Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Financial Stability 2 (2006) 173–193
Journal of FINANCIAL STABILITY
www.elsevier.com/locate/jfstabil

Credit risk transfer and financial sector stability

Wolf Wagner [a,*], Ian W. Marsh [b]

[a] *University of Cambridge and Tilburg University, Cambridge Endowment for Research in Finance (CERF), Judge Business School, Trumpington Street, University of Cambridge, Cambridge CB2 1AG, UK*
[b] *Cass Business School, 106 Bunhill Row, London EC1Y 8TZ, UK*

Received 11 May 2005; received in revised form 9 November 2005; accepted 11 November 2005
Available online 27 June 2006

Abstract

In this paper, we study credit risk transfer (CRT) in an economy with endogenous financing (by both banks and non-bank institutions). Our analysis suggests that the incentive of banks to transfer credit risk is aligned with the regulatory objective of improving stability, and so the recent development of credit derivative instruments is to be welcomed. Moreover, we find the transfer of credit risk from banks to non-banks to be more beneficial than CRT within the banking sector. Intuitively, this is because it allows for the shedding of aggregate risk which must otherwise remain within the relatively more fragile banking sector. Therefore, regulators should act to maximize the benefits from CRT by encouraging the development of instruments favorable to the cross-sectoral transfer of aggregate credit risk (including basket credit derivatives such as collateralized debt obligations). Finally, we derive the optimal regulatory stance for banks relative to non-bank financial institutions. We show that a level playing field approach is sub-optimal. Regulatory stances should be set to actively encourage cross-sector CRT, first because of the higher fragility of the banking sector and second to induce banks to incur the costs of CRT which otherwise lead them to undertake an insufficient amount of CRT.

JEL classification: E44; G21; G22; G28

Keywords: Credit risk transfer; Credit derivatives; Financial stability; Regulation

1. Introduction

Because of its inherent illiquidity, a bank loan was typically held on banks' balance sheets until maturity or default. A few credit risk management techniques were available. For example,

* Corresponding author. Tel.: +44 1223 760585; fax: +44 1223 339701.
E-mail addresses: w.wagner@cerf.cam.ac.uk, wagner@uvt.nl (W. Wagner), i.marsh@city.ac.uk (I.W. Marsh).

1572-3089/$ – see front matter
doi:10.1016/j.jfs.2005.11.001

a small loan sale market existed, and banks could securitize part of their portfolios. However, this was largely limited to non-commercial loans such as mortgages and credit card receivables.

More recently, the introduction of credit derivative markets has given banks a new risk management tool. These markets have shown a rapid growth. First traded in 1996, the outstanding value of credit derivatives is expected to reach US$ 7000 billions by the end of 2005 and growth is likely to continue at a similar pace.[1]

This tremendous development in credit risk transfer (CRT) markets has received the attention of policy makers worldwide. Most supranational and several national supervisors have issued reports on the topic (e.g., BBA, 2004; FSA, 2002; IMF, 2002; OECD, 2002; IAIS, 2003; Basel Committee on Banking Supervision, 2004). These reports are rather similar in tone. On one hand, they emphasize that credit risk transfer brings about benefits, in particular diversification gains. On the other, one can read a common concern that CRT may cause problems for the stability of the financial sector, for example, by destabilizing the institutions which buy credit risk or because diversification encourages the financial sector to take on new risks. Most of the arguments, however, are made on an informal basis, which is due to a lack of theoretical work on these issues.

This paper is a step towards filling this gap. Its first objective is to provide a framework which allows an analysis of the impact of developments in CRT markets on the stability of the financial system. A second objective of the paper is to consider the role of regulation in the market for CRT. We derive conditions under which CRT within and between financial sectors is harmful for stability, and consider the implications for regulation.

It is important to note that our analysis abstracts from a number of potentially important problems that still characterize current (immature) CRT markets and which may affect stability. We do not, for example, analyze regulators' concerns about the concentration of risks due to low numbers of market participants, the lack of transparency which makes it more difficult to evaluate risk exposures, or any mispricing because risk buyers may not fully understand the nature of risks involved.[2] Rather we take a prospective view and study the implications of CRT for when such market imperfections become small.

We develop a simple and tractable model of a financial system which allows us to study the effects of improvements in the efficiency of CRT markets. In this financial system, firms have access to both bank and market financing. There is a trade-off between both forms of financing because of an asymmetric information problem arising from entrepreneurial moral hazard. Bank financing helps to reduce this problem and gives firms 'certification' on the market. However, banks are fragile and need to be compensated for taking up credit risk, thus making bank financing more costly. The stability of the financial system is determined by the riskiness of the portfolios of the banking sector and other financial institutions, which are determined by the (endogenous) degree of financial intermediation and the overall amount of firm financing.

We represent increases in the efficiency of CRT markets by a fall in the (transaction-type) costs of CRT. We first consider a situation where CRT does not increase the fragility of the risk buyer. In this case we find that an increase in CRT efficiency unambiguously improves the stability of the financial system. This is because it facilitates the shifting of risk out of the banking sector into a sector which is not fragile. It is true that banks react to this by increasing their financing

[1] Sources: FitchRatings (2003), BBA (2004) and Basel Committee on Banking Supervision (2004). For an overview of the various credit risk transfer mechanisms see Kiff et al. (2003).

[2] Such issues are common to most developing derivative markets and may not require separate theoretical analysis.

activities and thus their primary risk-taking increases.[3] However, the main result is that all of this additional risk is transferred into the non-bank sector as well. Intuitively, this is because the higher efficiency of CRT reduces the marginal costs of transferring risks for banks, while the marginal costs of retaining risk remain the same. In the extreme case of costless CRT, the economy even achieves its first best outcome and banks become pure ‘originator–distributors’ of loans.

Subsequently, we acknowledge that in reality CRT may increase the fragility of the risk buyer. We find that more efficient CRT may then reduce stability, regardless of whether CRT takes place within the banking sector or across sectors. We argue that CRT across sectors is to be preferred and derive a condition for such CRT to be stability improving. We also address optimal regulation and find that it should encourage regulatory arbitrage across sectors. This is because differences in regulation are necessary to provide financial institutions with the correct incentives to choose a socially optimal allocation of risk across sectors.

As already mentioned, there has been little theoretical work on the impact of CRT on the stability of the financial system. Most contributions so far have focused on the riskiness of a single financial institution. Santomero and Trester (1998) consider innovations which reduce informational asymmetries when selling assets in a crisis. They find that such innovations, which can be interpreted as sophisticated CRT instruments, lead to increased bank risk-taking. Instefjord (2005) analyzes a bank which has access to credit derivatives for risk management purposes. He finds that innovations in credit derivatives markets increase the bank’s risk-taking. Wagner (in press) shows that innovations which enhance the liquidation of bank assets can reduce the stability of a bank by encouraging excessive risk-taking. Recent empirical evidence supports such risk-taking effects (e.g., Cebenoyan and Strahan, 2004; Franke and Krahnen, 2005).

There is already a significant literature on the implications of CRT for the efficiency of bank financing, mostly emphasizing informational issues. Duffee and Zhou (2001) show that credit derivatives are better equipped to deal with lemon problems compared to loan sales since they allow for buying protection on a shorter horizon, while informational asymmetries mostly arise in the longer term. Morrison (2005) shows that the introduction of credit derivatives can reduce welfare by undermining the certification value of bank financing, eventually leading to disintermediation.[4] Nicolo’ and Pelizzon (2005) demonstrate how different forms of credit derivatives can be used to signal the quality of bank loans under binding capital requirements. Arping (2005) and Goderis and Wagner (2005) show that CRT instruments can have beneficial implications by improving lenders’ bargaining position in negotiations with borrowers.

Several papers have also analyzed, more generally, the impact of financial innovations on welfare. Most of this work stresses that innovations complete markets and increase spanning (see the surveys in Allen and Gale, 1994; Duffie and Rahi, 1995). One would therefore expect financial innovations to be welfare enhancing. Allen and Gale (2004), in a financial crisis setting, find that the market completion function of financial innovations can indeed be welfare enhancing by increasing risk sharing possibilities. By contrast, Elul (1995) shows that new securities can have an almost arbitrary effect on welfare. Similar results are also obtained in Allen and Gale (1994), where innovations in an environment with short-selling constraints are considered.

[3] Risk-taking by banks may also go up indirectly because banks’ incentives to monitor firms may be reduced by CRT (since CRT partly passes on the benefits from firms’ success to the risk buyer), thus possibly making loans more risky. We have analyzed this channel in a previous version of this paper (Wagner and Marsh, 2004), however, without finding changes to the stability implications of CRT.

[4] We have analyzed similar effects in a previous version of this paper (Wagner and Marsh, 2004), finding that, perhaps surprisingly, a reduction in the certification value does not necessarily undermine intermediation.

Our analysis is based on the literature stressing banks' role as an intermediary which overcomes asymmetric information problems. In Diamond (1984), the reason for the existence of banks arises because it is optimal for depositors to delegate their monitoring activities. Diamond (1991) shows that bank monitoring can act as a substitute for firm reputation and has an efficiency enhancing role in this respect. Our model of financial intermediation is most closely related to the one in Besanko and Kanatas (1993), where entrepreneurs choose a mix of bank and market financing.

The remainder of the paper is organized as follows. In the next section we present our model of the financial system. In Section 3 we analyze the benchmark case of CRT into a sector which is not fragile. Section 4 deals with the impact of CRT when the buyers of risk are also fragile. The final section summarizes and presents the policy implications.

2. The model

Consider a two-period, one good, production economy. There is a continuum of regions of unit mass, indexed by $i \in [0, 1]$. Each region consists of a continuum of entrepreneurs of unit mass with index $j \in [0, 1]$. An entrepreneur j of region i operates a firm with production technology

$$y(k_{i,j}, e_{i,j}, \eta_i) = f(k_{i,j}, e_{i,j})\eta_i \tag{1}$$

where $k_{i,j}$ denotes capital, $e_{i,j}$ the entrepreneurial effort and η_i is a region-specific productivity shock. Capital depreciates completely in the production process. Exercising effort causes private costs $c(e_{i,j})$ to the entrepreneur. Entrepreneurs do not have a capital endowment and live only in the first period.[5]

We make the following (standard) assumptions about the production function $f(k, e)$ and the cost function $c(e)$ to ensure interior solutions: $f_k > 0$ and $f_e > 0$ (both inputs are productive), $f_{kk} < 0$ and $f_{ee} < 0$ (there are decreasing marginal returns), $f(k, 0) = \lim_{e\downarrow 0} f(k, e) = 0$ (effort is essential in the production process), $c_e > 0$ (exercising effort is costly) and $c_{ee} \geq 0$ (the marginal costs of effort are non-decreasing). Furthermore, we assume $f_{ek} > 0$ and $f_{ke} > 0$ (capital and effort are complements in production). For some of the results in the paper we need to assume more structure, in particular that $f = k^{\alpha}e^{1-\alpha}$ and $c_{ee} = 0$.

The productivity shock η_i has an idiosyncratic (region-specific) component ϕ_i and an aggregate component ϕ_w. We assume that $\eta_i = 1 + \phi_i + \phi_w$ with $E[\phi_i] = E[\phi_w] = 0$, and thus we have $E[\eta_i] = 1$. The ϕ_i's are identically and independently distributed across firms and are uncorrelated with the aggregate shock ϕ_w. We denote the variances of the shocks with $\sigma_i^2 = \text{var}(\phi_i)$, $\sigma_w^2 = \text{var}(\phi_w)$ and $\sigma^2 = \text{var}(\eta_i) = \sigma_i^2 + \sigma_w^2$.

Besides entrepreneurs, there are also investors in the economy. Investors are risk-neutral and are endowed with capital but have no productive opportunities. They have a storage technology available which secures them a (gross) return of 1. There is no direct assumption on the ratio of investors to entrepreneurs but it is assumed that the total amount of capital held by investors is sufficiently large to ensure that the economy is never capital-constrained, i.e., all worthwhile investments can be financed. Besides investing in the storage technology, investors can invest in banks (B) and non-bank financial institutions (NB). Bs and NBs in turn invest in the firms. While NBs have a pure channeling function, banks have additionally a monitoring technology.

[5] This assumption serves to simplify the analysis by ensuring that the entrepreneur sells his firm completely.

This technology enables them to (costlessly) observe entrepreneurial effort.[6] However, they cannot communicate an entrepreneur's effort choice to NBs. Furthermore, Bs are assumed to be regionally specialized and can only invest in firms of their region. For notational simplicity, NBs are also assigned to a region but they can invest across regions. We assume that this happens by NBs first investing in their respective regions and then diversifying risk among themselves. Each region has thus one B and one NB, which we accordingly index with $i \in [0, 1]$. We assume, moreover, that Bs and NBs behave competitively.

The timing in the economy is as follows. At $t = 0$, entrepreneurs decide on the production inputs in their firm and on the share of their firm sold to Bs and NBs of their region, and consume the revenues. Bs and NBs raise capital from the investors to finance purchases of the firms. At $t = 1$, the productivity shock realizes, production takes place and investors consume the proceeds of their investment.

More formally, at $t = 0$, after having decided upon $k_{i,j}$ and $e_{i,j}$, an entrepreneur j of region i sells claims to his firm's output to the B and the NB of his region.[7] We denote with $b_{i,j}$ $[1 - b_{i,j}]$ the fraction of the firm's output sold to Bs [NBs], with $w^{\rm B}_{i,j}$ $[w^{\rm N}_{i,j}]$ the amount of (expected) production sold to Bs [NBs] and with $V^{\rm B}_{i,j}(w^{\rm B}_{i,j})$ $[V^{\rm N}_{i,j}(w^{\rm N}_{i,j})]$ the revenue from selling production to Bs [NBs]. The entrepreneur's optimization problem can then be written as

$$\begin{aligned} &\max\nolimits_{e_{i,j},k_{i,j},b_{i,j}} U(e_{i,j}, k_{i,j}, b_{i,j}) = V^{\rm B}_{i,j}(w^{\rm B}_{i,j}) + V^{\rm N}_{i,j}(w^{\rm N}_{i,j}) - k_{i,j} - c(e_{i,j}), \text{ s.t.} \\ &(i)\, w^{\rm B}_{i,j} = b_{i,j} f(k_{i,j}, e_{i,j}) \\ &(ii)\, w^{\rm N}_{i,j} = (1 - b_{i,j}) f(k_{i,j}, \tilde{e}(k_{i,j}, b_{i,j})) \end{aligned} \tag{2}$$

where the entrepreneur's utility $U(e_{i,j}, k_{i,j}, b_{i,j})$ consists of consumption (equal to the revenues from selling the firm $V^{\rm B}_{i,j}(w^{\rm B}_{i,j}) + V^{\rm N}_{i,j}(w^{\rm N}_{i,j})$ net of investment $k_{i,j}$) minus the cost of effort $c(e_{i,j})$. The crucial difference between selling to Bs and to NBs is that when selling to a B, the entrepreneur is remunerated according to the actual effort chosen $e_{i,j}$ (since Bs can observe his effort choice), hence $w^{\rm B}_{i,j} = w^{\rm B}_{i,j}(e_{i,j})$. By contrast, when selling to NBs his remuneration will be according to the NB's expectation of his effort choice $\tilde{e}(k_{i,j}, b_{i,j})$, which is based on the entrepreneur's choice of $k_{i,j}$ and $b_{i,j}$, hence $w^{\rm N}_{i,j} = w^{\rm N}_{i,j}(\tilde{e}_{i,j})$. Since NBs are rational, they will correctly anticipate the entrepreneur's equilibrium effort choice, i.e., $\tilde{e}(k_{i,j}, b_{i,j})$ solves

$$\begin{aligned} &\tilde{e}(k_{i,j}, b_{i,j}) = \arg\max\nolimits_{e_{i,j}} V^{\rm B}_{i,j}(w^{\rm B}_{i,j}) + V^{\rm N}_{i,j}(w^{\rm N}_{i,j}) - k_{i,j} - c(e_{i,j}), \text{ s.t.} \\ &(i)\, w^{\rm B}_{i,j} = b_{i,j} f(k_{i,j}, e_{i,j}) \\ &(ii)\, w^{\rm N}_{i,j} = (1 - b_{i,j}) f(k_{i,j}, \tilde{e}(k_{i,j}, b_{i,j})) \end{aligned} \tag{3}$$

Next we turn to the optimization problem for investors and financial institutions. Investors require an expected return on capital of at least 1, because this is the return on the storage technology. Competitiveness of Bs and NBs implies then that the required expected return on capital for Bs and NBs is 1. Suppose that there is no CRT, i.e., banks cannot transfer the risks from their investments in firms. Then, bank i's (gross) return on its investment is $x^{\rm B}_i := w^{\rm B}_i \eta_i$ where $w^{\rm B}_i := \int_0^1 w^{\rm B}_{i,j}\, {\rm d}j$ is bank i's claim to firm output (analogous for NB i). In the absence of

[6] Costless monitoring for Bs but no monitoring for NBs is obviously an extreme assumption. However, what is essential here is only that Bs have an advantage over NBs in monitoring. Allowing, for example, for fixed monitoring costs would be straightforward as they do not affect marginal investment decisions.

[7] Thus, we focus on equity financing here. However, this is not essential for the results: any form of financing which is risky (in particular risky debt) will induce bank monitoring and thus make bank financing valuable.

fragility in the financial sector this implies that the required (expected) return on investing in a firm of region i is 1 because of the competitive behavior of Bs and NBs. In particular we would then have: $V_{i,j}^{\mathrm{B}}(w_{i,j}^{\mathrm{B}}) = w_{i,j}^{\mathrm{B}}$ and $V_{i,j}^{\mathrm{N}}(w_{i,j}^{\mathrm{N}}) = w_{i,j}^{\mathrm{N}}$.

However, we assume that Bs become bankrupt at $t = 1$ if the loss on their portfolio exceeds a certain threshold (consistent with a value-at-risk constraint). If bankrupt, they incur an exogenous cost. This may be, for example, because of administrative costs or because their portfolio cannot be liquidated at the fair value. It can be shown for a normally distributed portfolio that the expected loss from bankruptcy for a bank can be written as $(\alpha^{\mathrm{B}}/2)(\sigma_i^{\mathrm{B}})^2$ (see Danielsson and Zigrand, 2003), where $(\sigma_i^{\mathrm{B}})^2$ ($(\sigma_i^{\mathrm{B}})^2 = (\int_0^1 w_{i,j}^{\mathrm{B}}\,\mathrm{d}j)^2\sigma^2 = (w_i^{\mathrm{B}})^2\sigma^2$) is the variance of bank i's portfolio and $\alpha^{\mathrm{B}} \geq 0$ is a parameter which measures the expected cost of a unit of bank risk. The expected value of bank i's portfolio is hence $w_i^{\mathrm{B}} - (\alpha^{\mathrm{B}}/2)(\sigma_i^{\mathrm{B}})^2$. Since investors have to be compensated for the expected losses due to bankruptcy we have that bank i's total required return from financing all the firms of its region, $V_i^{\mathrm{B}}(w_i^{\mathrm{B}}) = \int_0^1 V_{i,j}^{\mathrm{B}}(w_{i,j}^{\mathrm{B}})\,\mathrm{d}j$, is

$$V_i^{\mathrm{B}}(w_i^{\mathrm{B}}) = w_i^{\mathrm{B}} - \frac{\alpha^{\mathrm{B}}}{2}(\sigma_i^{\mathrm{B}})^2 \tag{4}$$

Eq. (4) shows that B's return profile is identical to CARA utility with an (absolute) risk-aversion parameter of α and a risk premium of $(\alpha^{\mathrm{B}}/2)(\sigma_i^{\mathrm{B}})^2$. Hence, Bs will behave *as if* they were risk-averse.

Similarly, assuming that there is fragility in the NB sector implies

$$V_i^{\mathrm{N}}(w_i^{\mathrm{N}}) = w_i^{\mathrm{N}} - \frac{\alpha^{\mathrm{N}}}{2}(\sigma_i^{\mathrm{N}})^2 \tag{5}$$

where α^{N} is the expected cost of a unit of NB risk. The only difference to Bs is that because NBs can invest across regions, they can diversify their region-specific risk completely and we thus have that $(\sigma_i^{\mathrm{N}})^2 = (\int_0^1 w_{i,j}^{\mathrm{N}}\,\mathrm{d}j)^2\sigma_w^2 = (w_i^{\mathrm{N}})^2\sigma_w^2$.

Due to symmetry we will have $w_{i,j}^{\mathrm{B}} = w_i^{\mathrm{B}}$ and $w_{i,j}^{\mathrm{N}} = w_i^{\mathrm{N}}$ in equilibrium. It follows that in equilibrium the required return per firm $V_{i,j}^{\mathrm{B}}(w_{i,j}^{\mathrm{B}})$ will be equal to the total required return for the region $V_i^{\mathrm{B}}(w_i^{\mathrm{B}})$ (which equals the average required return because of a mass of unity)

$$\text{for } w_{i,j}^{\mathrm{B}} = w_i^{\mathrm{B}},\, w_{i,j}^{\mathrm{N}} = w_i^{\mathrm{N}} : V_{i,j}^{\mathrm{B}}(w_{i,j}^{\mathrm{B}}) = V_i^{\mathrm{B}}(w_i^{\mathrm{B}}) \text{ and } V_{i,j}^{\mathrm{N}}(w_{i,j}^{\mathrm{N}}) = V_i^{\mathrm{N}}(w_i^{\mathrm{N}}) \tag{6}$$

Competitiveness of financial institutions further ensures that for deviations from equilibrium we have that

$$\frac{\mathrm{d}V_{i,j}^{\mathrm{B}}(w_i^{\mathrm{B}})}{\mathrm{d}w_{i,j}^{\mathrm{B}}} = \frac{\mathrm{d}V_i^{\mathrm{B}}(w_i^{\mathrm{B}})}{\mathrm{d}w_{i,j}^{\mathrm{B}}} \text{ and } \frac{\mathrm{d}V_{i,j}^{\mathrm{N}}(w_i^{\mathrm{N}})}{\mathrm{d}w_{i,j}^{\mathrm{N}}} = \frac{\mathrm{d}V_i^{\mathrm{N}}(w_i^{\mathrm{N}})}{\mathrm{d}w_{i,j}^{\mathrm{N}}} \tag{7}$$

i.e., entrepreneurs are remunerated for additional output according to the risk premium of their region. Note that $\mathrm{d}V_{i,j}^{\mathrm{B}}(w_i^{\mathrm{B}})/\mathrm{d}w_{i,j}^{\mathrm{B}}$ is independent of $w_{i,j}^{\mathrm{B}}$ because a single firm cannot influence the risk premia ($(\alpha^{\mathrm{B}}/2)(\sigma_i^{\mathrm{B}})^2$ and $(\alpha^{\mathrm{N}}/2)(\sigma_i^{\mathrm{N}})^2$) of its region.[8]

Besides causing costs to the financial institution itself, bankruptcy also causes social costs. This may be, for example, due to knock-on effects on other financial institutions, such as through

[8] Formally: $\dfrac{\mathrm{d}^2 V_{i,j}^{\mathrm{B}}(w_i^{\mathrm{B}})}{\mathrm{d}^2 w_{i,j}^{\mathrm{B}}} = \dfrac{\mathrm{d}^2 V_i^{\mathrm{B}}(w_i^{\mathrm{B}})}{\mathrm{d}^2 w_{i,j}^{\mathrm{B}}} = \dfrac{\mathrm{d}^2 V_i^{\mathrm{B}}(w_i^{\mathrm{B}})}{\mathrm{d}^2 w_i^{\mathrm{B}}}\left(\dfrac{\mathrm{d}w_i^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}}\right)^2 = 0$ since $\left(\frac{\mathrm{d}w_i^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}}\right)^2 = 0$.

informational spillovers or counterparty risk, or by leading to a general failure of the payment system. These costs are assumed to be borne by all investors in the economy and simply reduce the return on their investments. They are not internalized by the financial institution and thus create a scope for regulation. Analogous to (4) and (5) we denote the expected social cost of a B failing and an NB failing by $\beta^{\mathrm{B}}(\sigma_i^{\mathrm{B}})^2$ and $\beta^{\mathrm{N}}(\sigma_i^{\mathrm{N}})^2$, where β^{B} and β^{N} measure the social cost of a bankruptcy (expressed in terms of reduced returns on investment). Since ex-ante all Bs are identical and all NBs are identical we have $(\sigma_i^{\mathrm{B}})^2 = (\sigma^{\mathrm{B}})^2$ and $(\sigma_i^{\mathrm{N}})^2 = (\sigma^{\mathrm{N}})^2$, hence we can write the expected social cost from bankruptcies in the financial sector (the social cost of financial instability) as

$$I = \beta^{\mathrm{B}}(\sigma^{\mathrm{B}})^2 + \beta^{\mathrm{N}}(\sigma^{\mathrm{N}})^2 \tag{8}$$

The following definition completes the description of the economy.

Definition 1. An equilibrium allocation in the economy is given by a triple of actions $(k_{i,j}^*, b_{i,j}^*, e_{i,j}^*)$ for each $i, j \in [0, 1]$, such that $(k_{i,j}^*, b_{i,j}^*, e_{i,j}^*)$ solve (2), with $\tilde{e}(k_{i,j}, b_{i,j})$, $V_{i,j}^{\mathrm{B}}(w_{i,j}^{\mathrm{B}})$, $V_{i,j}^{\mathrm{N}}(w_{i,j}^{\mathrm{N}})$ defined through (3), (6) and (7).

Moreover, since entrepreneurs are identical, we can write the average utility of entrepreneurs as $U = U_{i,j}$, which gives us that

Corollary 1. *Welfare in the economy is determined by I (decreasing) and U (increasing).*

Corollary 1 follows because an investor's welfare (which is equal to expected consumption because of risk-neutrality) depends only on the expected social costs in the economy since they earn an expected return of 1 on their investments in the absence of instability, while entrepreneurs' welfare is determined solely by U as defined in (2). We will refer to I in the following as the instability of the financial system and to U as the efficiency of the financial system (since increases in U arise only from efficiency gains in the financial sector which are passed on to the entrepreneurs).

Lemma 1. *The FOCs for $e_{i,j}$, $k_{i,j}$, $b_{i,j}$ in (2) are*

$$e_{i,j}^*: \quad \frac{\mathrm{d}V_{i,j}^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}} b_{i,j} \frac{\partial f_{i,j}}{\partial e_{i,j}} = c'(e_{i,j}) \tag{9}$$

$$k_{i,j}^*: \quad \left[b_{i,j} \frac{\mathrm{d}V_{i,j}^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}} + (1 - b_{i,j}) \frac{\mathrm{d}V_{i,j}^{\mathrm{N}}}{\mathrm{d}w_{i,j}^{\mathrm{N}}} \right] \frac{\partial f_{i,j}}{\partial k_{i,j}} + (1 - b_{i,j}) \frac{\mathrm{d}V_{i,j}^{\mathrm{N}}}{\mathrm{d}w_{i,j}^{\mathrm{N}}} \frac{\partial f_{i,j}}{\partial e_{i,j}} \frac{\partial \tilde{e}_{i,j}}{\partial k_{i,j}} = 1 \tag{10}$$

$$b_{i,j}^*: \quad (1 - b_{i,j}) \frac{\mathrm{d}V_{i,j}^{\mathrm{N}}}{\mathrm{d}w_{i,j}^{\mathrm{N}}} \frac{\partial f_{i,j}}{\partial e_{i,j}} \frac{\partial \tilde{e}_{i,j}}{\partial b_{i,j}} = \left(\frac{\mathrm{d}V_{i,j}^{\mathrm{N}}}{\mathrm{d}w_{i,j}^{\mathrm{N}}} - \frac{\mathrm{d}V_{i,j}^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}} \right) f_{i,j} \tag{11}$$

Proof. Straightforward from (2). □

Lemma 2. *Increases in an entrepreneur's bank financing $b_{i,j}$ and/or capital $k_{i,j}$ increase the entrepreneur's equilibrium effort choice, i.e., $\partial \tilde{e}_{i,j}/\partial b_{i,j} > 0$ and $\partial \tilde{e}_{i,j}/\partial k_{i,j} > 0$.*

Proof. See Appendix A. □

Equilibrium allocations in our economy are in general not efficient. First, this is because agents' optimization ignores the social cost of instability. Second, even in the absence of instability there is an inefficient amount of production. This can be seen from the FOCs of the entrepreneur's problem (which have been written such that marginal benefits on the left hand side are equated with marginal costs on the right). Eq. (9) reveals that chosen effort is inefficiently low (compared to the first best effort choice, which would require $\partial f_{i,j}/\partial e_{i,j} = c'(e_{i,j})$). There are two reasons for this. First, because of the fragility of the banking system, banks require a risk premium for buying firms' output, i.e., $\mathrm{d}V^{\mathrm{B}}_{i,j}/\mathrm{d}w^{\mathrm{B}}_{i,j} < 1$ (follows from (4) and (7) when $\alpha^{\mathrm{B}} > 0$). Second, the entrepreneur is only remunerated according to the bank's stake in the firm, adding another inefficiency when the firm is not fully financed by the bank ($b_{i,j} < 1$). Eq. (10) shows that the capital choice is also not first best (which would require $\partial f_{i,j}/\partial k_{i,j} = 1$ because the marginal cost of capital is foregone consumption). This is because risk premia required by Bs and NBs imply that the average outside value of a unit of the firm, $b_{i,j}(\mathrm{d}V^{\mathrm{B}}_{i,j}/\mathrm{d}w^{\mathrm{B}}_{i,j}) + (1 - b_{i,j})(\mathrm{d}V^{\mathrm{N}}_{i,j}/\mathrm{d}w^{\mathrm{N}}_{i,j})$, is smaller than 1, and thus the return on capital is inefficiently low (first term on the left hand side of (10)). However, there is a potentially compensating effect: a higher amount of capital invested in the firm 'signals' a higher effort choice to the NBs and thus increases the value at which the firm can be sold to NBs (second term). The reason for the positive signalling value of capital is that an increase in $k_{i,j}$ increases the productivity of effort (due to our assumption on the complementarity of inputs), thus increasing the equilibrium effort choice, which in turn increases the effort anticipated by NBs $\tilde{e}_{i,j}$ (Lemma 2).

Eq. (11) determines the extent of bank financing $b_{i,j}$. The right hand side represents the (marginal) cost of additional B-financing (compared to NB-financing): $\mathrm{d}V^{\mathrm{N}}_{i,j}/\mathrm{d}w^{\mathrm{N}}_{i,j} - \mathrm{d}V^{\mathrm{B}}_{i,j}/\mathrm{d}w^{\mathrm{B}}_{i,j}$. The left hand side gives the marginal gains from additional bank financing, which again is due to a signalling effect: an increase in $b_{i,j}$ increases an entrepreneur's remuneration for effort (from (9)) and thus increases equilibrium effort and anticipated effort (Lemma 2). This is the bank certification effect stressed in the intermediation literature: bank financing enhances the outside value of firms since NBs anticipate that because of banks' monitoring there will be a more efficient effort choice by the entrepreneur. Hence, Eq. (11) shows that optimal bank financing trades off the cost of bank financing (compared to NB-financing) with the benefits from bank financing due to the certification effect. Hence, in an equilibrium we have $\mathrm{d}V^{\mathrm{N}}_{i,j}/\mathrm{d}w^{\mathrm{N}}_{i,j} - \mathrm{d}V^{\mathrm{B}}_{i,j}/\mathrm{d}w^{\mathrm{B}}_{i,j} > 0$ (i.e., bank financing is more costly) as the marginal gains from bank financing are positive.

This concludes our exposition of the economy. In the remainder of the paper we analyze CRT by allowing Bs and NBs to trade firm risks. To save notation, we will thereby suppress the region index i and the entrepreneur index j wherever it does not create confusion. CRT has the effect of changing the portfolio variances of Bs and NBs, $(\sigma^{\mathrm{B}})^2$ and $(\sigma^{\mathrm{N}})^2$, and thus has a direct effect on stability (Eq. (8)). Furthermore, the changes in $(\sigma^{\mathrm{B}})^2$ and $(\sigma^{\mathrm{N}})^2$ affect the risk premia required by Bs and NBs (and thus change $\mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}}$ and $\mathrm{d}V^{\mathrm{N}}/\mathrm{d}w^{\mathrm{N}}$), which has an impact on entrepreneurial actions and the efficiency of the financial system. The model thus naturally allows us to study the impact of changes in CRT markets on the performance of the financial system.

3. Benchmark CRT

In this section we consider CRT which does not increase the fragility of the institution buying the risk. This can be done in our framework by studying CRT from Bs to NBs and assuming that the NB sector is not fragile, i.e., $\alpha^{\mathrm{N}} = \beta^{\mathrm{N}} = 0$. Hence, we have

$$V_{\mathrm{N}}(w_{\mathrm{N}}) = w_{\mathrm{N}} \tag{12}$$

$$\mathrm{d}V_{\mathrm{N}}/\mathrm{d}w_{\mathrm{N}} = 1 \tag{13}$$

showing that when the NB sector is not fragile, it does not require any risk premium (this assumption is relaxed in the next section). Moreover, we assume that CRT does not create incentive problems between the risk buyer and the seller. This would be the case if there were no information flow between the origination and portfolio risk management functions of a bank. Alternatively, reputational reasons may suffice. Specifically, we assume that firms are valued in the risk transfer according to their *actual* effort level (while direct financing by the NB is based on *inferred* effort).[9]

Our model of CRT is as follows. After Bs have bought claims to firms (and hence entrepreneurs have set k and e), they can sell these claims to NBs at price p. Selling the claim entails a proportional cost τ ($0 \leq \tau \leq 1$) on the value of the claim transferred, which is borne by the Bs. This cost has a broad interpretation in being related to imperfections in CRT markets (such as because of transaction costs or informational costs). The extreme cases where the asset cannot be sold at all and where the asset can be sold at zero cost can then be represented by $\tau = 1$ and $\tau = 0$, respectively.

B's return at $t = 1$, given that a share q ($0 \leq q \leq 1$) of its exposure $w^{\mathrm{B}}\eta$ is sold, is hence

$$x^{\mathrm{B}} = (1-q)w^{\mathrm{B}}\eta + qp - \tau pq \tag{14}$$

Since NBs are risk-neutral ($\alpha^{\mathrm{N}} = 0$) and behave competitively, and since it is assumed that firms are valued at actual effort, the price the bank is able to obtain is

$$p = w^{\mathrm{B}} = w^{\mathrm{B}}(e, k) \tag{15}$$

B's return can then be written as

$$x^{\mathrm{B}} = (1-q)w^{\mathrm{B}}\eta + qw^{\mathrm{B}} - \tau w^{\mathrm{B}}q \tag{16}$$

and analogous to (4) we have that the bank value of the firm's output is

$$V_{\mathrm{B}} = w^{\mathrm{B}} - \tau w^{\mathrm{B}}q - \frac{\alpha^{\mathrm{B}}}{2}(1-q)^2(w^{\mathrm{B}})^2\sigma^2 \tag{17}$$

Since w^{B} is given when banks choose q, from differentiating (17) with respect to q, the optimal proportion of the bank portfolio sold is

$$q^* = 1 - \frac{\tau}{\alpha^{\mathrm{B}}w^{\mathrm{B}}\sigma^2} \tag{18}$$

Hence, the extent of CRT, q^*, increases as τ falls.[10] For $\tau = 0$, all risk is transferred into the NB sector ($q^* = 1$). Moreover, and intuitively, we have that the extent of CRT increases if the bankruptcy costs of the banking sector α^{B} increase and when the risk the banking sector has sourced ($w^{\mathrm{B}}\sigma^2$) increases. As by Eq. (18) lower CRT costs lead to more CRT, we will from now on use 'increase in CRT' and 'reduction in the cost of CRT' interchangeably.

Inserting (18) into the expression for bank risk after CRT ($= (1-q)^2(w^{\mathrm{B}})^2\sigma^2$) we obtain that the equilibrium bank risk after CRT is

$$(\sigma^{\mathrm{B}})^2 = \frac{\tau^2}{(\alpha^{\mathrm{B}})^2\sigma^2} \tag{19}$$

Hence, banks retain less risk when transferring risk is cheap (τ is low) and when the bankruptcy costs α^{B} are high. Interestingly, the more risky firms are (σ^2 increases), the *less* risk the bank

[9] We analyzed incentive problems arising from CRT in a previous version of this paper (Wagner and Marsh, 2004). However, we found no qualititative changes to the stability results presented here. By contrast, efficiency can fall because banks' reduced incentives to monitor may outweigh the gains of better diversification.

[10] Note that because of our definition of the cost of CRT (Eq. (14)), q is restricted to be non-negative. Eq. (18) then imposes (implicitly) parameter restrictions.

retains. This is because when σ^2 increases, the banks' marginal benefits of CRT (for a given level of risk the bank retains) are higher because now a one unit increase in q achieves a higher risk reduction (while the CRT costs have stayed the same).[11]

Inserting (19) into (17) we find that the value of the firm's output for the bank V_{B} and the marginal bank value of output $\mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}}$ are decreasing in the cost of CRT τ

$$V^{\mathrm{B}} = (1-\tau)w^{\mathrm{B}} + \frac{\tau^2}{(\alpha^{\mathrm{B}})^2\sigma^2} \tag{20}$$

$$\mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}} = 1-\tau \tag{21}$$

From (21) we see that the possibility for Bs to sell risk subsequent to their investment in firms feeds back into the primary market: it reduces banks' required returns on investment (i.e., bank financing becomes cheaper) and hence will affect total bank financing. The reason why banks' required returns depend negatively on CRT costs τ is that lower CRT costs make it optimal for banks to transfer more risk, thus reducing their required risk premia.

Substituting $\mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}}$ and $\mathrm{d}V^{\mathrm{N}}/\mathrm{d}w^{\mathrm{N}}$ (Eqs. (21) and (13)) into the entrepreneur's FOCs (9)–(11) we get

$$e^*: \quad b(1-\tau)f_e = c_e \tag{22}$$

$$k^*: \quad (1-b\tau)f_k + (1-b)f_e\tilde{e}'(k) = 1 \tag{23}$$

$$b^*: \quad (1-b)f_e\tilde{e}'(b) = \tau f \tag{24}$$

Eqs. (22)–(24) enable us to study the impact of an increased efficiency of CRT markets on firms' decisions, where an increased efficiency is interpreted as a reduction in τ. For example, a fall in τ could arise from lower transaction costs due to a deepening of CRT markets.

Lemma 3. *A reduction in the CRT costs τ increases intermediation b and effort e but does not affect investment k. Banks' risk-taking ($bf\eta$) goes up.*

Proof. See Appendix A. □

The intuition for the results in Lemma 3 is as follows. A reduction in τ causes an increase of CRT to NBs and thus reduces Bs' exposure to firm risk. Hence, Bs reduce their risk premium and the marginal bank value of output $\mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}}$ goes up. This leads, first, to an increase in intermediation b because it reduces the marginal bank financing cost relative to NB-financing $\mathrm{d}V^{\mathrm{N}}/\mathrm{d}w^{\mathrm{N}} - \mathrm{d}V^{\mathrm{B}}/\mathrm{d}w^{\mathrm{B}}$ (from Eq. (24)). Second, it leads to a higher remuneration for effort e and thus a higher effort choice (Eq. (22)). Third, it increases the remuneration for capital to the extent that the firm is financed by the bank (Eq. (23)). However, there are two offsetting effects on the remuneration for capital, which come through the increase in intermediation b. First, the increase in b reduces the remuneration for capital since the marginal value of output for Bs is less than for NBs. Second, the increase in b reduces the value of the signalling effect of k, since the latter applies only to the share of the firm sold to NBs. In the proof of Lemma 3 it is shown that for our assumptions on the production function, all three effects on capital exactly offset each other

[11] This result relies on the assumption that CRT costs are proportional to the price p and thus independent of risk. By contrast, if CRT costs depend on risk, they will increase when firm risk rises, making the impact on bank risk after CRT ambiguous.

and thus the total impact of an increase in CRT on k is zero.[12] Finally, banks' risk-taking goes up because both intermediation b and output f go up (the latter because effort has increased).

Remark 1. A reduction in the CRT costs unambiguously raises welfare (i.e., it increases both stability and efficiency).

Proof. 1. Stability: from (8) and $\beta^{\mathrm{N}} = 0$ we have that stability increases if $\partial(\sigma^{\mathrm{B}})^2/\partial(-\tau) < 0$, which follows directly from (19). 2. Efficiency: we have

$$\frac{\mathrm{d}U(b,k,e,\tau)}{\mathrm{d}(-\tau)} \geq \frac{\partial U(b,k,e,\tau)}{\partial(-\tau)} = -\frac{\partial V^{\mathrm{B}}(w^{\mathrm{B}})}{\partial \tau} - \frac{\mathrm{d}V^{\mathrm{N}}}{\mathrm{d}w^{\mathrm{N}}}(1-b) f_e \frac{\partial \tilde{e}}{\partial \tau} > 0$$

where the first inequality follows because adjustments in the entrepreneur's choice variables (b, k, e) will not make him worse off; the second inequality follows from $\partial V^{\mathrm{B}}(w^{\mathrm{B}})/\partial\tau = -w^{\mathrm{B}}$ (from (20)) and $\partial\tilde{e}/\partial\tau \leq 0$ (from Lemma 3). □

The main insight from Remark 1 is that, although banks take up more risk, the stability of the financial system increases. In principal, there are two effects of lower CRT costs on the variance of banks' returns $(\sigma^{\mathrm{B}})^2$ (and thus financial stability as $\beta^{\mathrm{N}} = 0$). First, lower CRT costs increase CRT q and thus reduce $(\sigma^{\mathrm{B}})^2$. Second, bank risk-taking w^{B} goes up, potentially increasing $(\sigma^{\mathrm{B}})^2$. However, Eq. (19) shows that $(\sigma^{\mathrm{B}})^2$ is independent of w^{B}. This implies that any additional risk taken up by the bank is transferred to NBs, and hence the first effect prevails. The intuitive reason for the latter is that because the marginal costs of transferring risks have fallen (while the marginal costs of retaining risk in terms of bankruptcy costs have stayed the same) the bank finds it optimal to hold less risk.

Remark 2. For $\tau = 0$ the economy achieves the first best outcome.

Proof. Efficiency: from (24) we find that the relative marginal cost of bank financing is zero, while the benefits from bank financing are larger than zero, hence intermediation will be at its maximal level: $b = 1$. It then follows from (22) and (23) that $f_e = c_e$ and $f_k = 1$, implying that effort and capital are chosen efficiently. Stability: from (18) and (19) we have that $q^* = 1$ and $(\sigma^{\mathrm{B}})^2 = 0$ and it follows that $I = 0$. □

Remark 2 tells us that when CRT is costless, there are no longer any inefficiencies in the economy. The reason is that monitoring incentives and risk bearing can then be separated without cost. Banks become pure 'originator–distributors' of firm financing: they 'originate' all financing in the economy ($b = 1$) and thus ensure an efficient effort choice but distribute all claims to the NBs ($q = 1$). Thus, all firm risk is held by the institutions which are not fragile.

4. CRT and fragility of the risk buyer

4.1. CRT within the banking sector

The previous section has presumed that risk was transferred into a sector which is not fragile. From a stability perspective, the risk was simply disappearing. In reality, however, stability is likely be affected by increased risk-taking on the part of the risk buyer. We analyze this issue next. We first study CRT within the banking sector, and then CRT across sectors. The scant evidence on

[12] This result is not expected to generalize to other functional forms.

CRT suggests that both forms of CRT are relevant. For example, FitchRatings (2003) reports that there is a significant amount of net shifting of credit risk from the banking sector, most notably into the insurance sector. However, it also reports that around 90% of the outstanding amount of credit derivatives is held by banks.[13]

Since banks are ex-ante identical, incentives for CRT can only arise from diversification of idiosyncratic (region-specific) firm risk. We therefore directly consider CRT that allows banks to swap their regional claims against aggregate firm risk. Such a swap may be offered by an intermediary which buys the regional claims from all banks and then hedges the arising aggregate risk by buying (aggregate) risk protection from the banks. Since holding the assets of all banks creates only aggregate exposure by the law of large numbers, the intermediary can in this way create a riskless position.[14] The intermediary's expected return on one swap is

$$E[w^{\mathrm{B}}(\eta - \phi_w)] = w^{\mathrm{B}}$$

Competitiveness of the intermediary implies therefore that the price of the swap is $p = w^{\mathrm{B}}$.

Denoting with q ($0 \leq q \leq 1$) the fraction of its region-specific risk a B sheds using the swap, B's portfolio at $t = 1$ is given by

$$x^{\mathrm{B}} = w^{\mathrm{B}}\eta - qw^{\mathrm{B}}(\eta - \phi_w) + pq - \tau pq \tag{25}$$

$$= w^{\mathrm{B}}\eta - qw^{\mathrm{B}}\phi_i - \tau w^{\mathrm{B}}q \tag{26}$$

Hence, the bank value of firm output is

$$V_{\mathrm{B}} = w^{\mathrm{B}} - \tau w^{\mathrm{B}}q - \frac{\alpha^{\mathrm{B}}}{2}(w^{\mathrm{B}})^2(\sigma_w^2 + (1-q)^2\sigma_i^2) \tag{27}$$

Analogous to the previous section the optimal q can be derived as

$$q^* = 1 - \frac{\tau}{\alpha^{\mathrm{B}} w^{\mathrm{B}} \sigma_i^2} \tag{28}$$

Comparing (28) with the expression for CRT in the previous section (Eq. (18)) one can see that the extent of CRT is now lower for a given τ as $\sigma_i^2 < \sigma^2$. This is because CRT within the banking sector does not allow Bs to shed aggregate risk, so the benefits from one unit more CRT are now lower. If the costs of both types of CRT are the same, there is hence less CRT. Using (28) the portfolio variance $(\sigma^{\mathrm{B}})^2$ after CRT is

$$(\sigma^{\mathrm{B}})^2 = \frac{\tau^2}{(\alpha^{\mathrm{B}})^2\sigma_i^2} + (w^{\mathrm{B}})^2\sigma_w^2 \tag{29}$$

Bank risk is thus now higher than in the previous section (from comparing Eqs. (29) and (19)). This is for two reasons. First, aggregate risk cannot be shed in the banking sector, which is why bank risk has the additional term $(w^{\mathrm{B}})^2\sigma_w^2$. Second, there is now also less CRT because, as already discussed, the benefits from CRT are lower. Therefore, the first term in (29) is higher than (19) because $\sigma_i^2 < \sigma^2$.

Lemma 4. $w^{\mathrm{B}\prime}(\tau)/w^{\mathrm{B}} < -1$.

Proof. See Appendix A. □

[13] It should be noted that this number certainly overstates the influence of banks because of netting effects.

[14] In practice, such CRT may require two investments for a bank rather than one: selling its own assets (for example, in a loan sale market) and buying aggregate risk (for example, through investment in a credit risk index).

Proposition 1. *A reduction in the cost of CRT within the banking sector increases efficiency but will eventually reduce stability.*

Proof. Efficiency: analogous to the proof of Remark 1. Stability: total derivative of $(\sigma^{\mathrm{B}})^2$ with respect to τ is

$$\frac{\mathrm{d}\sigma_{\mathrm{B}}^2}{\mathrm{d}\tau} = \frac{2\tau}{(\alpha^{\mathrm{B}})^2\sigma_i^2} + 2w^{\mathrm{B}}w^{\mathrm{B}\prime}(\tau)\sigma_w^2 < \frac{2\tau}{(\alpha^{\mathrm{B}})^2\sigma_i^2} - 2(w^{\mathrm{B}})^2\sigma_w^2 \tag{30}$$

where Lemma 4 has been used to substitute for $w^{\mathrm{B}\prime}(\tau)$. If τ becomes sufficiently small, (30) will be negative. □

The efficiency result is obvious from the discussion of Remark 1: a reduction in τ increases CRT and thus reduces the risk premium required by banks. The impact on stability is less straightforward and generally ambiguous. As equation (29) shows, bank risk no longer depends on the region-specific risk of the initial portfolio $(w^{\mathrm{B}})^2\sigma_i^2$ (for similar reasons as in the previous section) but still fully contains the initial aggregate risk $(w^{\mathrm{B}})^2\sigma_w^2$. The latter is obviously because aggregate risk cannot be shed within the banking sector. If additional risk-taking by banks is sufficiently large, the increased aggregate risk may outweigh the stabilizing impact of the diversification of region-specific risk, making CRT destabilizing. Proposition 1 also states that if τ continues falling, stability is eventually reduced. Intuitively, this is because the impact of diversification on stability becomes smaller as τ decreases (and eventually becomes zero), while additional risk-taking does not fall below a certain (positive) level (as Lemma 4 shows).

4.2. CRT across Financial Sectors

We consider now CRT between the B and NB sectors and allow (in contrast to Section 3) for fragility in the NB sector: i.e., $\alpha^{\mathrm{N}} \geq 0$, $\beta^{\mathrm{N}} \geq 0$. Our model for CRT markets is as follows: banks sell their regional claims to an intermediary, which then holds only aggregate exposure and sells this in turn to the NBs.[15] Denoting with q^{B} the fraction of a bank's portfolio sold, we have for B's portfolio at $t = 1$ analogous to the benchmark case

$$x^{\mathrm{B}} = (1 - q^{\mathrm{B}})w^{\mathrm{B}}\eta + q^{\mathrm{B}}p - \tau w^{\mathrm{B}}q^{\mathrm{B}} \tag{31}$$

NB's portfolio consists of their initial holdings of firms (only consisting of the aggregate component, since region risk has been diversified away) plus the claims to firms bought in the CRT (again, containing only aggregate risk) minus the price

$$x^{\mathrm{N}} = (1 - b)f(1 + \phi_w) + w^{\mathrm{B}}(1 + \phi_w)q^{\mathrm{N}} - pq^{\mathrm{N}} = w^{\mathrm{B}}(\frac{1 - b}{b} + q^{\mathrm{N}})(1 + \phi_w) - pq^{\mathrm{N}} \tag{32}$$

where q^{N} refers to the amount of aggregate firm risk bought (expressed as a fraction of a bank's portfolio). Note that the price p is the same as in (31) as the intermediary makes zero profits.

[15] In practice, such CRT may, for example, take place through a simple loan sale market: Bs sell their exposure to regional firms to NBs, whereby NBs diversify their purchases across regions and thus only buy aggregate risk.

Lemma 5. *Equilibrium CRT is given by*

$$q^* = (q^{\mathrm{B}})^* = (q^{\mathrm{N}})^* = \frac{\alpha^{\mathrm{B}}\sigma^2 - \alpha^{\mathrm{N}}\sigma_w^2((1-b)/(b)) - \tau/w^{\mathrm{B}}}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2} \tag{33}$$

and resulting B and NB-risk is

$$(\sigma^{\mathrm{B}})^2 = \left(\frac{\alpha^{\mathrm{N}}\sigma_w^2 f + \tau}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2 \sigma^2 \tag{34}$$

$$(\sigma^{\mathrm{N}})^2 = \left(\frac{\alpha^{\mathrm{B}}\sigma^2 f - \tau}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2 \sigma_w^2 \tag{35}$$

Proof. See Appendix A. □

As with banking sector CRT, a reduction in the CRT costs increases CRT and reduces bank risk, however this comes now at the cost of increased NB risk. How does the extent of CRT compare to CRT within the banking sector? Suppose that there is no aggregate risk ($\sigma_w^2 = 0$). Inserting in (33) gives

$$q^* = \frac{\alpha^{\mathrm{B}}\sigma_i^2 - \tau/w^{\mathrm{B}}}{\alpha^{\mathrm{B}}\sigma_i^2} = 1 - \frac{\tau}{\alpha^{\mathrm{B}} w^{\mathrm{B}}\sigma_i^2}$$

which is identical to banking sector CRT (Eq. (28)). Hence, in the absence of aggregate risk, both transfer types lead to identical amounts of CRT. The intuition is straightforward: both types of CRT allow then only the shedding of region risk.

How do the two types of CRT compare in the presence of aggregate risk? Suppose now that region risk is absent (i.e., $\sigma_i^2 = 0$, hence $\sigma_w^2 = \sigma^2$). CRT within the banking sector is then obviously zero as there is no idiosyncratic risk to be shed (recall that q is restricted to be non-negative in (28)). For cross-sector CRT we obtain from Eq. (33) that

$$q^* = \frac{\alpha^{\mathrm{B}} b - \alpha^{\mathrm{N}}(1-b)}{(\alpha^{\mathrm{N}} + \alpha^{\mathrm{B}})b} - \frac{\tau}{(\alpha^{\mathrm{N}} + \alpha^{\mathrm{B}})\sigma_w^2 w^{\mathrm{B}}}$$

Hence, CRT depends on two terms. First, the sign of the first term depends on $\alpha^{\mathrm{B}} b - \alpha^{\mathrm{N}}(1-b)$, i.e., the *fragility weighted differences in exposures* to firms across sectors. If this difference is positive, there exists an incentive for transferring risk into the NB sector. Still, this does not imply that there will be positive CRT as there are also CRT costs: similar to the previous section, the second term shows that CRT falls with the costs of CRT. CRT will thus only be positive (and thus larger than CRT within the banking sector) if the difference in the fragility weighted exposures is sufficiently large compared to the costs of CRT.

What can be said about the merits of cross-sectoral CRT compared to CRT within the banking sector? Our analysis suggests that cross-sectoral CRT may be preferable. First, it allows for the transfer of risk into the less fragile NB sector. Second, it permits, in contrast to CRT among banks, for the complete shedding of additional risks taken on by banks.[16]

We simplify the analysis from now on by assuming that k and e (and hence f) are constant.

[16] There is a third effect which, however, does not arise in our setup: diversification gains across sectors may be relatively larger since across sectors the types of risks to which institutions are exposed are likely to be more different.

Proposition 2. *For fixed output f, a reduction in the CRT costs across sectors (i) improves stability for all τ if and only if*

$$\frac{\alpha^{\mathrm{N}}}{\alpha^{\mathrm{B}}} > \frac{\beta^{\mathrm{N}}}{\beta^{\mathrm{B}}} \tag{36}$$

(ii) always increases efficiency.

Proof. See Appendix A. □

Proposition 2 shows that for CRT to be stability-improving, it is not sufficient that risk flows from the socially more fragile sector into the less fragile sector (we refer to α and β as private and social fragility, respectively). To demonstrate this, suppose that $\beta^{\mathrm{B}} > \beta^{\mathrm{N}}$, $\alpha^{\mathrm{B}} > \alpha^{\mathrm{N}}$ and $b = 1 - b = 1/2$, i.e., the B sector is socially and privately more fragile than the NB sector and both sectors have the same size. Since $\alpha^{\mathrm{B}} > \alpha^{\mathrm{N}}$, there is a tendency for risk to flow into the NB sector as the fragility weighted exposure is then higher in the banking sector. However, as can be easily checked, condition (36) is not necessarily fulfilled. The reason is that because instability in the financial system is additive in the risks in each sector (Eq. (8)), minimizing instability is similar to standard portfolio optimization with two uncorrelated assets: the variance minimizing combination of both assets involves investment in both assets (though relatively more is invested in the less risky asset). The flow of risk into the NB sector may therefore lead to 'too little' risk in the banking sector.

What is the general intuition behind condition (36)? Ignoring the cost of CRT for the moment, when the ratio of private and social fragilities are identical across sectors (Eq. (36) with an equality sign), the private sector achieves a risk allocation across sectors that is also socially desirable. However, if the ratio of private to social fragility is larger in the NB sector condition ((36) is fulfilled), the NB sector behaves 'too risk-averse' from a social perspective, with the result that there is too much risk in the B sector. A reduction in CRT costs τ increases CRT into the NB sector (Eq. (33)) and is thus socially desirable.

Is condition (36) fulfilled in reality? The answer is not obvious. It is probably uncontroversial to assume that $\alpha^{\mathrm{B}} \geq \alpha^{\mathrm{N}}$ and $\beta^{\mathrm{B}} > \beta^{\mathrm{N}}$, i.e., the banking sector is privately and socially more fragile.[17] However, as mentioned above, this does not imply that condition (36) is fulfilled. Past experiences with banking crises have suggested that their costs can be very large compared to the failure of other institutions. Hence, β^{B} may be substantially larger than β^{N}, while the effective risk aversion in both sectors, α^{B} and α^{N} may not differ substantially. Thus, it is plausible to presume that condition (36) is met.

The previous discussion suggests that CRT is only stability improving because the relative private fragilities do not lead to a socially efficient outcome in the first place. Hence, there is scope for regulation. Regulation can be incorporated into our framework in the form of capital requirements: capital requirements, by forcing banks to hold more capital when they increase their risk, make risk costly for banks and thus affect banks' incentives in the same way as the bankruptcy costs do.[18] We can therefore interpret α^{B} and α^{N} as con-

[17] The empirical observation of net CRT from the banking to the NB sector (e.g., see FitchRatings, 2003) supports $\alpha_{\mathrm{B}} \geq \alpha_{\mathrm{N}}$ and the fact that the banks are typically considered as heavily regulated compared to other financial institutions indirectly supports $\beta_{\mathrm{B}} \geq \beta_{\mathrm{N}}$.

[18] The internal ratings-based approach of the new Basel accord allows banks to apply value-at-risk to determine their capital requirements. Since the functional form used for bankruptcy costs is consistent with value-at-risk (see Section 2), it is thus also consistent with capital requirements.

sisting of two components: private bankruptcy costs and costs arising from capital requirements. Since bankruptcy costs are independent of capital requirements, regulators can therefore, through capital requirements, set the level of α (where tighter capital requirements imply larger α's).

Proposition 3 shows how regulators should set the ratio of α's across sectors, $r = \alpha_R^{\mathrm{N}}/\alpha_R^{\mathrm{B}}$, if they want to maximize stability.

Proposition 3. *For fixed output f, optimal (stability maximizing) regulation r is given by*

$$r^* = \beta^{\mathrm{N}}/\beta^{\mathrm{B}} - \frac{\tau(\sigma_w^2\beta^{\mathrm{N}} + \sigma^2\beta^{\mathrm{B}})}{f(\alpha^{\mathrm{B}}\beta^{\mathrm{B}}\sigma_w^2\sigma^2)} \tag{37}$$

Hence, optimal regulation entails regulatory arbitrage ($\alpha^{\mathrm{N}}/\alpha^{\mathrm{B}} < 1$) *if* $\beta^{\mathrm{N}} < \beta^{\mathrm{B}}$. *Furthermore, optimal regulatory arbitrage exceeds the relative externalities* ($\alpha^{\mathrm{B}}/\alpha^{\mathrm{N}} > \beta^{\mathrm{B}}/\beta^{\mathrm{N}}$).

Proof. See Appendix A. □

Hence, as the previous discussion already suggested, in the absence of CRT costs ($\tau = 0$) we have $r^* = \beta^{\mathrm{N}}/\beta^{\mathrm{B}}$, i.e., optimal regulation sets the ratio of private fragility costs equal to the social costs. However, if CRT is costly, such regulation would lead to insufficient CRT from a social perspective since banks have to incur the costs of CRT and thus transfer less. Optimal, stability maximizing, regulation therefore sets r in excess of the ratio of social fragilities across sectors in order to compensate for banks' insufficient incentives to transfer credit risk in the presence of CRT costs.

In order to obtain these results we have assumed that k and e (and thus f) are exogenous; hence the overall amount of financing by the financial system was given. If f can adjust, then a reduction in the CRT costs increases total financing, as this allows better risk sharing in the financial sector and thus reduces risk premia (as in previous sections). This will enhance efficiency but comes at the cost of more risk in the financial system. In principle, regulation may respond to this by raising capital requirements in both sectors (while retaining regulatory arbitrage given by Proposition 3 to continue to ensure the optimal allocation across sectors) in order to discourage the taking of the additional risk. However, from a welfare perspective it is not clear whether such regulation is warranted as it will also limit the efficiency gains from CRT.

5. Summary and policy implications

This paper has analyzed the impact of an increased efficiency of CRT markets on the stability of the financial sector. Our results suggest that the incentive of banks to transfer credit risk is aligned with the regulatory objective of improving stability. Although banks ignore the social cost of a failure of their institution, due to their private costs of failure, they have an incentive to diversify and to shift risk out of the banking sector. This will increase stability if banks are more fragile than non-banks. Banks' aversion to risk due to their bankruptcy costs also ensures that they have an incentive to shed any additional risk they take up as a consequence of CRT. Improving the opportunities for CRT strengthens these incentives and is consequently socially desirable.

However, certain qualifications are in order. First, CRT within the banking sector may actually reduce stability if sufficient additional risk is taken up as a result of lower risk premia. This is because the aggregate component of the additional risk cannot be shed within the banking sector. This is a serious impediment since a large part of CRT actually seems to take place within the banking sector. Second, the optimality of CRT across sectors depends on the institutions' incentives relative to social incentives. That is, in order for increases in CRT efficiency to be stability improving, the social fragility of the banking sector relative to that of non-banks has to be higher than the relative private fragilities.

Moreover, in the presence of CRT across sectors, regulatory differences across sectors should be encouraged in order better to align private incentives to manage credit risk with the social objective of increased financial sector stability. In particular, a stricter regulation of the banking sector (e.g., higher capital requirements) encourages a socially optimal allocation of risk across sectors. A level playing-field for regulation is therefore likely to be sub-optimal. This runs counter to the harmonization of regulation across sectors which is the expressed goal of several policy makers. For example, the International Association of Insurance Supervisors (IAIS) concludes in their recent report on credit risk transfer: "... regulatory frameworks should ideally result in similar capital charges and risk management requirements for "like" risks" (page 5, IAIS, 2003). A similar opinion was expressed by Howard Davies when he was chairman of the Financial Services Authority (Davies, 2002).

Acknowledgements

We thank Gabriella Chiesa, Mathias Drehmann, Norvald Instefjord, Alan Morrison, Demosthenes Tambakis, and seminar participants at Cambridge University, Lancaster Business School, the Financial Service Authority, Lehman Brothers, Tilburg University, Fitch, the Bank of England, the Financial Intermediation Research Society Conference (Capri) and the CERF Conference on Financial Innovation (Cambridge) for helpful comments. We also thank the editor and a referee for suggestions.

Appendix A

Proof of Lemma 2. Partial differentiation of (9) with respect to $b_{i,j}$ (invoking $c_{ee} = 0$) gives

$$\frac{\mathrm{d}^2 V_{i,j}^{\mathrm{B}}}{\mathrm{d}^2 w_{i,j}^{\mathrm{B}}} f_{i,j} b_{i,j} \frac{\partial f_{i,j}}{\partial e_{i,j}} + \frac{\mathrm{d}V_{i,j}^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}} \frac{\partial f_{i,j}}{\partial e_{i,j}} + \frac{\mathrm{d}V_{i,j}^{\mathrm{B}}}{\mathrm{d}w_{i,j}^{\mathrm{B}}} b_{i,j} \frac{\partial^2 f_{i,j}}{\partial^2 e_{i,j}} \frac{\partial \tilde{e}_{i,j}}{\partial b_{i,j}} = 0.$$

Since $\mathrm{d}^2 V_{i,j}^{\mathrm{B}}/\mathrm{d}^2 w_{i,j}^{\mathrm{B}} = \mathrm{d}^2 V_i^{\mathrm{B}}/\mathrm{d}^2 w_{i,j}^{\mathrm{B}} = 0$ (see footnote 8) we obtain after dividing by $\mathrm{d}V_{i,j}^{\mathrm{B}}/\mathrm{d}w_{i,j}^{\mathrm{B}}$

$$\frac{\partial \tilde{e}_{i,j}}{\partial b_{i,j}} = \frac{\partial f_{i,j}/\partial e_{i,j}}{b_{i,j}(-\partial^2 f_{i,j}/\partial^2 e_{i,j})} > 0 \tag{38}$$

Similarly, partial differentiation of (9) with respect to $k_{i,j}$ gives

$$b_{i,j} \frac{\partial^2 f_{i,j}}{\partial e_{i,j} \partial k_{i,j}} + b_{i,j} \frac{\partial^2 f_{i,j}}{\partial^2 e_{i,j}} \frac{\partial \tilde{e}_{i,j}}{\partial k_{i,j}} = 0,$$

hence

$$\frac{\partial \tilde{e}_{i,j}}{\partial k_{i,j}} = \frac{\partial^2 f_{i,j}/(\partial e_{i,j} \partial k_{i,j})}{-\partial^2 f_{i,j}/\partial^2 e_{i,j}} > 0 \qquad \square \tag{39}$$

Proof of Lemma 3. Assume $c_{ee} = 0$ and $f(k, e) = k^{\alpha} e^{1-\alpha}$.

1. $b'(\tau) < 0$: inserting (38) in (11) and rearranging for f gives

$$\frac{1-b}{b}\frac{f_e}{f}\frac{f_e}{(-f_{ee})} = \tau \tag{40}$$

for $f(k, e) = k^{\alpha}e^{1-\alpha}$ we have

$$\frac{f_e}{f}\frac{f_e}{(-f_{ee})} = \bar{c} = \frac{1-\alpha}{\alpha} \tag{41}$$

inserting in (40) gives

$$b = \frac{\bar{c}}{\bar{c}+\tau} \tag{42}$$

and hence

$$b'(\tau) < 0 \tag{43}$$

2. $e'(\tau) < 0$: inserting (42) in (22) gives

$$\frac{\bar{c}(1-\tau)}{\bar{c}+\tau} f_e = c_e \tag{44}$$

hence

$$e'(\tau) = -\frac{1+\bar{c}}{(\bar{c}+\tau)(1-\tau)}\frac{f_e}{-f_{ee}} > 0 \tag{45}$$

3. $k'(\tau) = 0$: from (39) and $f = k^{\alpha}e^{1-\alpha}$ we get

$$f_e\tilde{e}'(k) = f_e\frac{f_{ek}}{-f_{ee}} = \frac{1-\alpha}{\alpha} f_k = \bar{c} f_k$$

inserting in (23) gives

$$[1 - \tau b + (1-b)\bar{c}]f_k = 1$$

and using (42) to substitute $\bar{c}$ gives

$$[1 - \tau b + (1-b)\tfrac{b\tau}{1-b}]f_k = 1$$
$$f_k = 1$$

hence

$$k'(\tau) = 0$$

4. $w^{\mathrm{B}\prime}(\tau) < 0$ ($w^{\mathrm{B}} = bf$): follows directly from $b'(\tau) < 0$, $e'(\tau) < 0$ and $k'(\tau) = 0$ □

Proof of Lemma 4. The total derivative of w^{B} with respect to τ is given by

$$w^{\mathrm{B}\prime}(\tau) = b'(\tau)f + bf_e e'(\tau) + bf_k k'(\tau)$$

Since $k'(\tau) = 0$ we get after dividing by w^{B}

$$\frac{w^{\mathrm{B}\prime}(\tau)}{w^{\mathrm{B}}} = \frac{b'(\tau)}{b} + \frac{f_e}{f}e'(\tau) \tag{46}$$

From (42) we have

$$\frac{b'(\tau)}{b} = \frac{-1}{\bar{c}+\tau}$$

and using (45) we get

$$\frac{f_e}{f} e'(\tau) = -\frac{(1+\bar{c})\bar{c}}{(\bar{c}+\tau)(1-\tau)}$$

Hence,

$$\frac{w^{\mathrm{B}\prime}(\tau)}{w^{\mathrm{B}}} = \frac{-1}{\bar{c}+\tau} - \frac{(1+\bar{c})\bar{c}}{(\bar{c}+\tau)(1-\tau)} = -\frac{(1-\tau)+(1+\bar{c})\bar{c}}{(\bar{c}+\tau)(1-\tau)} \tag{47}$$

Differentiating (47) with respect to τ gives

$$\partial \frac{w^{\mathrm{B}\prime}(\tau)}{w^{\mathrm{B}}} / \partial\tau < 0$$

Hence,

$$w^{\mathrm{B}\prime}(\tau)/w^{\mathrm{B}} \leq w^{\mathrm{B}\prime}(\tau=0)/w^{\mathrm{B}}(\tau=0) = -(1+\bar{c}) < -1 \qquad \square$$

Proof of Lemma 5. From (31) and (32) we have for B's and NB's portfolio volatility

$$(\sigma^{\mathrm{B}})^2 = (1-q^{\mathrm{B}})^2 (w^{\mathrm{B}})^2 \sigma^2 \tag{48}$$

$$(\sigma^{\mathrm{N}})^2 = \left(\frac{1-b}{b} + q^{\mathrm{N}}\right)^2 (w^{\mathrm{B}})^2 \sigma_w^2 \tag{49}$$

The FOC for q^{B} and q^{N} are then (after arranging for p/w^{B})

$$p/w^{\mathrm{B}} = 1 - \alpha^{\mathrm{B}} w^{\mathrm{B}} [(1-q^{\mathrm{B}})\sigma^2] + \tau \tag{50}$$

$$p/w^{\mathrm{B}} = 1 - \alpha^{\mathrm{N}} w^{\mathrm{B}} \left(\frac{1-b}{b} - q^{\mathrm{N}}\right) \sigma_w^2 \tag{51}$$

Equilibrium in CRT markets requires $q^{\mathrm{B}} = q^{\mathrm{N}} = q$. Using (50) and (51) to substitute p and solving for q gives

$$q = \frac{\alpha^{\mathrm{B}}\sigma^2 - \alpha^{\mathrm{N}}\sigma_w^2 \frac{1-b}{b} - \tau/w^{\mathrm{B}}}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2} \tag{52}$$

Inserting into (48) and (49) yields

$$(\sigma^{\mathrm{B}})^2 = \left(\frac{\alpha^{\mathrm{N}}\sigma_w^2 f + \tau}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2 \sigma^2 \tag{53}$$

$$(\sigma^{\mathrm{N}})^2 = \left(\frac{\alpha^{\mathrm{B}}\sigma^2 f - \tau}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2 \sigma_w^2 \qquad \square \tag{54}$$

Proof of Proposition 2. (i) Stability "if part" : $\alpha^{\mathrm{N}}/\alpha^{\mathrm{B}} > \beta^{\mathrm{N}}/\beta^{\mathrm{B}}$. Differentiating (8) with respect to τ (keeping f constant) gives

$$\frac{\partial I}{\partial \tau} = \left(\frac{\alpha^{\mathrm{N}}\sigma_w^2 + \tau/f}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)\left(\frac{1/f}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)\sigma^2\beta^{\mathrm{B}} - \left(\frac{\alpha^{\mathrm{B}}\sigma^2 - \tau/f}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)\frac{1/f}{\alpha^{\mathrm{N}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\sigma_w^2\beta^{\mathrm{N}}$$
$$= (\alpha^{\mathrm{N}}\beta^{\mathrm{B}} - \alpha^{\mathrm{B}}\beta^{\mathrm{N}})\sigma^2\sigma_w^2 + \tau/f(\sigma^2\beta^{\mathrm{B}} + \sigma_w^2\beta^{\mathrm{N}}) \quad (55)$$

Since $\alpha^{\mathrm{N}}/\alpha^{\mathrm{B}} > \beta^{\mathrm{N}}/\beta^{\mathrm{B}}$, it follows that

$$\frac{\partial I}{\partial \tau} > \tau/f(\sigma^2\beta^{\mathrm{B}} + \sigma_w^2\beta^{\mathrm{N}})$$

thus $\partial I/\partial \tau > 0$ for all τ. "Only if part" : $\partial I/\partial \tau > 0$ for all τ. In particular, for $\tau = 0$ we have

$$(\alpha^{\mathrm{N}}\beta^{\mathrm{B}} - \alpha^{\mathrm{B}}\beta^{\mathrm{N}})\sigma^2\sigma_w^2 > 0$$

from which it follows that $\alpha^{\mathrm{N}}/\alpha^{\mathrm{B}} > \beta^{\mathrm{N}}/\beta^{\mathrm{B}}$. (ii) Efficiency: proof analogous Remark 1. □

Proof of Proposition 3. Substituting $(\sigma^{\mathrm{B}})^2$ and $(\sigma^{\mathrm{N}})^2$ in I (Eq. (8)) using (53) and (54) and substituting α^{N} using $r = \alpha_R^{\mathrm{N}}/\alpha_R^{\mathrm{B}}$ gives

$$I = \left(\frac{r\alpha^{\mathrm{B}}\sigma_w^2 + b\tau/w^{\mathrm{B}}}{r\alpha^{\mathrm{B}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2\sigma^2\beta^{\mathrm{B}} + \left(\frac{\alpha^{\mathrm{B}}\sigma^2 - b\tau/w^{\mathrm{B}}}{r\alpha^{\mathrm{B}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)^2\sigma_w^2\beta^{\mathrm{N}} \quad (56)$$

The FOC for r is (for constant f)

$$0 = 2\left(\frac{r\alpha^{\mathrm{B}}\sigma_w^2 + b\tau/w^{\mathrm{B}}}{r\alpha^{\mathrm{B}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2}\right)\frac{\alpha^{\mathrm{B}}\sigma_w^2(r\alpha^{\mathrm{B}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2) - \alpha^{\mathrm{B}}\sigma_w^2(r\alpha^{\mathrm{B}}\sigma_w^2 + b\tau/w^{\mathrm{B}})}{(r\alpha^{\mathrm{B}}\sigma_w^2 + \alpha^{\mathrm{B}}\sigma^2)^2}\sigma^2\beta^{\mathrm{B}}$$
$$= (r\alpha^{\mathrm{B}}\sigma_w^2 + b\tau/w^{\mathrm{B}})\sigma^2\beta^{\mathrm{B}} - (\alpha^{\mathrm{B}}\sigma^2 - b\tau/w^{\mathrm{B}})\sigma_w^2\beta^{\mathrm{N}} \quad (57)$$

Solving for r gives

$$r = \frac{(\alpha^{\mathrm{B}}\sigma^2 - b\tau/w^{\mathrm{B}})\sigma_w^2\beta^{\mathrm{N}} - (b\tau/w^{\mathrm{B}})\sigma^2\beta^{\mathrm{B}}}{\alpha^{\mathrm{B}}\sigma_w^2\sigma^2\beta^{\mathrm{B}}} = \beta^{\mathrm{N}}/\beta^{\mathrm{B}} - \frac{\tau(\sigma_w^2\beta^{\mathrm{N}} + \sigma^2\beta^{\mathrm{B}})}{f(\alpha^{\mathrm{B}}\beta^{\mathrm{B}}\sigma_w^2\sigma^2)} \quad \square \quad (58)$$

References

Allen, F., Gale, D., 1994. Financial Innovation and Risk Sharing. MIT Press, London.

Allen, F., Gale, D., 2004. Financial intermediaries and markets. Econometrica 72, 1023–1061.

Arping, S., 2005. Credit Protection and Lending Relationships. University of Amsterdam, Mimeo.

Basel Committee on Banking Supervision, October 2004. Credit Risk Transfer.

Besanko, D., Kanatas, G., 1993. Credit market equilibrium with bank monitoring and moral hazard. Rev. Financ. Stud. 6, 213–232.

British Bankers Association, 2004. Credit Derivatives Report 2003/2004.

Cebenoyan, S., Strahan, P.E., 2004. Risk management, capital structure and lending at banks. J. Bank. Finance 28, 19–43.

Danielsson, J., Zigrand, J.P., 2003. What Happens When You Regulate Risk? Evidence from a Simple Equilibrium Model. London School of Economics (Working paper).

Davies, H., January 2002. Rational Expectations—What Should the Market, and Policy Holders, Expect from Insurance Regulation? (Speech given at the AIRMIC Annual Lecture).

Diamond, D.W., 1984. Financial intermediation and delegated monitoring. Rev. Econ. Stud. 51, 393–414.

Diamond, D.W., 1991. Monitoring and reputation: the choice between bank loans and directly placed debt. J. Polit. Econ. 99, 689–721.

Duffee, G.R., Zhou, C., 2001. Credit derivatives in banking: useful tools for managing risk? J. Monetary Econ. 48, 25–54.

Duffie, D., Rahi, R., 1995. Financial market innovation and security design: an introduction. J. Econ. Theory 65, 1–42.

Elul, R., 1995. Welfare effects of financial innovation in incomplete markets economies with several consumption goods. J. Econ. Theory 65, 43–78.

Financial Services Authority, May 2002. Cross-Sector Risk Transfer.

FitchRatings, September 2003. Global Credit Derivatives: A Qualified Success.

Franke, G., Krahnen, J.P., 2005. Default risk sharing between banks and markets: the contribution of collateralized debt obligations. In: Carey, M., Stulz, R.M. (Eds.), The Risks of Financial Institutions, Chicago University Press.

Goderis, B., Wagner, W., 2005. Credit Derivatives and Sovereign Debt Crises. Cambridge University, Mimeo.

IMF, March 2002. Global Financial Stability Report.

Instefjord, N., 2005. Risk and hedging: do credit derivatives increase bank risk? J. Bank. Finance 29, 333–345.

International Association of Insurance Supervisors, March 2003. Credit Risk Transfer Between Insurance, Banking and Other Financial Sectors.

Kiff, J., Michaud, F.-L., Mitchell, J., 2003. An Analytical Review of Credit Risk Transfer Instruments. Banque de France FSR, pp. 106–131.

Morrison, A.D., 2005. Credit derivatives, disintermediation and investment decisions. J. Business 78, 621–648.

Nicolo', A., Pelizzon, L., 2005. Credit Derivatives: Capital Requirements and Strategic Contracting. University of Venice, Mimeo.

OECD, June 2002. Risk Transfer Mechanisms: Converging Insurance, Credit and Capital Markets.

Santomero, A.M., Trester, J.J., 1998. Financial innovation and risk taking. J. Econ. Behav. Organ. 35, 25–37.

Wagner, W. The liquidity of bank assets and banking stability. J. Bank. Finance, in press.

Wagner, W., Marsh, I., 2004. Credit Risk Transfer and Financial Sector Performance. CEPR Discusssion Paper No. 4265.