99$ volatility estimators are presented in Figure 19. These results demonstrate that, for this fat-tailed distribution, the average of many conditional volatility estimates does not converge to the unconditional volatility. This phenomenon has been discussed in the context of GARCH(1,1) models by [Glasserman and Wu 2018].[17]

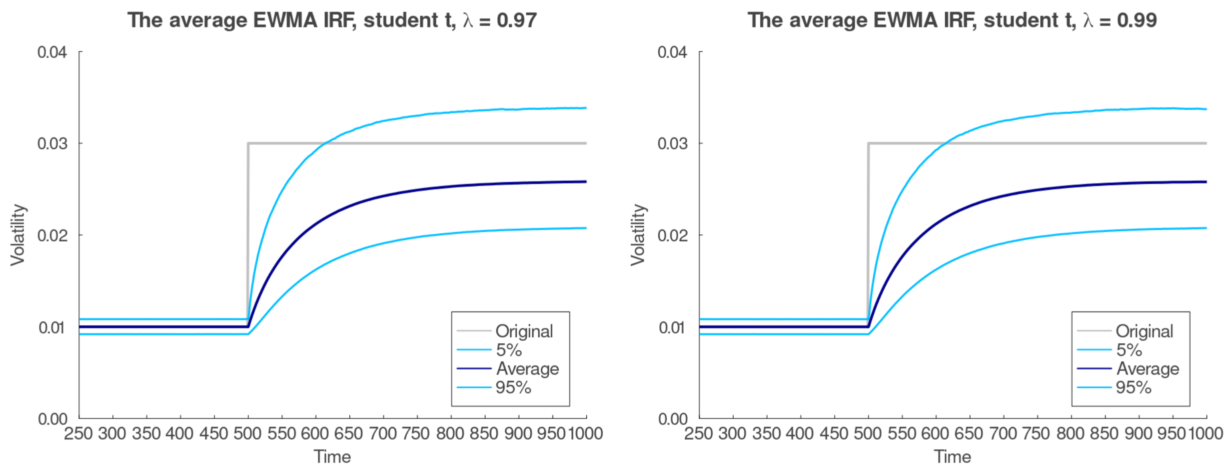

Figure 19: The IRFs for EWMA volatility estimators using student-t distributed returns after the step-up.

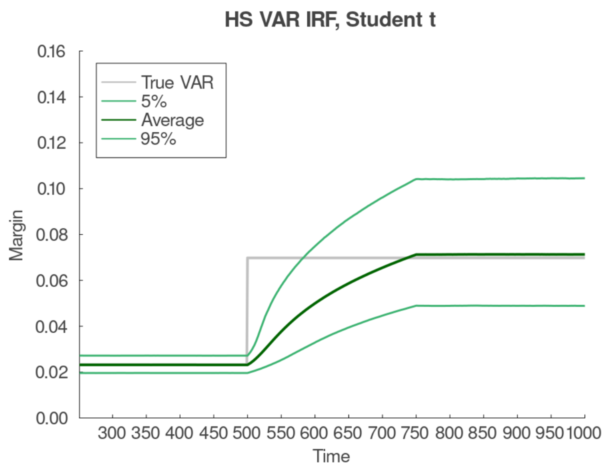

Figure 20: The IRFs for historical simulation VAR using student-t distributed returns after the step-up.

[16] Other moments of the distribution, such as skewness, can change too: these effects are not investigated.

[17] For a GARCH(1,1) model where conditional variance evolves by $\sigma_t^2 = \omega + \alpha r_{t-1}^2 + \beta \sigma_{t-1}^2$, [Glasserman and Wu 2018] introduces the parameter κ as the unique solution of $\mathbf{E}(\alpha Z^2 + \beta)^{\kappa/2} = 1$, where Z is the $\mathbf{N}(0,1)$ random variable driving returns. The authors show that the difference between the average of conditional quantile estimates and the unconditional quantile increases as κ falls. This is typical of stressed conditions: fitting a GARCH(1,1) model to the S&P 500 before and after the Covid stress, for instance, we find respectively $\kappa = 5.7$ and $\kappa = 3.2$. Thus, κ can be thought of as a parameter capturing the fatness of the tails. The increase in the bias of the average volatility estimate after the step-up observed in Figure 19, and similar albeit larger effects in margin estimation presented below, are exactly what would be expected for a fall in κ in [Glasserman and Wu 2018]'s setting.

5.2 IRFs for historical simulation VAR with a Student-t stressed episode

Figures 20 and 21 present the IRFs for historical simulation VAR using these fatter-tailed returns, respectively with and without APC tools. These figures are similar to the corresponding ones for normal returns, but the distribution of margin is significantly wider, and the margin estimates are slightly biased up.

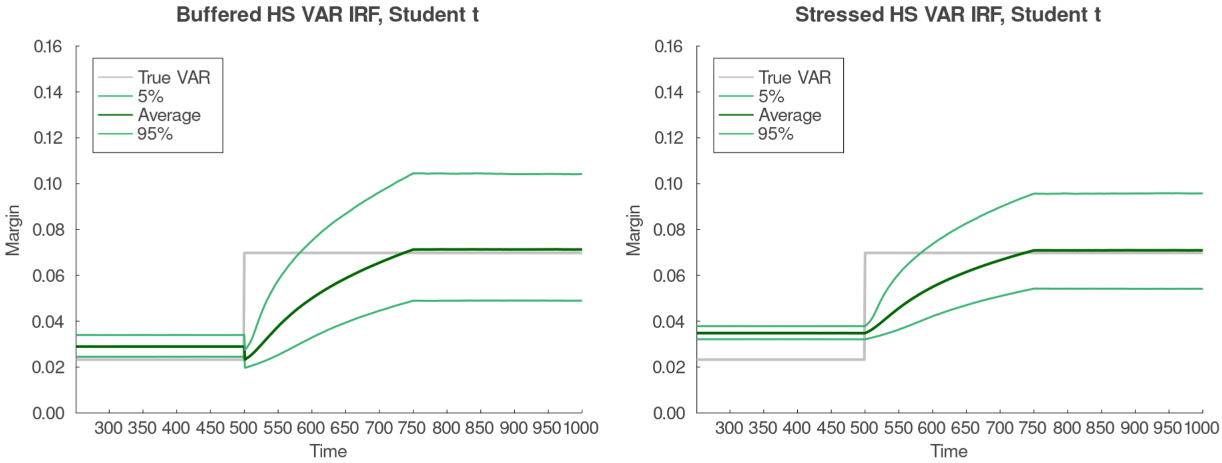

Figure 21: The IRFs for historical simulation VAR with the buffer and stressed period APC tools: student-t distributed returns after the step-up.

5.3 IRFs for parametric VAR with a Student-t distributed stressed episode

Figures 22 and 23 presents the IRFs for the parametric VAR models using the different volatility estimators using fatter-tailed returns after the step-up. No APC tools are applied here. There is significant bias, because the model is estimating the volatility of returns then scaling up assuming that the distribution is normal. This modelling error leads to substantial margin under-estimates and, likely, large numbers of backtesting exemptions. The confidence interval widens for the fatter-tailed returns, and model reactivity is similar.

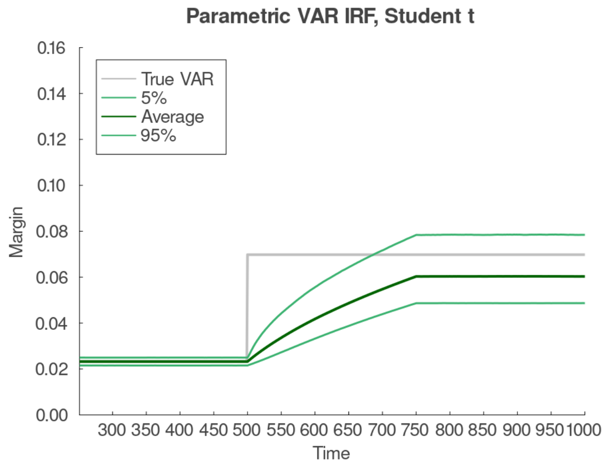

Figure 22: The IRFs for parametric VAR with an unweighted volatility estimator.

The result of the application of the APC tools is similar to the previous situation: Figures 24 and 25 shows the results of using the stressed period tool. This APC tool reduces the bias of the parametric models slightly and increases the speed of reaction, in both cases because it is calibrated to the correct level of margin. From the former effect, we can conclude that the parametric VAR models without an APC tool applied would show less bias if they were calibrated with a better estimate of the fatness of the tails. Such an approach would have an effect similar to the stressed period APC tool, lifting margin in the pre-step period, where the tails are not as fat as the model assumes, and reducing or eliminating the downwards bias after the step up. The buffer is ineffective, also for the same reasons as before: these IRFs are omitted for reasons of space.

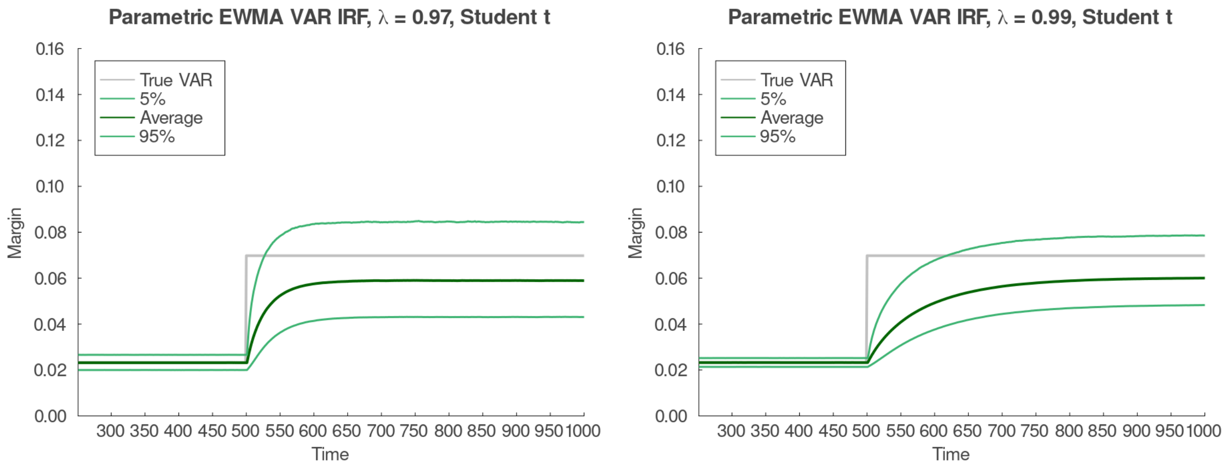

Figure 23: The IRFs for parametric VAR with an EWMA volatility estimators: student-t distributed returns after the step-up.

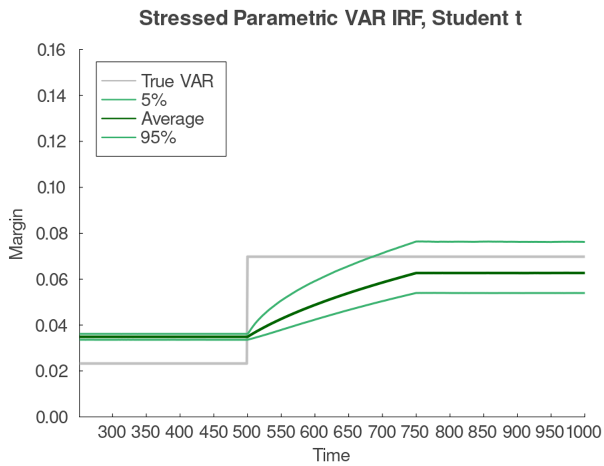

Figure 24: The IRFs for parametric VAR with an unweighted volatility estimator and the stressed period APC tool: student-t distributed returns after the step-up.

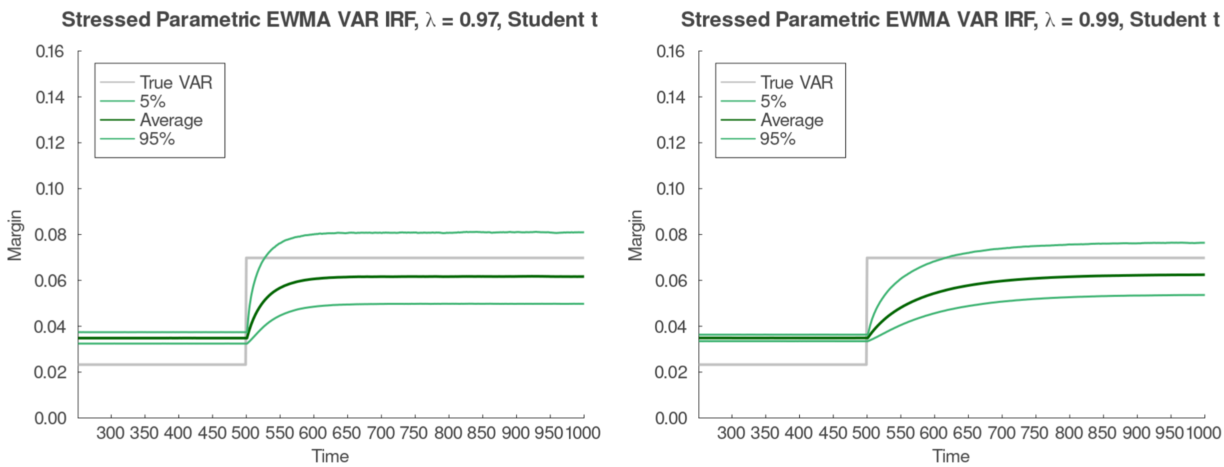

Figure 25: The IRFs for parametric VAR with the two EWMA volatility estimators and the stressed period APC tool: student-t distributed returns after the step-up.

5.4 IRFs for FHS VAR with a Student-t distributed stressed episode

The IRFs for FHS VAR with fatter tailed returns are shown in Figure 26. The general form of the response is similar to the normal case, but the confidence bounds are much wider, and the upward bias is larger.

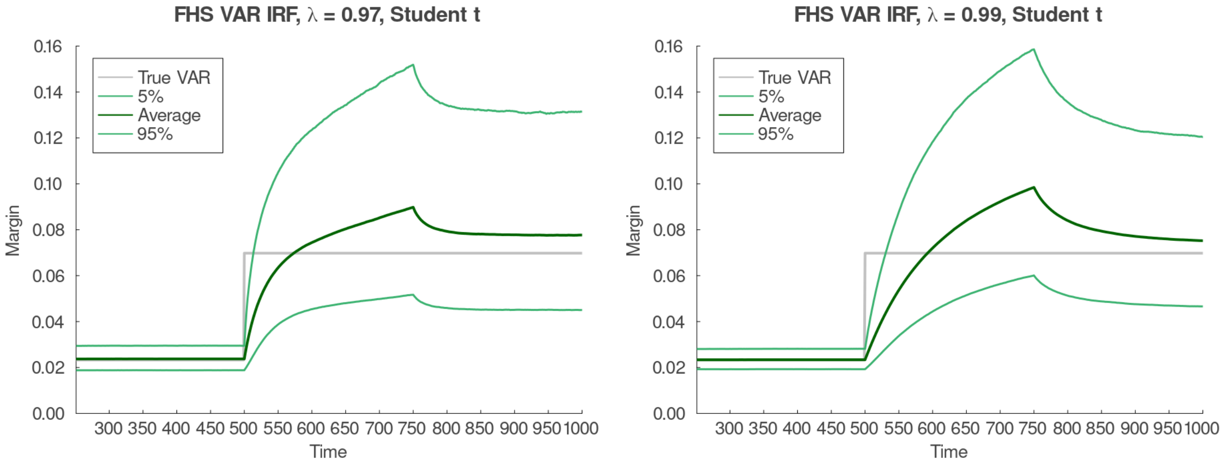

Figure 26: The IRFs for $\lambda = 0.97$ and $\lambda = 0.99$ FHS VAR: student-t distributed returns after the step-up.

It could be argued that the impulse we have used is quite extreme: a tripling of the true 99th percentile of the distribution, and a substantial fattening of the tails.[18] However, this is not unrealistic for the most affected risk factors in a crisis such as the early Covid period or the illegal Russian invasion of Ukraine. The over-reaction, on average, of the FHS VAR model analysed, and the huge spread in response depending on the risk factor path sampled are therefore concerning.

As before, the stressed period APC tool does mitigate these effects somewhat, but the effect is modest, as Figure 27 illustrates.

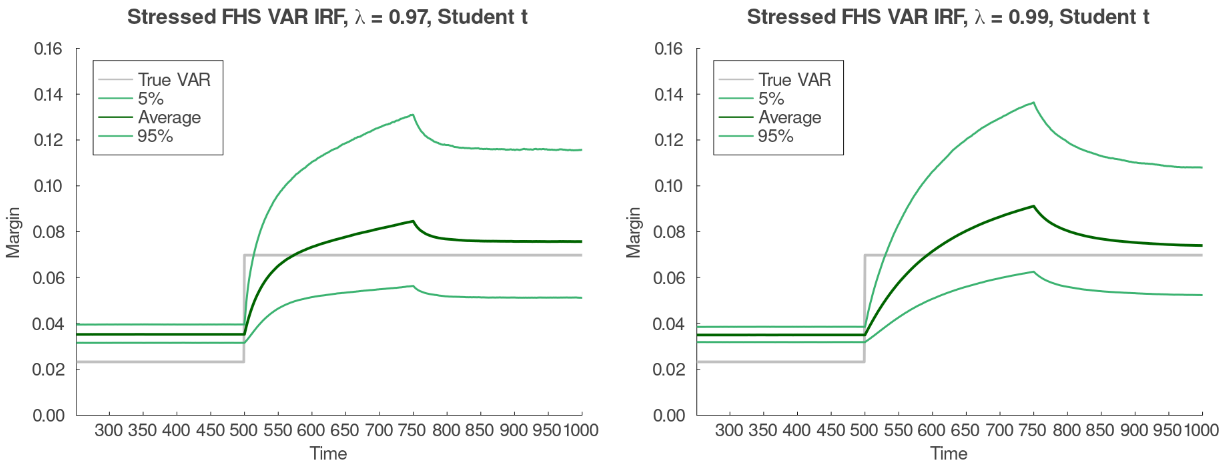

Figure 27: The IRFs for $\lambda = 0.97$ and $\lambda = 0.99$ FHS VAR with the stressed period APC tool: student-t returns after the step-up.

5.5 The relative peak-to-trough ratio and delay measures for student-t distributed returns

The impact of the use of student-t distributed returns after the step up on the relative P/T ratio and delay measures is illustrated in Figure 28. The average P/T ratios are either the same as the normal case or larger; significantly larger in the case of the FHS models. The spread in results between the 5^{th} and 95^{th} percentile is a little larger for the non-scaling models under the fatter-tailed distribution, and substantially larger for the FHS models. The margin estimates for the $\lambda = 0.97$ FHS model, in particular, are notable: from peak to trough, they vary on average almost three times more than the true margin requirement.

[18] The 99^{th} percentile of the Student-t distribution with 3 degrees of freedom is 4.5, compared to 2.33 for the normal distribution.

	Relative P/T			Delay (in days)		
Model	5%	Average	95%	5%	Average	95%
Historical simulation	0.93	1.36	2.00	63	231	N/A
Unweighted parametric VAR	0.81	1.03	1.37	141	369	N/A
$\lambda = 0.97$ EWMA parametric VAR	1.15	1.81	3.13	16	121	321
$\lambda = 0.99$ EWMA parametric VAR	0.86	1.17	1.77	61	306	N/A
$\lambda = 0.97$ FHS VAR	1.55	2.85	5.33	10	62	158
$\lambda = 0.99$ FHS VAR	1.26	2.12	3.26	24	97	216

Figure 28: The relative P/T and delay measures for the margin models without APC tools: student-t distributed returns after the step-up.

The delay measure is also interesting. Most models are slower to react in the fat-tailed case, but the effect for the FHS models is mild. As before, there is a trade-off between reactiveness and lower P/T ratios. Moreover, at the 95$^{\text{th}}$ percentile, several of the non-scaling models do not hit 90% of the true margin during the simulation at all – hence the 'N/A' value. Even when the models do eventually converge, it takes at least 150 days at the 95$^{\text{th}}$ percentile.[19]

The stressed period APC tool produces mild benefits, as before. Figure 29 presents the relative P/T ratio and delay measures using the tool. The improvement in the FHS models at the 95$^{\text{th}}$ percentile is quite encouraging. However, the tool does little to cure the downwards bias of the unweighted models.

	Relative P/T			Delay (in days)		
Model	5%	Average	95%	5%	Average	95%
Historical simulation	0.62	0.82	1.11	57	210	N/A
Unweighted parametric VAR	0.57	0.67	0.84	127	334	N/A
$\lambda = 0.97$ EWMA parametric VAR	0.71	1.00	1.60	15	104	266
$\lambda = 0.99$ EWMA parametric VAR	0.59	0.74	1.03	54	272	N/A
$\lambda = 0.97$ FHS VAR	0.89	1.45	2.52	9	56	141
$\lambda = 0.99$ FHS VAR	0.77	1.16	1.83	22	89	196

Figure 29: The relative P/T and delay measures for the margin models with the stressed period APC tool: student-t distributed returns after the step-up.

5.6 The 5- and 30-day call measures for the models with student-t distributed returns

Figure 30 presents the short and medium term funding measures using student-t returns. In this case, average margin calls are substantially higher, particular for the FHS models. As with the relative P/T ratios, the spread is wider, with unlucky paths – those at the 95$^{\text{th}}$ percentile of demands – creating large funding needs. The $\lambda = 0.97$ model is particularly procyclical: the 95$^{\text{th}}$ percentile of its 5-day demands is over nine times the margin in the low volatility period. The fact that the 30-day demands are only a small amount larger suggests that these large calls are driven by episodes in the (now fatter) tails of the distribution. These calls represent a substantial over-reaction versus the true VAR.

The stressed period APC tool improves matters, and notably so for the FHS models. However, the margin calls at the 95$^{\text{th}}$ percentile are still substantial multiples of the ones necessary to cover the risk.

[19] The failure to hit 90% of the true margin on some paths biases the other results up, as these paths create a delay of 500 days. A more sophisticated measurement approach would be to record what percentage of paths had hit 90% of the true margin after periods of, e.g., 30, 60, 90 and 120 days.

Model	Relative 5-day calls 5%	Average	95%	Relative 30-day calls 5%	Average	95%
Historical simulation	0.43	1.03	2.10	0.59	1.27	2.40
Unweighted parametric VAR	0.15	0.50	1.35	0.29	0.67	1.52
$\lambda = 0.97$ EWMA parametric VAR	0.78	2.25	5.47	1.09	2.57	5.77
$\lambda = 0.99$ EWMA parametric VAR	0.31	1.01	2.62	0.51	1.22	2.81
$\lambda = 0.97$ FHS VAR	1.24	3.71	9.01	1.64	4.22	9.62
$\lambda = 0.99$ FHS VAR	0.68	2.14	5.02	0.99	2.55	5.70

Figure 30: The relative 5- and 30-day call measures for the models: student-t distributed returns after the step-up. Above, without an APC tool applied; below, with the stressed period APC tool.

Model	Relative 5-day calls 5%	Average	95%	Relative 30-day calls 5%	Average	95%
Historical simulation	0.32	0.77	1.58	0.43	0.95	1.81
Unweighted parametric VAR	0.11	0.38	1.02	0.22	0.50	1.14
$\lambda = 0.97$ EWMA parametric VAR	0.59	1.69	4.11	0.82	1.93	4.32
$\lambda = 0.99$ EWMA parametric VAR	0.23	0.76	1.98	0.38	0.92	2.12
$\lambda = 0.97$ FHS VAR	0.93	2.79	6.75	1.23	3.17	7.21
$\lambda = 0.99$ FHS VAR	0.51	1.61	3.88	0.74	1.92	4.25

6 Conclusions and Policy Implications

Two fundamental issues have hampered the discussion about the procyclicality of margin models: the fact that model reactiveness can be very sensitive to the precise path of risk factors; and the absence of measures of procyclicality that are independent of the path. If this path-dependence is ignored, a model might be characterised as 'excessively' procyclical based only in its behaviour in a particular instance (either historical or hypothetical), without it being clear that this conclusion might be different in other, similar situations. As a result, there is always a probability that models deemed 'too procyclical' based in their performance in the last crisis may not be so in the next one, and vice versa. Without quantifying this probability, and without a tool to measure procyclicality across many similar scenarios, the evidence for any tools designed to mitigate procyclicality is incomplete and might even be misleading.

We have shown how specifying an impulse response function in a Monte Carlo simulation setting is a way of addressing these problems. The IRF approach provides a means of measuring the average margin model response to a change in volatility (and, more generally, to changes in the unconditional distribution), while also capturing the variability of this response. In this way, it provides a tool to compare models that does not depend just on the models' reaction in a particular scenario. The ability to incorporate uncertainty in procyclicality assessment allows a statistically robust evaluation of model behaviour.

These results have policy implications. First, they confirm the importance of quantifying the uncertainty arising from a model's sensitivity to the underlying path of returns. Models that produce the same procyclicality response *on average* may nevertheless have different probabilities of over- or under-reacting on particular future paths.

Second, they show that there is a trade-off between having a model that reacts quickly to changes in volatility and the uncertainty around the future behaviour of the model. In other words, as one chooses models that react more quickly to changes in volatility, the uncertainty of whether in the future the model will significantly under- or over-react increases.[20] This also means that policies to mitigate model

[20] This trade-off between speed of reaction and uncertainty in predictions adds an additional dimension to the trade-offs between speed of reaction and keeping central clearing economically efficient.

procyclicality must acknowledge the fact that, to the extent that margin models need to react to changes in volatility, there will always be many instances where the model will over- or under-react.

The work presented here again highlights the importance of acknowledging there is no 'correct' procyclicality value, but only acceptable choices given the situation (e.g., risk factor dynamics, set of portfolios being margined, participants' funding liquidity arrangements, etc.) and the desired trade-off between reactiveness, potential extent of over-reaction, and margin accuracy. These issues require expert judgement. The results also show that the impact of the choice of core margin model far exceeds the impact of the procyclicality mitigating measures analysed, supporting the adoption of an outcome-based approach to procyclicality, as one of us has previously advocated [Murphy and Vause 2021].

With regards to the comparison between different margin models compared to their non-filtered counterparts, the FHS models analysed offer higher speeds of reaction, but they tend to overreact to an abrupt change in volatility. The dispersion of their response around the average is also larger, meaning more uncertainty in procyclicality predictions. These results could be extended, for instance by considering extreme shortfall rather than VAR, and by analysing alternative types and calibrations of margin models.

Our analysis has also shown that, among volatility-filtered models, increasing the speed of reaction to a change in volatility regime (i.e., decreasing the value of the decay factor λ) decreases on average the propensity to overreact but increases the dispersion around the average.

Finally, the IRF approach is also useful to measure the impact of APC measures. Under the IRF, the two APC measures analysed (the stressed period and the buffer) perform poorly, even under the assumption of an omniscient risk manager who calibrates the buffer to release immediately after volatility increases. In fact, they may even have undesirable effects. Critically, these APC measures do not significantly reduce the probability of the model over- or under-reacting.

In summary, our analysis has shown the usefulness of the IRF approach as a tool to assess the reactiveness of a model in a statistically robust way, allowing the CCP to apply mitigation tools in a meaningful way only when a model tends to over-react. It has also uncovered further issues to explore, such as the bias observed in EWMA-volatility-filtered FHS VAR quantile estimates, the behaviour of other classes of margin model, and the analysis of the IRFs of margin models applied to portfolios sensitive to multiple risk factors, which should be the subject of future research.

References

[Avramidis and Wilson 1998] Avramidis, A., and Wilson, J., *Correlation-Induction Techniques for Estimating Quantiles in Simulation Experiments*, Operations Research, Vol. 46, No. 4, pp. 574-591, 1998, available at `https://www.jstor.org/stable/223133`

[Bakoush et al. 2019] Bakoush, M., Gerding, E., and Wolfe, S., *Margin requirements and systemic liquidity risk*, Journal of International Financial Markets, Institutions & Money, Vol. 58, 2019

[Barone-Adesi et al. 1998] Barone-Adesi, G., Burgoin, F., and Giannopoulos, K., *Don't Look Back*, Risk Magazine, Vol. 11, 1998

[BCBS and IOSCO 2013] Basel Committee on Banking Supervision, Board of the International Organization of Securities Commissions, *Margin requirements for non-centrally cleared derivatives*, 2013, available at `https://www.bis.org/publ/bcbs261.pdf`.

[BCBS et al. 2022] Basel Committee on Banking Supervision, Committee on Payments and Market Infrastructures, Board of the International Organization of Securities Commissions, *Review of margining practices*, 2022, available at `https://www.bis.org/bcbs/publ/d537.pdf`

[Boudoukh et al. 1998] Boudoukh, J., Richardson, M., and Whitelaw, R., *The Best of Both Worlds*. Risk Magazine, May 1998, available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=51420`

[Brunnermeier and Pedersen 2009] Brunnermeier, M., and Pedersen, L., *Market Liquidity and Funding Liquidity*, The Review of Financial Studies, Vol. 22, No. 6, pp. 2201-2238, 2009, available at `https://doi.org/10.1093/rfs/hhn098`

[Buczyński and Chlebus 2020] Buczyński, M., and Chlebus, M., *Old-fashioned parametric models are still the best: a comparison of value-at-risk approaches in several volatility states*, Journal of Risk Model Validation, Vol. 14, No. 2, 2020

[Cominetta et al. 2019] Cominetta, M., Grill, M., and Jukonis, A., *Investigating initial margin procyclicality and corrective tools using EMIR*, Macroprudential Bulletin, European Central Bank, 2019, available at https://www.ecb.europa.eu/pub/financial-stability/macroprudential-bulletin/html/ecb.mpbu201910_5~6c579ba94e.en.html

[CPMI and IOSCO 2012] Committee on Payment and Settlement Systems, Technical Committee of the International Organization of Securities Commissions, *Principles for financial market infrastructures*, 2012, available at https://www.bis.org/cpmi/publ/d101a.pdf

[ESMA 2018] European Securities and Markets Authority, *Guidelines on EMIR anti-procyclicality margin measures for central counterparties*, 2018, available at https://www.esma.europa.eu/sites/default/files/library/esma70-151-1496_guidelines_on_ccp_apc_margin_measures.pdf

[ESMA 2022] European Securities and Markets Authority, *Consultation paper – Review of RTS No 153/2013 with respect to procyclicality of margin*, 2022, available at https://www.esma.europa.eu/sites/file/122618.

[EU 2012] European Union, *Regulation (EU) No 648/2012*, 2012, available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32012R0648

[EU 2013] European Union, *Commission Delegated Regulation (EU) No 153/2013*, 2013, available at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013R0153

[FSB 2018] Financial Stability Board, *Incentives to centrally clear over-the-counter (OTC) derivatives*, 2018, available at https://www.fsb.org/2018/11/incentives-to-centrally-clear-over-the-counter-otc-derivatives-2/

[FSB 2023] Financial Stability Board, *The Financial Stability Aspects of Commodities Markets*, 2023, available at https://www.fsb.org/wp-content/uploads/P200223-2.pdf

[Glasserman and Wu 2018] Glasserman, P., and Wu, Q., *Persistence and Procyclicality in Margin Requirements*, Management Science, Vol. 64, No. 12, 2018, available at https://doi.org/10.1287/mnsc.2017.2915

[Goldman and Shen 2020] Goldman, E., and Shen, X., *Procyclicality mitigation for initial margin models with asymmetric volatility*, Journal of Risk, Vol. 22, No. 5, pp. 1-41, 2020 available at DOI:10.21314/JOR.2020.435

[Gurrola Perez 2018] Gurrola Perez, P., *The validation of filtered historical value-at-risk models*, Journal of Risk Model Validation, Vol. 12, No. 1, 2018, available at https://ssrn.com/abstract=3037099

[Gurrola Perez 2021] Gurrola Perez, P., *Procyclicality of CCP margin models: systemic problems need systemic approaches*, Journal of Financial Market Infrastructures, Vol. 10, No. 1, pp. 23-55, 2021, available at https://ssrn.com/abstract=3779896

[Gurrola Perez 2023] Gurrola Perez, P., *How confident are we of margin model procyclicality measurements?*, Working Paper, World Federation of Exchanges, 2020, available at https://ssrn.com/abstract=4394462

[Gurrola Perez and Murphy 2015] Gurrola Perez, P., and Murphy, D., *Filtered historical simulation Value-at-Risk models and their competitors*. Staff Working Paper No. 525, Bank of England, 2015, available at https://ssrn.com/abstract=2574769

[Hull and White 1998] Hull, J., and White, A., *Incorporating volatility updating into the historical simulation method for Value-at-Risk*. Journal of Risk, Vol. 1, 1998, available at https://doi.org/10.21314/JOR.1998.001

[ISDA 21] International Swaps and Derivatives Association, *ISDA SIMM Methodology, version 2.4*, 2021, available at https://www.isda.org>ISDA-SIMM-v2.4-PUBLIC.pdf

[Kahros and Weissler 2022] Kahros, A., and Weissler, M., *Mitigating margin procyclicality: the effectiveness of anti-procyclicality measures during the Covid-19 stress event*, Journal of Financial Market Infrastructures, Vol. 10, No. 3, pp. 69-97, 2022

[King et al. 2020] King, T., Nesmith, T., Paulson, A., and Prono, T., *Central Clearing and Systemic Liquidity Risk*, Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System, Report No. 2020-009, 2020, available athttps://doi.org/10.17016/FEDS.2020.009

[Lewandowska and Glaser 2017] Lewandowska, O., and Glaser, F., *The recent crises and central counterparty risk practices in the light of procyclicality: empirical evidence*, Journal of Financial Market Infrastructures, Vol. 5, No. 3, pp. 1–24, 2017

[Maruyama and Cerezetti 2019] Maruyama, A., and Cerezetti, F., *Central counterparty anti-procyclicality tools: a closer assessment*, Journal of Financial Market Infrastructures, Vol. 7, No. 4, pp. 1-25, 2019, available at https://ssrn.com/abstract=3302127

[Murphy et al. 2014] Murphy, D., Vasios, M., and Vause, N., *An investigation into the procyclicality of risk-based initial margin models*, Financial Stability Paper, Bank of England, No. 29, 2014, available at https://ssrn.com/abstract=2437916

[Murphy et al. 2016] Murphy, D., Vasios, M., and Vause, N., *A comparative analysis of tools to limit the procyclicality of initial margin requirements*, Staff working paper No. 597, Bank of England, 2016, available at https://ssrn.com/abstract=2772569.

[Murphy and Vause 2021] Murphy, D., and Vause, N., *A cost–benefit analysis of anti-procyclicality: analyzing approaches to procyclicality reduction in central counterparty initial margin models*, Journal of Financial Market Infrastructures, Vol. 9, No. 4, pp. 27-50, 2021, available at DOI:10.21314/JFMI.2021.013

[Najim 2006] Najim, M., *Digital Filters Design for Signal and Image Processing*, Wiley, 2006

[Parrish 1990] Parrish, R., *Comparison of Quantile Estimators in Normal Sampling*, Biometrics, Vol. 46, No. 1., pp. 247-257, 1990, available at https://www.jstor.org/stable/2531649f

[Rached and Larsson 2019] Rached, I., and Larsson, E., *Tail Distribution and Extreme Quantile Estimation Using Non-parametric Approaches*, in High-Performance Modelling and Simulation for Big Data Applications, Kołodziej, J., and González-Vélez, H., eds., Springer-Verlag Lecture Notes in Computer Science Vol. 11400, 2019

[Raykov 2018] Raykov, R., *Reducing margin procyclicality at central counterparties*, Journal of Financial Market Infrastructures, Vol. 7, No. 2, pp. 43-59, 2018, available at https://ssrn.com/abstract=3950564

[Wang et al. 2022] Wang, J., Capponi, A., and Zhang, W., *A Theory of Collateral Requirements for Central Counterparties*, Management Science, Vol. 68, No. 9, 2022, available at https://ssrn.com/abstract=3290397

[Wong and Zhang 2021] Wong, L., and Yang Zhang, *Procyclicality control in risk-based margin models*, Journal of Risk, Vol. 23, No. 5, pp. 1-24, 2021, available at DOI:10.21314/JOR.2021.010

1 Appendix on Quantile Estimation

The selection of an estimator of high quantiles from sample distributions is (much) less straightforward than it appears. In general, there are no unbiased finite sample α-quantile estimators for general α and unknown underlying distribution, and in practice, unhelpfully, the bias of α-quantile estimators tends to be larger for αs close to 0 or 1: see [Avramidis and Wilson 1998]. Moreover, while there are large sample results for the distribution of quantile estimators in various situations – see, for example, [Rached and Larsson 2019] – VAR windows are often in practice not large enough for these results to be highly accurate. Thus, care is needed in selection of a quantile estimator.

In this appendix, the central issues are illustrated theoretically and experimentally. To begin, consider the problem in more detail. We have N samples assumed to be from the same unknown underlying distribution with CDF, F. We will assumed that these samples are ordered and write the i^{th} sample as $Y_{(i)}$ so $Y_{(1)} \leqslant Y_{(2)} \leqslant \dots Y_{(N)}$. These samples have an empirical distribution with CDF:

$$\Pr(Y \leqslant \zeta) = \begin{cases} 0 & \text{if } \zeta < Y_{(1)} \\ i/N & \text{if } Y_{(i)} \leqslant \zeta < Y_{(i+1)} \\ 1 & \text{if } Y_{(N)} \leqslant \zeta \end{cases}$$

As N increases, this empirical distribution converges on F. However, N is fixed by the design of the VAR model, so the question is how to 'best' use the available information $Y_{(i)}$ to estimate the inverse CDF of F, i.e. the smallest ζ such that

$$F(\zeta) \equiv \Pr\{Y \leq \zeta\} \geq \alpha,$$

given the inevitable sampling error. We will write $F^{-1}(\alpha)$ for the inverse CDF and sometimes suppress the α where it is obvious from context. $F(\zeta)$ is assumed to be continuous and everywhere twice-differentiable.

A simple estimator of F^{-1} based on the observed samples $Y_{(i)}$ is

$$\widetilde{F_S}^{-1}(\alpha) = Y_{(\lceil \alpha N \rceil)}$$

where $\lceil x \rceil$ is the largest integer bigger than x. Thus, this estimator would estimate the 99^{th} percentile from 250 observations as the $\lceil 0.99 \times 250 \rceil = 248^{\text{th}}$ largest.

There are two obvious kinds of error in a quantile estimator: bias and mean squared error. Bias is a deviation between the expected value of the estimator from the true value. For instance, if the underlying distribution is symmetrically distributed with mean zero and variance one, and $N = 49$, the simple median estimator $\widetilde{F_S}^{-1}(0.5) = Y_{(25)}$ is biased up because in expectation, the median lies midway between $Y_{(24)}$ and $Y_{(25)}$: for $N = 50$, $Y_{(25)}$ is unbiased. The mean squared error for this case is, by definition, the expected value of the squared deviation of the estimator from the true value. It is equal to the variance of the estimator plus the square of the bias, as [Parrish 1990], in a helpful comparison of various quantile estimators for normal distributions, discusses.

In order to analyse the issues at a typical VAR quantile, $\alpha = 0.99$, consider a different estimator of F^{-1} defined as follows:

$$\widetilde{F_D}^{-1}(\alpha) = \begin{cases} Y_{(1)} & \text{if } \alpha < 1/2N \\ \theta\, Y_{(\lceil \alpha N + 0.5 \rceil - 1)} + (1-\theta)\, Y_{(\lceil \alpha N + 0.5 \rceil)} & \text{if } 0.5/N < \alpha < (N-0.5)/N \\ Y_{(N)} & \text{if } 0.5/N \leqslant \alpha \end{cases}$$

where $\theta = \lceil \alpha N + 0.5 \rceil - (\alpha N + 0.5)$. As illustration, suppose $N = 240$ and $\alpha = 0.99$. Then $\alpha N = 237.6$, $\lceil \alpha N + 0.5 \rceil = 239$ so $\theta = 0.9$. The estimator for the 99^{th} percentile is $0.9Y_{(238)} + 0.1Y_{(239)}$. This estimator assumes that $Y_{(i)}$ is the best estimator of the $(2i-1)/2N^{\text{th}}$ percentile and, via θ, uses linear interpolation between these values.

Under certain regularity conditions on the inverse CDF,[21] almost always met in practice, [Avramidis and Wilson 1998] gives the bias of the two estimators at quantile α for underlying distribution with CDF

[21] These are conditions RC_1 to RC_4 of [Avramidis and Wilson 1998].

F. If F' and F'' are respectively the first and second derivatives of F, then

$$\text{Bias}\left(\widetilde{F_S}^{-1}(\alpha)\right) \;=\; \frac{1}{N}\left(\frac{\lceil \alpha N \rceil - \alpha N - \alpha}{F'(\zeta)} - \frac{\alpha(1-\alpha)F''(\zeta)}{2(F'(\zeta))^3}\right) + \mathcal{O}(N^{-1})$$

A problem is immediately evident: in the tails of the distribution, F' maybe small if the CDF is flat, so the denominator of the second term could be large. For the standard normal distribution at 99%, i.e. $\zeta = 2.33$, $F'(\zeta)$ is 0.027 and F'' is -0.06. This gives a bias of the order of -0.16 using $\widetilde{F_S}^{-1}$ to estimate the 99$^{\text{th}}$ percentile, which is tolerable but not negligible.

The mean squared error of the two estimators, $\widetilde{F_S}^{-1}$ and $\widetilde{F_D}^{-1}$ is the same. It is

$$\text{Var}\left(\widetilde{F_D}^{-1}(\alpha)\right) \;=\; \text{Var}\left(\widetilde{F_S}^{-1}(\alpha)\right) \;=\; \frac{\alpha(1-\alpha)}{N(F'(\zeta))^2} + \mathcal{O}(N^{-1})$$

For the same situation as before, $\alpha = 0.99$, $N = 240$, standard normal distribution, this is 0.058. The second estimator has different bias properties, however:

$$\text{Bias}\left(\widetilde{F_D}^{-1}(\alpha)\right) \;=\; \frac{1}{N}\left(\frac{0.5 - \alpha}{F'(\zeta)} - \frac{\alpha(1-\alpha)F''(\zeta)}{2(F'(\zeta))^3}\right) + \mathcal{O}(N^{-1})$$

For the example, this is -0.14. Thus, a different quantile estimator has the same mean squared error but lower bias, and hence it is preferred. This illustrates the importance of selecting an optimal quantile estimator given the α being estimated and the likely shape of the underlying distribution. A more sophisticated approach would consider not just elementary quantile estimators, as surveyed in [Parrish 1990], but also kernel estimators, as in [Rached and Larsson 2019].

2 Appendix on the Buffer APC Tool with Gradual Release

Figure 31 presents the IRFs for a buffer where 2% of the initial buffer is released each day after the step-up in volatility. It can be seen that for the first four models – which do not over-react – the gradually-released buffer does mitigate procyclicality somewhat. It does not have the undesirable imprudence of the suddenly-released buffer. However, because the buffer is only 25% of the pre-step-up margin, it does little to address the over-reaction of the FHS models.

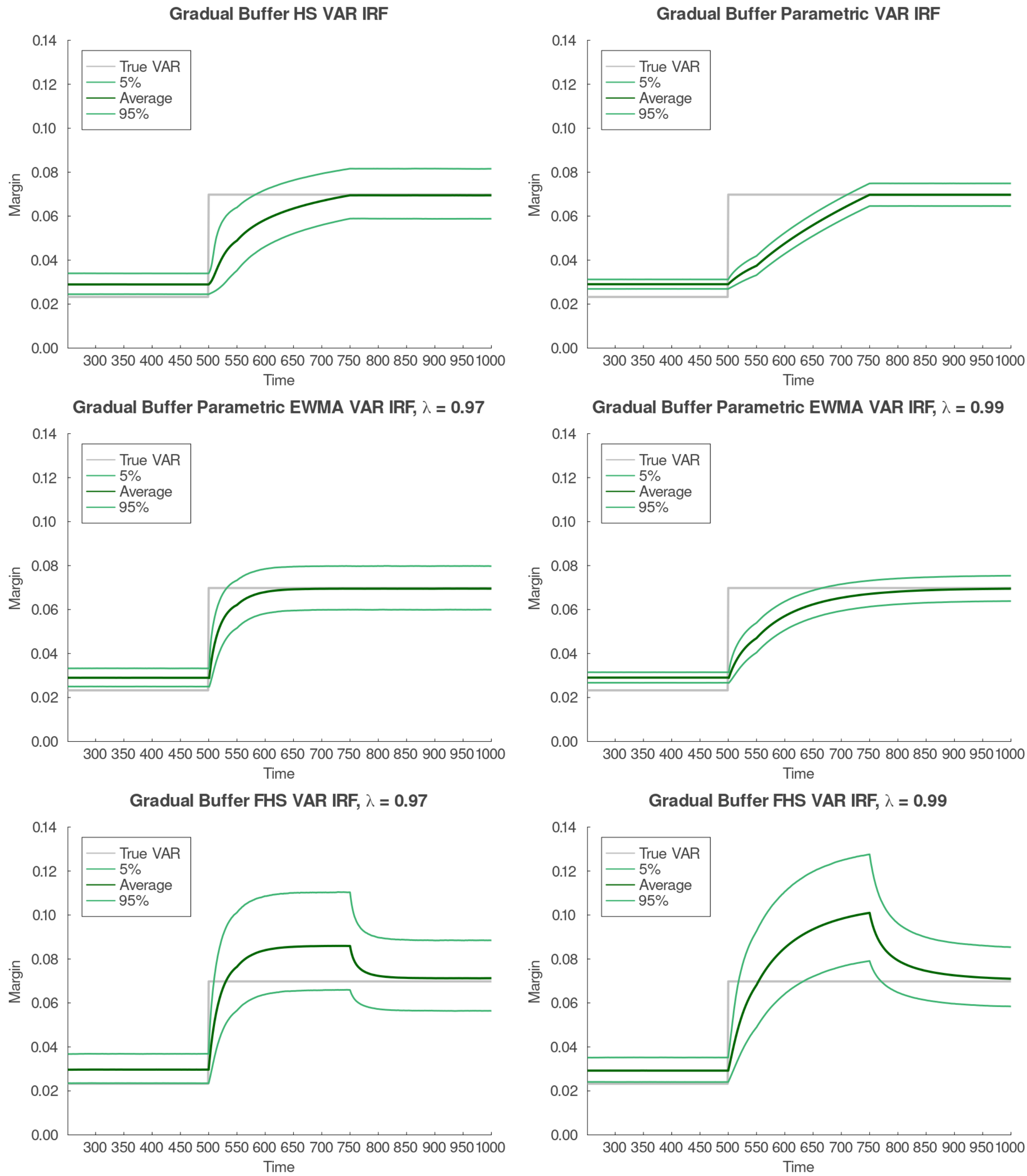

Figure 31: The IRFs for all six margin models with a gradually released buffer.