Bank Diversification: Incentives, Fallacies and a Remedy*

Andreas Blöchlinger†
Zürcher Kantonalbank

Basile Maire‡
Zürcher Kantonalbank, University of Zurich

May 31, 2011

*The content of this paper reflects the personal view of the authors. In particular, it does not necessarily represent the opinion of Zürcher Kantonalbank. We thank the Treasury department of Zürcher Kantonalbank, especially Martin Bardenhewer for valuable discussions. We would like to thank the audience and discussants at the conference of the Swiss Society for Financial Market Research in Zurich 2011, and at the Australasian Finance and Banking Conference in Sydney 2010, for helpful comments and suggestions on a previous version of this paper. Further, we thank Markus Leippold from the University of Zurich for valuable comments and suggestions.

†**Correspondence Information:** Andreas Blöchlinger, Zürcher Kantonalbank, Josefstrasse 222, CH-8010 Zurich, Switzerland; tel: +41 44 292 45 80, `mailto:andreas.bloechlinger@zkb.ch`

‡**Correspondence Information:** Basile Maire, Zürcher Kantonalbank, Josefstrasse 222, CH-8010 Zurich, Switzerland; tel: +41 44 292 45 79, `mailto:basile.maire@zkb.ch`

Bank Diversification: Incentives, Fallacies and a Remedy

Abstract

When bank liabilities are implicitly guaranteed, creditors do not charge the full default premium. The undercharged debt creates an incentive for leverage. A more diversified bank can employ more debt. Since many banks have the incentives for leverage and diversification, they end up owning similar assets creating a system with highly correlated failures: systemic risk is generated despite capital constraints. Larger banks have higher guarantee values, this creates an incentive to grow big. We derive the insurance premium depending on default probability and degree of diversification. As a remedy we propose to charge this premium to correct the wrong incentives.

JEL Classification Codes: G21, G28, G38

Key Words: Systemic Risk, Implicit guarantee, Too big to fail, Incentives

1. Introduction

Governments have rescued failed banks regularly during times when closures could exacerbate losses. We argue that the wrong incentives which emerged due to this implicit government guarantees is one key driver of the crisis of 2007-2009. A bank which is considered too systemic to fail by the market can issue its debt for a discount compared to a non-systemic institution. Brewer and Jagtiani (2007) found that banks are willing to pay a considerable premium to grow too big to fail. Figure 1 shows that from 1990 to 2010 bank asset values increased much more than the asset values of non-financial firms. The increasing gap between the median and the mean for banks suggests that there are increasingly fewer but bigger mega banks.

[Figure 1 about here.]

Another instrument to become more systemic is diversification. Diversifying the asset portfolio provides advantages for at least two reasons. First, a more diversified bank is allowed to employ more debt under a capital constraint. Second, if the individual bank ends up with almost no idiosyncratic risk, the bank is more likely to fail when others fail, and thereby increases the probability for government support in case the bank is in need of it. Therefore, the value of the guarantee increases with diversification allowing the bank to issue even cheaper debt, as we show. A prime example at the micro-level are triple-A rated mortgage backed securities for which banks had to hold almost no capital against. They contained virtually no idiosyncratic risk but a lot of systematic risk. In the crisis, all these securities simultaneously lost value. As we show, the process of diversifying and levering the asset portfolio increases the probability for many simultaneous bank failures. Eventually, the outcome is not as outlined in the following quote:

> "As a consequence of greater diversification of risks and of sources of funds, problems in the financial sector are less likely to intensify shocks hitting the economy and financial market."
>
> Kohn (2005)

Regulators strengthened the capital constraints as a response to the recent crisis. There is however a widespread disagreement on how much capital banks need. Many regulators consider contingent capital to increase the capital reserves. These contingent capital bonds have a forced debt-for-equity conversion whenever a pre-specified threshold of distress is met, see e.g. Doherty and Harrington

(1997). As argued by Acharya, Pedersen, Philippon, and Richardson (2010a), the issuance of contingent capital does not correct the wrong incentives, since there is no link between the bank's own contribution to the aggregate losses and the interest it must pay on the convertible bond. The Basel Committee proposes that big banks hold an additional amount of capital. If this surcharge is set high enough, it offsets the subsidized borrowing costs. However, it is unlikely that a global surcharge for all too big to fail banks will correct the incentives. A large bank which is focused on its home market is likely to be charged too much with this scheme. On the other hand, small banks may be undercharged. Spain's troubled savings banks showed that the failure of many small banks can cause as much trouble as one big bank failure. The remedy we propose is to charge the value of the implicit guarantee ex-ante, in the form of a Pigovian tax. Under this scheme, banks are free to choose their asset size, leverage, and degree of diversification, but they may have to pay a higher insurance premium. The premium increases with asset size in a non-linear way, hence growing too big to fail is expensive. The incentive for diversification would be at the level of non-financial firms, since extensive diversification becomes costly when the banks' returns become too correlated with the market.

Systemic risk is considered as the failure of a significant part of the financial sector, see e.g. Acharya, Pedersen, Philippon, and Richardson (2010b). Traditionally, the emergence of systemic risk has been attributed to cascading defaults, see e.g. Rochet and Tirole (1996), Kiyotaki and Moore (1997), Allen and Gale (2000), or Freixas, Parigi, and Rochet (2000). We focus on systemic risk arising from simultaneous defaults, since default correlation is the outcome of the wrong incentives as we show. An example in the literature is Wagner (2009), who shows that systemic risk can arise from correlated asset returns due to diversification. In our setting, banks diversify by swapping assets in the interbank market. The interbank market integration determines the degree of diversification a bank can reach. We therefore use the terms interbank market integration and diversification interchangeably. Besides the diversification fallacy, there is the fallacy to equate interbank market integration with cost reductions for society. Financial integration is widely perceived to reduce costs. Popov and Ongena (2010) find that the integration of interbank markets results in substantially lower loan rates. We define social costs as the costs the society would face if they paid the insurance fee for the implicit government guarantee. We show that social costs increase with interbank market integration, even though the bank's probability of failure does not increase due to the capital constraint. Popov and Ongena (2010) find that higher interbank market integration leads to over-leveraged firms during the built-up of the recent crisis, a finding which supports our theory. We present a market based way to compute the insurance premia. The study of Acharya, Pedersen, Philippon, and Richardson (2010a) is related to our paper in the approach to charge a tax to decrease systemic risk. They propose to charge a contingent capital insurance. Similar to our proposal, this tax penalizes banks which are likely to need capital in a bad state of the economy, and rewards banks which are less correlated with the market.

The rest of the paper is structured as follows. Section 2 introduces the model, presents the prevailing incentives and fallacies and how they create systemic risk. In Section 3 we show, based on a theoretical example with a large market of small and large banks, why a tax can align the wrong incentives. The last section concludes.

2. The Model

There exists a banking industry with n banks and a government. There is one period. The bank is default if the value of the assets falls below the nominal amount of debt. Historically, governments have rescued troubled banks in times when closures were likely to intensify losses. In our model, the government may rescue the failed banks. In case of rescue, each defaulted bank receives assets with a value equal to the lost asset value, which is the difference of the end of period asset value and the asset value at the beginning of the period. The government rescues the failed banks if the sum of the failed banks' lost assets exceeds a threshold value.

The government imposes a capital constraint on the banks, that is, each bank has to hold enough capital such that the probability for the asset value to fall below the nominal amount of debt does not exceed a certain value. There are local markets and there is an interbank market. The banks extend loans such as mortgages, student loans, consumer credits, or corporate credits in their local markets. In the interbank market, banks can swap assets to diversify their portfolios. Banks maximize their firm values by choosing the nominal amount of debt they issue and the degree of diversification subject to the capital constraint.

2.1. Incentives

In this section we formally define the implicit government guarantee and show that this guarantee creates incentives for leverage, diversification, and a large balance sheet.

Uncertainty is induced by the probability space $(\omega, \mathcal{F}, \mathbb{P})$. The return of bank i's assets is given by the random variable $\tilde{r}_i$:

$$\tilde{r}_i = r_f + \beta_i \tilde{f} + \upsilon_i \tilde{\varepsilon}_i, \tag{2.1}$$

where r_f is the risk-free rate, $\tilde{f}$ is a factor, β_i is bank i's factor loading and ε_i is mean zero diversifiable risk, υ_i is the volatility of bank i's diversifiable risk. We assume a normally distributed factor, and normally distributed diversifiable risk terms:

$$\tilde{f} \sim \mathcal{N}(\lambda, 1) \tag{2.2}$$

$$\tilde{\varepsilon}_i \sim \mathcal{N}(0, 1), \text{ for all } i. \tag{2.3}$$

By definition, factor risk is independent of diversifiable risk. In a well diversified portfolio, the influence of the diversifiable risk terms becomes negligible, that is, the sample mean of the diversifiable risk terms converges in probability to zero:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \upsilon_i \tilde{\varepsilon}_i \xrightarrow{\mathbb{P}} 0. \tag{2.4}$$

For the exposure to the non-diversifiable risk, banks require a risk-premium, λ, in analogy to the arbitrage pricing theory of Ross (1976).

Bank i is default if the asset value falls below the nominal amount of debt, N_i. Hence, bank i's probability of default is:

$$p_i = \mathbb{P}\left\{A_i(1 + \tilde{r}_i) < N_i\right\}, \tag{2.5}$$

where A_i is the value of bank i's assets. Note that we define the default probability as the probability of hitting the default threshold, regardless whether the bank is bailed out or becomes bankrupt. The capital constraint is that bank i must choose the nominal amount of debt issued such that its probability of default does not exceed π, $p_i \leq \pi$. We define the leverage as the nominal amount of debt over assets, $\ell_i = N_i/A_i$. Now, using the normality assumption of the return distribution, the capital constraint is:

$$\ell_i \leq \Phi^{-1}(\pi)\sigma_i + 1 + r_f + \beta_i \lambda, \tag{2.6}$$

where $\Phi^{-1}(\cdot)$ is the inverse of the cumulative normal distribution, and σ_i is bank i's asset return volatility. Hence if the probability of default exceeds π, the bank must choose a lower leverage ℓ_i. We define the loss on assets, κ_i, in the terminology of James (1991), as loss on the value of the assets after default in percentage terms of the asset value A_i. Having an estimate of the recovery rate, ω_i, the loss on assets κ_i is $\kappa_i = 1 - \omega_i \ell_i$. This follows from the definitions of the recovery rate, and the loss on assets: $A_i \kappa_i = (1 - \omega_i) N_i + (A_i - N_i)$. For simplicity, we assume κ_i to be the same for all banks throughout the paper and drop the index i. Now, the loss of bank i is given by

$$\tilde{C}_i = \mathbf{1}_{\{\tilde{r}_i < \ell_i - 1\}} \kappa A_i. \tag{2.7}$$

where $\mathbf{1}_{\{\tilde{r}_i < \ell_i - 1\}}$ is the random indicator which equals one if the the bank is default, zero otherwise. If the government rescues the failed banks, it has to pay κA_i for bank i. The government's policy is to bail out the failed banks if the economy-wide loss,

$$\tilde{K} = \sum_{i=1}^{n} \tilde{C}_i, \tag{2.8}$$

exceeds a threshold K^*. In this bailout scheme the government bails out all failed banks if a significant proportion of the banking system fails. Hence, not only the failure of large, too big to fail banks triggers a government intervention, but also the simultaneous failure of many small banks. Such guarantees have been provided regularly for banks during crisis. This too many to fail guarantee is addressed e.g. by Penati and Protopapadakis (1988), Acharya and Yorulmazer (2007), and Acharya and Yorulmazer (2008).

The value of the guarantee for each bank therefore depends also on the economy-wide loss, $\tilde{K}$, and the threshold c. Brennan (1979), and Stapleton and Subrahmanyam (1984) show that risk-neutral valuation can be applied to value contingent claims that are functions of cash flows and not traded assets.[1] We denote the $\mathbb{P}$-equivalent risk-neutral measure by $\mathbb{Q}$. Now, the value of the guarantee bank i receives is given by the expected value of the bailout costs under the risk-neutral measure:

$$G_i = \mathbb{E}^Q \left[\tilde{C}_i \middle| \tilde{K} > K^* \right] \mathbb{Q} \left\{ \tilde{K} > K^* \right\}, \tag{2.9}$$

We write this expression as the joint probability of default and rescue:

$$G_i = \kappa A_i \mathbb{Q} \left\{ \tilde{r}_i < \ell_i - 1, \tilde{K} > K^* \right\}. \tag{2.10}$$

In analogy to equation (2.5), we define q_i as the risk-neutral probability of default:

$$q_i = \mathbb{Q}\left\{A_i(1+\tilde{r}_i) < N_i\right\}. \tag{2.11}$$

We are now ready to consider the optimization of the firm value each bank faces. The firm value is the value of its financial claims, see Anderson and Sundaresan (1996). We consider the firm value added compared to a purely equity financed bank. Debt provides advantages in terms of taxes and agency costs. We express these advantages as an increasing function $f(.)$ of leverage.[2] Now, the firm value added compared to a purely equity financed bank, relative to the asset value A_i, is given by:

$$\Delta_i = f(\ell_i) - \mathbb{E}^Q[\tilde{C}_i]/A_i + G_i/A_i \tag{2.12}$$

$$= f(\ell_i) - q_i\kappa + \kappa\mathbb{Q}\left\{\tilde{r}_i < \ell_i - 1, \tilde{K} > K^*\right\} \tag{2.13}$$

consisting of the firm value-added due to the advantages of debt $f(\ell_i)$, minus the expected loss on assets, $q_i\kappa$, plus the value of the implicit guarantee relative to the asset value. We obtain the second line using equation (2.7). Banks maximize the firm value added, equation (2.13), by choosing the degree of diversification (or equivalent the asset return volatility σ_i) and leverage, such that the capital constraint is satisfied:

$$\max_{\sigma_i, \ell_i} \Delta_i, \text{ s.t. } p_i \leq \pi. \tag{2.14}$$

Figure 2 plots the firm value added, equation (2.13), as a function of leverage for a higher asset return volatility (left panel), and a lower asset return volatility (right panel).[3] The firm value added of a non-financial firm does not include the government guarantee. Therefore the bank's firm value added lies above the value of the firm value added of a non-financial firm as you can see in Figure 2. Furthermore, due to the implicit guarantee, the bank has a higher optimal leverage than a non-financial firm.

[Figure 2 about here.]

Choosing the leverage, ℓ_i, which maximizes Δ_i is a trade-off. Issuing no debt at all yields no value added, since the value of the advantages of debt is zero ($f(0) = 0$) and the value of the guarantee and bankruptcy costs are both zero because there is no default. Increasing the leverage increases the value of the advantages of debt and the value of the guarantee, but it also increases bankruptcy costs. Finally, the bank is not allowed to increase the leverage above a certain level, since the probability of default has to remain below π, due to the capital constraint. However, if the tax shield is steep enough, the optimal amount of debt of a bank lies above the optimal amount of a non-financial firm, since the bank enjoys the free government guarantee, while a non-financial firm does not. Choosing a low asset return volatility by diversifying is value-adding for two reasons. First, the capital constraint allows for a higher leverage with a lower volatility. Second, the value of the guarantee increases relative to bankruptcy costs as diversification increases, the reason is as follows. By diversifying the asset portfolio, the correlation with the market return increases because idiosyncratic risk is reduced. In consequence the states of the world where the bank is default and bailed out relative to the states of the world where the bank is default increases with diversification, and therefore the value of the guarantee increases relative to bankruptcy costs. If the bank can choose the asset size which maximizes the firm value, then clearly the bank will decide to grow. The reason is that larger banks face a larger value of the implicit guarantee relative to the asset size than smaller banks, since fewer additional banks have to fail in order to trigger the bailout. To present these incentives in mathematical terms, we describe how banks diversify in our model in the next section.

2.2. *Diversification*

In this section, we present our model for diversification that yields returns of the form of equation (2.1). The return of bank i's asset portfolio, generated in its local market, is given by

$$\tilde{a}_i = \beta \tilde{f} + \nu \tilde{\epsilon}_i, \tag{2.15}$$

where $\tilde{f}$ is the factor, $\tilde{\epsilon}_i$ reflects idiosyncratic risk, β is the factor loading of the assets, and ν is the volatility of the asset returns attributable to idiosyncratic risk. The idiosyncratic risk terms are normally distributed with mean zero, and unit volatility:

$$\tilde{\epsilon}_i \overset{iid}{\sim} \mathcal{N}(0, 1). \tag{2.16}$$

By definition, the idiosyncratic risk terms $\tilde{\epsilon}_i$ and the factor are independent. The initial endowment of each bank is the same, $A_i = A_j$, for all i, j. Banks i and j swap the fraction w_{ij} of their assets in the interbank market to diversify. The fraction swapped is symmetric, that is, $w_{ij} = w_{ji}$. Now, the return of bank i's asset portfolio is given by

$$\tilde{r}_i = \sum_{j=1}^{n} w_{ij}\tilde{a}_j \tag{2.17}$$

$$= \beta\tilde{f} + \nu \sum_{j=1}^{n} w_{ij}\tilde{\epsilon}_j. \tag{2.18}$$

The second line, equation (2.18), writes the return in the form of equation (2.1), with $r_f = 0$, $\beta_i = \beta$ for all i, and $\upsilon_i\tilde{\varepsilon}_i = \nu \sum_{j=1}^{n} w_{ij}\tilde{\epsilon}_j$. The diversifiable risk terms satisfy equation (2.4), that is, they converge in probability to zero, we prove this in Appendix A, Lemma A.1

The interbank market is characterized by a distance measure between the banks, and the degree of interbank market integration. Two banks with a small distance ("close banks") swap a larger fraction of their assets than two distant banks. The term "close" may refer to the geographical distance, or to the business connection of the two banks. If the interbank market integration is low, banks only swap assets with close neighbors. If it is high, banks swap a larger amount of assets with distant banks. To represent these characteristics, we resort to the following model for the fraction swapped between bank i and bank j:

$$w_{ij} = \frac{e^{-d(i,j)/\gamma}}{\sum_{k=1}^{n} e^{-d(i,k)/\gamma}}, \tag{2.19}$$

where $d(i,j)$ the distance between bank i, and bank j. We term the natural logarithm of γ the interbank market integration. The distance, $d(i,j)$, is given by the difference of the bank indices:

$$d(i,j) = \min(|i-j|, n-|i-j|), \tag{2.20}$$

where n is the number of banks in the market. Figure 3 illustrates the asset portfolios, determined by equation (2.19), in a market with 4 banks for different degrees of interbank market integration. We see that close banks swap larger fractions, and that with high integration distant banks swap larger fractions than with a low integration.

[Figure 3 about here.]

When swapping assets, the value of the bank's asset portfolio remains constant, since a fraction of the assets worth $A_i w_{ij}$ is traded against a fraction of the same value. The return variance, however, is reduced when swapping assets. Calculating the covariance of bank i and bank j's asset returns yields

$$\mathbb{COV}(\tilde{r}_i, \tilde{r}_j) = \beta^2 + \nu^2 \left(\sum_{k=1}^{n} w_{ik} w_{jk} \right), \text{ for all } i, j. \tag{2.21}$$

Consequently, the volatility is:

$$\sigma_i(\gamma) = \left[\beta^2 + \nu^2 g(\gamma) \right]^{1/2}, \text{ for all } i, \tag{2.22}$$

where $g(\gamma) = \sum_{k=1}^{n} w_{ik}^2$. The function $g(\gamma)$ has the same value for all banks, and is monotonically decreasing in γ.

[Figure 4 about here.]

Figure 4 shows correlations and variances for different degrees of interbank market integrations, as given by equations (2.21) and (2.22), and it illustrates the following Propositions.

Proposition 2.1: *In large markets, banks are able to completely diversify away idiosyncratic risk.*

Proposition 2.2: *By increasing the interbank market integration, banks are able to completely correlate their asset returns.*

We prove the two Propositions in the Appendix A. If banks are, in the limit, able to completely correlate their asset returns by diversification, they maximize the value of the implicit guarantee since all banks fail simultaneously. We address this issue in the next section.

2.3. Fallacies

In this section we show that diversification increases systemic risk. In the second part of this section, we show that the costs for society increase with market integration. Ultimately the tax payers bear the costs to bail out the failed banks. We are therefore interested in the costs the tax payers would face if they insured the banks against bankruptcy. We term these insurance costs the *social costs*.

To assess the systemic risk for different degrees of market integration, we consider a system of 100 equally sized A-rated banks and compute the default probability distribution. We compute the probability distribution for the banking system to illustrate how systemic risk is generated. The literature provides several approaches to measure systemic risk. Acharya, Pedersen, Philippon, and Richardson (2010b) compute an individual bank's contribution to systemic risk by estimating the expected undercapitalization, given the system as a whole is undercapitalized. Adrian and Brunnermeier (2009) propose a measure called CoVaR, based on quantile regression, to determine an individual bank's systemic risk contribution. Huang, Zhou, and Zhu (2009) suggest to use the insurance premium of insuring the a banking system for contingent capital as an indicator for systemic risk. Lehar (2005) proposes, among other measures, to compute the probability of more than a threshold number of bank failures.

All banks in the market have a one-year probability of default of 0.1%. This corresponds to a banking system of A rated banks, see e.g. Jarrow, Lando, and Turnbull (1997), or Crosbie and Bohn (2003). Because an economic crisis is the realization of a systematic shock, we compute the probability of observing m defaults, conditional on the realization of a systematic shock that occurs on average once every 100 years:

$$\mathbb{P}\left\{ \sum_{j=1}^{n} \mathbf{1}_{\{\tilde{r}_i < \ell_i - 1\}} = m \,\middle|\, \tilde{f} < \Phi^{-1}(0.01) \right\}, \tag{2.23}$$

where $\mathbf{1}_{\{\tilde{r}_i < \ell_i - 1\}}$ is the random default indicator which equals one if the bank is default, zero otherwise. The parameters for the asset returns, equation (2.15), are

$$\lambda = 1.5,\ \beta = 1\%,\ \nu = 5\%. \tag{2.24}$$

Hence, we have an expected return on assets of 1.5%, and an asset return volatility of 11.2% without diversification. Figure 5 graphs the conditional probability of default, equation (2.23), for three levels of market integration. First, a banking system with low interbank market integration (top panel, $\ln \gamma = -1$), where banks are weakly diversified. Second, a banking system with an integrated interbank market (middle panel, $\ln \gamma = 0.5$), and third, a high interbank market integration (bottom, $\ln \gamma = 2$). Since with a highly diversified bank portfolio the banks contain almost no idiosyncratic risk, a system of less diversified banks is able to better absorb a systematic shock. The top panel shows that the probability for no bank default in a crisis at the given low level of diversification is 26%, it is 3% for the system of more diversified banks (middle panel), and it is 2% for the system of highly diversified banks (bottom panel). We observe a large probability for just a few bank defaults in the top panel, e.g. the probability for exactly two bank defaults is 33%, while probabilities stay below 0.5% for more than 5 bank defaults. In the middle panel we see that the system with better diversified banks has lower probabilities for small numbers of bank defaults, but higher probabilities for large numbers of bank defaults, e.g. we now have a probability of almost 6% for exactly 10 bank defaults, while this probability is close to zero in the less diversified system. With a probability of 38% there will now be more than 10 defaults. The bottom panel completes the picture: high diversification increases the probability for a large number of bank defaults in an economic crisis. The probability of observing more than 10 defaults now sums up to above ninety percent. Due to the increased correlation of the bank asset returns (as illustrated in Figure 4), there is now a large probability for the whole system to collapse, that is, the probability for 100 bank defaults is 4%. We conclude that systemic risk increases with the interbank market integration.

[Figure 5 about here.]

The social costs an individual bank creates, increases with diversification. Consider a digital option that pays one consumption unit if the bank defaults, zero otherwise. Using risk-neutral valuation, the price of the digital option is given by the risk-neutral probability of default, q_i, equation (2.11). As the market integration increases, a bank can increase its leverage while maintaining the same probability of default. To see this, we resort to the following Lemma.

Lemma 2.1: *Bank i's risk-neutral probability of default is given by*

$$q_i = \Phi\left(\Phi^{-1}(p_i) + \frac{\beta_i}{\sigma_i}\lambda\right), \text{ for all } i, \tag{2.25}$$

where p_i is bank i's probability of default, β_i its factor exposure, σ_i bank i's asset return volatility, and λ the market price of risk.

Proof. See Appendix A. ■

Table 1 shows this process of diversification and leverage. For every degree of market integration, $\ln\gamma$, listed in Table 1, we compute the asset return variance σ_i using equation (2.22) in a market with 100 banks, and evaluate the risk-neutral probability of default, q_i, as given in Lemma 2.1. We use the parameters given in equation (2.24), and a probability of default of $p_i = 0.1\%$ for all i.

[Table 1 about here.]

The risk-neutral probabilities, q_i, increase with market integration. Hence if the bank would insure its debt, the insurance premium increases with market integration. Table 1 also shows the value of the implicit government guarantee, G_i, in percent of the asset value. To compute G_i we evaluate equation (2.10) using Monte Carlo simulations. The bailout is triggered if more than 20% of the equally sized banks default. The value of the implicit government guarantee increases with the interbank market integration. Above all, the value of the guarantee provided for each bank increases with market integration relative to the full insurance premium (see column "Ratio" of Table 1). Hence the costs which are borne by society increase with market integration. These are the numbers for the incentives presented in Section 2.1.

The implicit government guarantee is not charged, hence the government and ultimately the tax payers bear the costs for the guarantee. While each individual bank meets the capital constraint, the social costs increase with market integration.

3. The remedy

In this section we discuss the remedy to align the incentives. The remedy is to charge the value of the implicit guarantee. We show in an illustrative market, how asset size influences the value of the implicit guarantee.

To align the incentives, each bank has to pay the insurance premium for the implicit government guarantee it receives:

$$G_i = \kappa A_i \mathbb{Q}\left\{\tilde{r}_i < \ell_i - 1, \tilde{K} > K^*\right\}, \tag{3.1}$$

To gain some insights into the Pigovian tax, we construct a market where banks can merge. There are n banks and $h \leq n$ bank holdings, that is, each holding consists of one or several banks. The capital constraint must hold for each bank holding. The banks diversify as described in Section 2.2. The return of each bank belongs to its holding, hence the return of a bank holding is the average of its banks' returns. The return volatility of a bank holding is lower than that of a single bank. Therefore, bank holdings are allowed to have a higher leverage by the capital constraint, equation (2.6), than single banks.

We now consider a market with 1000 banks. There are 662 bank holdings, and four different sizes of bank holdings: Two bank holdings consist of 50 banks each, 10 holdings consist of 10 banks each, 150 holdings consist of 2 banks each, and 500 holdings consist of 1 banks each. The first bank holding consists of the banks 1 to 50, the second bank holding of the banks 51 to 100 et cetera. The return variance of bank holding 1 is given by

$$\mathbb{V}\left(\frac{1}{50}\sum_{j=1}^{50}\tilde{r}_j\right) = \beta^2 + \nu^2\frac{1}{50^2}\sum_{k=1}^{n}\left(\sum_{j=1}^{50}w_{jk}\right)^2, \tag{3.2}$$

accordingly for the remaining banks.

We set the degree of market integration equal to $\ln\gamma = 0$. Each bank sets its leverage such that the probability of default equals $p_i = 0.1\%$. We compute the value of an explicit insurance, $q_i\kappa A_i$, as well as the value of the government guarantee, equation (3.1). We set the loss on assets equal to 50% of the asset value A_i. The government bails out the failed banks, if the loss, $\tilde{K}$, exceeds 5% of total assets in the economy. Table 2 shows the results.

[Table 2 about here.]

The bank pays this Pigovian tax ex-ante to the government. To charge this insurance premium is a market based remedy since the banks pay for the insurance they receive. The insurance is not an explicit insurance, hence the bank is not guaranteed to be bailed out in case of failure. Whether the bank is bailed out or not depends on the state of the economy: if the economy-wide loss is large, the government bails out the banks. However, if a bank chooses its exposure such that it tends to fail when others fail, the insurance premium is higher. The premium increases with the size of the bank in a non-linear way as Table 2 demonstrates. Therefore, growing too big to fail is expensive. With the Pigovian tax, diversification still exists. However, the additional frictions imposed by the implicit government guarantee which lead to higher leverage and further diversification cease to exist. Therefore, the tax leads the banking system to a level of diversification of non-financial firms. The rationale for the insurance premium is not to grow reserves for the event of a crisis but to align incentives.

4. Conclusion

The implicit government guarantee banks enjoy imposes an additional friction for the Modigliani and Miller (1958) theorem of capital structure irrelevancy to hold. This results in higher leverage ratios and the incentive to diversify away idiosyncratic risk, ultimately creating systemic risk despite capital constraints on the individual bank. To align the incentives we propose to charge the insurance premium for the implicit guarantee. The tax depends on the asset size, the degree of diversification, and the probability of default. In our proposal, the banks are free to choose their asset size, their leverage, and degree of diversification, but they may have to pay a higher tax.

A. Proofs

Lemma A.1: *If bank i's return is given by*

$$\tilde{r}_i = \beta \tilde{f} + \nu \sum_{j=1}^{n} w_{ij} \tilde{\epsilon}_j, \tag{A.1}$$

as in equation (2.18), *the sample mean of the diversifiable risk terms*

$$\tilde{x}_n = \frac{1}{n} \sum_{i=1}^{n} \nu \sum_{j=1}^{n} w_{ij} \tilde{\epsilon}_j \tag{A.2}$$

converges in probability to zero, for all i.

Proof. We want to show that

$$\lim_{n \to \infty} \mathbb{P} \{ |\tilde{x}_n| \geq b \} = 0, \tag{A.3}$$

for all $b > 0$. First we note that the random variable $\tilde{x}_n$ has mean 0, and variance ν^2/n. Now, recall that Chebishev's inequality states for any $b > 0$ and a random variable $\tilde{y} \in L_2$ with mean μ and variance σ^2, that $\mathbb{P} \{ |\tilde{y} - \mu| \geq b \} \leq \sigma^2/b^2$. Using Chebishev's inequality we have that:

$$\mathbb{P} \{ |\tilde{x}_n| \geq b \} \leq \frac{\nu^2}{n b^2}, \tag{A.4}$$

so that in the limit, $\tilde{x}_n \xrightarrow{\mathbb{P}} 0$, from the definition (A.3). ■

Proposition (2.1): *In large markets, banks are able to completely diversify away idiosyncratic risk.*

Proof. We have that $\lim_{\gamma \to \infty, n \to \infty} g(\gamma) = 0$. Hence the asset return variance of a bank is

$$\lim_{\gamma \to \infty, n \to \infty} \left(\beta^2 + \nu^2 g(\gamma) \right) = \beta^2, \tag{A.5}$$

for all i, which equals the asset return variance of a bank with no idiosyncratic risk. ■

Proposition (2.2): *By increasing the interbank market integration, banks are able to completely correlate their asset returns.*

Proof. Since $\lim_{\gamma \to \infty} w_{ij} = \frac{1}{n}$ for all i, j, the covariance of bank i and bank j's returns

$$\mathbb{COV}(\tilde{r}_i, \tilde{r}_j) = \beta^2 + \nu^2 \left(\sum_{k=1}^{n} w_{ik} w_{jk} \right), \tag{A.6}$$

equals the variance

$$\sigma_i^2 = \beta^2 + \nu^2 \sum_{k=1}^{n} w_{ik}^2, \tag{A.7}$$

for any i, j, in the limit. ■

Lemma (2.1): *Bank i's risk-neutral probability of default is given by*

$$q_i = \Phi\left(\Phi^{-1}(p_i) + \frac{\beta_i}{\sigma_i}\lambda\right), \text{ for all } i, \tag{A.8}$$

where p_i is bank i's probability of default, β_i its factor loading, σ_i bank i's asset return volatility, and λ the market price of risk.

Proof. The risk-neutral probability of default of bank i is $q_i = \mathbb{Q}(\tilde{r}_i < \ell_i - 1)$. Knowing the physical probability of default, $p_i = \mathbb{P}(\tilde{r}_i < \ell_i - 1)$, we can solve for the leverage:

$$\ell_i = \Phi^{-1}(p_i)\,\sigma_i + 1 + r_f + \beta_i\lambda. \tag{A.9}$$

Using ℓ_i and normalizing with the mean $\mathbb{E}^Q[\tilde{r}_i] = r_f$ and variance σ_i, we can write q_i as

$$q_i = \mathbb{Q}\left\{\frac{\tilde{r}_i - r_f}{\sigma_i} < \frac{\Phi^{-1}(p_i)\,\sigma_i + r_f + \beta_i\lambda - r_f}{\sigma_i}\right\} \tag{A.10}$$

$$= \Phi\left(\Phi^{-1}(p_i) + \frac{\beta_i}{\sigma_i}\lambda\right), \tag{A.11}$$

which completes the proof. ■

References

Acharya, V., L. Pedersen, T. Philippon, and M. Richardson (2010a): "A tax on systemic risk," *Working Paper, New York University, Stern School of Business.*

——— (2010b): "Measuring systemic risk," *Working Paper, New York University, Stern School of Business.*

Acharya, V., and T. Yorulmazer (2007): "Too many to fail–An analysis of time-inconsistency in bank closure policies," *Journal of Financial Intermediation*, 16(1), 1–31.

——— (2008): "Cash-in-the-market pricing and optimal resolution of bank failures," *Review of Financial Studies*, 21(6), 2705.

Adrian, T., and M. Brunnermeier (2009): "CoVaR," *Federal Reserve Bank of New York Working Paper, August.*

Allen, F., and D. Gale (2000): "Financial contagion," *Journal of Political Economy*, 108(1), 1–33.

Anderson, R., and S. Sundaresan (1996): "Design and valuation of debt contracts," *Review of Financial Studies*, 9(1), 37.

Brennan, M. (1979): "The pricing of contingent claims in discrete time models," *Journal of Finance*, pp. 53–68.

Brewer, E., and J. Jagtiani (2007): "How much would banks be willing to pay to become" too-big-to-fail" and to capture other benefits?," .

Crosbie, P., and J. Bohn (2003): "Modeling default risk," *Modeling Methodology, Moodys KMV documentation, San Francisco.*

Doherty, N., and S. Harrington (1997): "Managing corporate risk with reverse convertible debt," Discussion paper, Working Paper, Wharton.

Freixas, X., B. Parigi, and J. Rochet (2000): "Systemic risk, interbank relations, and liquidity provision by the central bank," *Journal of Money, Credit and Banking*, 32(3), 611–638.

Huang, X., H. Zhou, and H. Zhu (2009): "A framework for assessing the systemic risk of major financial institutions," *Journal of Banking & Finance*, 33(11), 2036–2049.

James, C. (1991): "The losses realized in bank failures," *The Journal of Finance*, 46(4), 1223–1242.

Jarrow, R., D. Lando, and S. Turnbull (1997): "A Markov model for the term structure of credit risk spreads," *Review of Financial Studies*, 10(2), 481.

Kiyotaki, N., and J. Moore (1997): "Credit cycles," *Journal of Political Economy*, 105(2), 211–248.

Kohn, D. (2005): "Remarks by Governor Donald L. Kohn," Symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, http://www.federalreserve.gov/boarddocs/speeches/2005/20050827/default.htm.

Lehar, A. (2005): "Measuring systemic risk: A risk management approach," *Journal of Banking and Finance*, 29(10), 2577–2603.

Modigliani, F., and M. Miller (1958): "The cost of capital, corporation finance and the theory of investment," *The American Economic Review*, 48(3), 261–297.

Penati, A., and A. Protopapadakis (1988): "The effect of implicit deposit insurance on banks portfolio choices, with an application to international overexposure," *Journal of Monetary Economics*, 21(1), 107–126.

Popov, A., and S. Ongena (2010): "Interbank market integration, loan rates, and firm leverage," *Journal of Banking & Finance*.

Rochet, J.-C., and J. Tirole (1996): "Interbank Lending and Systemic Risk," *Journal of Money, Credit and Banking*, 28(4), 733–762.

Ross, S. (1976): "The arbitrage theory of capital asset pricing," *Journal of Economic Theory*, 13(3), 341–360.

Stapleton, R., and M. Subrahmanyam (1984): "The valuation of multivariate contingent claims in discrete time models," *Journal of Finance*, pp. 207–228.

Wagner, W. (2009): "Diversification at financial institutions and systemic crises," *Journal of Financial Intermediation*.

Notes

[1]As shown by Brennan (1979), and Stapleton and Subrahmanyam (1984), the necessary and sufficient condition for a risk-neutral valuation relationship to hold if the joint distribution of returns of the underlying stochastic variables and aggregate wealth is multivariate normal, is that the representative investor has constant absolute risk aversion.

[2]We make the following assumptions for $f(.)$: $\partial f(\ell_i)/\partial \ell_i > 0$, and $f(0) = 0$, but do not further specify the function $f(.)$.

[3]To plot Figure 2 we chose a market with 50 banks. The returns have a factor loading of 1.5% and normal idiosyncratic risk. The normal factor has zero mean and volatility 1. In the left panel the idiosyncratic risk has a volatility of 6%, in the low volatility scenario (right panel) a volatility of 1.25%. All banks in the market have the same size and set the same leverage. The expected asset return is 1.5%. The government bails out the banks, if more than one bank defaults. The capital constraint is chosen such that $p_i = 0.5\%$.

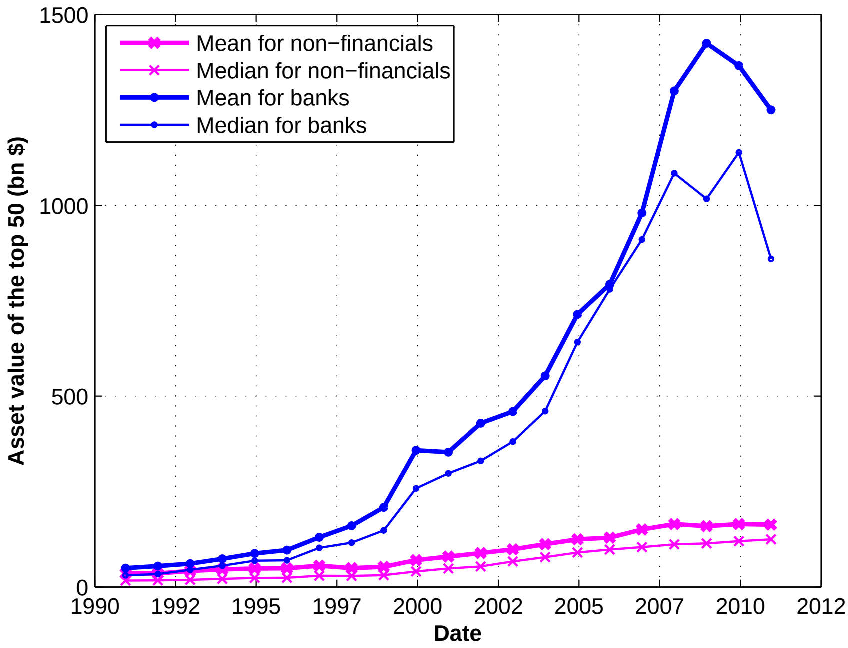

Figure 1: Mean and median asset values of banks and non-financials

This figure plots the mean and median asset values of the largest fifty S&P rated banks and the largest fifty S&P rated non-financials. The asset values of banks increased clearly more than the asset values of non-financial firms. Furthermore, the increasing gap between the mean and median suggests an increasing skewness in asset size, that is, big banks become even bigger. Source: Bloomberg, company reports.

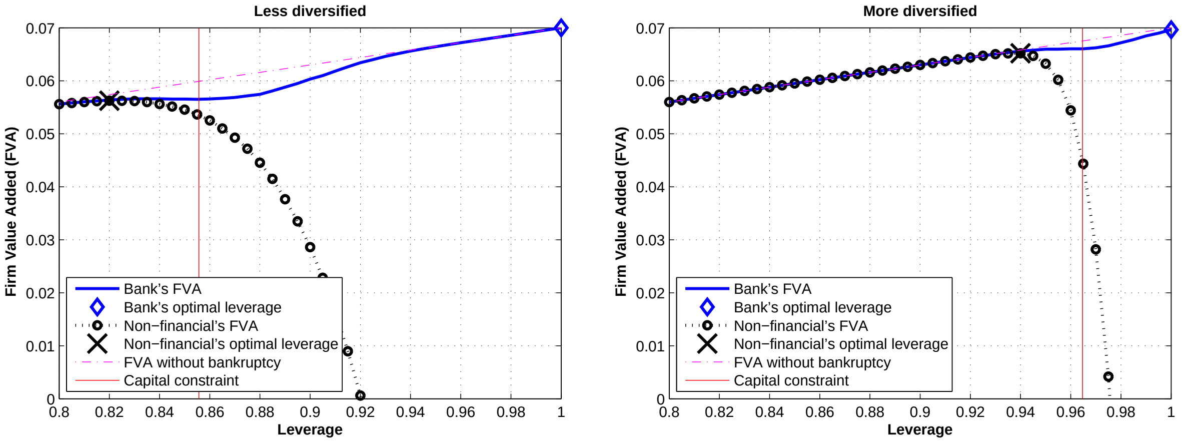

FIGURE 2: INCENTIVES
This plot shows the firm value added (FVA) relative to the value of a purely equity financed firm. The line with bullets is the FVA for a non-financial firm, and the solid thick line is the FVA for a financial firm (bank). The firm value added contains the advantages of debt minus expected bankruptcy costs. In contrast to the non-financial firm's FVA, the bank's FVA additionally contains the value of the implicit government guarantee. Debt advantages are modeled as a linear function of leverage (dashed upward sloping line). The left panel shows the FVA with high asset return volatilities, the right panel with low volatilities. With a lower volatility (more diversified), expected bankruptcy costs decrease. Further, the capital constraint (solid vertical line) allows the firm a higher leverage with a lower return volatility. Due to the implicit guarantee, the bank's FVA is above the non-financials FVA. If the advantages of debt are sufficiently increasing with leverage, the bank has a higher optimal leverage (diamond), than the non-financial firm (cross). In this example, the bank chooses the maximal leverage allowed by the capital constraint, where the solid vertical line crosses the FVA line. In contrast, the non-financial firm chooses a leverage below the maximum in the low-volatility scenario (right panel).

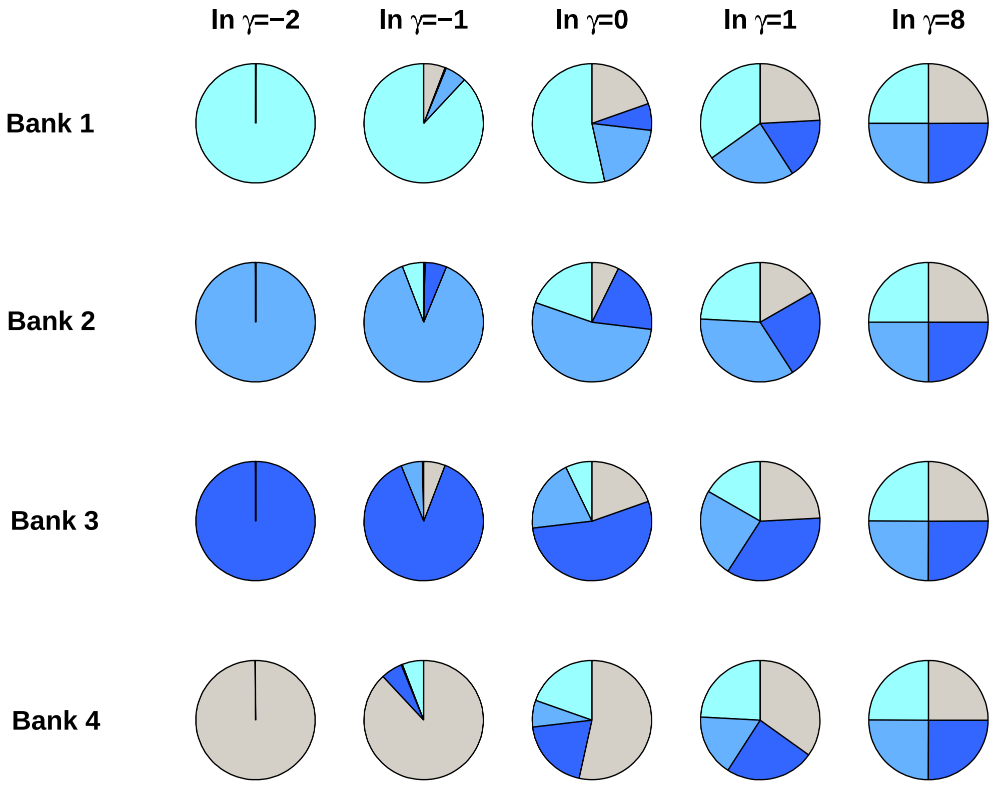

Figure 3: The effect of interbank market integration on the bank asset portfolios
The Figure depicts the asset portfolios of four banks for different degrees of interbank market integration, $\ln \gamma$. The colors denote the assets generated in the four local markets. If the interbank market integration is very low ($\ln \gamma = -2$), each bank's asset portfolio solely consists of the assets generated in its local market. If the integration is small ($\ln \gamma = -1$), banks swap assets with their neighboring banks. As the integration increases further ($\ln \gamma = 0$, $\ln \gamma = 1$), banks swap an increasing amount of assets with more distant banks. If the integration is very high ($\ln \gamma = 8$), each bank ends up with the same asset portfolio.

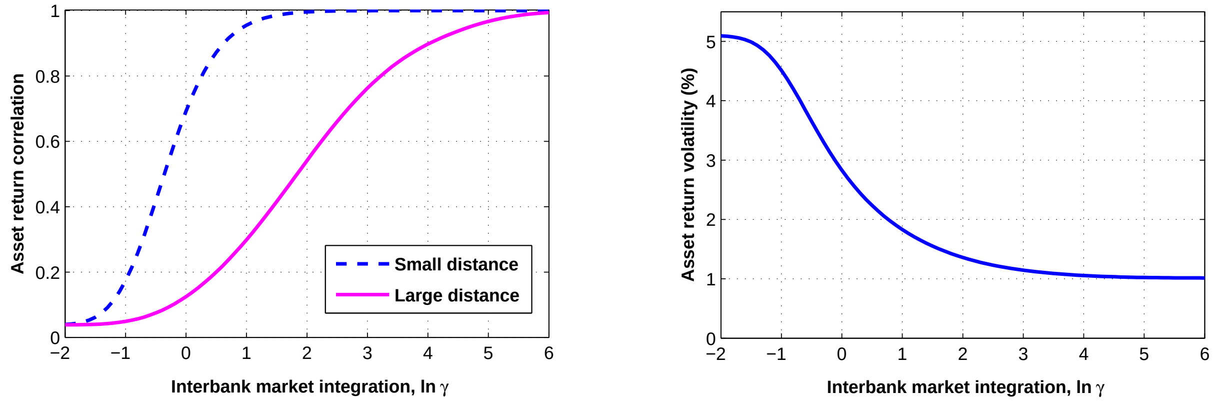

Figure 4: Correlation and volatility of asset returns

The left panel plots the asset return correlation for different degrees of interbank market integration, $\ln\gamma$, in an economy with 1000 equally sized banks. "Close" banks have a higher return correlation. The correlation approaches 1, as the interbank market integration increases. The right panel plots the return volatility of the bank assets, which is the same for every bank. The return volatility decreases with an increasing interbank market integration.

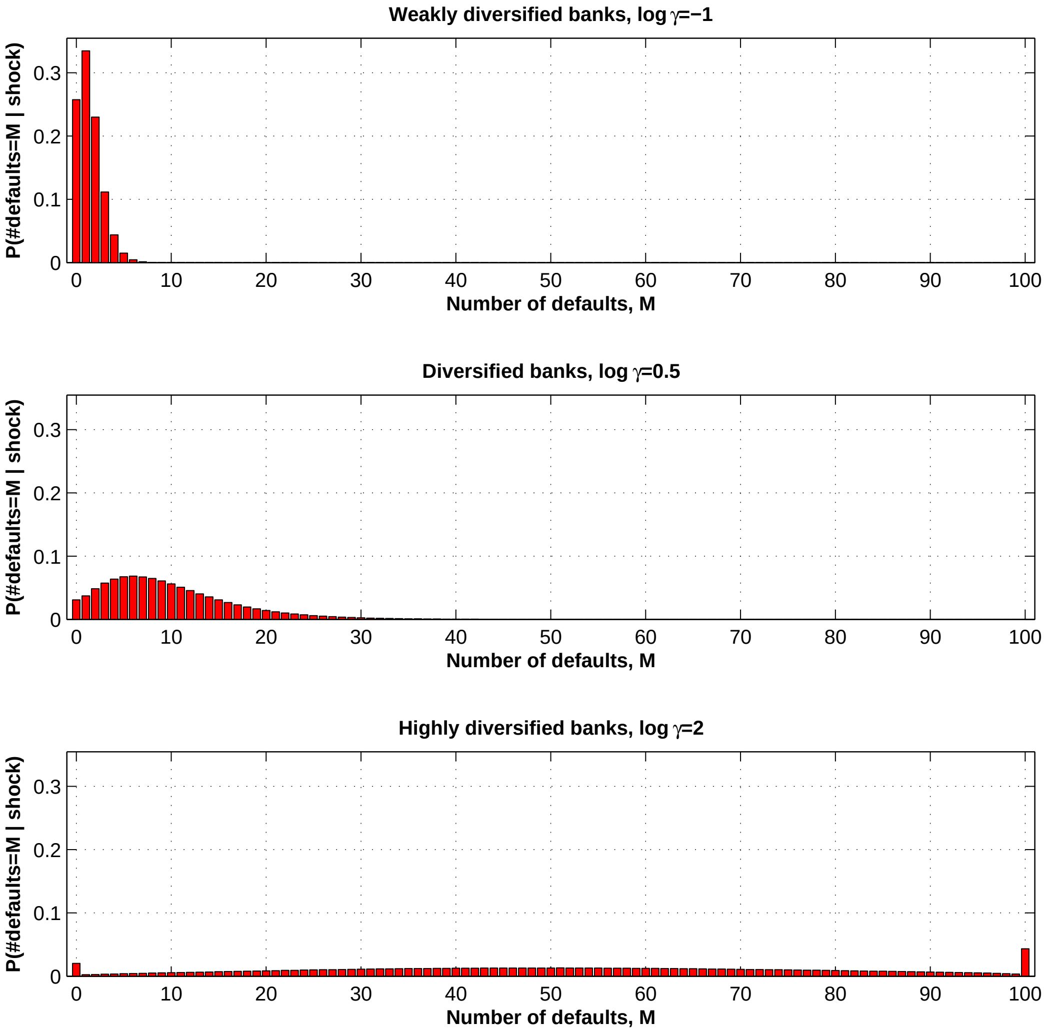

Figure 5: Default probability distribution of a banking system in a crisis

This Figure shows default probability distributions, given that there is a systematic shock. The size of the systematic shock corresponds to a 100 year event. The variable $\ln \gamma$ determines the degree of interbank market integration. We choose the integration equal to -1, 0.5, and 2 from the top panel to the bottom panel. If the interbank market integration (and diversification) is low, the system is able to survive a systematic shock with a few bank defaults (top panel). If it is higher, the probability for a few bank defaults decreases, but the probability for a large fraction of bank defaults increases (middle panel). If the interbank market integration increases to $\ln \gamma = 2$, the probability distribution flattens, and we have large probabilities of observing multiple bank defaults (bottom panel).

Table 1: Social costs. The table shows risk-neutral probabilities of default (RNPD), and physical probabilities of default (PD) in percent, for different degrees of interbank market integration, $\ln \gamma$, in a market with 100 equally sized banks. The government bails out the failed banks in this example, if more than 20% of the banks default. We compute the risk-neutral probability of a default and a bailout (RNPDB), and the ratio RNPDB/RNPD in percent (Ratio). The probability of default remains constant as the market integration increases, but the risk-neutral probability of default, which reflects the costs for an insurance against bankruptcy, increases. Moreover, the risk-neutral probability of defaulting and being bailed out (RNPDB) increases, which reflects the value of the free government guarantee. At the degree of market integration $\ln \gamma = 3$, 93% of the bankruptcy insurance costs are borne by the society.

$\ln \gamma$	RNPD	RNPDB	Ratio	PD
-1.0	0.29	0.00	0	0.10
-0.5	0.37	0.00	0	0.10
0.0	0.52	0.00	0	0.10
0.5	0.78	0.01	2	0.10
1.0	1.16	0.16	14	0.10
1.5	1.69	0.64	38	0.10
2.0	2.34	1.52	65	0.10
2.5	3.02	2.52	83	0.10
3.0	3.53	3.29	93	0.10

TABLE 2: AN ILLUSTRATIVE MARKET. This Table shows the insurance premium (G_i) for a market with 662 banks and the price of an explicit insurance (Ins.), in percent of the asset value. There are four types of banks which differ in the asset size (A_i). All banks have the same one-year probability of default of 0.1%. The number of banks of each type are given in the column labeled "Nmbr.". The asset return volatility (σ_i in percent) decreases with the size of the bank, since the larger banks are more diversified. The government bails out all failed banks if the sum of all loss on assets exceeds 5% of the bank assets in the economy. The loss on assets is 50% of the asset value A_i. The two big banks of type I contribute more to systemic risk than its smaller rivals and have to pay the largest premium, which is more than one third of the price of an explicit insurance.

Type	Nmbr.	A_i	σ_i	G_i	Ins.
I	2	50	1.22	0.58	1.57
II	10	10	1.78	0.11	0.62
III	150	2	2.60	0.03	0.30
IV	500	1	2.83	0.02	0.26