BANK OF CANADA
BANQUE DU CANADA

Working Paper/Document de travail
2014-29

Optimal Margining and Margin Relief in Centrally Cleared Derivatives Markets

by Radoslav S. Raykov

www.bank-banque-canada.ca

Bank of Canada Working Paper 2014-29

June 2014

Optimal Margining and Margin Relief in Centrally Cleared Derivatives Markets

by

Radoslav S. Raykov

Financial Stability Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
RRaykov@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

The author would like to thank (in alphabetical order): James Chapman, Corey Garriott, Thorsten Koeppl, Hector Perez Saiz, Joao Santos and Joshua Slive for useful research suggestions, and Carol Ann Northcott, Darcey McVanel, Stacey Anderson and members of the Bank of Canada’s oversight divisions for discussing CCP regulation.

Abstract

A major policy challenge posed by derivatives clearinghouses is that their collateral requirements can rise sharply in times of stress, reducing market liquidity and further exacerbating downturns. Smoothing sharp changes in collateral requirements – an approach known as through-the-cycle margining – however, has its own disadvantages, one of which is increased risk sharing among clearinghouse members when financial risk is high. This can give rise to undesirable side effects, including distorted incentives, which can reverse the conventional knowledge about collateral policy. In contrast to the existing literature, I show that through-the-cycle margining can increase as well as *reduce* trading in volatile markets. Due to increased risk sharing, clearinghouse members may prefer to overcollateralize transactions, leading to lower than socially optimal trading. This creates a challenge for policy-makers, since it may be challenging to push for lower collateral standards than deemed proper by the industry. For such cases, I propose an alternative policy tool – increasing default penalties – to align private and social incentives.

JEL classification: G19, G21
Bank classification: Economic models; Payment clearing and settlement systems

Résumé

Un grand défi en matière de politiques que posent les chambres de compensation des produits dérivés est le fait que leurs exigences en matière de garanties peuvent monter en flèche en périodes de tensions, réduisant ainsi la liquidité du marché et exacerbant les ralentissements économiques. Le lissage des brusques variations de ces exigences – que l'on appelle mode de calcul des marges en fonction du cycle intégral – présente cependant ses propres désavantages, dont le partage accru des risques entre les membres des chambres de compensation lorsque le risque financier est élevé. Cela peut avoir des effets secondaires indésirables, notamment la distorsion des incitations, qui peuvent contredire les connaissances généralement admises au sujet des politiques de garantie. Contrairement aux études précédentes, celle-ci montre que le calcul des marges en fonction du cycle intégral peut entraîner une augmentation aussi bien qu'une *réduction* du volume des opérations sur des marchés volatils. En raison du partage accru des risques, les membres des chambres de compensation préfèrent parfois procéder à un surnantissement, qui donne lieu à un volume d'opérations inférieur au niveau socialement optimal. Cela complique les choses pour les décideurs publics, étant donné qu'il pourrait être difficile de tenter d'imposer des normes de garanties inférieures à celles que le secteur juge adéquates. Dans de tels cas, l'auteur propose une autre mesure de politique – l'augmentation des pénalités en cas de défaut – afin d'harmoniser les incitations d'ordre privé et social.

Classification JEL : G19, G21
Classification de la Banque : Modèles économiques; Systèmes de compensation et de règlement des paiements

Non-Technical Summary

A major policy challenge posed by derivatives clearinghouses is that their collateral requirements can rise sharply in times of stress, reducing market liquidity and exacerbating economic downturns. For this reason, international regulatory bodies have asked clearinghouses to reduce procyclical collateral movements. However, smoothing spikes in collateral requirements – an approach known as through-the-cycle (TTC) margining – inherently results in clearinghouses mutualizing more risk in times of stress, which is likely to affect members' behavior and distort their incentives. This paper studies how to design a less procyclical collateral policy while minimizing the undesirable effects associated with increased risk sharing.

The paper contains several sets of results. In contrast to the existing literature, I show that TTC margining in stressed markets can increase as well as reduce trading activity; the outcome depends on financial market volatility and on clearinghouse members' risk aversion. When banks have low risk aversion, they are more concerned with the collateral cost of trading than with their mutualized risk exposure, and are willing to take advantage of margin relief even though it implies an increase in mutualized risk. To compensate for their higher mutual exposures through the clearing fund, banks instead reduce their trading positions. Thus, through-the-cycle margining can reduce trading activity during stressed periods.

Conversely, when banks are sufficiently risk averse, they may find it more beneficial *not* to take advantage of the margin relief associated with the TTC approach, even when this could benefit the market as a whole. In this case, member banks' influence on the clearinghouse's risk management could slow down the adoption of TTC margining. Banks with high risk aversion will focus predominantly on their mutualized risk exposure, and will therefore tend to see collateral more as a form of protection than as a trading cost. This can create a challenge for policy-makers, since it may be politically challenging for regulators to push for lower collateral standards than desired by the industry. I find that this dilemma is more likely to occur when the social benefit from trade exceeds the expected social costs of default, and the banks' risk aversion is sufficiently high.

One way of motivating banks to agree to less procyclical margins is to increase the amount of "skin in the game" they have in the CCP: that is, increase their stakes to lose when they bring too much risk to the clearinghouse. This provides a way to control bad incentives and reduce the need for protection in the form of collateral. Increasing the "skin in the game" can be done by raising the clearinghouse's default penalties (such as the severity of fines and suspensions) or by applying deeper haircuts to the pool of eligible collateral, either of which will increase the opportunity cost of default and motivates members to avoid risky activities outside the CCP. I show this solution works best when the default probability of member banks is sufficiently elastic to changes in the default penalty.

1 Introductory Remarks

Financial market reforms after the 2008 crisis have resulted in a large number of changes in the financial and regulatory landscape. One of the most significant changes is the G-20 commitment to clear standardized over-the-counter derivatives through clearinghouses known as central counterparties. Central counterparties (CCPs) provide clearing services, guarantee member trades by pooling participant credit risks, and impose market discipline by collecting collateral, called margin.

Because of the large growth of the bilateral OTC market in the past decade, channelling standardized derivatives through central counterparties is likely to significantly increase centrally cleared volumes.[1] This is likely to also increase the systemic importance of CCPs in mitigating (or propagating) financial shocks and impart on them the role of systemic risk managers in addition to that of clearing service providers (Tucker, 2013).

A major policy challenge posed by derivatives clearinghouses is that their collateral requirements can rise sharply in times of stress, reducing market liquidity and potentially exacerbating economic downturns. Brunermeier and Pedersen (2009) have shown that procyclical margin movements can lead to "margin spirals" – situations in which financial market volatility triggers additional margin calls, which in turn worsen market liquidity and precipitate further defaults, creating a vicious circle. For example, an $8.6 billion collateral call on Lehman Brothers by J.P. Morgan accelerated Lehman's demise, and collateral calls on Merrill Lynch likely contributed to its sale to Bank of America (Carney 2008). Collateral calls are also known to have played a role in the defaults of Metallgesellschaft, Long-Term Capital Management (LTCM), and more recently Bear Stearns and Lehman Brothers (Gibson and Murawski, 2013). The procyclical nature of margin requirements remains a significant concern for policy-makers and market participants, since it can cause market liquidity to dry up when it is already scarce, and therefore push fundamentally solvent, but illiquid institutions over the brink.

In response to the systemic risk implications of procyclical collateral policy, international regulatory bodies such as the CPSS-IOSCO[2] have formulated explicit directives asking clearinghouses to look for ways to reduce the procyclicality of their collateral requirements.[3] However, curbing margin growth in volatile periods implies an increase in

---

[1] For example, less than half of the trades in the $250 trillion interest rate swap market, and only one tenth of the $25 trillion credit default swap market were centrally cleared by the end of 2011 (Tucker, 2013)

[2] Committee on Payment and Settlement Systems and International Organization of Securities Commissions.

[3] For example, see the CPSS-IOSCO (2012) *Principles for Financial Market Infrastructures*, Principle 6,

the mutualized risk held at the clearinghouse relative to the pre-crisis period. A CCP has two main resource buffers to cover losses from member default: the pool of collateral ("defaulter-pays resources"), and a mutualized clearing fund ("survivors-pay resources"). Risks not covered by the margin requirement therefore end up being redistributed across members via the clearing fund.[4] Smoothing sharp spikes in collateral requirements – an approach known as through-the-cycle (TTC) margining – therefore inherently transfers risk from non-mutualized to mutualized resources in times of stress. This change in exposures and risk-sharing is likely to also affect clearing members' behavior, for example, by altering their trading activity, as well as distort incentives precisely when the CCP's pre-funded resources are at their lowest and financial risk is at its highest. Thus, instead of stabilizing derivatives markets, through-the-cycle margining could play a destabilizing role if implemented without discretion. This paper studies how to design a less procyclical collateral policy while minimizing the undesirable effects associated with increased risk sharing.

The paper contains several sets of results. First, in contrast to the existing literature, I show that TTC margining in stressed markets can increase as well as *reduce* trading activity, depending on financial market volatility and clearinghouse members' risk aversion. The reason for this is that smoothing out margin spikes, while making it less costly to trade in volatile periods, also increases member banks' exposure to each other via the CCP's clearing fund. Regulatory bodies are mainly concerned with procyclical margin movements because they believe that spikes in collateral requirements dry up market liquidity and inhibit trading in times of stress. But I show that, in volatile periods, TTC margining can *depress* trading, provided that volatility is sufficiently high and that clearinghouse members are not too risk averse. At low risk-aversion levels, banks are more concerned about the collateral cost of trading than about their mutual exposures within the clearinghouse, and are willing to take advantage of margin relief even though it implies an increase in mutualized risk. To compensate for their higher mutual exposures through the clearing fund, banks instead reduce their trading positions. Thus, TTC margining can ultimately lead to a *reduction* in trading activity during stressed periods.

The opposite side of the same phenomenon is that when banks are sufficiently risk averse, they may find it more beneficial *not* to take advantage of the margin relief associated with the TTC approach, even when this could benefit the market (i.e. when margin relief

Key consideration 6.3. See also Committee on the Global Financial System (2010).

[4] In practice, CCPs typically have a modest amount of own capital pledged before they can resort to the clearing fund, but such resources are not required by international risk-management standards and are inconsequential for this model.

is socially optimal). In this case, member banks' influence on the clearinghouse's risk management could slow down the adoption of TTC margining.[5] The reason for this is that banks with high risk aversion will focus predominantly on their mutualized risk exposures and will therefore tend to see margin more as a form of protection than as a trading cost. This can create a challenge for policy-makers, since it may be politically challenging to push for lower collateral standards than deemed appropriate by the industry, without incurring a loss in reputation or credibility. I find that this dilemma is more likely to occur when the social benefit from trade exceeds the expected social costs of default, and the banks' risk aversion is sufficiently high. If TTC margining results in an uncomfortably large risk being mutualized through the clearing fund in stressed times, then something else must be used to motivate banks to agree to less procyclical margins. One solution is to increase member banks' "skin in the game" – that is, their stakes to lose when they bring too much risk to the clearinghouse – as a way to control bad incentives and reduce the need for protection in the form of collateral. This can take different forms: from increasing the CCP's penalties on defaulting members (such as the severity of fines and suspensions) to applying deeper haircuts to the pool of eligible collateral, either of which increases the opportunity cost of default and reduces member incentives to engage in risky activities outside the CCP's control. In Section 5, I show this solution works best when the member banks' default probability is sufficiently elastic to changes in the default penalty.

The remainder of the paper is structured as follows. Section 2 presents the related literature, and section 3 the model. Section 4.1 considers the effect of margin on trading volume, section 4.2 focuses on the relationship between the bank-optimal margin and trading, and section 4.3 compares the private and social optimality of through-the-cycle margins. Section 5 discusses the case when banks are too risk averse to want to take advantage of margin relief and proposes alternate policy tools to motivate them. Section 6 concludes.

2 Related Literature

Margin evolved as a risk-mitigation mechanism jointly with the development of derivatives markets. As derivatives transactions, where each party can potentially be a creditor or a debtor, became more common, so did the need for a mechanism to mitigate default losses from non-performance (Moser, 1994). Margining reduces a derivative's default risk by increasing the minimum amount received from the counterparty, thus reducing the credit

[5] An actual example is discussed in section 4.3.

risk associated with the transaction. It is likely that margining practices arose primarily out of credit risk concerns, since it is known that early derivatives traders (e.g., farmers trying to secure crop prices by entering futures contracts) recognized the role of margin as a credit risk mitigant. It is less clear whether early market participants appreciated the role of collateral as a device to control participant incentives in addition to credit risk.

Several papers have studied margining and mutualized risk in the context of derivatives markets. Santos and Scheinkman (2001) build a stylized model with asymmetric information to study whether competition between exchanges could lead to excessively low margin standards, a concern periodically voiced by the industry, but find no evidence to support this scenario. Gibson and Murawski (2013) develop an optimal margining model in a similar spirit, but for bilateral trades. Nahai-Williamson, Ota, Vital and Wetherilt (2013) explore the optimal trade-off between clearing fund and margin requirements, and find that the use of the default fund is preferable as a cost-effective insurance against large losses when member credit quality is good. Unlike the results obtained here, however, they do not consider endogenous credit quality changes induced by moral hazard or by non-linearities in the clearing fund – margin trade-off. Haene and Sturm (2009) outline the main factors affecting the trade-off between margins and the CCP's default fund. Similar to my findings here, Haene and Sturm find that a CCP's margin requirement should be balanced against the participants' chance of default, making high margins optimal only when participants have a high chance of default or when the incentive effect of margin is sufficiently high. However, they do not consider the effect of financial market volatility and banks' risk aversion on derivatives trading, a result that is key in formulating the right margining policy for a stressed market. Consistent with their framework, however, I consider the incentive effect of margin as a default-risk mitigant.

As Koeppl (2013) writes, collateral can be used to control counterparty risk in two different ways: as a prepayment (insurance) and as a bond that provides incentives to keep counterparty risk low. The latter incentive arises because a defaulter's margin is confiscated in the event of default, which implies that default is more costly for a party with some pledged margin compared to a party with no margin pledged. This provides incentives to avoid risks that exacerbate the chance of default. The converse is true as well: in stressed markets, liquid collateral becomes scarce and has a higher opportunity cost; if the margin requirement is too low, trading parties may be tempted to tolerate higher counterparty risk, because collateral is expensive. As summarized by Koeppl (2013), "When collateral is costly, the contracting parties have an incentive to economize on it: they can opt for a trade that requires little collateral at the expense of higher counterparty risk. In particular, in transactions with bad counterparty quality, collateral requirements could be too high to make it (privately) optimal to control [...] all counterparty risk exposure." The model presented here admits such incentive effects as one possible source of endogenous changes in counterparty credit quality. However, my results are not exclusively driven by the incentive effects of margin, although they can play a role. Instead, I argue more generally that there is a non-linear (convex) relationship between expected clearing fund losses and the margin requirement, arising both from the CCP clearing fund sizing methodology and from possible incentive effects such as moral hazard. This non-linearity implies that margin and clearing fund contributions are not perfect substitutes, contrary to the assumptions of simplified models, and therefore the trade-off between the two can be optimized depending on banks' willingness to take risks and on financial market conditions. In what follows, I discuss the choice of optimal point in this trade-off and how the privately optimal point could differ from the socially optimal one.

3 Model

To study the optimal trade-off between collateralized and mutualized risk, I build a stylized partial equilibrium model of banks trading through a representative central counterparty. To model member banks' behavior in periods with different margin requirements, I explore the comparative statics of the model by varying margin in conjunction with other parameters such as financial market volatility and the banks' risk aversion. I consider two groups of risk-averse banks, who are subject to privately observed random shocks to their endowments. To hedge their endowment uncertainty, banks trade in a stylized futures contract by choosing an optimal trading position subject to the margin requirement set by the CCP. Banks can default on their positions; the clearinghouse pools the banks' default risks and redistributes default losses across the remaining survivors. The higher the CCP's margin requirement, the more protection it offers to member banks, but at the cost of increased collateral opportunity cost. Recognizing that in stressed periods, TTC margining reduces the amount of prepayment relative to the pre-crisis period and can introduce moral hazard, I study optimal margining from both a private and a social standpoint. I characterize the cases where private and social incentives diverge due to the social costs of default and positive spillovers from trading, and suggest alternative policy tools to align social and private incentives.

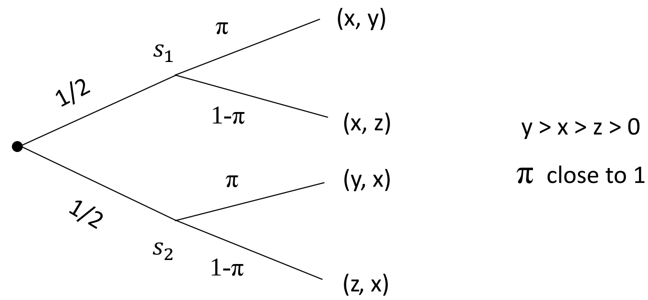

Figure 1: Stochastic endowments for Bank 1 and Bank 2

3.1 Economic Environment

I consider an economy similar to Santos and Scheinkman (2001) and Gibson and Murawski (2013), consisting of two groups of risk-averse banks who maximize the expected utility of their uncertain endowments of a consumption good, evaluated by a strictly concave, thrice-differentiable vonNeumann-Morgenstern utility function u. The model considers two periods. In the first period, banks trade assets; in the second period, endowment realization and consumption occur. At date 1, banks choose their trading positions in a derivative contract subject to a margin requirement, and at date 2, random shocks to banks' endowments are realized. In the second period, two equally likely, observable states of the world can occur: s_1 or s_2. In state s_1, the banks from group 1 receive a certain endowment of x units of a consumption good, while group 2 receives a lottery between a good endowment realization of y and a bad one of z, where $y > x > z$. In state s_2, the roles of the two groups are reversed: group 2 receives the certain endowment, while group 1 receives the uncertain one, so that either side can potentially default, depending on the endowment outcome; this allows the contract to transfer wealth between states of the world. In the uncertain endowment scenario, with high likelihood π close to 1, a good realization occurs, but there is also a small probability $1 - \pi$ of a bad realization, so that a bank could default if it receives the bad endowment realization z. The occurrence of the bad shock is private information and is assumed to be independent across banks. Since each bank (group) mirrors the other, the aggregate expected endowment is constant in this economy. Banks within each group are identical, so it suffices to consider the optimization of one representative from each group, denoted with Bank 1 and Bank 2. The stochastic endowment process is summarized in Figure 1.

Several additional assumptions are needed to fully specify the contract, since the optimal buy/sell behavior of banks depends on the relative size of the endowment outcomes x, y and z. To introduce the possibility of default, it makes sense to calibrate the endowments so that, in equilibrium, banks buy assets that transfer consumption from the uncertain-endowment state to the state with certain endowment x (otherwise default would not be probabilistic). Therefore I assume that the marginal utility from the certain endowment exceeds the expected marginal utility from the uncertain endowment:

ASSUMPTION 1: $u'(x) > \pi u'(y) + (1-\pi)u'(z)$.

This calibration assumption, which is standard in the literature,[6] permits me to normalize the derivatives contract to a generic futures contract in which the bank holding a long position is entitled to receive one unit of the consumption good in state s_1, and must deliver one unit of the consumption good in state s_2. Banks can default on the contract if the disutility of default is smaller than the disutility from getting the idiosyncratic uninsurable shock z. As in Gibson and Murawski (2013), defaulting allows the payoffs to co-vary, albeit imperfectly, with the idiosyncratic shock, and in doing so it permits some relief from the uninsurable shock z. Similar to Diamond (1984) and Dubey et al. (2005), I assume that default is costly and is associated with a disutility proportional to the defaulted amount. Specifically, if a bank must deliver ℓ dollars on a position but ends up delivering a smaller amount D, then it suffers a utility penalty equal to

$$\lambda \max\{\ell - D, 0\}$$

subtracted directly from the utility obtained in the state triggering the default. The literature interprets the parameter $\lambda > 0$ as the economy-wide bankruptcy code, or the marginal disutility from a dollar defaulted. In the context of a central counterparty, however, λ can be broadly interpreted as anything that increases the disutility to a member from not fulfilling its obligations to the CCP. For example, CCPs typically foresee specific penalties for defaulting members (such as force-reducing or closing out member positions, assessing the member's clearing fund and margin deposit, or imposing fines or suspensions). The clearinghouse typically has some discretion regarding what penalty to impose, so λ here is assumed to be under the CCP's control; it cannot be changed by the banks. As I show later, adjusting the default penalty λ can be used as an alternate policy tool to control private risk-taking incentives by clearinghouse members.

[6] Very similar assumptions are made in Santos and Scheinkman (2001) and Gibson and Murawski (2013).

Since defaulting makes economic sense only if it provides some relief from the idiosyncratic bad shock z, it is meaningful to assume that marginal utility in the bad state, $u'(z)$, is sufficiently higher than the default penalty λ in order to warrant default. Specifically, I assume that $u'(z) > \lambda + u'(x)$. It is also assumed that the default penalty is not too small and exceeds the bank's expected marginal utility from the lottery between y and z:

Assumption 2: $u'(z) - u'(x) \ > \ \lambda \ > \ \pi u'(y) + (1 - \pi)u'(z)$.

As shown in Lemma 1 in the Appendix, Assumption 2 guarantees that, on average, banks will not default when faced with a lottery, since they receive the good outcome y with a much higher probability than the bad outcome z. This assumption makes sure that the endowment outcomes are calibrated so that default is the exception and not the norm in this economic environment. It also implies that the banks receiving z will default and deliver only the posted collateral.

Finally, I assume that the endowment outcomes are calibrated so that banks have no incentives to default strategically when receiving the good shock y, and therefore a bank with position θ receiving y will meet its full obligation:

Assumption 3: $u'(y - \theta) < \lambda$.

Assumption 3 rules out strategic defaults not triggered by a bad endowment realization. In this environment, such defaults are unrealistic: they make sense only when the marginal disutility from default is so low that the utility gain from the good endowment shock y outweighs it. In practice, none of the real-world defaults discussed in the paper can be described as strategic, as they ultimately led to the demise of the affected firms. Assumption 3 is not particularly restrictive either, since the model can always be calibrated so that the assumption holds regardless of the bank's position and endowment size.

3.2 The banks' optimization problem

In what follows, each bank solves an optimization problem taking into account the CCP's margin requirement and the bank's exposure to the CCP's mutualized clearing fund. Equilibrium is attained when each member bank has chosen an optimal trading position and an optimal amount to deliver on a trade, and all applicable resource constraints are satisfied (for example, each bank always delivers at least the posted collateral).

Banks are subject to a margin requirement $\Phi \in [0,1]$ set by the CCP, indicating the fraction of the position θ to be collateralized. A bank can take either a positive (long) or a negative (short) position θ in the contract, pledging $|\theta|\Phi$ dollars of margin to the clearinghouse. In the event of default, $|\theta|\Phi$ is therefore the minimum delivered amount to the CCP. Losses from realized defaults in excess of the posted collateral accrue to the CCP, which redistributes them across the large number of remaining survivors. (This assumption is equivalent to assuming the existence of a mutualized clearing fund.) To finance the cost of defaults, the CCP pays less than one unit for each unit of the consumption good claimed. A bank that is long one unit of the contract receives $K < 1$ units of the consumption good from the CCP in state s_1. K is interpreted as the amount delivered by the CCP per unit of contract, net of clearing fund losses. The CCP pools member banks' credit risks and redistributes default losses across survivors, which implies that it faces a standard long-run zero-profit condition requiring total revenue to equal total expected loss:[7]

$$0 = \frac{N}{2}[\pi\theta + (1-\pi)\Phi\theta - K\theta].$$

Assumptions 1-3 imply that the $N/2$ banks that owe consumption good units to the CCP either deliver their positions θ in full with probability π, or else default and deliver just the collateral $\Phi\theta$; from this revenue, the CCP transfers $K\theta$ to their original counterparties. Since the number of banks is large, laws of large numbers and the zero-profit equation permit one to say that the amount delivered by the CCP per unit of contract (net of clearing fund losses) equals

$$K = \mathbb{E}K = \pi + (1-\pi)\Phi.$$

Thus, each member due to receive one unit of the consumption good from its original counterparty instead receives $K < 1$ units from the CCP with certainty.

When determining their trading positions, banks have to weigh the benefit of trading against the shadow cost of collateral. Since the two bank groups are symmetric, it suffices to consider Bank 1's optimization problem. A bank from group 1 maximizes its expected utility from trading, subject to choosing an optimal trading position θ and an optimal delivered amount $D(\cdot)$ in each state x, y or z. Since the objective function is strictly concave and is optimized over a convex set, the optimization problem has a unique solution determined by

[7] It is assumed that the CCP has no operating costs.

solving

$$\max_{\theta,D} EU = \frac{1}{2}\Big[u(c_x) - \lambda \max\{-\theta - D(x), 0\}\Big] + \frac{\pi}{2}\Big[u(c_y) - \lambda \max\{\theta - D(y), 0\}\Big] + \\ + \frac{1-\pi}{2}\Big[u(c_z) - \lambda \max\{\theta - D(z), 0\}\Big],$$

where

$$\begin{aligned} c_x &= x - D(x) + \max\{\theta, 0\}K; \\ c_y &= y - D(y) + \max\{-\theta, 0\}K; \\ c_z &= z - D(z) + \max\{-\theta, 0\}K; \end{aligned}$$

subject to the constraints

$$\begin{aligned} z &\geq \max\{\theta, 0\}\Phi; \qquad & x &\geq \max\{-\theta, 0\}\Phi; \\ D(y), D(z) &\geq \max\{\Phi\theta, 0\}; \qquad & D(x) &\geq \max\{-\Phi\theta, 0\}. \end{aligned}$$

The equations for c_x, c_y and c_z define Bank 1's final wealth conditional on the endowment realization. For example, in state s_1 Bank 1 receives endowment x with certainty and, if it holds a long position ($\theta > 0$), is entitled by contract to receive one unit of the consumption good from Bank 2 via the CCP, which withholds $1 - K$ and delivers K units to Bank 1. Since Bank 1 is the one receiving the good, its delivery rate is then naturally $D(x) = 0$ and the wealth equation becomes $c_x = x + \theta K$. Conversely, if Bank 1's position is short ($\theta < 0$), Bank 1 is obliged to deliver one unit of the consumption good to the CCP, and its wealth in state s_1 is $c_x = x - D(x)$, where $D \geq \theta\Phi$ is the amount it finds optimal to deliver, conditional on obtaining x. The constraints on $|\theta|\Phi$ and $D(\cdot)$ ensure that the bank delivers at least the posted collateral, regardless of how good or bad the endowment realization is.

To simplify the problem, Lemma 1 in the Appendix shows that Assumptions 1 and 2 guarantee that it is optimal for Bank 1 to go long ($\theta > 0$) and for Bank 2 to go short ($\theta < 0$), so that, by the symmetry of the two banks, $\theta_1 = -\theta_2$ and the market clears. In addition, Assumption 2 implies that a bank receiving the bad shock z will default and deliver only the collateral, so $D(z) = |\theta|\Phi$, while Assumption 3 implies that $D(y) = |\theta|$, since strategic defaults are ruled out. This determines the delivered amount D in all states of the world and allows one to rewrite the maximization as a problem over the trading position θ only:

$$\max_{\theta} EU = \frac{1}{2}u(x + \theta K) + \frac{\pi}{2}u(y - \theta) + \frac{1-\pi}{2}[u(z - \theta\Phi) - \lambda(1 - \Phi)\theta].$$

The associated Lagrangian is:

$$\mathscr{L} = \frac{1}{2}u(x+\theta K) + \frac{\pi}{2}u(y-\theta) + \frac{1-\pi}{2}[u(z-\theta\Phi) - \lambda(1-\Phi)\theta] - \gamma(\theta\Phi - z),$$

where the multiplier γ reflects the shadow cost of collateral, reflecting that at least the margin must be delivered in the event of default.

Bank 1's first-order condition with respect to its trading position θ is

$$\frac{d\mathscr{L}}{d\theta} = \frac{K}{2}u'(x+\theta K) - \frac{\pi}{2}u'(y-\theta) - \frac{1-\pi}{2}[\Phi u'(z-\theta\Phi) + \lambda(1-\Phi)] - \gamma\Phi = 0.$$

Denote this equation with $G(\theta|K,\Phi) = 0$. The equation $G = 0$ implicitly defines the optimal trading position θ as a function of the CCP's delivery rate K and margin Φ. But since K is already a function of Φ, the optimal choice of θ is ultimately a function of margin Φ. I will use this relationship extensively in order to derive the results that follow.

3.3 Concavity of the clearinghouse's zero-profit location

One of the main problems with TTC margining is that it causes more risk sharing among members in financially stressed periods, because limiting margin spikes transfers risk from the margin pool to the mutualized clearing fund. Moreover, the margin – clearing fund relationship is likely to be non-linear for a number of reasons, including both the methodology used by CCPs to size their clearing funds and conventional moral hazard considerations. This non-linearity is at the heart of much of the results that follow. First I discuss how a typical CCP clearing fund is structured, and then, how increased risk sharing could change member behavior in times of stress.

A CCP has two main resource buffers to cover losses from member non-performance: the pool of posted collateral and a mutualized clearing fund, maintained by pooled member contributions. Any default loss in excess of the defaulter's margin therefore typically spills over to the mutualized fund.[8] CCPs typically compute their margin requirements to guarantee that the risk of spillover is acceptably low, for example, by ensuring that the margin requirement covers expected losses with a high level of statistical confidence (usually 99 or 99.87 per cent). However, even when margin is computed at a high confidence level, dropping members' margin requirements by a dollar across the board translates into less than a dollar increase in the clearing fund. The reason for this is that a CCP clearing fund typically covers only the largest exposure created by any single participant, or at most the top two

[8] A third resource buffer, CCP capital to be used before assessing the clearing fund, is typically negligible relative to the CCP's total activity and is not required by regulatory standards.

largest exposures (for systemically important CCPs).[9] To size the clearing fund, the CCP estimates the top exposure and then determines individual member contributions to cover it fully, proportional to each member's clearing activity over the year. Hence, permitting members to take advantage of margin relief does not trigger increases in the clearing fund contributions for any other members but the one creating the largest exposure. Therefore, in any CCP with more than a couple of members, reducing the margin requirement by a dollar results in less than a dollar increase in the clearing fund. This observation suggests a diminishing marginal rate of substitution between clearing fund losses and the margin requirement, which implies a convex relationship between the expected clearing fund loss $(1 - K)$ and margin Φ. Clearing fund size is decreasing and convex in margin because mutualized and pre-funded resources are not perfect substitutes. Reversing the viewpoint, one can equivalently say that the CCP's delivery rate K is increasing and *concave* in margin: protection naturally increases with margin, but the incremental contribution of collateral to additional safety becomes vanishingly small when most of the position is already collateralized. Hence, the structure and clearing fund sizing methodology already imply a concave relationship between K and Φ.

The protection K provided by the clearinghouse, however, can also be concave in margin for other reasons, such as moral hazard. For example, the increased risk sharing in stressed periods associated with TTC margining can distort incentives by encouraging unobserved risky behaviors. In the case of a CCP, this can manifest itself in strategic behaviors where each individual bank benefits from playing it safe only if the remaining members are safe, but is better off being risky if the remaining members are risky, since it cannot control its mutualized exposure. Such incentive effects of collateral can also result in a non-linear relationship between clearing fund and margin. Moral hazard effects are likely to be negligible at high margin levels (when Φ is close to 1) but would become increasingly severe if collateral requirements were to drop near zero, so the banks' default risk is likely to be higher at near-zero margins, and increasingly so the lower the margin requirement. For concreteness, in the following analysis I will adhere to the moral hazard interpretation by assuming that counterparty quality worsens because of unobserved risk taking when the collateral requirement is too low, and that this fact is common knowledge, so the concavity of $K(\Phi)$ is driven by changes in the banks' default probability. However, it is worth emphasizing that this is mostly a point of interpretation: whether one attributes the concavity of $K(\Phi)$ to moral hazard or to the clearing fund sizing methodology does not make much

[9] CPSS-IOSCO (2012), Principles for Financial Market Infrastructures, Principle 4 (Credit risk).

difference to the results that follow; both factors result in a concave relationship.

To reflect the fact that counterparty quality progressively worsens at low margins, I assume that the bank's non-default probability $\pi(\Phi)$ is increasing in margin Φ, concave and twice differentiable with $\pi'(0) \to \infty$, $\pi'(1) = 0$. This assumption implies that when most of the position is collateralized, moral hazard is negligible, but when margin approaches 0 per cent, moral hazard becomes increasingly severe. The incentive effect of margin can vary the probability π within the bounds[10] $[\underline{\pi}, \bar{\pi}]$, so $\pi(0) = \underline{\pi}$ and $\pi(1) = \bar{\pi}$. Recalling that the CCP's delivered amount K is linear in the delivery probability π, but π is non-linear in margin, we recognize that the level of protection K provided by the CCP varies non-linearly with margin Φ so that the concavity of $\pi(\Phi)$ translates to $K(\Phi)$. The curve $K(\Phi)$ shown in Figure 2 is the clearinghouse's *zero-profit location.* The clearinghouse can offer its customers any (Φ, K) combination located on this curve.

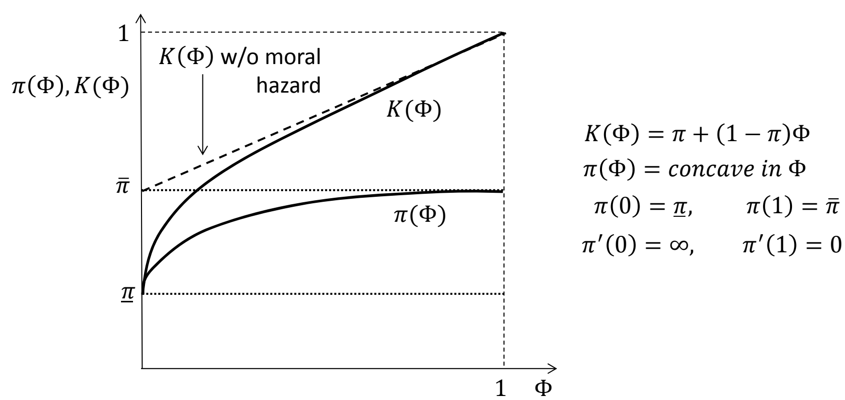

Figure 2: The relationship between protection level K and margin Φ is concave

4 Results

I analyze optimal margining in this stylized economy and obtain three broad sets of results: on the relationship between trading volume and margin, on optimal margining, and on policy implications.

4.1 Relationship between trading volume and margin

By now, the view that high collateral requirements reduce clearing activity has been uncritically accepted by many authors as intuitive. For example, Gibson and Murawski (2013,

[10]It is also assumed that $\underline{\pi}$ is sufficiently high to satisfy Assumption 1, but the assumption is immaterial since the results depend on the *rate of change* in π, not on its level.

Proposition 2) show that a bank's trading position in a derivative asset is a decreasing function of the margin requirement. As I demonstrate here, this reasoning is correct only if one ignores the concave relationship between the protection afforded by the CCP and its margin requirement, likely to arise from clearing fund sizing rules or from moral hazard. For instance, undercollateralizing a trade reduces the amount recovered after default, but at the same time also worsens the incentives to avoid risks leading to a default. For this reason, the relationship between margin requirements and trading volume is not likely to be monotonic. The next proposition shows that this non-monotonicity causes trading volume to peak at a certain level of margin. This implies that margin relief in stressed periods, as implied by TTC margining, can increase as well as *depress* trading, depending on the peak's location.

Proposition 1. *The relationship between trading volume θ and margin Φ is non-monotonic.*

a) Trading volume θ is increasing in margin Φ at low levels of margin and decreasing at high levels of margin.

b) There exists an interior level of margin $\Phi_M \in (0,1)$ where trading position θ is maximal and where $(d\theta/d\Phi)|_{\Phi_M} = 0$.

Proof: See the appendix.

Proposition 1 implies that the relationship between market volume θ and margin Φ is hump-shaped and has an interior peak, whose location is crucial in determining whether changes to the margin requirement will dampen or boost trading. Regulatory bodies are concerned with margin jumps in stressed markets because they believe that such jumps dry up market liquidity and inhibit trading, which TTC margining should ideally help avoid. But if the CCP's margin requirement is to the left of the peak-inducing level Φ_M, reducing it further will not result in more trading; on the contrary, it will depress market activity.

The intuition is that when margin requirements are low, banks have little collateral pledged with low associated opportunity cost, which lowers the economic cost of trading. Under the moral hazard interpretation of the concavity of K, as the margin requirement drops, undercollateralizing induces risk taking, because there is too little bond to secure the performance. Rather than keeping the would-be collateral in its safe, the bank could invest it in high-risk, high-return projects associated with a higher risk of non-performance. Proposition 1 suggests that banks see such undercollateralized trades as risky and try to reduce their exposure by limiting trading instead. However, when margin requirements become high, the opportunity cost of collateral becomes too high relative to the risk of non-performance, and this likewise dampens trading, resulting in a hump-shaped relationship.

4.2 Bank-optimal margin and trading volume

In this simple model, the clearinghouse redistributes all default losses to members, so it can offer them any combination of margin and protection (Φ, K) located on its zero-profit curve (curve $K(\Phi)$ in Figure 2). The actual combination is set by the CCP, typically in consultation with its members. CCP member banks often have a significant say over the clearinghouse's margin policy, and it is reasonable to assume that a through-the-cycle margining policy cannot be adopted without their consent.[11] For this reason, the bank-optimal margin is a useful benchmark to consider.

Since margin relief reduces trade volume when the actual margin is set to the left of Φ_M, the location of the peak in the volume-margin relationship is critical in determining the response of trading volume (increase or decrease) to margin relief. It is therefore natural to ask whether the bank-optimal margin lies to the left or to the right of Φ_M. Below, I show that the answer to this question depends on two variables: the spread of the endowment distribution, which I interpret as financial market volatility, and on risk aversion.

Recall that banks maximize their expected utility with respect to the trading position θ, taking the margin requirement as given. To find the bank-optimal margin, banks optimize over both trading position and margin. The first-order condition with respect to margin is given by:

$$\frac{\partial \mathscr{L}}{\partial \Phi} = \frac{\theta K'(\Phi)}{2} u'(x + \theta K) + \frac{\pi'(\Phi)}{2} \left[u(y - \theta) - u(z - \Phi\theta) + \lambda(1 - \Phi)\theta \right] + \\ + \frac{1 - \pi(\Phi)}{2} \left[-\theta u'(z - \Phi\theta) + \lambda\theta \right] - \gamma\theta = 0.$$

The solution to this equation is the bank-optimal margin Φ_B for a given trading position θ. However, the size of the trading position itself is a function of margin via the hump-shaped volume-margin relationship defined implicitly by banks' first-order condition in θ. Recognizing this, I substitute the latter first-order condition into the former. To determine whether the bank-optimal margin Φ_B corresponds to below-peak or above-peak trading volume, in the next step I evaluate the sign of $d\mathscr{L}/\partial\Phi$ at the point (Φ_M, θ_M). As it turns out, risk aversion together with the spread of the endowment distribution jointly determine

[11] For example, the Principles for Financial Market Infrastructures (key consideration 3.2.14) require that, as a minimum, a CCP should have a risk committee composed of a majority of non-executive members, which ensures that any substantive changes to the CCP's risk-management policy cannot be approved without a majority member vote.

the locations of Φ_B and Φ_M with respect to each other. When market volatility is high and risk aversion is low, banks are less concerned about moral hazard than about the shadow cost of collateral driven by volatility spikes, and so have no problem trading at lower margins below Φ_M. However, above a certain critical value of risk aversion, banks will prefer high margins above Φ_M, regardless of volatility, because moral hazard concerns outweigh the shadow cost of collateral. These results are summarized in Proposition 2.

Proposition 2. *Define a model calibration as a choice of endowment outcomes x, y, z, a choice of a concave utility function $u(\cdot)$, and a choice of a concave function $\pi(\Phi)$, as described in section 3. Then:*

a) Given sufficiently high volatility $|y - z|$, CARA utility and sufficiently low absolute risk aversion $R_A < 1/z$, there exists a model calibration for which the bank-optimal margin is to the left of the volume-maximizing margin level ($\Phi_B < \Phi_M$).

b) Given CARA utility and sufficiently high absolute risk aversion $R_A > 1/z$, there exists a model calibration for which the bank-optimal margin is to the right of the volume-maximizing margin level ($\Phi_B > \Phi_M$).

Proof: See the appendix.

This proposition shows that in high-volatility times, margin relief can depress trading, provided that volatility is sufficiently high and clearinghouse members are not too risk averse.

When the banks' risk aversion is low, they are more concerned about the collateral cost of trading rather than about their mutual exposures within the clearinghouse; therefore, banks are willing to take advantage of margin relief, even though it implies an increase in mutualized risk. To compensate for their higher mutual exposures through the clearing fund, banks instead reduce their trading positions. Thus TTC margining can ultimately lead to a reduction of trading activity during stressed periods.

Technically, the result can be decomposed into a "volatility effect" and a "risk aversion effect." Market volatility determines whether margin relief boosts or suppresses trading activity, but risk aversion can override the effect of volatility when risk aversion is high. Let us consider the two effects in turn. A higher spread of the endowment outcomes y and z increases y and reduces z, and, all else held equal, makes it harder to deliver the pledged collateral $\Phi|\theta|$, since it must be delivered in all cases, including a bad draw z. This increases the shadow cost of collateral, reflected by the Lagrange multiplier γ, and increases the utility cost associated with trading. If banks are not too concerned about mutualized risk, margin-related cost considerations prevail and their optimal margin is low enough to be to the left of Φ_M. However, if risk aversion exceeds the critical level $(1/z)$, default fund risk considerations prevail *regardless* of volatility and banks find it best to protect themselves with a high margin to the right of Φ_M. The implication is that when the bank-optimal margin is to the left of Φ_M, margin relief will trigger a slowdown in trading activity; and when the bank-optimal margin is to the right of Φ_M, margin relief boosts trading.

The proposition suggests that margin relief is more likely to dampen trading in periods with high volatility, when the collateral opportunity cost is high, provided that banks are not overly sensitive to the risk mutualized through the clearing fund. In contrast, margin relief will boost trading if banks are sufficiently risk averse, because they will demand high margins to the right of Φ_M.

This finding's main policy implication is that margin relief policy should have a discretionary component. Since much depends on the mutual locations of the bank-optimal and the volume-maximizing level of margin, policy-makers must assess the situation on a case-by-case basis; margin relief, if granted in the wrong circumstances, could lower trading in stressed markets contrary to the policy-maker's objective.

4.3 Privately optimal versus socially optimal margin

When daily stock market swings of 200-300 points became common at the height of the 2008 crisis, higher market volatility quickly translated into sharp spikes in CCP margin requirements. For example, the Canadian Derivatives Clearing Corporation's margin requirement on 27 October 2008 reached a level seven times higher than at the beginning of 2003 (Chande and Labelle, 2013). Some clearinghouses recognized that coming up with additional liquid collateral to satisfy extra margin calls might cause difficulties for members and could put additional liquidity strain on the system. For example, one derivatives CCP[12] proposed to its members a methodology to smooth out the sharpest volatility-driven spikes in its margin requirements, effectively offering members a variant of the TTC margin approach. Surprisingly, members rejected the proposal quoting concerns about CCP safety and the extra tail risk involved. Member banks were much more concerned with the extra mutualized risk resulting from TTC margining than with the extra cost of collateral. If such concerns became widespread, they could become a serious obstacle to reducing margin procyclicality.

[12]The institution has been anonymized to preserve its confidentiality.

This section explores situations where the socially optimal margin can diverge from the banks' privately optimal margin, with a special focus on cases where margin relief may be optimal from a social standpoint but privately undesirable, as in the example above. Such situations are a challenge for policy-makers, since it may be politically difficult to lower collateral standards without credibility risk. The political risk stems from the fact that regulators usually mandate minimum risk-management standards, which could be eroded by margin relief and made more difficult to reinforce in the future. For this reason, I assume that margin relief cannot be approved as a CCP policy without support from its member banks, even if deemed desirable by regulators. This is consistent with major international regulatory standards, which require that changes to the CCP's risk management be approved by a majority vote of its participants.

This partial equilibrium model focuses on the banking sector, which represents only a fraction of the economy. In reality, many other financial agents – traders, brokers, asset managers, insurance companies, custodians, and securities depositories – benefit directly or indirectly from derivatives trading, but are not modelled here because of complexity limitations. Likewise, bank defaults often have ripple effects reaching beyond the banking sector and spilling over to other institutions in the financial system, such as insurance companies.[13] Recognizing the partial equilibrium nature of the model, I include such external spillovers in the form of reduced-form externalities. Despite the stylized treatment, this approach can still generate some valuable insights, because the social benefits from trading and the social costs of default do not simply cancel each other out: they are not linear in each other. In the next step, I show that it is possible for the privately optimal margin Φ_B to exceed the socially optimal margin Φ_S, and identify the necessary conditions for this.

It makes sense to assume that social benefits from trading are proportional to trading volume, so define $SB = b\theta$. Likewise, it makes sense to assume that social costs of default are proportional to the expected size of the defaulted position $(1-\pi)\theta(1-\Phi)$, so I define the social costs of default as $SC = c\theta(1-\pi)(1-\Phi)$. This implies net social benefits equal $\theta[b - c(1-\pi)(1-\Phi)]$ and allows one to define the social welfare function W as

$$W = N\mathbb{E}u(\Phi) + \theta(\Phi)[b - c(1-\pi(\Phi))(1-\Phi)].$$

The first term is the sum of the expected private utilities of the N banks, defined in section 3.2; the second term is the expected net social benefit. I am interested in when the privately

[13] An example is AIG's 2008 near-collapse brought about by its excessively large exposure to credit default swaps during the crisis.

optimal margin Φ_B can exceed the socially optimal one Φ_S. For this to hold, it must be true that $\frac{dW}{d\Phi}\big|_{\Phi_B} < 0$. The next Proposition specifies when this is the case.

Proposition 3. *If the following conditions hold:*

1) Risk aversion is sufficiently high so that $R_A > 1/z$.

2) The social benefit from trade exceeds the expected social cost of default on a per-dollar basis, so that $b > c(1-\pi)(1-\Phi)$.

Then there exists a sufficiently large value of $b > 0$ for which $\Phi_B > \Phi_S$.

Proof: See the appendix.

Proposition 3 implies that at sufficiently high levels of risk aversion, mutualized risk exposure concerns can drive banks to an excessively high margin requirement. Since banks do not factor in third-party benefits from interbank trading, the privately optimal margin exceeds the socially optimal level. A precondition for this is that social benefits from derivatives trading must exceed the expected social cost from defaults – a realistic condition in view of the low default probability forming the expectation. (Recall that earlier we assumed $1-\underline{\pi}$ is close to zero, reflecting that in the banking world, defaults are plausible yet sufficiently rare events.)

Situations where the privately optimal margin exceeds the socially optimal one are challenging for policy-makers, since they may not be able to implement TTC margin relief when needed from a system-wide perspective. As illustrated by the derivatives CCP example, implementing the margin relief associated with TTC margining may turn out impossible in stressed periods if deemed risky by the financial industry. Proposition 3 replicates this situation and provides some additional insight as to why it occurred. For privately optimal margins to diverge from the social optimum, banks must be sufficiently risk averse, while at the same time generating sufficient positive spillovers from trading to the remainder of the economy. This is exactly the case with the derivatives CCP from the example. This CCP's concentrated membership of a limited number of large, conservative banks with substantial links to the remaining economy implies that risk aversion was high and the parameter b was large. These are precisely the conditions predicted by the model to generate high privately optimal margin.

5 Alternative Policy Tools

Since it could become politically challenging for authorities to push for TTC margining, they could seek alternative tools to influence markets. One alternative way to align the privately optimal margin closer to the socially optimal level is to increase default penalties. A CCP's typical selection of penalty measures includes limiting the member's transactions, requiring the member to reduce or close out its positions, assessing the member's clearing fund and margin deposit, preventing withdrawal of excess margin deposits, or fining a non-conforming member. Permanent suspension from the CCP is usually the last and most severe penalty. By making defaults more costly, higher default penalties could assure banks that mutualized risk is under control, motivating them to agree to lower margins in stressed periods. Higher default penalties decrease external risk-taking incentives, since it becomes more costly to default on a trade, and reduce the banks' mutualized exposures through the clearing fund.

The following analysis shows that this intuition is indeed correct if higher default penalties reduce the default probability, all else being equal; that is, regardless of the amount of margin posted. In this specification of the model, which is likely more realistic, higher default penalties can indeed motivate member banks to agree to margin relief; this is shown in Proposition 4 below.

Recall that in section 4, we posited that the non-default probability $\pi(\Phi)$ is increasing in margin and strictly concave with a range of $[\underline{\pi}, \bar{\pi}]$. For simplicity, so far $\underline{\pi}$ and $\bar{\pi}$ were treated as fixed. In reality, higher default penalties are likely to increase π (and reduce the default probability $1-\pi$) for any given level of margin. It is therefore realistic that the minimum and maximum default probabilities $1-\bar{\pi}$ and $1-\underline{\pi}$ are functions of λ. To reflect this, the function $\pi(\Phi)$ can be decomposed into an additively separable intercept in λ and a concave function $v(\Phi)$:

$$\pi(\Phi) = v(\Phi) + \underline{\pi}(\lambda),$$

where the intercept $\underline{\pi}(\lambda)$ is increasing and differentiable in λ and $v(\Phi)$ retains the remaining desirable properties[14] of the function $\pi(\Phi)$. This way, $\bar{\pi}$ defined as $\bar{\pi} \equiv v(1) + \underline{\pi}(\lambda)$ also becomes a function of λ, and the parameter λ acts as a shifter of the default probability, therefore subsuming all analysis performed so far as a partial case.[15] With this specification,

[14]Strict concavity, $v(0) = 0, v'(0) = \infty, v'(1) = 0$, $v(1) \equiv \bar{\pi}(\lambda_0) - \underline{\pi}(\lambda_0)$, where λ_0 is the initial value of λ. Also observe that increasing default penalties under this specification can only increase $\underline{\pi}$ and therefore Assumption 1, which requires $\underline{\pi}$ to be sufficiently large, continues to hold.

[15]Notice that since values of the derivative of $\pi(\Phi)$ remain the same for each value of Φ, introducing this shifter does not change any of the first-order conditions discussed previously.

it is possible to show the following result:

Proposition 4. *Increasing the default penalty λ will motivate member banks to agree to margin relief (a reduction in Φ_B) provided that*

$$\underline{\pi}'(\lambda) \geq \frac{1}{\lambda}.$$

Proof: See the appendix.

This result implies that the default probability must be be sufficiently elastic to changes in λ for banks to agree to margin relief, as reflected by a sufficiently large value of $\underline{\pi}'(\lambda)$. Clearinghouses have a variety of practical ways to increase their default penalties: from increasing the length of time a member is suspended following a default, to imposing fines on defaulters that would permit the clearinghouse to recover *ex post* the amount undercollected as margin owing to margin relief. The latter option appears especially attractive as a practical solution. For example, a CCP could size its default penalty proportional to the dollar amount of margin relief granted to each bank. In the event of default, the CCP could impose a fine proportional to the margin relief times a surcharge large enough to ensure that the policy has enough "bite." If the member bank does not enter bankruptcy, it will suffer a large financial penalty, while in a bankruptcy scenario, the CCP will retain a valid claim on the defaulting member's assets. The CCP therefore may recover most of the fine in the long run, and the surcharge coefficient could be sized to anticipate for its unrecovered portion. This should create better incentives to control risk seeking while preserving the long-term health of the CCP.

In order to implement this process, communication between CCPs and regulators is essential. Changing the process of penalizing defaulters requires changes to CCP default rules and procedures, which typically involve a lengthy regulatory approval. Without an established communication process, obtaining such approval may take too long and delay the implementation of policy beyond the point when it is needed.

6 Conclusion

A major policy challenge in centrally cleared derivatives markets is that collateral requirements can rise sharply in times of stress, inhibiting trading and causing liquidity pressures that exacerbate the crisis. Margin-smoothing approaches to avoid such spikes, such as through-the-cycle margining, however, inevitably increase the risk mutualized through the clearinghouse in times of stress. In this paper, I study the optimal trade-off between margin relief and mutualized risk in the context of a centrally cleared derivatives market.

I begin with the observation that the CCP's zero-profit location is concave in the margin requirement, for reasons having to do both with how CCPs structure and size their clearing funds as well as with how participant incentives respond to increased risk sharing. Margin is viewed as both a form of prepayment and an incentive device to reduce moral hazard, and thus can have ambiguous effects on trading volume and risk. In contrast to the existing literature, I show that TTC margining in stressed markets can increase as well as *decrease* trading volume, depending on financial market volatility and member banks' risk aversion. This implies that TTC margining should not be implemented without a careful survey of market volatility and clearing members' risk attitudes. Since TTC margining increases the share of mutualized risk in stressed periods, clearing members may prefer higher margin requirements and lower trading volume than is socially optimal, creating a challenge for regulators, since it may be politically challenging to enforce lower collateral standards than deemed appropriate by the industry.

This challenge is more likely to occur when banks' risk aversion is high and the social benefits from trade exceed the expected social cost of default. Favorably for less procyclical margin policy, such privately optimal overmargining is less likely to occur in times of high volatility, when the shadow cost of collateral is high. Unfavorably, when it does happen, policy-makers may have limited tools to remedy it. When the privately optimal and the socially optimal margin diverge, regulators may need to pursue alternate policy tools and instead work with the CCP to increase the penalties on defaulting members, which can align the privately optimal with the socially optimal level of margin.

7 Appendix: Proofs

Lemma 1. *The optimal position θ_1 of a bank from group 1 (henceforth denoted Bank 1) is nonnegative.*

Proof. This auxiliary result, which is standard in the literature, is proved by contradiction in the same way as in Santos and Scheinkman (2001). Suppose, counterfactually, that instead $\theta_1 < 0$. Let D^* be the solution to $u'(x - D^*) = \lambda$; denote D^* as the bank's *critical delivery rate*. There are three possibilities for the location of D^* with respect to the bank's position θ:

1. $-D^* \leq \theta_1$;
2. $\theta_1 < -D^* < \theta_1 \Phi$;
3. $-D^* \geq \theta_1 \Phi$.

Consider case 1 first. Here the critical delivery rate exceeds the bank's position in absolute value, so Bank 1 delivers fully and $D(x) = -\theta_1$. Applying Bank 1's first-order condition for θ, this yields

$$\begin{aligned} u'(x) < u'(x + \theta_1) = u'(x - D) = \lambda = \\ = K[\pi u'(y - \theta_1 K) + (1 - \pi) u'(z - \theta_1 K)] < \\ < \pi u'(y) + (1 - \pi) u'(z), \end{aligned}$$

in contradiction to Assumption 1; therefore case 1 leads to a contradiction.

In case 2, the critical delivery rate D^* is smaller than the bank's position (in absolute value), so there is a partial default. The bank delivers only $D(x) = D^*$ and defaults on the rest. Substituted in the first-order condition for θ_1, this yields

$$\begin{aligned} u'(x) < u'(x - D^*) = \lambda = \\ = K[\pi u'(y - \theta_1 K) + (1 - \pi) u'(z - \theta_1 K)] < \\ < \pi u'(y) + (1 - \pi) u'(z), \end{aligned}$$

again in contradiction to Assumption 1.

In case 3, $D(x) = -\theta_1 \Phi$, because the bank always delivers the pledged collateral $|\theta|\Phi$ as a minimum. Once again using the first order condition for θ_1, obtain

$$\begin{aligned} \lambda = u'(x - D^*)\Phi + \lambda(1 - \Phi) \leq u'(x + \theta_1 \Phi)\Phi + \lambda(1 - \Phi) = \\ = K[\pi u'(y - \theta_1 K) + (1 - \pi) u'(z - \theta_1 K)] < \\ < \pi u'(y) + (1 - \pi) u'(z), \end{aligned}$$

in contradiction to Assumption 2.

Since all three cases are impossible, it follows that $\theta_1 \geq 0$. This implies that $D(x) = 0$; moreover, observe that Assumption 1 also guarantees that $D(z) = |\theta|\Phi$, implying that a bank receiving z will deliver only the collateral. $\square$

Proposition 1. *The relationship between trading volume θ and margin Φ is non-monotonic.*

a) *Trading volume θ is increasing in margin Φ at low levels of margin and decreasing at high levels of margin.*

b) *There exists an interior level of margin $\Phi_M \in (0,1)$ where trading position θ is maximal and where $(d\theta/d\Phi)|_{\Phi_M} = 0$.*

Proof. a) By the implicit function theorem, the relationship between the optimal θ and Φ is characterized by

$$\frac{d\theta}{d\Phi} = -\frac{\partial G/\partial \Phi}{\partial G/\partial \theta},$$

provided that $\partial G/\partial \theta$ is non-zero and preserves its sign in a neighborhood of θ_{opt}. Observe that $\partial G/d\theta \equiv d^2EU/d\theta^2 < 0$, since the expected utility function is strictly concave, so all requirements of the implicit function theorem are satisfied. Therefore the sign of $\frac{d\theta}{d\Phi}$ is determined by that of $\partial G/\partial \Phi$:

$$\frac{\partial G}{\partial \Phi} = \frac{K'(\Phi)}{2}u'(x+\theta K) + \frac{\pi'(\Phi)}{2}[\Phi u'(z-\theta\Phi) + \lambda(1-\Phi) - u'(y-\theta)] + \\ + \frac{1-\pi(\Phi)}{2}[u'(z-\theta\Phi) - \theta\Phi u''(z-\theta\Phi) - \lambda] - \gamma.$$

I evaluate the expression $\frac{dG}{d\Phi}$ at two points $\Phi = 0$ and $\Phi = 1$:

$$\left.\frac{\partial G}{\partial \Phi}\right|_{\Phi=0} = \frac{1-\underline{\pi}}{2}u'(x+\theta\underline{\pi}) + \frac{\pi'(0)}{2}[u'(x+\theta\underline{\pi}) - u'(y-\theta) + \lambda] - \frac{1}{2}[u'(z) - \lambda] - \gamma$$

$$\left.\frac{\partial G}{\partial \Phi}\right|_{\Phi=1} = \frac{1-\overline{\pi}}{2}[u'(x+\theta) - u'(z-\theta) + \theta u''(z-\theta) + \lambda] - \gamma,$$

using the fact that $\pi'(1) = 0$.

The first expression is strictly positive when

$$(1-\underline{\pi})u'(x+\theta\underline{\pi}) + \pi'(0)[u'(x+\theta\underline{\pi}) - u'(y-\theta) + \lambda] > [u'(z) - \lambda] + 2\gamma,$$

which is guaranteed when $\pi'(0)$ is large enough. Since it is assumed that $\lim_{\Phi\to 0}\pi'(\Phi)=\infty$, this is always true, so $(dG/d\Phi)|_{\Phi=0} > 0$ as claimed.

The second expression, $(dG/d\Phi)|_{\Phi=1}$, is strictly negative whenever

$$u'(x+\theta) - [u'(z-\theta)-\lambda] + \theta u''(z-\theta) < 2\gamma/(1-\overline{\pi}).$$

By Assumption 3, $u'(z) > u'(x)+\lambda$, which implies $u'(x+\theta)-[u'(z-\theta)-\lambda] < 0$. Since $u'' < 0$, the left-hand side is negative and the inequality is satisfied as claimed. □

b) I use the fact that $d\theta/d\Phi$ exists and is continuous. Since $(d\theta/d\Phi)|_{\Phi=0} > 0$ while $(d\theta/d\Phi)_{\Phi=1} < 0$, by continuity there exists a point Φ_M s.t. $(d\theta/d\Phi)|_{\Phi_M} = 0$. This point is a maximum, since to the left of it, the derivative is positive, and to the right of it, negative.

Existence of $d\theta/d\Phi$ follows directly from the implicit function theorem. Its continuity follows from the fact that it is differentiable. This can be verified by taking the derivative

$$\frac{d^2\theta}{d\Phi^2} = -\left[\frac{\partial G}{\partial\theta}\right]^{-2}\left[\frac{\partial^2 G}{\partial\Phi^2}\frac{\partial G}{\partial\theta} - \frac{\partial G}{\partial\Phi}\frac{\partial^2 G}{\partial\theta^2}\right],$$

which exists whenever the partials $\frac{\partial^2 G}{\partial\Phi^2}$ and $\frac{\partial^2 G}{\partial\theta^2}$ exist. It is easy to verify that

$$\frac{\partial^2 G}{\partial\theta^2} = \frac{K^3}{2}u'''(x+\theta K) - \frac{\pi(\Phi)}{2}u'''(y-\theta) - \frac{(1-\pi(\Phi))}{2}\Phi^3 u'''(z-\theta\Phi),$$

and this second derivative therefore exists whenever $u'''(\cdot)$ exists. Likewise, since $\partial^2 G/\partial\theta^2 \equiv d^3EU/d\theta^3$, the existence of u''' also guarantees that of $\partial^2 G/\partial\theta^2$. □

Proposition 2. *Define a model calibration as a choice of endowment outcomes x, y, z, a choice of a concave utility function $u(\cdot)$, and a choice of a concave function $\pi(\Phi)$, as described in section 3. Then:*

a) Given sufficiently high volatility $|y-z|$, CARA utility and sufficiently low absolute risk aversion $R_A < 1/z$, there exists a model calibration for which the bank-optimal margin is to the left of the volume-maximizing margin level ($\Phi_B < \Phi_M$).

b) Given CARA utility and sufficiently high absolute risk aversion $R_A > 1/z$, there exists a model calibration for which the bank-optimal margin is to the right of the volume-maximizing margin level ($\Phi_B > \Phi_M$).

Proof.

a) To compare the relative magnitudes of Φ_B and Φ_M, I evaluate $\frac{\partial \mathscr{L}}{\partial \Phi}$ at $\Phi = \Phi_M, \theta = \theta_M$. If its sign is positive, then $\Phi_B > \Phi_M$. If negative, then $\Phi_B < \Phi_M$.

The value of Φ_M, in turn, is implicitly defined by the equation $dG/d\Phi = 0$:

$$\frac{dG}{d\Phi} = \frac{K'(\Phi)}{2} u'(x + \theta K) + \frac{\pi'(\Phi)}{2} [\Phi u'(z - \theta\Phi) + \lambda(1 - \Phi) - u'(y - \theta)] + \\ + \frac{1 - \pi(\Phi)}{2} [u'(z - \theta\Phi) - \theta\Phi u''(z - \theta\Phi) - \lambda] - \gamma = 0.$$

Using the fact that Φ_M satisfies $dG/d\Phi = 0$, I express the shadow cost of collateral γ from equation $dG/d\Phi = 0$ in terms of all remaining variables and substitute in the expression $\frac{\partial \mathscr{L}}{\partial \Phi}$. This substitution implies that the bank is optimizing both over its trading position and over the optimal margin, resulting in

$$\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M} = \frac{\pi'(\Phi_M)}{2} [u(y - \theta_M) - u(z - \Phi_M\theta_M) + \\ + \theta_M [u'(y - \theta_M) - \Phi_M u'(z - \theta_M\Phi_M)]] - \frac{1 - \pi(\Phi_M)}{2} \theta_M^2 \Phi_M u''(z - \theta_M\Phi_M).$$

To simplify matters, select a calibration of π such that $\pi'(\Phi_M) = \frac{1-\pi(\Phi_M)}{\Phi_M}$. By the mean value theorem, such a calibration always exists, since π is twice-differentiable and the continuous first derivative π' changes from infinity to zero. This allows me to factor out $\frac{1-\pi(\Phi_M)}{\Phi_M}$ and rewrite

$$\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M} = \frac{1 - \pi(\Phi_M)}{\Phi_M} \left[u(y - \theta_M) - u(z - \theta_M\Phi_M) + \right. \\ \left. + \theta_M u'(y - \theta_M) - \theta_M\Phi_M u'(z - \theta_M\Phi_M) - \theta_M^2 \Phi_M^2 u''(z - \theta_M\Phi_M)\right].$$

First I will show that there exists a model calibration that makes this expression negative. First observe that, since utility is ordinal, the utility function's units are arbitrary and hence the difference $u(y - \theta_M) - u(z - \theta_M\Phi_M)$ can be assumed to be arbitrarily small. The sign of $\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M}$ therefore depends on the remaining terms. For simplicity, I denote $\tilde{y} = y - \theta_M$ and $\tilde{z} = z - \theta_M\Phi_M$. Therefore

$$\operatorname{sgn}\left\{\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M}\right\} = \operatorname{sgn}\left\{\theta_M[u'(\tilde{y}) - \Phi_M u'(\tilde{z}) - \theta_M\Phi_M^2 u''(\tilde{z})]\right\}.$$

Further, I assume that utility is CARA with risk aversion coefficient $R_A = r$, so that $u(w) = -e^{-rw}$. By factoring out r and θ_M I obtain

$$\operatorname{sgn}\left\{\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M}\right\} = \operatorname{sgn}\left\{e^{-r\tilde{y}} + \Phi_M e^{-r\tilde{z}}[r\theta_M\Phi_M - 1]\right\}.$$

To obtain $\frac{\partial \mathscr{L}}{\partial \Phi}\big|_{\Phi_M} < 0$, the following inequality must hold:

$$e^{-r\tilde{y}} < \Phi_M e^{-r\tilde{z}}[1 - r\Phi_M\theta_M],$$

which reduces to

$$\exp\{-r[y - z - \theta_M(1 - \Phi_M)]\} < \Phi_M[1 - r\Phi_M\theta_M].$$

For this to be true, as a minimum, the right-hand side must be positive. The optimization problem constraint $\theta_M\Phi_M \leq z$ implies that $[1 - r\Phi_M\theta_M]$ is bounded from below by the quantity $1 - rz$, and therefore

$$1 - r\Phi_M\theta_M \geq 1 - rz,$$

so the right-hand side is positive whenever the absolute risk aversion $r < (1/z)$. Now if the left-hand side $\exp\{-r[y - z - \theta_M(1 - \Phi_M)]\}$ is sufficiently small, I will obtain $d\mathscr{L}/d\Phi < 0$. Intuitively, this can be done by making the difference $|y - z|$ sufficiently large. However, notice that the peak volume θ_M and volume-maximizing margin Φ_M are endogenously determined based partly on x, y, z. Therefore, changing $|y - z|$ could also change θ_M and Φ_M and thus offset the increase in $|y - z|$. The technical argument that follows establishes rigorously that z and y can be expanded around x in such a way that Φ_M remains constant, while θ_M is bounded, so that the expression $y - z - \theta_M(1 - \Phi_M)$ can be made arbitrarily large. Since this argument is highly technical, it is presented as a separate lemma below and can be omitted by non-technical readers without loss of understanding.

Finally, since $e^{-r[y-z-\theta_M(1-\Phi_M)]}$ is continuous and approaches 0 at infinity, continuity implies that there exists a large enough $y - z$ so that

$$e^{-r[y-z-\theta_M(1-\Phi_M)]} < \Phi_M[1 - r\Phi_M\theta_M] \text{ and therefore } \left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M} \Rightarrow \Phi_B < \Phi_M.$$

b) To find a calibration for which $\frac{\partial \mathscr{L}}{\partial \Phi}\big|_{\Phi_M} < 0$, observe that it is sufficient to calibrate CARA utility's risk-aversion coefficient to

$$r > \frac{1}{z}.$$

This calibration implies $\Phi_M e^{-r\tilde{z}}[r\theta_M\Phi_M - 1] > 0$ and therefore

$$\operatorname{sgn}\left\{\left.\frac{\partial \mathscr{L}}{\partial \Phi}\right|_{\Phi_M}\right\} = \operatorname{sgn}\left\{\underbrace{e^{-r\tilde{y}}}_{+} + \underbrace{\Phi_M e^{-r\tilde{z}}[r\theta_M\Phi_M - 1]}_{+}\right\} > 0,$$

which in turn implies $\Phi_B > \Phi_M$. $\square$

Lemma 2. *Existence of a Φ_M-preserving spread of the distribution (x, y, z).*

I prove the Lemma in a sequence of two steps, recalling the notation $\tilde{y} \equiv y - \theta_M$ and $\tilde{z} \equiv z - \theta_M \Phi_M$. In step 1, I show that increases in y result in increases in $\tilde{y}$, and reductions in z result in reductions in $\tilde{z}$ so that one can expand the distance $\tilde{y} - \tilde{z}$ to any desired value. In step 2, I show that it is possible to increase $\tilde{y}$ while reducing $\tilde{z}$ in such a way as to maintain Φ_M constant. This establishes that there is a way to expand the spread of y and z around x without changing Φ_M.

Step 1. From the first-order condition $G = 0$ it is easy to show that, via application of the implicit function theorem,

$$\frac{d\theta}{dy} = \frac{\frac{\pi}{2}u''(y-\theta)}{\frac{K}{2}u''(x+\theta K)K + \frac{\pi}{2}u''(y-\theta) + \frac{1-\pi}{2}\Phi^2 u''(z-\theta\Phi)} \in (0,1)$$

and

$$\frac{d\theta}{dz} = \frac{\frac{1-\pi}{2}\Phi u''(z-\theta\Phi)}{\frac{K}{2}u''(x+\theta K)K + \frac{\pi}{2}u''(y-\theta) + \frac{1-\pi}{2}\Phi^2 u''(z-\theta\Phi)} \in (0,1).$$

This implies that increasing y by δ causes θ_M to grow, but by less than δ, while reducing z by ε causes θ_M to decrease, but by an amount less than ε. The existence of the derivatives $\frac{d\theta}{dy}$ and $\frac{d\theta}{dz}$ implies that, as a function of y and z, θ_M is continuous. Therefore, through sequential increases in y by the amounts $\delta_1, \delta_2, \ldots$ coupled with sequential reductions in z by $\varepsilon_1, \varepsilon_2, \ldots$, one can expand the spread of $\tilde{y}$ and $\tilde{z}$ to any desired values $\bar{y} > \bar{z}$.

Step 2. In order to maintain Φ_M fixed, $\partial G/\partial\Phi = 0$ must continue to hold when I increase y and reduce z. Therefore,

$$-\frac{\pi'}{2}u'(y+\delta-\theta) + \frac{\Phi\pi' - (1-\pi)}{2}u'(z-\varepsilon-\theta\Phi) + \frac{1-\pi}{2}\theta\Phi u''(z-\varepsilon-\theta\Phi) =$$
$$= \frac{\pi'}{2}u'(y-\theta) + \frac{\Phi\pi' - (1-\pi)}{2}u'(z-\theta\Phi) + \frac{1-\pi}{2}\theta\Phi u''(z-\theta\Phi).$$

To simplify things, let's select a calibration of $\pi(\Phi)$ such that $\pi'(\Phi_M) = [1-\pi(\Phi_M)]/\Phi_M$, where Φ_M is the volume-maximizing level of margin set by the starting calibration selected. By the mean value theorem, such a calibration always exists, since π is twice-differentiable and the continuous first derivative π' changes from infinity to zero. The middle term then disappears and one only needs to compare how small changes in $\tilde{y}$ and $\tilde{z}$ affect

$$\frac{-\pi'}{2}u'(\tilde{y}+\delta) \text{ vs. } \frac{1-\pi}{2}\theta\Phi u''(\tilde{z}-\varepsilon),$$

where the functions π and π' are evaluated at $\Phi = \Phi_M$, and likewise $\theta = \theta_M$ and $\Phi = \Phi_M$. Since ε and δ vary independently and u' and u'' are continuous, given an $\varepsilon > 0$, it is always possible to select a $\delta > 0$ such that

$$\frac{-\pi'}{2} u'(\tilde{y} + \delta) = \frac{1 - \pi}{2} \theta \Phi u''(\tilde{z} - \varepsilon).$$

This implies that the equation $\partial G / \partial \Phi = 0$, which determines Φ_M, continues to hold, and therefore its solution Φ_M remains unchanged.

This completes the construction a Φ_M-preserving spread of y and z around x. $\square$

Proposition 3. *If the following conditions hold:*

1) *Risk aversion is sufficiently high so that $R_A > 1/z$.*

2) *The social benefit from trade exceeds the expected social cost of default on a per-dollar basis, so that $b > c(1 - \pi)(1 - \Phi)$.*

Then there exists a sufficiently large value of $b > 0$ for which $\Phi_B > \Phi_S$.

Proof. The question of interest is when $\Phi_B > \Phi_S$, which is identified by the condition $\frac{dW}{d\Phi}\big|_{\Phi_B} < 0$. Recall that

$$W = N\mathbb{E}u(\Phi) + \theta(\Phi)[b - c(1 - \pi(\Phi))(1 - \Phi)].$$

Since at Φ_B, $\frac{d\mathbb{E}u}{d\Phi}\big|_{\Phi_B} = 0$, the derivative of the social welfare function evaluated at Φ_B equals

$$\left.\frac{dW}{d\Phi}\right|_{\Phi_B} = \theta'(\Phi_B)[b - c(1 - \pi(\Phi_B))(1 - \Phi_B)] + c\,\theta(\Phi_B)[\pi'(\Phi_B)(1 - \Phi_B) + 1 - \pi(\Phi_B)].$$

For the condition $\frac{dW}{d\Phi}\big|_{\Phi_B} < 0$ to hold, it must be the case that

$$\theta'(\Phi_B)[b - c(1 - \pi(\Phi_B))(1 - \Phi_B)] < -c\,\theta(\Phi_B)[\pi'(\Phi_B)(1 - \Phi_B) + 1 - \pi(\Phi_B)].$$

Recall from Proposition 2 that when $R_A > 1/z$, the privately optimal margin is above the margin that is generating maximum volume ($\Phi_B > \Phi_M$), so that $\theta'(\Phi_B) < 0$. Since the term on the right-hand side is negative, the inequality can hold only if $b - c(1 - \pi(\Phi_B))(1 - \Phi_B) > 0$. Having acknowledged that, observe that the left-hand side expression is a continuous function of b, while b does not enter the right-hand side; therefore, it is always possible to select a value of b large enough so that the above inequality holds. This implies that $\frac{dW}{d\Phi}\big|_{\Phi_B} < 0$

and therefore $\Phi_B > \Phi_S$. $\square$

Proposition 4. *Increasing the default penalty λ will motivate member banks to agree to margin relief (reduce Φ_B) provided that*

$$\underline{\pi}'(\lambda) \geq \frac{1}{\lambda}.$$

Proof. Consider the last term of Bank 1's utility function, the utility cost associated with default:

$$\frac{1-\pi(\Phi)}{2}\lambda(1-\Phi)\theta.$$

For the bank to agree to margin relief when λ increases, it must maintain the same or higher utility. When default penalties do not act as a shifter of the default probability ($\underline{\pi}$ and $\bar{\pi}$ are constants), the utility term above suggests only two possible reactions by the bank to an increase in λ, holding all else equal: 1) reduce trading positions θ; or 2) increase margin Φ. Clearly, banks will agree to margin relief only at the expense of reducing trading, which is unlikely to serve the regulator's policy goal.

However, when default penalties reduce the default probability independent of margin, so that $\pi(\Phi) = v(\Phi) + \underline{\pi}(\lambda)$, observe that the utility loss associated with default becomes

$$\frac{1-v(\Phi)-\underline{\pi}(\lambda)}{2}\lambda(1-\Phi)\theta,$$

so $-\underline{\pi}(\lambda)$ falls and the bank can afford some reduction in Φ without losing utility or reducing its trading position. The exact extent of the reduction depends on the elasticity of $\underline{\pi}(\lambda)$ with respect to λ. Formally, let $A \equiv [1-v(\Phi)-\underline{\pi}(\lambda)]/2$; then for banks to agree to margin relief, it must be the case that $\%\Delta(A\lambda) \leq 0$ given an increase in λ, which implies $\%\Delta A + \%\Delta\lambda \leq 0$ or $\epsilon_{A,\lambda} \leq -1$. Using the fact that $\%\Delta A = dA/A$ and $\%\Delta\lambda = d\lambda/\lambda$, the previous inequality implies

$$\frac{\underline{\pi}'(\lambda)d\lambda}{1-v(\Phi)-\underline{\pi}(\lambda)} \geq \frac{d\lambda}{\lambda},$$

which is equivalent to

$$\underline{\pi}'(\lambda) \geq \frac{1-v(\Phi)-\underline{\pi}(\lambda)}{\lambda}.$$

Taking into account that the numerator $1-v(\Phi)-\underline{\pi}(\lambda) < 1$, we conclude that banks will not lose utility and will agree to margin relief, coupled with an increase in default penalties, whenever

$$\underline{\pi}'(\lambda) \geq \frac{1}{\lambda}.$$

References

[1] Brunnermeier, Markus and Hasse Pedersen. (2009.) "Market Liquidity and Funding Liquidity," *Review of Financial Studies*, 22(6).

[2] Carney, John. (2008.) "Blood on the Street: JP Morgan Pushed Merrill Lynch into Bank of America Merger," *Business Insider*.

[3] Chande, Nikil and Nicholas Labelle. (2013.) "Proposal to Address the Procyclicality of Margins at Central Counterparties." Bank of Canada, mimeo.

[4] Committee on the Global Financial System. (2010.) "The Role of Margin Requirements and Haircuts in Procyclicality," CGFS Paper, 36.

[5] Committee on Payment and Settlement Systems and Technical Committee of the International Organization of Securities Commissions. (2012.) *Principles for Financial Market Infrastructures*. Bank of International Settlements and International Organization of Securities Commissions: Basel, Switzerland.

[6] Diamond, Douglas. (1984.) "Financial intermediation and delegated monitoring," *Review of Economic Studies*, 51(3), 393-414.

[7] Dubey, Pradeep, John Geanakoplos, and Martin Shubik. (2005.) "Default and punishment in general equilibrium," *Econometrica*, 73(1), 1-37.

[8] Gibson, Rajna and Carsten Murawski. (2013.) "Margining in derivatives markets and the stability of the banking sector," *Journal of Banking & Finance*, 37, 1119-1132.

[9] Haene, Philip and Andy Sturm (2009.) "Optimal Central Counterparty Risk Management," Swiss National Bank Working Paper 2009-7.

[10] Koeppl, Thorsten (2013.) "The Limits of Central Counterparty Clearing: Collusive Moral Hazard and Market Liquidity," Queen's University, working paper.

[11] Moser, James T. (1994.) "Origins of the modern exchange clearinghouse: a history of early clearing and settlement methods at futures exchanges." Working paper, Federal Reserve Bank of Chicago.

[12] Nahai-Williamson, Paul, Tomohiro Ota, Mathieu Vital, and Anne Wetherlit (2013.) "Central Counterparties and their Financial Resources - a numerical approach." Bank of England Financial Stability Paper No. 19, April 2013.

[13] Santos, Tano and Jose Scheinkman. (2001.) "Competition among Exchanges," *Quarterly Journal of Economics*, 116(3), 1027-1061.

[14] Tucker, Paul. (2013.) "Central counterparties in evolving capital markets: safety, recovery and resolution," *Banque de France Financial Stability Review*, 17.