OFR WORKING PAPER
OFFICE OF FINANCIAL RESEARCH

17-01 | February 21, 2017

Persistence and Procyclicality in Margin Requirements

Paul Glasserman
Office of Financial Research
paul.glasserman@ofr.treasury.gov

Qi Wu
Chinese University of Hong Kong
qwu@se.cuhk.edu.hk

Persistence and Procyclicality in Margin Requirements

Paul Glasserman* and Qi Wu†

Abstract

Margin requirements for derivative contracts serve as a buffer against the transmission of losses through the financial system by protecting one party to a contract against default by the other party. However, if margin levels are proportional to volatility, then a spike in volatility leads to potentially destabilizing margin calls in times of market stress. Risk-sensitive margin requirements are thus procyclical in the sense that they amplify shocks. We use a GARCH model of volatility and a combination of theoretical and empirical results to analyze how much higher margin levels need to be to avoid procyclicality while reducing counterparty credit risk. Our analysis compares the tail decay of conditional and unconditional loss distributions to compare stable and risk-sensitive margin requirements. Greater persistence and burstiness in volatility leads to a slower decay in the tail of the unconditional distribution and a higher buffer needed to avoid procyclicality. The tail decay drives other measures of procyclicality as well. Our analysis points to important features of price time series that should inform "anti-procyclicality" measures but are missing from current rules.

1 Introduction

Regulatory changes following the financial crisis of 2007-08 have led to much greater use of collateral in the over-the-counter derivatives market. Collateral serves as a buffer against the transmission of losses through the financial system by protecting one party to a contract against default by the other party. The Dodd-Frank Wall Street Reform and Consumer Protection Act in the United States and similar mandates in other jurisdictions require that standard derivatives be traded through central counterparties (CCPs). A key function of a CCP is to collect collateral from its members in the form of *initial margin* and contributions to a guarantee fund. For customized derivatives that do not trade through CCPs, new rules require an exchange of bilateral margin, meaning that each party to the contract posts collateral to guarantee its potential future payments to the other party with high probability; see Margin and Capital Requirements for Covered Swap Entities [25].

Margin requirements are typically designed to cover potential price changes over a period of 5–10 days with a probability of 99 percent or higher. If one party were to default, this level

*Office of Financial Research and Columbia Business School. This paper was produced while Paul Glasserman was under contract with the Office of Financial Research and not an employee. He also serves as an independent member of the risk committee of a swaps clearinghouse.

†Assistant Professor, Department of Systems Engineering and Management, Chinese University of Hong Kong

of collateral would likely cover losses incurred by the other party in replacing or unwinding the contract. In times of higher market volatility, price changes are larger, so the minimum level of margin required to cover potential price changes with high confidence must also be larger. To put it another way, risk-sensitive margin requirements will tend to be lower in quiet times and higher in turbulent times.

This dynamic can, however, have a destabilizing effect on financial markets. With risk-sensitive margin requirements, a spike in volatility leads to margin calls on the firms trading through a CCP or bilaterally. The increase in volatility is likely to be correlated with other indicators of market stress, in which case these firms would need to post additional collateral precisely when it becomes most difficult to raise cash or other liquid assets. Firms short on cash may be forced to sell assets, driving down prices, or pull back funding to other firms, spreading a liquidity shortage. Risk-sensitive margin requirements are thus *procyclical* in the sense that they can amplify shocks. They provide a buffer against counterparty credit risk, but they create a channel for the spread of funding liquidity risk.

The alternative is to set "through the cycle" margin levels that are less sensitive to current conditions and therefore less prone to spike at the onset of market stress. The cost of this added stability is that margin levels need to be higher in quiet times and may seem unnecessarily high as memories of past volatility fade. The goal of this paper is to investigate how much higher margin levels need to be to avoid procyclicality and to examine specific rules for limiting procyclicality.

The tension between the merits of greater risk-sensitivity on one hand and the amplifying effects of the same risk-sensitivity on the other challenges other areas of financial regulation and risk management as well. Adrian and Shin [1] find that in the lead up to the financial crisis, banks and broker-dealers increased leverage in inverse proportion to their value-at-risk estimates; the sharp increase in volatility starting in mid-2007 thus led firms to delever, driving down prices of less liquid assets and further increasing volatility. Bank capital requirements under the Basel II international agreement were designed to be more risk-sensitive than requirements under Basel I, but it was recognized early on that this feature could amplify macroeconomic shocks: in a recession, borrowers are more likely to default, leading banks to hold more capital and reduce lending, thus exacerbating the downturn. As Repullo and Suarez [34] put it, "[microprudential regulators] think that procyclicality is a necessary evil, whereas others with a more macroprudential perspective think that it should be explicitly corrected." Basel III introduced a countercyclical capital buffer to have banks hold extra capital in good times so that they can continue to supply credit in bad times while maintaining adequate capital by draining the buffer.

Gorton and Metrick [21] point to a run in the repo market in 2007–08 as a key factor precipitating the financial crisis. Lenders sought to reduce their risk by demanding wider "haircuts" in repurchase agreements, effectively increasing the collateral demanded of borrowers and shrinking their access to funding. A widening haircut is thus analogous to a margin call. Regulators have sought to dampen procyclicality by setting stable — and therefore higher — minimum haircuts, trading higher collateral requirements for greater stability; see Committee on the Global Financial System [9] and Financial Stability Board [19]. Higher minimum haircuts for repurchase agreements are similar to the higher margin requirements for derivatives on which we focus.

As noted above, changes in financial regulation now require that most over-the-counter (OTC) derivatives trade through CCPs. Under normal conditions, a CCP has a "matched book" of buyers and sellers of each contract and thus no exposure to the market. However, if a clearing member defaults, the CCP has to continue to meet the payment obligations on the other side of the failed member's contracts. To protect itself against losses as it seeks to restore a matched book, the CCP collects initial margin from its members, typically in the form of cash or other high-quality liquid assets.

Initial margin requirements are commonly set to cover losses with probability at least 99 percent or 99.5 percent over a period of 5–10 days, depending on the type of derivative contract and the legal jurisdiction. These potential losses are primarily based on changes in market prices and thus driven by market volatility. Each member's margin requirements are updated periodically (daily or weekly) as the member's positions and market conditions change. The CCP may make intraday margin calls if volatility increases sharply.

A CCP also credits or debits a clearing member's *variation* margin account for gains and losses on the member's positions; variation margin is thus based on realized price changes rather than potential price changes. In addition, CCPs usually collect contributions to a guarantee fund that would be tapped to cover losses if the failed member's collateral and the CCP's own capital contribution proved inadequate to cover losses. Guarantee fund contributions are commonly based on stressed market conditions rather than current market conditions. Initial margin — henceforth simply margin — is thus the most procyclical component of the total collateral provided by a clearing member to the CCP and the focus of our analysis.

Procyclical margin has attracted the concern of regulators and international standard-setting bodies. The Committee on the Global Financial System [9] studied the causes and consequences of procyclicality in margin requirements and proposed countercyclical measures. The Committee on Payment and Settlement Systems [10] recommends that financial market infrastructures (including CCPs) adopt stable and conservative margin requirements to avoid procyclical increases in margin. The European Union has adopted specific requirements on CCPs to mitigate procyclicality in margin requirements (European Commission [18]), to which we return later. These rules also apply to U.S.-based CCPs seeking to clear trades for European market participants.

Our objective is to analyze procyclicality in margin requirements through a combination of theoretical and empirical results. Our analysis is based primarily on a GARCH model (Engle [16], Bollerslev [5]) of price changes — not because the model is a perfect description of market data, but because it provides a valuable lens through which to highlight important features. The GARCH setting captures volatility clustering and persistence in volatility, which, we will argue, drive the size of the buffer needed to counter procyclicality; they also drive the length of historical data needed to estimate stable margin levels. The GARCH setting provides a natural distinction between current market conditions, described by conditional volatility, and long-run conditions, described by unconditional distributions. The unconditional distribution of volatility has heavy tails even when conditional changes in volatility do not, and this leads the stable non-procyclical margin level to be much larger than the average (risk-sensitive) conditional margin level.

The GARCH setting offers two parameters that neatly summarize important features of price data which, we will argue, should inform measures to limit procyclicality: one is the exponent describing the tail of the unconditional distribution, and the other is the *extremal index* of the process. The tail exponent governs the size of the buffer needed to counter procyclicality; it also helps quantify the degree of procyclicality that results from risk-sensitive margin requirements. The extremal index governs the length of the historical lookback period required to set adequate margin levels that are resistant to procyclicality.

Both of these parameters reflect the clustering and persistence of volatility, which we see as a robust features of market data, beyond the specific structure of GARCH models. The GARCH setting offers a convenient setting in which to highlight these features and their consequences, but these properties of volatility are not limited to the GARCH setting. A shortcoming of current proposals to mitigate procyclicality is that they fail to address these features or incorporate them in differentiating across different types of assets.

The rest of this paper is organized as follows. Section 2 formulates our model and presents our first results relating the tail exponent to stable margin levels. Section 3 studies the probability of sharp increases in required margin under a risk-sensitive rule and shows how these measures of procyclicality are related to the tail exponent. Section 4 studies the probability of margin breaches and the transition from light-tailed price changes over short horizons to heavy-tailed changes over long horizons. Section 5 uses the framework of earlier sections to analyze rules to counter procyclicality adopted by the European Union. One of these rules requires a 10-year lookback for estimation; Section 6 provides some theoretical guidance for setting the length of the lookback period. Most proofs are deferred to appendices, Sections A–E.

2 Persistence and Procyclicality

2.1 A Simple Margin Model

In this section, we present a simple model of margin requirements and use it to illustrate and quantify how persistence and clustering in volatility widens the gap between through-the-cycle and conditional margin levels.

As discussed in the introduction, initial margin is intended to cover potential losses from price changes over the period it takes the CCP to restore a matched book, following the failure of a clearing member. This interval is called the margin period of risk and is customarily 5–10 days. We will index time $t, t+1, \ldots$ in multiples of this basic time unit. We denote by X_t the price change for a swap contract or portfolio from t to $t+1$, and we adopt the convention that, in setting the margin level for a particular clearing member, positive values of X_t reflect losses to the CCP. (In bilateral trading, each party undertakes a similar analysis with respect to its counterparty.)

We model these price changes using a GARCH(1,1) process,

$$\begin{aligned} X_t &= \sigma_t Z_t & (1) \\ \sigma_t^2 &= \omega + \alpha X_{t-1}^2 + \beta \sigma_{t-1}^2 & (2) \\ &= \omega + (\alpha Z_{t-1}^2 + \beta)\sigma_{t-1}^2, & (3) \end{aligned}$$

where the innovations $\ldots, Z_{-1}, Z_0, Z_1, \ldots$ are independent and identically distributed (i.i.d.) with $\mathbb{E}Z = 0$ and $\mathbb{E}Z^2 = 1$, and $\omega > 0$, $\alpha > 0$ and $\beta \geq 0$ are constants. A key feature of this model is that the conditional variance σ_t^2 depends on both the prior conditional variance σ_{t-1}^2 and the squared price change X_t^2. The sum $\alpha + \beta$ controls the persistence of volatility, and the parameter α controls the burstiness of volatility — with larger α, large price moves in one period fuel large price moves in the next period.

As an illustration of these effects, Figure 1 shows daily returns on the S&P 500 stock index in 2004-07. The index moved by more than 2 percent on only two days during 2004–2006. It moved by more than 2 percent on 17 days in 2007, and 15 of those moves occurred in the last six months of the year. A large price move in either direction is often followed by another large price move, producing clustering and persistence in volatility. These effects are well known in financial risk management and have long informed the estimation of value-at-risk; see, for example, the books by Alexander [2] and McNeil, Frey, and Embrechts [26].

We will use the following conditions:

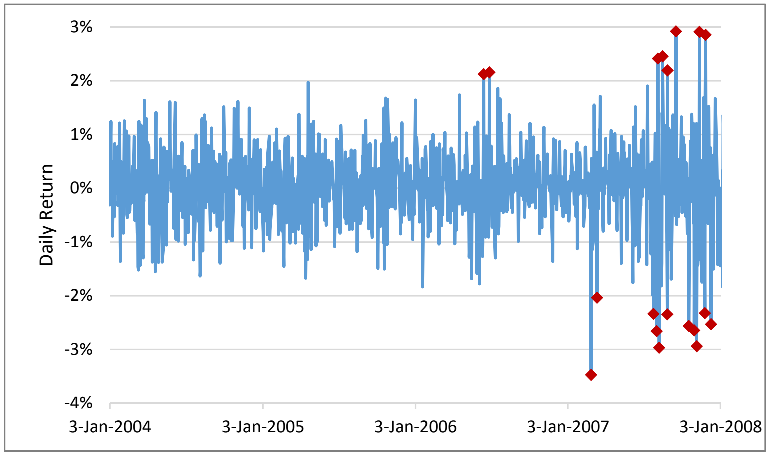

Figure 1: Daily returns on the S&P 500 index illustrate volatility clustering. The index moved by more than 2 percent on only two days during 2004–2006. It moved by more than 2 percent on 17 days in 2007, and 15 of those moves occurred in the last six months of the year. Source: Standard & Poor's Compustat and authors' analysis.

(A1) The parameters α and β satisfy $\alpha + \beta < 1$.

(A2) The innovations Z_t have a symmetric distribution, and their cumulative distribution function F_Z is continuous with $0 < F_Z(x) < 1$ for all $x \in \mathbb{R}$.

(A3) For some $0 < r < \infty$, we have $1 < \mathbb{E}[(\alpha Z^2 + \beta)^r] < \infty$.

Condition (A1) ensures that (X_t, σ_t) admits a stationary ergodic distribution, by Theorem 2 of Nelson [31]. The notation $(X_\infty, \sigma_\infty)$ denotes a stochastic vector with this stationary distribution, which we also refer to as the unconditional distribution. Condition (A2) simplifies the statement of several of our results. Examples in which (A3) is violated are rare. In particular, (A2) and (A3) hold when the distribution of Z is standard normal, Laplace (double-exponential), or Student t with more than two degrees of freedom (scaled to have unit variance). For several results, conditions (A1)–(A3) are somewhat stronger than necessary. We adopt them because they are simple to state and widely applicable.

Given market conditions at time t, the margin required to ensure that the probability of a loss is not greater than $p > 0$ (with, e.g., $p = 0.005$) is the conditional quantile $M_{p,t}$, defined by

$$M_{p,t} = \inf\{x \in \mathbb{R} : \mathbb{P}_t(X_t \geq x) \leq p\},$$

where $\mathbb{P}_t$ denotes conditional probability given the history $\{(\sigma_s, Z_s), s \leq t-1\}$. This conditional margin level is given by

$$M_{p,t} = \sigma_t \bar{F}_Z^{-1}(p) \equiv \sigma_t k_p \tag{4}$$

where $\bar{F}_Z^{-1}(p)$ is the $(1-p)$th quantile of F_Z; in particular $F_Z(k_p) = 1-p$, under our assumption that F_Z is continuous. For example, if F_Z is the standard normal and $p = 0.005$, then $M_{p,t} \approx 2.58\sigma_t$. Setting the margin requirement at this level yields the target rate of exceedances in the following sense:

Proposition 1 *If (A1)–(A2) hold, then, with probability 1,*

$$\frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{X_t > M_{p,t}\} \to p, \quad \text{as } T \to \infty.$$

Proof: The existence of the limit follows from the ergodicity of the GARCH process, which follows from (A1) through Nelson [31], Theorem 2. Because the indicator variables are bounded, we get the same limit if we take expectations on the left. But $\mathbb{P}(X_t > M_{p,t}) = \mathbb{E}[\mathbb{P}_t(X_t > M_{p,t})] = p$, by definition. □

As a consequence of (4), the conditional margin level $M_{p,t}$ increases or decreases together with the conditional volatility σ_t. The corresponding through-the-cycle or unconditional margin level is the corresponding percentile of the distribution of X_∞,

$$u_p = \inf\{x \in \mathbb{R} : \mathbb{P}(X_\infty \geq x) \leq p\}. \tag{5}$$

To justify this definition, suppose the CCP demands a fixed margin of u_p each period, regardless of the current level of volatility. Then the long-run average fraction of periods in which price changes exceed u_p is given by

$$\lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}\{X_t > u_p\} = \mathbb{P}(X_\infty > u_p) = p.$$

The limit follows from the ergodicity of the GARCH process, which follows from (A1), and the last equality follows from the continuity assumed in (A2). Thus, the fixed margin level u_p is exceeded exactly a fraction p of the time: a lower margin would be exceeded more often, and a higher margin would be exceeded less often. This makes u_p the appropriate constant level of margin for a confidence level of $1-p$. The difference between $M_{p,t}$ and u_p is illustrated schematically in Figure 2.

We want to contrast this fixed, unconditional margin level with the average margin level m_p under the procyclical rule (4), defined by

$$\frac{1}{T}\sum_{t=1}^{T} M_{p,t} \to k_p\mathbb{E}[\sigma_\infty] \equiv m_p. \tag{6}$$

Here, m_p is average level of the risk-sensitive, conditional margin requirement $M_{p,t}$ at confidence level $1-p$, whereas u_p is the stable, unconditional margin requirement at the same confidence level.

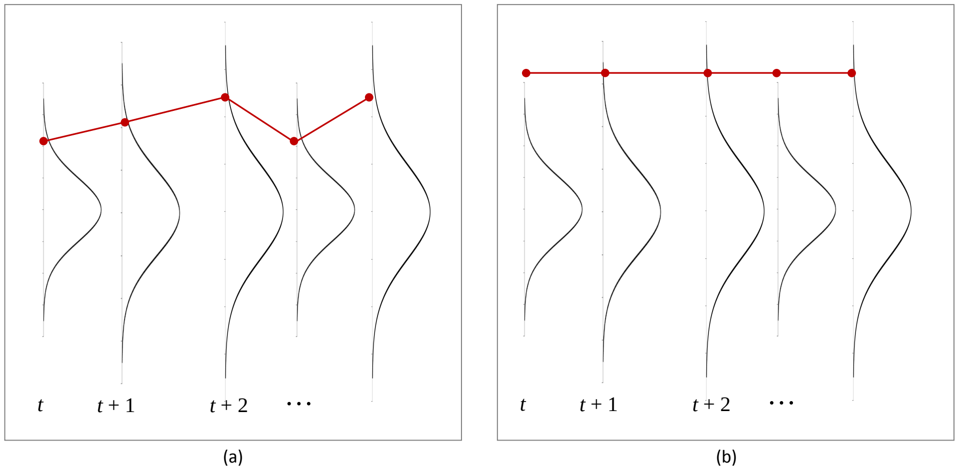

Figure 2: Contrast between (a) conditional and (b) unconditional margin levels. In (a), the margin level is set based on the conditional distribution of price changes given current market conditions. In (b), the margin level is set based on the unconditional distribution and thus remains constant. Source: authors' analysis.

The conditional margin $M_{p,t}$ will be higher than u_p in times of high volatility and lower than u_p in times of low volatility, so one might expect the average m_p to be close to u_p, but we will see that this is not the case.

The comparison uses a parameter κ defined as the strictly positive solution to the equation

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

Conditions (A1)–(A3) ensure the existence of a unique solution $\kappa > 0$. Through Theorem 2.1 of Mikosch and Stărică [29] and related work cited there, κ describes the tail of X_∞. We can now compare the unconditional margin u_p and the average conditional margin m_p at small values of p, beginning with the case of normal innovations. The symbol $\sim$ between two expressions indicates that the ratio of the two expressions converges to 1.

Proposition 2 *With normally distributed innovations Z_t and (A1),*

$$\frac{\text{unconditional margin}}{\text{average conditional margin}} = \frac{u_p}{m_p} \sim c_N \frac{p^{-1/\kappa}}{\sqrt{\log(1/p)}}, \quad \textit{as } p \to 0,$$

where $c_N > 0$ is a constant and $\kappa > 0$ is defined by (7).

This result makes two points. First, it shows that at high confidence levels (small values of p) avoiding procyclicality requires a surprisingly high level of margin. For small p, the numerator $p^{-1/\kappa}$ will be quite large compared with the denominator. Figure 3 plots $p^{-1/\kappa}$ and $\sqrt{\log(1/p)}$ as

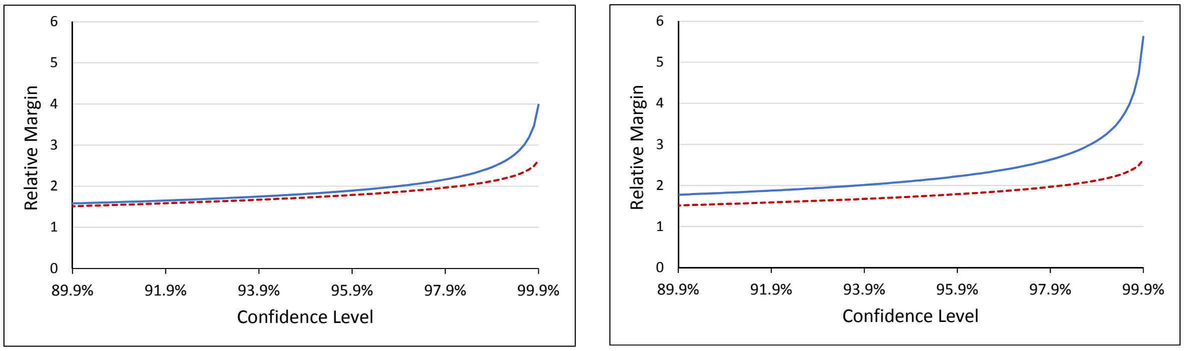

Figure 3: Comparison of $p^{-1/\kappa}$ (solid) and $\sqrt{\log(1/p)}$ (dashed) as functions of the confidence level $1-p$, for $\kappa = 5$ (left) and $\kappa = 4$ (right). Source: authors' analysis.

functions of the confidence level $1-p$ at $\kappa = 5$ and $\kappa = 4$. The ratio between the two curves becomes very large at high but practically relevant confidence levels. At realistic parameter values, we find that c_N is in the range of 0.5–2, so the ratio in Proposition 2 gives the right order of magnitude for the comparison.

The second point we learn from Proposition 2 is that to understand when "acyclicality" is costliest, in the sense that u_p/m_p is largest, we need to understand the parameter κ. Moreover, as κ may vary substantially across different types of assets, the buffer required to mitigate procyclicality may also vary widely.

We record some properties of κ in the following proposition and formulate these without assuming normality in the definition (7).

Proposition 3 *Under conditions (A1)–(A3), the following properties hold:*

(a) For each α and β, equation (7) has a unique strictly positive solution $\kappa = \kappa(\alpha, \beta)$;

(b) κ is a decreasing function of α and β;

(c) As α and β increase with $\alpha + \beta \uparrow 1$, we have $\kappa \downarrow 2$;

(d) As α and β decrease with $\alpha + \beta \downarrow 0$, κ increases without bound.

In light of these properties, we may interpret κ as a reflection of persistence (through $\alpha + \beta$) and burstiness (through α), with smaller values of $\kappa > 2$ resulting from larger values of these parameters, and $\kappa \approx 2$ corresponding to the nearly integrated case $\alpha + \beta \approx 1$. Proposition 2 then says that *the higher the persistence the wider the gap between stable and procyclical margin levels.*

It may be tempting to read Proposition 2 as just reflecting the familiar notion that the light tails of the normal distribution make it unsuitable for modeling market returns. But that would be

Series	α	β	$\alpha+\beta$	κ
S&P 500	0.075	0.915	0.990	5.2
EUROSTOXX 50	0.083	0.904	0.987	5.4
CDX NA IG	0.257	0.731	0.988	2.4
US 10yr IR Swap	0.047	0.951	0.998	3.8
USD-BRL	0.118	0.878	0.996	2.6
Brent Crude Oil	0.045	0.952	0.997	4.6

Table 1: Estimates of α, β, and κ for the S&P 500 index (1/90–4/16), the EUROSTOXX 50 index (1/90–4/16), the CDX North American Investment Grade credit default swap index (11/04–4/16), the 10-year U.S. dollar interest rate swap rate (1/90–4/16), the exchange rate between the U.S. dollar and the Brazilian real (12/92–4/16), and the price of Brent crude oil (1/99–4/16). Source: The Volatility Laboratory of the NYU Stern Volatility Institute (http://vlab.stern.nyu.edu) and authors' analysis.

a misinterpretation of the result. The proposition holds under the hypothesis that the innovations are in fact normal, so the conditional margin levels $M_{p,t}$ are correctly determined: the CCP is not misjudging the market in setting this margin requirement. Indeed, the conditional margin level performs as intended in the sense that the frequency of exceedances equals the target, as we saw in Proposition 1. The shortcoming of $M_{p,t}$ is its variability — in particular, because an increase in volatility is often accompanied by other types of market stress. Proposition 2 shows that avoiding this pattern with a stable margin requires a substantially higher level of margin.

Table 1 reports estimates of α, β, and κ for financial time series drawn from different markets — equities, credit, interest rates, foreign exchange, and crude oil. In all cases, the estimates of $\alpha+\beta$ are close to 1, as is typical of financial data. Given estimates of these parameters, we solve for κ numerically using (7). The κ values show noteworthy variation. The smallest values (corresponding to the heaviest tails) correspond to the CDX credit index and the exchange rate between the U.S. dollar and the Brazilian real.

The smallest values of κ in Table 1 coincide with the largest values of α. From the GARCH dynamics in (1)–(2), we see that α controls the degree of burstiness or self-excitation in the model: with a larger α, a larger squared return produces a larger increase in volatility. Holding $\alpha+\beta$ fixed, the stronger this feedback mechanism, the smaller the value of κ, and the wider the gap between the average conditional margin level and the unconditional margin level.

2.2 Innovations with Heavier Tails

We next examine how the gap identified in Proposition 2 changes if the innovations Z_t have heavier tails than the normal distribution. We consider two cases: innovations with a Laplace density, $z :\mapsto \exp(-\sqrt{2}|z|)/\sqrt{2})$ on $\mathbb{R}$, and innovations with the density of $\sqrt{\nu-2}t_\nu/\sqrt{\nu}$, $\nu > 2$, where t_ν

has a Student t distribution with ν degrees of freedom. The Laplace distribution has exponential tails whereas the tails of the scaled t distribution decay like a power law. (More precisely, the tails are regularly varying; see, for example, Section A3 of Embrechts et al. [15] for background on regular variation.) In both cases, the innovations have mean zero and are scaled to have unit variance.

Proposition 4 *With Laplace distributed innovations,*

$$\frac{u_p}{m_p} \sim c_L \frac{p^{-\frac{1}{\kappa}}}{\log(1/p)},$$

and with scaled t-distributed innovations,

$$\frac{u_p}{m_p} \sim c_\nu p^{\frac{1}{\nu}-\frac{1}{\kappa}},$$

as $p \to 0$, where c_L and c_ν are positive constants.

As we move from the normal distribution to the Laplace distribution to the t distribution, the tails of the innovations become heavier. Holding all else fixed, heavier-tailed innovations lead to higher conditional margin levels $M_{p,t}$. However, the unconditional margin u_p will generally also be higher. For fixed $r > 0$, the rth moment $\mathbb{E}[(\alpha Z^2 + \beta)^r]$ typically increases as the tails of Z become heavier, so, holding α and β fixed, κ will typically be smaller for innovations with heavier tails. As a consequence, we cannot directly compare the ratios in Proposition 4 with each other or with the normal case in Proposition 2. The main point of Proposition 4 is that we continue to have a large wedge between the unconditional margin and the average conditional margin at high confidence levels, even with heavier-tailed innovations.

Our analysis deals mainly with margin levels based on quantiles, but we close this section by showing that similar results hold using expected shortfall. Let $M_{p,t}^{ES} = \mathbb{E}[X_t | X_t \geq M_{p,t}, \sigma_t]$ denote the expected shortfall of X_t, given σ_t. We can express $M_{p,t}^{ES}$ as

$$M_{p,t}^{ES} = \mathbb{E}[X_t | X_t \geq M_{p,t}, \sigma_t] = \int_0^p M_{q,t}\, dq = \sigma_t \int_0^p k_q\, dq = \sigma_t \int_0^p \bar{F}_Z^{-1}(q)\, dq \equiv \sigma_t k_p^{ES}.$$

We can similarly define the expected shortfall based unconditional margin level, u_p^{ES}, as the unconditional expected shortfall of X_∞,

$$u_p^{ES} = \mathbb{E}[X_\infty | X_\infty \geq u_p].$$

The next result shows that the comparison of margin levels based on expected shortfall is very similar to the comparison based on quantiles:

Proposition 5 *As $p \to 0$, with normally distributed innovations,*

$$\frac{u_p^{ES}}{m_p^{ES}} \sim c_N \frac{\kappa}{\kappa - 1} \frac{p^{-1/\kappa}}{\sqrt{\log(1/p)}},$$

with Laplace distributed innovations,

$$\frac{u_p^{ES}}{m_p^{ES}} \sim c_L \frac{\kappa}{\kappa - 1} \frac{p^{-1/\kappa}}{\log(1/p)},$$

and with scaled t_ν innovations,

$$\frac{u_p^{ES}}{m_p^{ES}} \sim c_\nu \frac{\kappa}{\kappa - 1} \frac{\nu - 1}{\nu} p^{\frac{1}{\nu} - \frac{1}{\kappa}},$$

where c_N, c_L, and c_ν are the constants appearing in Propositions 2 and 4.

2.3 Alternative Models

For simplicity, we limit our analysis primarily to GARCH(1,1) models. The results of Sections 2.1 and 2.2 are driven by the fact that the GARCH dynamics turn light-tailed conditional distributions into heavy-tailed unconditional distributions. In Section B, we show that this property extends to the asymmetric GARCH model of Glosten, Jagannathan, and Runkle [20], GARCH(p,q) models, regime-switching models, and a GARCH diffusion process.

3 Margin Calls

In Section 2, we investigated how much higher unconditional margin requirements need to be (compared with average conditional margin requirements) to meet a target confidence level. In this section, we analyze the dynamics of margin levels under a procyclical (conditional) margin requirement. These dynamics reflect the stress imposed on market participants by a sudden run-up in margin requirements in response to climbing volatility. We analyze the probability of a large cumulative margin call in Section 3.1; in Section 3.2 we estimate the probability that margin climbs to a high level before returning to a more typical level. These two measures are similar to measure of procyclicality proposed independently in the simulation study of Murphy, Vasios, and Vause [30].

3.1 Margin Run-Ups

As discussed in Section 1, the problem with risk-sensitive margin rules is that they lead to sharp increases in margin requirements at the onset of market stress, when market participants may have the most difficulty raising cash. As a measure of this hazard, we consider the probability of a large climb in margin over a fixed number of periods n. For example, Murphy et al.[30] note that cumulative increases over 30-day periods are particularly important because the Basel III liquidity coverage ratio requires banks to plan for cash outflows over 30 days in periods of stress. We consider both the net increase and the cumulative margin over n periods when margin $M_{p,t}$ is set proportional to current volatility, as in (4).

Proposition 6 *Suppose $M_{p,t} = k_p\sigma_t$, with σ_t given by a stationary GARCH(1,1) model satisfying (A1)–(A3). Let κ be as in (7). Then the n-period margin increase satisfies*

$$\mathbb{P}(M_{p,t+n} - M_{p,t} > x) \sim c_n x^{-\kappa}, \quad \text{as } x \to \infty,$$

for some constant $c_n > 0$, and the cumulative n-period margin satisfies

$$\mathbb{P}\left(\sum_{i=1}^{n} M_{p,t+i-1} > x\right) \sim c'_n x^{-\kappa}, \quad \text{as } x \to \infty$$

for some constant $c'_n > 0$, for any fixed $n \geq 1$.

This result shows that the parameter κ, which controlled the size of the buffer in Section 2 required to dampen procyclicality, also controls the probability of large run-ups in margin requirements. All else being equal, an asset category with a smaller κ is more prone to large run-ups in margin requirements when margin is proportional to current volatility.

3.2 The Storm Before the Calm

Section 2 compared unconditional and average conditional levels of margin and showed that persistence in volatility leads to a wide gap between the two. In this section, we turn to a more dynamic measure that describes the risk of sudden run-ups in margin levels. Such run-ups are potentially destabilizing because they stress funding liquidity at times of elevated volatility.

Starting from a period of low or moderate (conditional) margin, we consider the probability that margin climbs to a high level before returning to a low level, which corresponds to the peak-to-trough measure in Murphy et al. [30]. Through the relation $M_{p,t} = \sigma_t k_p$ in (4), this probability reduces to a corresponding statement about the path of volatility. The main result of this section is that, for high thresholds, this probability is controlled by κ.

Denote by σ_0 the initial or current level of volatility in the GARCH model (1)–(2). Let a and b satisfy

$$0 < \sqrt{\frac{\omega}{1-\beta}} < a < \sigma_0 < b < \infty;$$

think of a and b as low and high levels of volatility, respectively. Define stopping times

$$\tau_a \quad = \quad \inf\{t \geq 1 : \sigma_t < a\}$$

$$\tau_b \quad = \quad \inf\{t \geq 1 : \sigma_t > b\}.$$

We are interested in $\mathbb{P}_{\sigma_0}(\tau_b < \tau_a)$, for large b and fixed a, where the subscript σ_0 indicates the initial level of volatility. This is the probability of a sharp rise in volatility, particularly if we take a close to σ_0.

Proposition 7 *Under conditions (A1)–(A3), there are constants $c_\ell, c_u > 0$ such that*

$$c_\ell b^{-\kappa} \leq \mathbb{P}_{\sigma_0}(\tau_b < \tau_a) \leq c_u b^{-\kappa}, \tag{8}$$

for all sufficiently large b.

To interpret this result, suppose that margin is set according to the conditional rule in (4), so $M_{p,t} = k_p \sigma_t$. Then the probability in (8) is also the probability that margin climbs to $k_p b$ starting from $k_p \sigma_0$ before declining to $k_p a$. Think of $k_p b$ as a high, fixed level of margin and an alternative to the conditional margin rule. The probability that the conditional margin climbs above $k_p b$ is then roughly $cb^{-\kappa}$. For the stable margin level to be effective, this probability should be small, which is to say that b should be large. Just how large depends on κ.

Figure 4 plots simulation estimates of $\log \mathbb{P}_{\sigma_0}(\tau_b < \tau_a)$ against $\log b$ using parameters from Table 1 for the S&P 500 and the CDX index. In each case, we set a equal to 1.4 times the minimum value of σ_t, $\sqrt{\omega/(1-\beta)}$, and $\sigma_0 = \mathbb{E}[\sigma_\infty]$. The straight lines in the figure have slope $-\kappa = -5.2$ (solid) for the S&P 500 and $-\kappa = -2.4$ (dashed) for the CDX index. As predicted by Proposition 7, the decay rate of the log probability is well approximated by κ in each case. The vertical offset between the simulation estimates and the asymptotic slopes results from the constants c_ℓ, c_u in (8).

4 The Probability of a Margin Breach

In Section 2, we saw that the margin level required to meet a target exceedance probability unconditionally can be much larger than the average level required to meet the target conditional on the current market environment. This conclusion holds whenever the unconditional, long-run distribution of volatility has a heavier tail than the conditional distribution.

In this section, we examine the transition from the lighter tails associated with the changes in volatility over a single time step to the heavier tails produced over a long horizon. We continue to work within the GARCH(1,1) setting of Section 2.1, and we show that the transition is, in a certain sense, abrupt. The main result of this section associates with each tail probability p a time T_p. If the margin requirement is set at its unconditional level u_p, then over horizons longer than T_p the

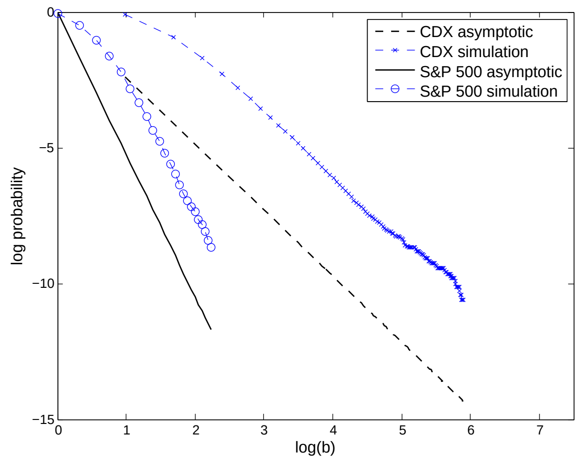

Figure 4: The decrease in $\log \mathbb{P}_{\sigma_0}(\tau_b < \tau_a)$, as estimated by simulation, is well approximated by a straight line of slope $-\kappa$, using GARCH parameters from Table 1. Source: authors' analysis.

probability of a margin breach is approximately p; over horizons shorter than T_p, the probability is much smaller than p. The critical time T_p depends on the parameter κ.

Given a GARCH model (1)–(2), let

$$\psi(s) = \log \mathbb{E}[(\alpha Z^2 + \beta)^s], \quad s \geq 0; \tag{9}$$

we assume that $\psi(s)$ is finite for all $s < \bar{s}$, for some $\bar{s} > 0$. If Z is normal or Laplace, we may take $\bar{s} = \infty$, but if Z has a scaled t_ν distribution, then the largest such $\bar{s}$ is $\nu/2$.

The function ψ is convex on $[0, \bar{s})$, with $\psi(0) = 0$, $\psi'(0) = \mathbb{E}[\log(\alpha Z^2 + \beta)] \leq \log(\alpha + \beta) < 0$. Under (A3), we have $\psi(s) > 0$ for some $s \in (0, \bar{s})$, and this ensures the existence of the parameter κ defined by the condition $\psi(\kappa/2) = 0$. See Figure 5. We write $s_{\min}$ for the point at which ψ is minimized.

For each point $s \in (0, \bar{s})$ with $\psi'(s) > 0$, we may interpret $\psi'(s)$ as a potential rate of increase for $\log \sigma_t^2$. In other words, to assess the probability that σ_t exceeds some high level v, which is the probability that $\log \sigma_t$ exceeds $\log v$, we check the slope $(\log v)/t$. Similarly, the time it takes the process $\log \sigma_t$ to reach $\log v$, growing at an average rate of $\psi'(s) > 0$, is

$$t_v \approx \log v / \psi'(s). \tag{10}$$

By varying s over $(s_{\min}, \bar{s})$, we obtain any slope between 0 and $\lim_{s \to \bar{s}} \psi'(s)$, which is typically infinite. We will see, building on Buraczewski et al. [7], that if the slope $\psi'(s)$ is smaller than $\psi'(\kappa/2)$, then the probability that σ_t exceeds v is asymptotically equivalent to the probability that σ_∞ exceeds v; if this slope is greater than $\psi'(\kappa/2)$, then the probability that σ_t exceeds v is

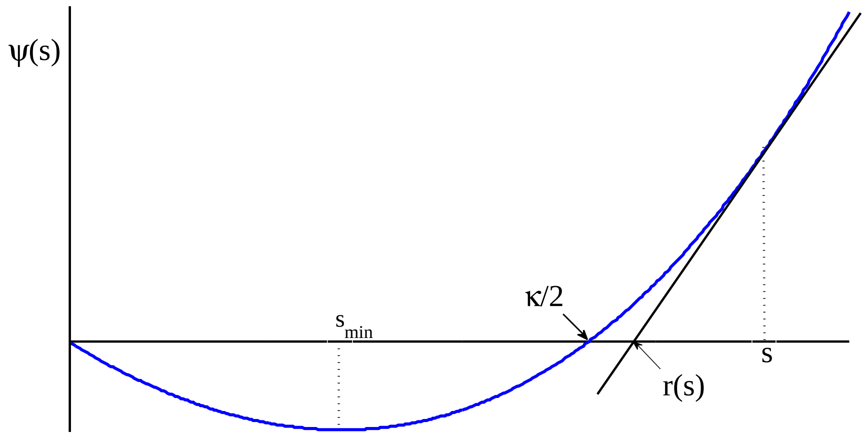

Figure 5: Illustration of $\psi(s)$, $s_{\min}$, $\kappa/2$, and $r(s)$. Source: authors' analysis.

asymptotically negligible compared to the probability that σ_∞ exceeds v. In this sense, $\psi'(\kappa/2)$ determines the natural growth rate of volatility, conditional on reaching a large level.

This discussion presupposes that the volatility process starts at a moderate level — if σ_0 is very large, then even a large level u may be reachable in a short amount of time. We will make the simplifying assumption that volatility starts at a low level $\sigma_0^2 \le \omega/(1-\beta)$. (The interval $[0, \omega/(1-\beta))$ is transient for σ_t^2.) This simplification yields the following property:

Lemma 1 *If $0 < \sigma_0 \le \sqrt{\omega/(1-\beta)}$, then $\{\sigma_t, t = 0, 1, 2, \dots\}$ is a stochastically increasing sequence. In other words, $\mathbb{P}(\sigma_t > x)$ is increasing in t, for all $x > 0$.*

Let c_0 be the constant in (13), and let $c_1 = c_0\mathbb{E}[|Z|^\kappa]/2$, as in (15); our assumption that $\psi(s) < \infty$ for some $s > \kappa/2$ ensures that Z^2 has a finite moment of order $\kappa/2$. For $p > 0$ and $s \in (s_{\min}, \bar{s})$, set

$$t_p = \left\lfloor \frac{\log(c_1/p)}{\kappa\psi'(s)} \right\rfloor. \tag{11}$$

Consistent with the discussion of (10), we interpret t_p as the time to reach $(1/\kappa)\log(c_1/p)$ for a process growing at rate $\psi'(s)$. Moreover, through (16), $(1/\kappa)\log(c_1/p)$ approximates the log of the unconditional quantile u_p of X_∞, so t_p approximates the time to reach u_p at rate $\psi'(s)$. For the special case $s = \kappa/2$, set

$$T_p = \left\lfloor \frac{\log(c_1/p)}{\kappa\psi'(\kappa/2)} \right\rfloor. \tag{12}$$

If we adopt the unconditional margin level u_p in (5), then a margin breach occurs at time t if $X_t > u_p$. By definition, the stationary probability of a margin breach is $\mathbb{P}(X_\infty > u_p) = p$. The following result approximates this probability for finite t by considering a sequence of times t_p that increase without bound as p decreases to zero. Observe that with s fixed we have $t_p < T_p$, $t_p = T_p$ or $t_p > T_p$, for all sufficiently small p.

Theorem 1 *Suppose $\sigma_0 = \sqrt{\omega}$. (i) If $t_p > T_p$, then $\mathbb{P}(X_{t_p} > u_p) \sim p$, as $p \to 0$. (ii) If $t_p = T_p$, then $\mathbb{P}(X_{t_p} > u_p) \sim p/2$. (iii) If $t_p < T_p$ and $\mathbb{E}[\exp(\gamma|Z|)] < \infty$ for some $\gamma > 0$, then $\mathbb{P}(X_{t_p} > u_p) = p^{\delta+o(1)}$, for some $\delta > 1$.*

This result should be interpreted as follows. Part (i) says that the unconditional margin yields the target exceedance probability over horizons longer than T_p. The last statement of the theorem states that over horizons shorter than T_p, the exceedance probability will be much smaller than p. Part (ii) is a knife-edge case: at T_p itself, the probability of a margin breach is $p/2$, hence smaller than p but of the same order of magnitude. The overall conclusion is that the unconditional margin is conservative over horizons shorter than T_p but not over longer horizons.

We can be more explicit about the exponent δ appearing in part (iii) of the theorem. For any $s \in (s_{\min}, \bar{s})$ with $\psi'(s) > 0$, define

$$r(s) = s - \frac{\psi(s)}{\psi'(s)}.$$

Graphically, $r(s)$ is the point at which the line tangent to ψ at s crosses the horizontal access. In particular, $r(\kappa/2) = \kappa/2$, and $r(s)$ is increasing for $s \in [\kappa/2, \bar{s})$. This representation, which is also used in Buraczewski et al. [7], differs from the expression used in Nyrhinen [33] but is equivalent to it and yields the following result:

Proposition 8 *Let t_p be as in (11) for some $s \in (\kappa/2, \bar{s})$. Then the exponent in part (iii) of Theorem 1 is given by $\delta(s) = 2r(s)/\kappa > 1$, which is increasing for $s \in (\kappa/2, \bar{s})$.*

This result provides a rough measure of the probability $\mathbb{P}(X_t > u_p)$ of a margin breach, starting from a low initial volatility σ_0, as follows. A breach occurs at time $t < T_p$ if the growth rate of $\log \sigma$ is $\log u_p/t$. The value of s that yields this growth rate satisfies $\psi'(s) = \log u_p/t$. The probability of a margin breach is then $p^{\delta(s)+o(1)}$. As t increases to T_p, s decreases to $\kappa/2$, $\delta(s)$ decreases to 1, and the probability of a margin breach approaches p.

5 Analysis of Countermeasures

In this section, we apply the framework developed in previous sections to evaluate specific rules adopted by the European Union to counter procyclicality in CCP margin requirements. These measures were enacted by the European Commission in December 2012 (European Commission [17]). As of 2016, these rules are effectively binding on U.S. central counterparties seeking to clear trades for European firms. An agreement between the primary U.S. regulator of CCPs, the Commodity Futures Trading Commission, and European regulators determined that rules in the two jurisdictions would be considered equivalent under certain conditions, and one of these conditions is adherence to the European Union's rules on limiting procyclicality.[1]

The European Union rules (European Commission [17], Article 28) require a CCP to adopt one of three measures to counter procyclicality, stating that "the CCP shall employ at least one of the following options:

(a) applying a margin buffer at least equal to 25% of the calculated margins which it allows to be temporarily exhausted in periods where calculated margin requirements are rising significantly;

(b) assigning at least 25% weight to stressed observations in the lookback period calculated in accordance with Article 26;

(c) ensuring that its margin requirements are not lower than those that would be calculated using volatility estimated over a 10 year historical lookback period."

These rules are vague and therefore open to interpretation, despite the specificity suggested by the numerical values associated with the three options. The rules also fail to differentiate between features of different asset classes. We examine each of the options more closely through the lens of the GARCH model.

5.1 The 25 Percent Buffer

The first of the three options above presupposes that the CCP allows clearing members to fall short of their required margin levels when these levels are rising quickly. This rule is difficult to evaluate because the permitted shortfall is unspecified. Indeed, a CCP that permitted no shortfall could in theory comply with this rule with no additional buffer.

As a conservative analysis, we can investigate whether adding a 25 percent buffer to the average level of a procyclical margin requirement approximates the level of an unconditional margin requirement. This comparison is conservative because it implies a buffer at least as large would be calculated under E.U. rules.

In the notation of Section 2.1, we want to compare the ratio u_p/m_p to 1.25. We already know that this ratio increases without bound as p approaches zero, so the threshold of 1.25 is eventually exceeded. For a practical comparison, we take $p = 0.01$ or $p = 0.005$, corresponding

[1]See "The U.S. Commodity Futures Trading Commission and the European Commission: Common Approach for Transatlantic CCPs," February 10, 2016, at www.cftc.gov/PressRoom/PressReleases/cftc_euapproach021016.

		One-Day Period				*Five-Day Period*		
		Margin Ratio		Stress		Margin Ratio		Stress
Series	κ	99%	99.5%	Threshold	κ	99%	99.5%	Threshold
S&P 500	5.2	1.17	1.26	6%	6.2	1.12	1.17	7%
EUROSTOXX 50	5.4	1.17	1.24	6%	6.6	1.11	1.16	8%
CDX NA IG	2.4	1.48	1.73	3%	2.7	1.40	1.65	3%
US 10yr IR Swap	3.8	1.24	1.38	5%	4.1	1.24	1.35	4%
USD-BRL	2.6	1.45	1.69	3%	2.8	1.18	1.26	5%
Brent Crude Oil	4.6	1.19	1.26	6%	5.1	1.18	1.26	5%

Table 2: The third and fourth columns show estimated margin ratios u_p/m_p for six time series at $1-p = 0.99$ and $1-p = 0.995$. Estimates of κ are reproduced from Table 1 for comparison. Smaller values of κ yield higher margin ratios. The fifth column shows the stress threshold required to match the margin ratio at a 99 percent confidence level. The last four columns report corresponding results for a five-day margin period of risk. Source: authors' analysis.

to confidence levels of 99 percent and 99.5 percent, respectively. Using the parameter values in Table 1, we simulate GARCH(1,1) processes and estimate u_p/m_p from a large number of simulated observations.

The results are summarized in Table 2 for the time series reported in Table 1. The left panel corresponds to a one-day horizon. The two columns under "Margin Ratio" report u_p/m_p at $p = 0.01$ and $p = 0.005$. Interestingly, with margin levels set at 99.5 percent confidence, a buffer of 25 percent looks about right for several of the asset classes. However, the 25 percent buffer falls significantly short of the margin ratio *precisely when κ is small*. This pattern is consistent with the asymptotic results of earlier sections, which emphasize the importance of the index κ; but the numerical results in the table confirm that these insights are relevant at practical parameter values. In fact, in the left panel the correlation between the κ values and each of the margin ratio columns is -0.96, indicating that κ very effectively explains the size of the required buffer.

In the right panel of Table 2, we report corresponding results for a five-day horizon. We convert the daily GARCH parameters in Table 1 to five-day parameters using equations (9) and (10) of Drost and Nijman [13] and then calculate the resulting κ values. As expected, the κ values are larger over the longer horizon. But the main qualitative conclusion remains unchanged: the margin ratios and stress thresholds vary across assets, and the κ values help explain the variation.

The main implication of these numerical examples is that the first option in the E.U. rules could be improved by making the size of the buffer depend on characteristics of specific derivative contracts or asset classes. In particular, asset classes with greater persistence and burstiness in volatility require larger buffers to mitigate procyclicality.

5.2 Stressed Observations

The second option under the E.U. rules suffers from two significant indeterminacies: it does not identify which observations should be considered stressed, and it does not define what it means to give these observations a weight of 25 percent. (Article 26, to which the rules refer, is titled "Time horizons for the liquidation period" and deals with the number of days a CCP requires to liquidate the portfolio of a clearing member in default.) Within the framework of Section 2, we can interpret the 25 percent weight to mean that the conditional margin $M_{p,t}$ in (4) should instead be set according to the rule

$$\tilde{M}_{p,t} = k_p \left(0.25\sigma_{\text{stress}} + 0.75\sigma_t\right),$$

with σ_{stress} estimated from stressed observations in some as yet unspecified manner. Such a rule creates some stickiness in margin levels by putting less weight on the current level of volatility σ_t. Asymptotically, however, this specification fails to address the qualitative behavior of u_p/m_p we observed before: Propositions 1 and 4 apply to $\tilde{M}_{p,t}$ as they did to $M_{p,t}$, with only the constants c_N, c_L, and c_ν changing. Asymptotically, the adjustment has a negligible effect.

At fixed values of p, we can again use simulation estimates to evaluate the potential improvement using $\tilde{M}_{p,t}$ rather than $M_{p,t}$. To do so, we need to specify how the stressed volatility is determined. We assume that the stressed volatility is estimated from the largest (in absolute value) x percent of observations. A smaller value of x yields a larger stressed volatility because it relies on more extreme observations.

Rather than fix a particular value of x and estimate margin ratios, we take the opposite perspective: we evaluate how small x would have to be in order to get $\tilde{m}_p = u_p$, where $\tilde{m}_p$ is the long-run average value of $\tilde{M}_{p,t}$, in the sense of (6). The results appear in the last column of Table 2, based on $p = 0.01$. For example, in order to get $\tilde{m}_p = u_p$ for the GARCH parameters of the S&P 500 index, we need to calculate σ_{stress} from the most extreme six percent of observations. For the CDX index, with a smaller κ, the required threshold drops to three percent. This difference is large: it implies that only half as many observations should be used to calculate a stressed volatility for the CDX index as would be used for the S&P 500 index.

The numerical examples suggest that putting 25 percent weight on stressed observations is a potentially reasonable way to counter procyclicality, under our interpretation of the rule. However, the examples again point to significant heterogeneity across asset classes in the appropriate implementation of the rule, as a consequence of different levels of persistence and burstiness in volatility.

5.3 The 10-Year Lookback

The third option in the E.U. rules allows a CCP to set margin levels "using volatility estimated over a 10 year historical lookback period." Our framework suggests that, taken literally, this option is ineffective. Volatility estimated from a long time series would come close to the stationary mean $\mathbb{E}[\sigma_\infty]$. Scaling this long-run average level of volatility by the multiplier k_p in (4) would produce a stable (in fact, constant) margin level; but this level would be precisely the average m_p which we have argued is insufficient. As stated, the rule fails to specify how a long-run average level of volatility $\mathbb{E}[\sigma_\infty]$ should be translated into an effective level of margin. In particular, there is no simple relationship between $\mathbb{E}[\sigma_\infty]$ and the unconditional margin level u_p.

The rule would be more effective if it referred to estimating a *quantile* over a long lookback period, rather than a *volatility*. Given a history of price changes, the sample quantile at a tail probability of p converges to u_p in great generality. (We formulate the sample quantile precisely in the next section.) So, within our framework, using a sample quantile eventually yields the right answer. The only remaining question is how long a history one needs to estimate u_p accurately.

Here again we can gain some insight at practical parameter values through simulation experiments. For each of the asset classes in Table 1, we simulate long histories of GARCH processes at the estimated parameter values. We delete an initial segment of 50,000 observations to ensure stationarity of the simulated data. We then plot the ratio of the sample quantile estimated from n years to our estimate from a simulation of 50 years, which we take as an accurate approximation to u_p. We use $p = 0.005$ in our examples.

Results for two cases are shown in Figure 6; the other cases are very similar. The sample quantiles (marked with circles) converge slowly and require a surprisingly long lookback to reach their limiting values. A 10-year history looks like a good rule of thumb, and a history of, say, two years is clearly inadequate. The estimates in the figure are shown in multiples of the stationary volatility $\mathbb{E}[\sigma_\infty]$ for each case. The figure also shows estimates of the average conditional quantile m_p (marked with asterisks), which converge very quickly.

These examples (and similar results for other cases) suggest that the 10-year lookback specified in the E.U. rules is effective, provided the rule is modified to refer to a quantile estimate rather than a volatility estimate.

For comparison, the plots in Figure 6 include a third line (marked with crosses), labeled "Unconditional, IID." For this sequence, we create an i.i.d. sequence of price changes with the same marginal distribution as the simulated GARCH price changes by resampling randomly from the empirical distribution of the simulated GARCH data. Because the two time series have the same marginal distributions, they have the same quantile, and their quantiles have the same dependence

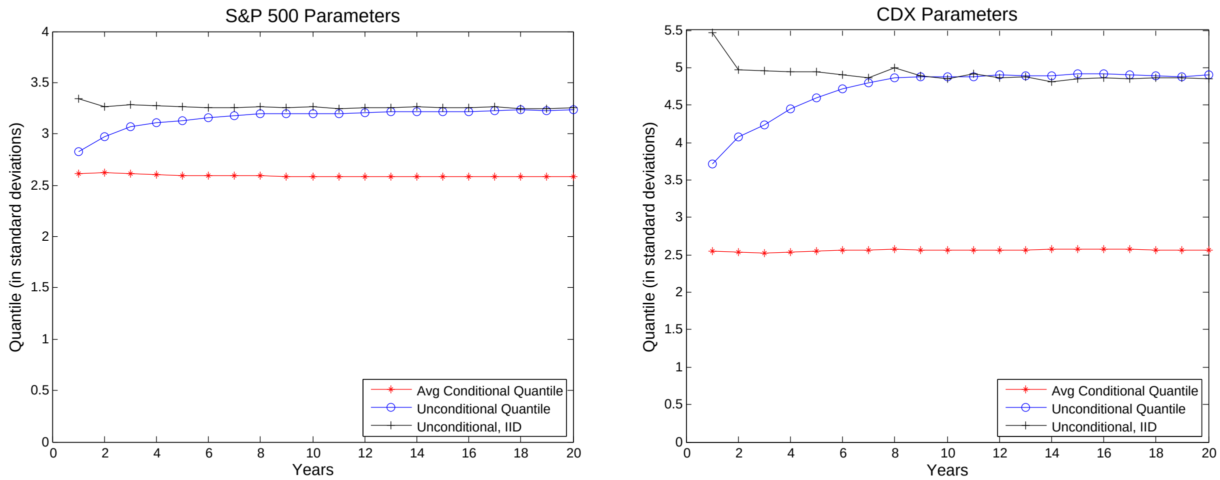

Figure 6: Convergence of quantile estimates in GARCH simulations using parameters from Table 1. Source: authors' analysis.

on κ with the same value of κ. Nevertheless, the sample quantiles estimated from i.i.d data converge much more quickly than the sample quantiles from the GARCH data. In other words, the slow convergence of the sample quantiles in the GARCH data cannot be ascribed to κ or to heavy tails — those features are shared by the i.i.d. data; the slow convergence must instead be due to serial dependence in the GARCH data. We make this conclusion precise in the next section.

6 Quantile Estimation

In this section, we seek to explain the long lookback periods required in Section 5.3 to obtain accurate estimates of unconditional margin levels u_p. As before, let $X_1, X_2, \ldots, X_n$ denote price changes, and suppose that this sequence is stationary. In our earlier notation, this means that X_1 has the distribution of X_∞. Let

$$X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$$

denote the ordered observations. The sample quantile associated with upper tail probability p is

$$\hat{U}_p = X_{\lceil n(1-p) \rceil : n},$$

where $\lceil x \rceil$ denotes the smallest integer greater than or equal to x. If $p < 1/n$, then $\hat{U}_p$ is simply $X_{n:n}$, the largest value observed.

Alternative estimators, based on extreme value theory, seek to extract more information from the data to provide better estimates of u_p, particularly at small values of p. If the tail of X decays like a constant times $u^{-\kappa}$, for some $\kappa > 0$, then one can try to extrapolate from a less extreme quantile to a more extreme quantile using an estimator of the form (see Drees [12], p.619)

$$\hat{V}_p = X_{n-k_n:n} \left(\frac{np}{k_n} \right)^{-\hat{\kappa}}.$$

Here, k_n is an intermediate value that grows with n but more slowly than n, and $\hat{\kappa}$ is an estimate of the tail decay power κ. A popular choice is the Hill estimator

$$\hat{\kappa} = \frac{1}{k_n} \sum_{i=1}^{k_n} \log \frac{X_{n-i+1:n}}{X_{n-k_n:n}}.$$

The sample quantile estimator $\hat{U}_p$ requires almost no assumptions on the underlying data. The extreme value estimator $\hat{V}_p$ requires that the tail of X have a power law. In the case of a GARCH(1,1) process, the power κ can be calculated directly from (7), given estimates of α, β, and the distribution of Z; however, this requires imposing stronger assumptions on the data. Thus, we focus on the sample quantile, which is probably the most widely used estimator in practice. We examine the behavior of the sample quantile when the data generating process is GARCH(1,1) but without using knowledge of this process to estimate the quantile. In other words, we consider a case in which the CCP estimates the unconditional margin level from historical data which, unbeknownst to the CCP, comes from a GARCH process. (In their simulation study, Dutta and Biswas [14] find that the sample quantile estimator performs about as well as the extreme value estimator in GARCH models; see their Table 10 and especially their model (ix) which has parameters similar to those we find in Table 1. This lends further support to our focus on $\hat{U}_p$.)

We discuss two results that shed light on the performance of the sample quantile. Drees [12], p.627, gives conditions under which the relative error of the estimator $\hat{U}_p$ is asymptotically normal; more precisely,

$$\sqrt{np} \left(\frac{\hat{U}_p}{u_p} - 1 \right) \Rightarrow N(0, \sigma_U^2)$$

where $\Rightarrow$ denotes convergence in distribution, $N(0, \sigma_U^2)$ denotes the normal distribution with mean zero and some variance $\sigma_U^2 > 0$, and the limit holds as $np \to \infty$.

This result implies that the sample size n needs to be substantially larger than $1/p$. If, for example, a sample size $n_{.01}$ is sufficient to provide an adequately accurate estimate of the unconditional margin u_p at 99 percent confidence ($p = .01$), then the sample size required to estimate u_p at 99.9 percent margin ($p = 0.001$) is greater than $10n_{.01}$. A sample size requirement of 1–2 years can thus quickly grow to a requirement of 5–10 years at high confidence levels.

For $\hat{U}_p$ to be close to u_p, we need a fraction p of the observations to exceed u_p and the remaining fraction $1 - p$ to fall below u_p. Clearly, we need at least one observation that exceeds u_p to get an accurate estimate of the quantile. We consider the time required until the first such exceedance, taking this duration as a measure of the difficulty in estimating u_p. To state the result, let

$$\tau_p = \inf\{n \geq 0 : X_n > u_p\}.$$

Let ξ denote an exponential random variable with mean 1.

Proposition 9 *In a stationary GARCH(1,1) model satisfying (A1)–(A3),*

$$p\tau_p \Rightarrow \xi/\theta, \quad \text{as } p \to 0,$$

with $\theta \in (0, 1)$.

The parameter θ in this result comes from extreme value theory and is known as the *extremal index* of the process X_t; see Section 8.1 of Embrechts et al. [15] for a general discussion of the extremal index, and see Mikosch and Stărică [29] for the specific case of a GARCH(1,1) process. The extremal index measures the clustering of exceedances. Roughly speaking, once X_t exceeds u_p, we expect to see a cluster of exceedances with a mean of $1/\theta$.

To interpret Proposition 9, contrast it with the case of an i.i.d. sequence. By definition, $\mathbb{P}(X_t > u_p) = p$; so, in the i.i.d. case, the time τ_p until the first exceedance has a geometric distribution with mean $1/p$. For geometrically distributed τ_p, the product $p\tau_p$ converges in distribution to a standard exponential random variable as $p \to 0$. Proposition 9 shows that the time until the first exceedance is much longer if θ is small, because then ξ/θ is much larger than ξ. If exceedances occur on average with frequency p, and if they occur in clusters when they occur, then the time between clusters must on average be longer. If, for example, each cluster had exactly $1/\theta$ exceedances, then the average time between clusters would have to be p/θ in order that the overall exceedance rate remain p.

Estimating the extremal index from historical data is difficult; see Mikosch and Stărică [29] for an example. One can fit a model to the data, such as a GARCH model, and then calculate θ for the model, but this requires imposing additional assumptions. The main point of this discussion is that greater burstiness in volatility will lead to greater clustering of exceedances and therefore slower convergence of quantile estimates. This qualitative relationship should inform the choice of lookback period in setting margin requirements to mitigate procyclicality.

7 Concluding Remarks

The growing use of margin requirements in over-the-counter derivative transactions reduces counterparty credit risk but may amplify shocks to the financial system through margin calls when volatility spikes. We have examined the problem of setting margin levels that are high enough to reduce credit exposure yet stable enough to avoid these amplifying, procyclical effects. Using a simple GARCH model as a lens, we have given a precise formulation of the problem of procyclicality through conditional and unconditional margin levels. The GARCH setting yields two key insights: the buffer required to offset procyclicality depends on the tail exponent κ, and the lookback period required to estimate the quantile accurately depends on the extremal index θ. Both of these parameters depend on the persistence and burstiness of volatility. The qualitative impact these features of volatility have on margin procyclicality should hold well beyond the GARCH setting and should inform current practice in mitigating procyclicality.

A Comparison of Margin Levels

Proof of Proposition 2

Proof: The unconditional distribution of X_∞ has the distribution of $\sigma_\infty Z$, with Z independent of σ_∞, so the unconditional distribution inherits three properties from Z: if Z is symmetric then so is X_∞; if Z has unbounded support then so does X; and if Z has a continuous distribution then so does X_∞. All three properties apply in the normal case.

Symmetry implies that $\mathbb{P}(X_\infty > u) = \mathbb{P}(|X_\infty| > u)/2$, for any $u > 0$. By Theorem 2.1 of Mikosch and Stărică [29],

$$\mathbb{P}(\sigma_\infty > u) \sim c_0 u^{-\kappa}, \quad \text{as } u \to \infty, \tag{13}$$

and

$$\mathbb{P}(|X_\infty| > u) \sim c_0 \mathbb{E}[|Z|^\kappa] u^{-\kappa}, \quad \text{as } u \to \infty, \tag{14}$$

with $c_0 > 0$ a constant. Unbounded support implies that the unconditional quantile u_p satisfies $u_p \to \infty$, as $p \to 0$. Continuity implies that $\mathbb{P}(X_\infty > u_p) = p$. Thus,

$$p = \mathbb{P}(X_\infty > u_p) \sim \frac{c_0}{2} \mathbb{E}[|Z|^\kappa] u_p^{-\kappa}, \quad \text{as } p \to 0,$$

and then, with

$$c_1 = \frac{c_0}{2} \mathbb{E}[|Z|^\kappa], \tag{15}$$

$$u_p \sim (p/c_1)^{-1/\kappa} \equiv \hat{u}_p. \tag{16}$$

For the normal distribution function Φ and density ϕ, we have $1 - \Phi(x) \sim \phi(x)/x$, as $x \to \infty$. With $p = 1 - \Phi(k_p)$, this yields $p \sim \phi(k_p)/k_p$, as $p \to 0$, and then $\log p + \log k_p \sim \log \phi(k_p)$, which we can write as

$$\frac{1}{2} k_p^2 \sim -\log p - \log k_p - \log \sqrt{2\pi}.$$

Since $\log k_p/k_p^2 \to 0$, this yields $k_p^2 \sim -2\log p$ and $k_p \sim \sqrt{2\log(1/p)}$. Combining this limit for k_p with the limit of u_p proves the result. The constant c_N is given by $c_N = c_1^{1/\kappa}/(\sqrt{2}\mathbb{E}[\sigma_\infty])$, with c_1 as in (15). $\square$

Proof of Proposition 4

Proof: The proof follows the same argument as that of Proposition 2; only the analysis of k_p changes. With the Laplace distribution,

$$p = \mathbb{P}(Z > k_p) = \frac{1}{2}e^{-\sqrt{2}k_p},$$

so $k_p = (1/\sqrt{2})\log(1/2p) \sim (1/\sqrt{2})\log(1/p)$. The constant c_L is given by $c_L = c_1^{1/\kappa}/(\sqrt{2}\mathbb{E}[\sigma_\infty])$, with c_1 as in (15). In the case of a scaled t distribution, we have $\mathbb{P}(Z > z) \sim \tilde{c}z^{-\nu}$, for some $\tilde{c} > 0$, so $k_p \sim \tilde{c}^{1/\nu}p^{-1/\nu}$. Using the density of the scaled t distribution, one finds that

$$\tilde{c} = \frac{\Gamma\left(\frac{\nu+1}{2}\right)(\nu-2)^{\nu/2}}{\nu\sqrt{\pi}\Gamma(\nu/2)},$$

and then $c_\nu = c_1^{1/\kappa}/(\tilde{c}^{1/\nu}\mathbb{E}[\sigma_\infty])$. $\square$

Proof of Proposition 5

Proof: If the distribution of Z is normal or Laplace, then $k_p^{ES}/k_p \to 1$ as $p \to 0$. If Z has a scaled t_ν distribution, then its tail distribution is regularly varying with index $-\nu$, and this implies $k_p^{ES}/k_p \to \nu/(\nu-1)$. The unconditional distribution X_∞ is regularly varying with index $-\kappa$ (see Mikosch and Stărică [29]), so $u_p^{ES}/u_p \to \kappa/(\kappa-1)$ The proof now follows by combining these limits with Propositions 2 and 4. $\square$

B Alternative Models

In this section, we review some alternative models that yield the key feature of the GARCH(1,1) model discussed in the previous section — namely, the qualitative divergence between the average conditional margin and the unconditional margin.

B.1 Asymmetric and Higher Order GARCH Models

In the GJR-GARCH model of Glosten, Jagannathan, and Runkle [20], negative returns amplify volatility more than positive returns. The model replaces the variance equation (2) with

$$\sigma_t^2 = \omega + \alpha X_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma X_{t-1}^2\mathbf{1}_{\{X_{t-1}<0\}}, \quad \omega, \alpha, \beta > 0.$$

If $\gamma > 0$, then a large squared return X_{t-1}^2 will have a greater effect on volatility when $X_{t-1} < 0$ than when $X_{t-1} > 0$.

If we continue to assume that the innovations Z_t have a symmetric, standardized distribution, then a sufficient condition for the existence of a stationary distribution of (X_t, σ_t) is $\alpha+\beta+\gamma/2 < 1$. The equation

$$\mathbb{E}[(\alpha Z^2 + \beta + \gamma Z^2 \mathbf{1}_{\{Z<0\}})^{\kappa/2}] = 1 \tag{17}$$

has a unique positive solution $\kappa > 0$ if we modify (A3) accordingly. The results of Sections 2.1 and 2.2 then hold for the GJR-GARCH model with κ defined by (17).

In a GARCH(p,q) model, $p, q \geq 1$, the variance equation becomes

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2. \tag{18}$$

We assume that the parameters α_i and β_j are not all identically zero and that they satisfy

$$\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1,$$

which is sufficient for the existence of a stationary distribution; see Basrak et al. [4], p.107.

The vector $Y_t = (\sigma_{t+1}^2, \ldots, \sigma_{t-q+2}^2, X_t^2, \ldots, X_{t-p+2}^2)^\top$ satisfies a linear difference equation of the form

$$Y_t = A_t Y_{t-1} + B_t,$$

where $B_t = (\omega, 0, \ldots, 0)^\top$, and, to encode (18), the first row of the matrix A_t is given by

$$(\alpha_1 Z_t^2 + \beta_1, \beta_2, \ldots, \beta_q, \alpha_2, \ldots, \alpha_p).$$

In particular, each A_t is a function only of Z_t and the model parameters, so the A_t are i.i.d. matrices. See equation (3.2) of Basrak et al. [4] for a complete specification of A_t.

For any $d \times d$ matrix A, let $\|A\| = \sup_{\|x\|=1} \|Ax\|$, with the supremum taken over vectors $x \in \mathbb{R}^d$ and with $\|x\|$ denoting the Euclidean norm of the vector x. In this setting, it follows from Kesten [23] that the appropriate generalization of κ is defined by

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\|A_n \cdots A_1\|^\kappa] = 0. \tag{19}$$

We now have the following:

Proposition 10 *Propositions 2 and 4 hold under the GJR-GARCH model with κ defined by (17), and they hold under the GARCH(p,q) model with κ defined by (19).*

Proof: For the GJR-GARCH model, we can write

$$\sigma_t^2 = \omega + (\alpha Z_{t-1}^2 + \beta + \gamma Z_{t-1}^2 \mathbf{1}_{\{Z_{t-1}<0\}})\sigma_{t-1}^2 \equiv \omega + a_t \sigma_{t-1}^2.$$

Then $\mathbb{E}[\log a_t] \leq \log \mathbb{E}[a_t] = \log(\alpha + \beta + \gamma/2) < 0$. Under the normal, Laplace, and Student distributions for Z, we also have $\mathbb{P}(a_t > 1) > 0$. It follows from Theorem 2.1 of Mikosch and Stărică [29] that (17) has just one positive solution. The rest of the analysis then follows as in the GARCH(1,1) setting. For the GARCH(p,q) model, existence and uniqueness of κ in (19) follows from Theorem 3.1(B) of Basrak et al. [4], under our parameter and distributional assumptions. The same result also shows that this κ determines the tail decay of the stationary distribution of σ_t and X_t. The rest of the proof then follows as in the GARCH(1,1) case. □

B.2 Regime-Switching Model

In some cases, the onset of elevated volatility is well described through a change in regime that produces a structural break in the dynamics of volatility; the transition in the second half of 2007 illustrated in Figure 1 might, for example, be modeled this way.

In a regime-switching GARCH model, the parameters of the volatility equation depend on the state of an underlying Markov chain S_t, so we write $\omega(S_t)$, $\alpha(S_t)$, and $\beta(S_t)$ to indicate this dependence. The Markov chain S_t is irreducible with finite state space $\{1, 2, \ldots, N\}$. We assume that some state (which we may take to be N) is reachable in a single step from every other state,

$$\mathbb{P}(S_{t+1} = N | S_t = i) > 0, \quad i = 1, \ldots, N-1, \tag{20}$$

which implies, in particular, that the chain is aperiodic. We denote by $\pi_1, \pi_2, \ldots, \pi_N$, the Markov chain's stationary probabilities. The innovations Z_t are i.i.d. and, in particular, do not depend on S_t.

We assume that $\omega(i) > 0$ and $\beta(i) > 0$, for $i = 1, \ldots, N$, and we impose the stability condition

$$\sum_{i=1}^{N} \pi_i \mathbb{E}[\log(\alpha(i)Z^2 + \beta(i))] < 0. \tag{21}$$

This holds, for example, if $\sum_i \pi_i(\alpha(i) + \beta(i)) < 1$, even if $\alpha(i) + \beta(i) > 1$, for some i. We suppose the existence of $r_0 > 0$ such that

$$1 < \min_i \mathbb{E}[(\alpha(i)Z^2 + \beta(i))^{r_0/2}] \leq \max_i \mathbb{E}[(\alpha(i)Z^2 + \beta(i))^{r_0/2}] < \infty. \tag{22}$$

For any $r \geq 0$, define the $N \times N$ matrix $B(r)$ by setting

$$B_{ij}(r) = \mathbb{P}(S_{t+1} = j | S_j = i)\mathbb{E}[(\alpha(j)Z^2 + \beta(j))^{r/2}],$$

whenever the expectations are finite for all $j = 1, \ldots, N$. The matrix $B(r)$ has nonnegative entries, and it inherits irreducibility from the transition probabilities of the Markov chain. By the Perron-Frobenius theorem, the spectral radius $\rho(r)$ of $B(r)$ is real and positive. At $r = 0$, $B(0)$ is the transition matrix of S_t, so $\rho(0) = 1$.

Proposition 11 *Suppose (20), (21), and (22) hold. There exists a unique solution $\kappa > 0$ to the equation $\rho(\kappa) = 1$, and Propositions 2 and 4 hold for the regime-switching GARCH model with this κ.*

Proof: Write the variance equation as

$$\sigma_{t+1}^2 = \omega(S_t) + (\alpha(S_t)Z_t^2 + \beta(S_t))\sigma_t^2 \equiv \omega(S_t) + a_t(S_t)\sigma_t^2.$$

The existence of a stationary distribution for σ_t^2 (and then for $(X_t, \sigma_t) = (\sigma_t Z_t, \sigma_t)$) follows from Theorem 1 of Brandt [6] under condition (21). It follows from Kingman [24] that $\log \rho(r)$ is convex in r. As already noted, $\rho(0) = 1$, so $\log \rho(0) = 0$. Condition (20) implies (see Collamore [8], p.1418) that $(\log \rho(0))' = \sum_{i=1}^N \pi_i \mathbb{E}[\log(\alpha(i)Z^2 + \beta(i))] < 0$, in light of (21). For r satisfying (22), every row sum of $B(r)$ is strictly greater than 1, so $\rho(r) > 1$. But if $\log \rho$ is a convex function with $\log \rho(0) = 0$, $(\log \rho(0))' < 0$, and $\log \rho(r) > 0$, for some $r > 0$, then there is a unique $\kappa > 0$ at which $\log \rho(\kappa) = 0$. To show that $\mathbb{P}(\sigma_\infty > x) \sim cx^{-\kappa}$, for some $c > 0$, we may apply Theorem 2.2 of Collamore [8]. Under (21), his condition (M) and (H3) are satisfied (see Collamore [8], p.1418). His conditions (H1) and (H2) are then implied by (22). $\Box$

B.3 Diffusion Models

We mainly work in discrete time, where the notion of conditional margin is well-defined. However, in this section we consider continuous-time models. Essentially any diffusion model looks approximately Gaussian over a short time period. Our focus then is on settings that lead to heavy-tailed stationary distributions.

Building on the regime-switching model of Section B.2, we may consider a Markov-modulated diffusion model of the form

$$d\sigma_t = a(S_t)(\sigma_t - \bar{\sigma})\, dt + b(S_t)\, dW_t,$$

in which W is a standard Brownian motion, $\bar{\sigma}$ is a constant, and S_t is now a continuous-time Markov chain on $\{1, \ldots, N\}$ with stationary distribution π.

For fixed $a < 0$ and $b > 0$, this equation defines an Ornstein-Uhlenbeck process with a Gaussian stationary distribution. The tails of the unconditional (stationary) distribution are thus qualitatively the same as those of the conditional distribution of changes in σ_t over short time interval.

In the Markov-modulated case, Guyon, Iovleff, and Yao [22] show that σ_t has a stationary distribution if

$$\sum_{i=1}^{N} \pi_i a(i) < 0.$$

De Saporta and Yao [11] show that this stationary distribution is light-tailed if $a(i) < 0$, for all $i = 1, \ldots, N$. However, if $a(i) > 0$ for some i, then the tail of the stationary distribution is regularly varying with index $\kappa > 0$. As in Section B.2, the parameter κ is determined through the spectral radius of a matrix associated with the transition probabilities of the underlying Markov chain; see Theorem 2 of De Saporta and Yao [11]. In particular, then, burstiness in volatility, in the form of occasional periods of rapid growth in volatility (regimes in which $a(i) > 0$) create a wedge between the tail behavior of conditional and unconditional volatility. This feature is qualitatively similar to what we see in the discrete-time models considered previously.

Nelson [32] proved a continuous-time limit for GARCH(1,1) processes in which the limiting process takes the form

$$\begin{aligned} dX_t &= \sigma_t \, dW_t^{(1)} \\ d\sigma_t^2 &= (\omega - \theta\sigma_t^2)\, dt + \alpha\sigma_t^2 \, dW_t^{(2)}, \end{aligned}$$

where $W^{(1)}$ and $W^{(2)}$ are independent Brownian motions, and θ is the limiting value of $1-(\alpha+\beta)$ in a sequence of processes approaching the diffusion limit. He showed that if $2\theta/\alpha^2 > -1$, then σ_t^2 has a stationary distribution, and this distribution is an inverse gamma distribution with shape parameter $\kappa = 1 + 2\theta/\alpha^2$. It follows that the tail of the stationary distribution is regularly varying with parameter κ. Moreover, a higher degree of persistence, in the form of a smaller $\theta > 0$, and higher degree of self-excitation, in the form of a larger $\alpha > 0$, lead to a smaller κ and a heavier tail for the stationary distribution, consistent with what we saw in discrete time.

To summarize, we have shown in the section that a variety of models beyond the GARCH(1,1) process share the main property underlying the results of Section 2: their unconditional distributions are heavy-tailed even when their conditional distributions are not. This feature arises from persistence or burstiness in volatility. It leads to unconditional margin requirements that are substantially higher than average conditional margin levels.

C Analysis of Margin Calls

C.1 Proof of Proposition 6

Proof: By Theorem 2.3(c) of Mikosch and Stărică [29], the vector $(\sigma_t, \sigma_{t+1}, \ldots, \sigma_{t+n})$ satisfies multivariate regular variation with index κ, for any fixed n. It then follows from Theorem 1.1(i) of

Basrak et al. [3] that, for any $w = (w_1, \dots, w_n)$, not identically zero,

$$\mathbb{P}\left(\sum_{i=1}^{n} w_i \sigma_{t+i-1} > x\right) \sim c(w)L(x)x^{-\kappa}, \quad \text{as } x \to \infty,$$

for some slowly varying function L not depending on w. From the special case $w = (1, 0, \dots, 0)$, we see by comparison with (13) that L converges to a constant and can therefore be absorbed into $c(w)$. The two claims in the proposition are now the special cases $w = (-k_p, 0, \dots, 0, k_p)$ and $w = (k_p, \dots, k_p)$. $\Box$

C.2 Proof of Proposition 7

We separate the proof of Proposition 7 into separate arguments for the lower and upper bounds in (8).

Lower Bound

Define

$$\xi_t = \frac{1}{2}\log(\alpha Z_t^2 + \beta), \quad t = 0, 1, 2, \dots,$$

and, with $Y_0 = \log \sigma_0$, set

$$Y_{t+1} = Y_t + \xi_t, \quad t = 0, 1, 2, \dots.$$

Whenever $\exp(Y_t) \le \sigma_t$,

$$\begin{aligned} \exp(Y_{t+1}) &= \exp(Y_t + \xi_t) \\ &\le \sigma_t \sqrt{\alpha Z_t^2 + \beta} \\ &\le \sqrt{\omega + \sigma_t^2(\alpha Z_t^2 + \beta)} = \sigma_{t+1}. \end{aligned}$$

With $\exp(Y_0) = \sigma_0$, we conclude that $\exp(Y_t) \le \sigma_t$ for all t.

Define

$$\begin{aligned} \tau_a^Y &= \inf\{t \ge 1 : Y_t < \log a\} \\ \tau_b^Y &= \inf\{t \ge 1 : Y_t > \log b\}. \end{aligned}$$

Then $\sigma_t \ge \exp(Y_t)$ for all t implies that

$$\mathbb{P}_{\sigma_0}(\tau_b < \tau_a) \ge \mathbb{P}_{Y_0}(\tau_b^Y < \tau_a^Y).$$

By the definition of κ in (7), we have $\mathbb{E}[\exp(\kappa \xi_t)] = 1$. We may therefore define an i.i.d. sequence $\tilde{\xi}_t$, $t = 0, 1, \dots$, in which each $\tilde{\xi}_t$ has the distribution

$$\mathbb{P}(\tilde{\xi}_t \in \cdot) = \mathbb{E}[\exp(\kappa \xi_t)\mathbf{1}\{\xi \in \cdot\}].$$

Set $\tilde{Y}_0 = Y_0 = \log \sigma_0$ and $\tilde{Y}_{t+1} = \tilde{Y}_t + \tilde{\xi}_t$, $t = 0, 1, \ldots$. Equation (39) of Siegmund [35] yields

$$\mathbb{P}_{Y_0}(\tau_b^Y < \tau_a^Y) = \mathbb{P}_{Y_0}(\tau_b^{\tilde{Y}} < \tau_a^{\tilde{Y}})\mathbb{E}_{Y_0}\left[\exp(-\kappa(\tilde{Y}_T - \log b))|\tilde{Y}_T \geq \log b\right]\left(\frac{b}{\sigma_0}\right)^{-\kappa}, \tag{23}$$

where $T = \min\{\tau_a^{\tilde{Y}}, \tau_b^{\tilde{Y}}\}$.

The random walk $\tilde{Y}_t$ has positive drift because $\mathbb{E}[\tilde{\xi}_t] = \mathbb{E}[\xi_t \exp(\kappa\xi_t)] > 0$. It follows that $\mathbb{P}(\tau_b^{\tilde{Y}} < \infty) = 1$ and $\mathbb{P}(\tau_a^{\tilde{Y}} = \infty) > 0$. Then

$$\mathbb{P}_{Y_0}(\tau_b^{\tilde{Y}} < \tau_a^{\tilde{Y}}) \geq \mathbb{P}_{Y_0}(\tau_b^{\tilde{Y}} < \tau_a^{\tilde{Y}}, \tau_a^{\tilde{Y}} = \infty) = \mathbb{P}(\tau_a^{\tilde{Y}} = \infty) \equiv q,$$

independent of b. From Remark (i) in Siegmund [35],

$$\lim_{b\to\infty} \mathbb{E}_{Y_0}\left[\exp(-\kappa(\tilde{Y}_T - \log b))|\tilde{Y}_T \geq \log b\right] \equiv c' > 0.$$

The lower bound in (8) now follows from (23) for all sufficiently large b, for any $c_\ell < \sigma_0^\kappa q c'$.

Upper Bound

Under the stability condition (A1), σ_t has a stationary distribution, which we represent through a probability measure π on $\mathbb{R}$. We begin by recording a simple property of π.

Lemma 2 *Under conditions (A1)–(A2), the support of* π *is* $[\sqrt{\omega/(1-\beta)}, \infty)$.

Proof: It suffices to show that the stationary distribution π_2 of σ_t^2 has support $[\omega/(1-\beta), \infty)$. The mapping $(x, y) \mapsto \omega + x(\alpha y + \beta)$ is monotone and continuous on $[0, \infty) \times [0, \infty)$. So, for any x in the support of π_2 and any y in the support of the distribution of Z_t^2, $\omega + x(\alpha y + \beta)$ is also in the support of π_2. Because the support of Z_t^2 is $[0, \infty)$, it follows that if x is in the support of π_2, then so is every $x' > x$. In other words, the support of π_2 is a set of the form $[x_0, \infty)$. It now follows from Proposition 2.1(b) of Meister and Kreiss [27] that x_0 is the fixed point of the mapping $x \mapsto \omega + x\beta$, which is $\omega/(1-\beta)$. $\square$

For any set $A \subseteq [\sqrt{\omega/(1-\beta)}, \infty)$ with $\pi(A) > 0$, write π_A for the restriction of π to A. Define

$$\tau_A = \inf\{t \geq 1 : \sigma_t \in A\}.$$

For any $B \subseteq \mathbb{R}$, we have (see Theorem 10.0.1 of Meyn and Tweedie [28])

$$\pi(B) = \frac{\mathbb{E}_{\pi_A}\left[\sum_{t=1}^{\tau_A} \mathbf{1}\{\sigma_t \in B\}\right]}{\mathbb{E}_{\pi_A}[\tau_A]}, \tag{24}$$

where $\mathbb{E}_{\pi_A}$ indicates expectation when σ_0 is drawn from the stationary distribution restricted to A. In particular, if we take $B = (b, \infty)$ and $A = [\sqrt{\omega/(1-\beta)}, a)$, then

$$\pi((b, \infty)) = \frac{\mathbb{E}_{\pi_A}\left[\sum_{t=1}^{\tau_A} \mathbf{1}\{\sigma_t > b\}\right]}{\mathbb{E}_{\pi_A}[\tau_A]} \geq \frac{\mathbb{P}_{\pi_A}(\tau_b < \tau_a)}{\mathbb{E}_{\pi_A}[\tau_A]},$$

so

$$\mathbb{P}_{\pi_A}(\tau_b < \tau_a) \leq c\pi((b,\infty))) \tag{25}$$

for a constant c not depending on b. We also have, using the Markov property of σ_t,

$$\begin{aligned}\mathbb{P}_{\pi_A}(\tau_b < \tau_a) &\geq \mathbb{P}_{\pi_A}(\sigma_1 \in [\sigma_0, \sigma_0 + \epsilon]) \cdot \inf_{x \in [\sigma_0, \sigma_0+\epsilon]} \mathbb{P}_x(\tau_b < \tau_a) \\ &\geq \mathbb{P}_{\pi_A}(\sigma_1 \in [\sigma_0, \sigma_0 + \epsilon]) \cdot \mathbb{P}_{\sigma_0}(\tau_b < \tau_a), \end{aligned} \tag{26}$$

and $\mathbb{P}_{\pi_A}(\sigma_1 \in [\sigma_0, \sigma_0 + \epsilon]) > 0$ because the support of Z_t^2 is all of $\mathbb{R}_+$. In (26), we have used the fact that $\mathbb{P}_x(\tau_b < \tau_a)$ is increasing in x, as σ_{t+1} is increasing in σ_t. (In other words, σ_t is a stochastically monotone Markov chain.) Combining (25) and (26), we get

$$\mathbb{P}_{\sigma_0}(\tau_b < \tau_a) \leq \frac{c}{\mathbb{P}_{\pi_A}(\sigma_1 \in [\sigma_0, \sigma_0 + \epsilon])}\pi((b,\infty)).$$

But (13) yields $\pi((b,\infty)) = \mathbb{P}(\sigma_\infty > b) \sim c_0 b^{-\kappa}$, and this establishes the upper bound.

D Analysis of Margin Breaches

Proof of Lemma 1

Proof: If we set $A_t = (\alpha Z_t^2 + \beta)$, then we can write the GARCH recursion (3) as $\sigma_t^2 = \omega + A_t\sigma_{t-1}^2$, which yields

$$\sigma_t^2 = \omega + A_t\omega + A_tA_{t-1}\omega + \cdots + A_tA_{t-1}\cdots A_1\sigma_0^2.$$

Because the A_t are i.i.d., σ_t^2 has, for each t, the same distribution as

$$\tilde{\sigma}_t^2 = \omega + A_1\omega + A_1A_2\omega + \cdots + A_1A_2\cdots A_t\sigma_0^2. \tag{27}$$

Now we have

$$\tilde{\sigma}_{t+1}^2 - \tilde{\sigma}_t^2 = A_1A_2\cdots A_t(\omega + (A_{t+1}-1)\sigma_0^2) \geq 0,$$

because $A_{t+1} \geq \beta$ and $0 < \sigma_0^2 \leq \omega/(1-\beta)$. For any $x \geq 0$,

$$\mathbb{P}(\sigma_t^2 > x) = \mathbb{P}(\tilde{\sigma}_t^2 > x) = \mathbb{P}(\max_{0\leq n\leq t} \tilde{\sigma}_n^2 > x), \tag{28}$$

are all increasing in t. □

D.1 Proof of Theorem 1

The (standard) fact noted in (28) that σ_t^2 has the same marginal distribution as the running maximum of (27) allows us to apply results for processes like (27) with $\sigma_0^2 = \omega$ (in particular, Buraczewski et al. [7] and Nyrhinen [33]) to analyze tail probabilities of σ_t^2, $t \geq 0$.

Proof: (i). For any $y > 0$ and $a > 0$, set

$$n_{u,y} = \left\lfloor \frac{\log u}{\psi'(\kappa/2)} + ay\sqrt{\log u} \right\rfloor.$$

Because the process σ_t^2 is nonnegative, we can apply Theorem 2.5 of Buraczewski et al. [7] to conclude that a may be chosen so that

$$u^{-\kappa/2}\mathbb{P}(\sigma^2_{n_{u,y}} > u) \to c_0\Phi(y),$$

with c_0 as in (13) and Φ the standard normal distribution function. See, in particular, equation (2.44) of Buraczewski et al. [7].

Set $n_u = \lfloor \log u/\psi'(s) \rfloor$, with s as in the definition of t_p in (11). The condition $t_p > T_p$ implies that $\psi'(s) < \psi'(\kappa/2)$. Then for all sufficiently large u,

$$\frac{\psi'(\kappa/2)}{\psi'(s)} > 1 + \frac{\psi'(\kappa/2)ay}{\sqrt{\log u}},$$

in which case $n_{u,y} \le n_u$. Lemma 1 implies that

$$u^{-\kappa/2}\mathbb{P}(\sigma^2_{n_{u,y}} > u) \le u^{-\kappa/2}\mathbb{P}(\sigma^2_{n_u} > u) \le u^{-\kappa/2}\mathbb{P}(\sigma^2_{\infty} > u).$$

As $u \to \infty$, the leftmost expression converges to $c_0\Phi(y)$ and, by (13), the rightmost expression converges to c_0. Because we can take y arbitrarily large, we conclude that the expression in the middle converges to c_0, which implies $\mathbb{P}(\sigma_{n_u} > u) \sim c_0 u^{-\kappa}$.

If we take $u = u_p$ then, in particular, we have $\mathbb{P}(\sigma_{n_{u_p}} > u_p) \sim c_0 u_p^{-\kappa}$, as $p \to 0$. We claim that this limit holds with n_{u_p} replaced by t_p. To see why, notice from (11) that $t_p = n_{\hat{u}_p} \equiv \lfloor \log \hat{u}_p/\psi'(s) \rfloor$, with $\hat{u}_p$ as in (16). From (16) we know that $u_p/\hat{u}_p \to 1$, so $\log u_p - \log \hat{u}_p \to 0$. Thus, for all sufficiently small p,

$$\frac{\log \hat{u}_p}{\log u_p}\frac{\psi'(\kappa/2)}{\psi'(s)} > 1 + \frac{\psi'(\kappa/2)ay}{\sqrt{\log u_p}}$$

which implies $n_{u_p,y} \le t_p$; so, the argument used above for n_u applies as well to t_p, and we have

$$\mathbb{P}(\sigma_{t_p} > u_p) \sim c_0 u_p^{-\kappa}.$$

To consider margin breaches, we need to extend this limit from σ_{t_p} to X_{t_p}, which has the distribution of $\sigma_{t_p}Z$, with Z independent of σ_{t_p}. Conditioning on Z and then integrating over Z yields

$$\mathbb{P}(|X_\infty| > u) = \mathbb{E}[\mathbb{P}(\sigma_\infty|Z| > u|Z)],$$

and, for $Z \neq 0$, (13) gives

$$\mathbb{P}(\sigma_\infty|Z| > u|Z) \sim (c_0|Z|^\kappa)u^{-\kappa}.$$

Thus, (14) is an interchange of limit and expectation:

$$\lim_{u\to\infty} u^\kappa \mathbb{E}[\mathbb{P}(\sigma_\infty|Z| > u|Z)] = \mathbb{E}[\lim_{u\to\infty} u^\kappa \mathbb{P}(\sigma_\infty|Z| > u|Z)] = c_0\mathbb{E}[|Z|^\kappa]. \tag{29}$$

The stochastic monotonicity of σ_t in Lemma 1 implies that for all Z,

$$u_p^\kappa \mathbb{P}(\sigma_{t_p}|Z| > u_p|Z) \le u_p^\kappa \mathbb{P}(\sigma_\infty|Z| > u_p|Z). \tag{30}$$

As (29) permits us to interchange the limit in u_p with the expectation over Z for the upper-bounding sequence on the right side of (30), the Lebesgue convergence theorem allows us to do so for the sequence on the left side of (30) to conclude that

$$\mathbb{P}(|X_{t_p}| > u_p) \sim \mathbb{E}[|Z|^\kappa]\mathbb{P}(\sigma_{t_p} > u_p) \sim c_0\mathbb{E}[|Z|^\kappa]u_p^{-\kappa}.$$

By the symmetry of Z,

$$\mathbb{P}(X_{t_p} > u_p) = \frac{1}{2}\mathbb{P}(|X_{t_p}| > u_p) \sim \frac{c_0}{2}\mathbb{E}[|Z|^\kappa]u_p^{-\kappa} = c_1 u_p^{-\kappa}.$$

Applying (16), we conclude that $\mathbb{P}(X_{t_p} > u_p) \sim p$.

For part (ii), set

$$n_{p,\pm\epsilon} = \left\lfloor \frac{\log u_p}{\psi'(\kappa/2)} \pm a\epsilon\sqrt{\log u_p} \right\rfloor.$$

Because $\log u_p - \log \hat{u}_p \to 0$,

$$\frac{\log u_p}{\psi'(\kappa/2)} - a\epsilon\sqrt{\log u_p} \le \frac{\log \hat{u}_p}{\psi'(\kappa/2)} \le \frac{\log u_p}{\psi'(\kappa/2)} + a\epsilon\sqrt{\log u_p}$$

for all sufficiently small p, for any $\epsilon > 0$. In this case, $n_{p,-\epsilon} \le T_p \le n_{p,\epsilon}$ and Lemma 1 implies

$$u_p^{-\kappa/2}\mathbb{P}(\sigma^2_{n_{p,-\epsilon}} > u_p) \le u^{-\kappa/2}\mathbb{P}(\sigma^2_{T_p} > u_p) \le u_p^{-\kappa/2}\mathbb{P}(\sigma^2_{n_{p,\epsilon}} > u_p).$$

Letting $p \to 0$ and applying Theorem 2.5 of Buraczewski et al. [7], we find that

$$c_0\Phi(-\epsilon) \le \liminf_{p\to 0} u_p^{-\kappa/2}\mathbb{P}(\sigma^2_{T_p} > u_p) \le \limsup_{p\to 0} u_p^{-\kappa/2}\mathbb{P}(\sigma^2_{T_p} > u_p) \le c_0\Phi(\epsilon).$$

As $\epsilon > 0$ can be taken arbitrarily close to zero and $\Phi(0) = 1/2$, we conclude that

$$u_p^{-\kappa}\mathbb{P}(\sigma_{T_p} > u_p) \to c_0/2.$$

The extension from σ_{T_p} to X_{T_p} proceeds exactly as in part (i).

For case (iii), set $n_u = \lfloor \log u/\psi'(s) \rfloor$. Theorem 2 of Nyrhinen [33] yields

$$\lim_{u\to\infty} (\log u)^{-1} \log \mathbb{P}(\sigma^2_{n_u} > u) = -r(s), \tag{31}$$

where $r(s) > \kappa/2$ depends on the value of s in n_u. Nyrhinen's [33] result applies to (27), and (31) then follows from (28). Our $r(s)$ corresponds to his $R(1/\psi'(s))$.

Next, we extend (31) to X_{n_u}. The condition that $\mathbb{E}[\exp(\gamma|Z|)] < \infty$ implies that

$$\mathbb{P}(|Z| > x) = o(e^{-\gamma x}), \quad \text{as } x \to \infty.$$

Let $m_u = a \log u$, for any $a > 2r(s)/\gamma$; then $\mathbb{P}(|Z| > m_u) = o(u^{-a\gamma}) = o(u^{-2r(s)})$. Now

$$\begin{aligned}
\mathbb{P}(\sigma_{n_u}|Z| > u) &\leq \mathbb{P}(\sigma_{n_u}|Z| > u, |Z| \leq m_u) + P(|Z| > m_u) \\
&\leq \mathbb{P}(\sigma_{n_u} > u/m_u) + o(u^{-2r(s)}). \qquad (32)
\end{aligned}$$

To analyze the first term on the right, set $n(s) = \lfloor \log(u/m_u)/\psi'(s) \rfloor$. For any sufficiently small $\epsilon > 0$ and all sufficiently large u, $n(s+\epsilon) \leq n_u \leq n(s-\epsilon)$, so Lemma 1 implies

$$\mathbb{P}(\sigma_{n(s+\epsilon)} > u/m_u) \leq \mathbb{P}(\sigma_{n_u} > u/m_u) \leq \mathbb{P}(\sigma_{n(s-\epsilon)} > u/m_u).$$

Then (31) implies

$$\begin{aligned}
-2r(s+\epsilon) &\leq \liminf_{u\to\infty} (\log u/m_u)^{-1} \log \mathbb{P}(\sigma_{n_u} > u/m_u) \\
&\leq \limsup_{u\to\infty} (\log u/m_u)^{-1} \log \mathbb{P}(\sigma_{n_u} > u/m_u) \leq -2r(s-\epsilon).
\end{aligned}$$

Because r is continuous (see Nyrhinen [33], p.268) and $\epsilon > 0$ can be made arbitrarily small,

$$\lim_{u\to\infty} (\log u/m_u)^{-1} \log \mathbb{P}(\sigma_{n_u} > u/m_u) = -2r(s).$$

But $\log(u/m_u)/\log u \to 1$, so we also have

$$\lim_{u\to\infty} (\log u)^{-1} \log \mathbb{P}(\sigma_{n_u} > u/m_u) = -2r(s).$$

Applying this limit in (32), where the last term is negligibly small, we get

$$\liminf_{u\to\infty} -(\log u)^{-1} \log \mathbb{P}(\sigma_{n_u}|Z| > u) \geq 2r(s).$$

On the other hand,

$$\begin{aligned}
\mathbb{P}(\sigma_{n_u}|Z| > u) &\geq \mathbb{P}(\sigma_{n_u}|Z| > u, |Z| > 1) \\
&\geq \mathbb{P}(\sigma_{n_u} > u, |Z| > 1) \\
&= \mathbb{P}(\sigma_{n_u} > u)\mathbb{P}(|Z| > 1).
\end{aligned}$$

It follows that

$$\limsup_{u\to\infty} -(\log u)^{-1} \log \mathbb{P}(\sigma_{n_u}|Z| > u) \leq \lim_{u\to\infty} -(\log u)^{-1} \log \mathbb{P}(\sigma_{n_u} > u) = 2r(s).$$

We have shown that

$$\lim_{u\to\infty} -(\log u)^{-1} \log \mathbb{P}(|X_{n_u}| > u) = 2r(s).$$

Replacing $|X_{n_u}|$ with X_{n_u} cuts the probability in half, but the factor of 1/2 disappears in the limit after we take logs. Replacing u with u_p yields

$$\lim_{p\to 0} -(\log u_p)^{-1} \log \mathbb{P}(X_{n_{u_p}} > u) = 2r(s).$$

Moreover, from (16) we have $\kappa \log u_p \sim -\log p$. Thus,

$$\lim_{p\to 0} (\log p)^{-1} \mathbb{P}(X_{n_{u_p}} > u_p) = 2r(s)/\kappa > 1.$$

In other words, $\mathbb{P}(X_{n_{u_p}} > u_p) = p^{\delta+o(1)}$, with $\delta = r(s)/(\kappa/2)$. Following the same argument used in part (i), we can replace n_{u_p} with t_p. $\square$

E Proof of Proposition 9

Proof: From Theorem 4.1 of Mikosch and Stărică [29], we have that for any $x > 0$,

$$\mathbb{P}(\max_{1\le i\le n} X_i \le (c_1 n)^{1/\kappa} x) \to e^{-\theta x^{-\kappa}}, \quad \text{as } n \to \infty,$$

with c_1 as in (15). If we set

$$n_p = \left\lfloor \left(\frac{u_p^\kappa}{c_1}\right) x^{-\kappa} \right\rfloor,$$

then $(c_1 n_p)^{1/\kappa} x \le u_p$. Furthermore, $c_1 n_p / u_p^\kappa \to x^{-\kappa}$ as $p \to 0$; so, for any $\epsilon > 0$ and all sufficiently small p, $(c_1 n_p)^{1/\kappa}(x+\epsilon) \ge u_p$. These inequalities imply

$$\mathbb{P}(\max_{1\le i\le n_p} X_i \le (c_1 n_p)^{1/\kappa} x) \le \mathbb{P}(\max_{1\le i\le n_p} X_i \le u_p) \le \mathbb{P}(\max_{1\le i\le n_p} X_i \le (c_1 n_p)^{1/\kappa}(x+\epsilon)).$$

Letting $p \to 0$, we get

$$e^{-\theta x^{-\kappa}} \le \liminf_{p\to 0} \mathbb{P}(\max_{1\le i\le n_p} X_i \le u_p) \le \limsup_{p\to 0} \mathbb{P}(\max_{1\le i\le n_p} X_i \le u_p) \le e^{-\theta (x+\epsilon)^{-\kappa}}.$$

Because $\epsilon > 0$ is arbitrary, we conclude that these inequalities hold as equalities at $\epsilon = 0$. In other words, we have shown that

$$\mathbb{P}(\max_{1\le i\le n_p} X_i \le u_p) \to e^{-\theta x^{-\kappa}}.$$

For any n, the event $\{\tau_p > n\}$ coincides with the event $\{\max_{1\le i\le n} X_i \le u_p\}$. Set $t = x^{-\kappa}$ and write $n_p = n_p(t)$ for emphasis. Then

$$\mathbb{P}(\tau_p > n_p(t)) = \mathbb{P}(\max_{1\le i\le n_p(t)} X_i \le u_p) \to e^{-\theta t}, \quad \text{as } p \to 0.$$

For any $t > 0$ and any $\epsilon > 0$ for which $n_p(t+\epsilon) \geq n_p(t) + 1$,

$$\frac{c_1}{u_p^\kappa} n_p(t) \leq t \leq \frac{c_1}{u_p^\kappa} n_p(t+\epsilon). \tag{33}$$

The condition $n_p(t+\epsilon) \geq n_p(t)+1$ is satisfied if $\epsilon u_p^\kappa / c_1 \geq 1$, and this holds for all sufficiently small p. Thus, for any $t > 0$ and $\epsilon > 0$, (33) holds for all sufficiently small p. Then

$$\mathbb{P}(c_1 u_p^{-\kappa} \tau_p > c_1 u_p^{-\kappa} n_p(t+\epsilon)) \leq \mathbb{P}(c_1 u_p^{-\kappa} \tau_p > t) \leq \mathbb{P}(c_1 u_p^{-\kappa} \tau_p > c_1 u_p^{-\kappa} n_p(t)),$$

so

$$\mathbb{P}(\tau_p > n_p(t+\epsilon)) \leq \mathbb{P}(c_1 u_p^{-\kappa} \tau_p > t) \leq \mathbb{P}(\tau_p > n_p(t)).$$

Letting $p \to 0$,

$$e^{-\theta(t+\epsilon)} \leq \liminf_{p \to 0} \mathbb{P}(c_1 u_p^{-\kappa} \tau_p > t) \leq \limsup_{p \to 0} \mathbb{P}(c_1 u_p^{-\kappa} \tau_p > t) \leq e^{-\theta t}.$$

Because $\epsilon > 0$ is arbitrary, we conclude that these inequalities hold as equalities at $\epsilon = 0$. In other words, we have shown that

$$c_1 u_p^{-\kappa} \tau_p \Rightarrow \xi / \theta.$$

But from (16) we know that $c_1 u_p^{-\kappa}/p \to 1$, so the same limit in distribution holds for $p\tau_p$. $\square$

References

[1] Adrian, T., and Shin, H. S. (2010). Liquidity and leverage, *Journal of Financial Intermediation* 19:418-437.

[2] Alexander, C. (2001) *Market Models: A Guide to Financial Data Analysis*, Wiley.

[3] Basrak, B., Davis, R. A., and Mikosch, T. (2002). A characterization of multivariate regular variation, *Annals of Applied Probability* 12(3), 908-920.

[4] Basrak, B., Davis, R.A., and Mikosch, T. (2002) Regular variation of GARCH processes, *Stochastic Processes and their Applications* 99, 95–115.

[5] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity *Journal of Econometrics* 31(3), 307–327.

[6] Brandt, A. (1986) The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients, *Advances in Applied Probability*, 211–220.

[7] Buraczewski, D., Collamore, J.F., Damek, E., and Zienkiewicz, J. (2016) Large deviation estimates for exceedance times of perpetuity sequences and their dual processes, *Annals of Probability*, to appear.

[8] Collamore, J.F. (2009) Random recurrence equations and ruin in a Markov-dependent stochastic economic environment. *Annals of Applied Probability* 19(4), 1404–1458.

[9] Committee on the Global Financial System. (2010) *The Role of Margin Requirements and Haircuts in Procyclicality.* CGFS Papers 36. Available at http://www.bis.org/publ/cgfs36.pdf.

[10] Committee on Payment and Settlement Systems (2012) *Principles for Financial Market Infrastructures.* Available at http://www.bis.org/cpmi/publ/d101a.pdf.

[11] De Saporta, B. and Yao, J.F. (2005) Tail of a linear diffusion with Markov switching, *Annals of Applied Probability* 15(1B), 992–1018.

[12] Drees, H. (2003) Extreme quantile estimation for dependent data, with applications to finance, *Bernoulli* 9(1), 617–657.

[13] Drost, F.C., and Nijman, T.E. (1993) Temporal aggregation of GARCH processes, *Econometrica* 61(4), 909–927.

[14] Dutta, S., and Biswas, S. (2015) Extreme quantile estimation based on financial time series, *Communications in Statistics – Simulation and Computation*, accepted.

[15] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events: For Insurance and Finance*, Springer, Berlin.

[16] Engle, R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, 987-1007.

[17] European Commission (2013) Commission Delegated Regulation No 153/2013 of 19 December 2012 supplementing Regulation (EU) No 648/2012 of the European Parliament and of the Council with Regard to Regulatory Technical Standards on Requirements for Central Counterparties, *Official Journal of the European Union* 52, 41–74, February 23, 2013.

[18] European Commission (2012) Regulation (EU) No. 648/2012 of the European Parliament and of the Council, of 4 July 2012, on OTC derivatives, central counterparties and trade repositories, *Official Journal of the European Union* 55, 1–59.

[19] Financial Stability Board (2015) *Transforming Shadow Banking into Resilient Market-based Finance: Regulatory framework for haircuts on non-centrally cleared securities financing transactions.*

[20] Glosten L.R., Jagannathan R., and Runkle D.E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48(5), 1779–1801.

[21] Gorton, G. and Metrick, A. (2012) Securitized banking and the run on repo, *Journal of Financial Economics* 104(3), 425–451.

[22] Guyon, X., Iovleff, S. and Yao, J.F. (2004) Linear diffusion with stationary switching regime, *ESAIM: Probability and Statistics* 8, 25–35.

[23] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices, *Acta Mathematica* 131, 207–248.

[24] Kingman, J.F.C. (1961) A convexity property of positive matrices, *Quarterly Journal of Mathematics* 12(1), 283–284.

[25] Margin and Capital Requirements for Covered Swap Entities (2015) *Federal Register* 80 (November 30, 2015), 74840–74914.

[26] McNeil, A.J., Frey, R., and Embrechts, P. (2005) *Quantitative Risk Management*, Princeton University Press.

[27] Meister, A., and Kreiss, J.-P. (2016) Statistical inference for nonparametric GARCH models, *Stochastic Processes and their Applications*, in press.

[28] Meyn, S. P., and Tweedie, R. L. (1993) *Markov Chains and Stochastic Stability*, Springer-Verlag, London.

[29] Mikosch, T., and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process, *Annals of Statistics* 28(5), 1427–1451.

[30] Murphy, D., Vasios, M., and Vause, N. (2016) A comparative analysis of tools to limit the procyclicality of initial margin requirements, Staff Working Paper 597, Bank of England.

[31] Nelson, D.B. (1990) Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory* 6(3), 318–334.

[32] Nelson, D.B. (1990) ARCH models as diffusion approximations, *Journal of Econometrics* 45(1), 7–38.

[33] Nyrhinen, H. (2001) Finite and infinite time ruin probabilities in a stochastic economic environment, *Stochastic Processes and their Applications* 92(2), 265–285.

[34] Repullo, R. and Suarez, J. (2013) The procyclical effects of bank capital regulation, *Review of Financial Studies* 26(2), 452–490.

[35] Siegmund, D. (1975) Error probabilities and average sample number of the sequential probability ratio test, *Journal of the Royal Statistical Society. Series B* 37(3), 394–401.