Credit Default Swap Spreads and Systemic Financial Risk

Appendix

Stefano Giglio

April 2014

Appendix A - Collateral agreements and the pricing of counterparty risk

In the text, I argue that the collateral agreements used for CDS contracts during the financial crisis were unlikely to eliminate counterparty risk. Buyers of CDS protection were aware of this and priced it into the spreads. Here I report some evidence for the main points of the argument.

An initial question is whether counterparty risk was perceived at all by market participants. The growth of the percentage of OTC derivative contracts covered by some form of collateral confirms this indirectly: for credit derivatives, the volume-weighted percentage of collateralized contracts went from 39% in 2004 to 58% in 2005, to 66% in 2007 and 2008 (ISDA Margin Survey 2006, 2008). Besides, documents and interviews from practitioners directly confirm that the issue was taken into account by financial participants throughout the crisis. Robert McWilliam, head of Counterparty Risk management at ABN Amro, reports in January 2008[1]*: "The golden rule is to start early. If you start worrying about the counterparty when they are under duress your options are fairly limited"*. A document from Barclays dated February 2008[2] states:*"While the maximum potential loss to the seller of protection is the contract spread for the rest of the contract duration, the buyer of protection could arguably lose the full notional of the contract (in case of simultaneous defaults by counterparty and the reference credit and zero recovery). Thus, counterparty risk is evidently more of a concern for buyers of protection."*

Even if agents were aware of counterparty risk, it was standard practice to ask for relatively little collateral, especially from the largest counterparties. ISDA reports that only about $\frac{2}{3}$ of the contracts were covered by a collateral agreement, up to 2009. Besides, calculations by Singh and Aitken (2009) and Singh (2010) show that, even at the end of 2009, large financial

[1]Reuters,"Banks move to guard against counterparty failures", Jan 24, 2008.
[2]Barclays Capital, 2008, "Counterparty Risk in Credit Markets".

institutions still carried large under-collateralized derivative liabilities. In particular, they compute the total value of "residual derivative payables" - liabilities from derivative positions after netting under master netting agreements and in excess of the collateral posted. For the 5 largest US dealers this amount was more than $250bn. Even though these numbers include all derivative contracts, and not only CDSs, they suggest a general under-collateralization of derivative positions from these counterparties. As an example of this, in 2008 Goldman Sachs had received collateral for 45% of the value of its receivable OTC derivatives, but posted only 18% of the value of payables. Similarly, JP Morgan in the same year had received collateral for 47% of receivables, but posted only 37% of the payables[3]. Finally, as reported in the text, even the most active dealer in counterparty risk management, Goldman Sachs, failed to cover the full value of exposure on its CDS position with AIG.

Even when a collateral agreement is in place and actively managed, residual counterparty risk cannot be eliminated when the value of the derivative is subject to jumps. While during the crisis we did see gradual increases in CDS spreads of banks, a crucial episode - the Lehman bankruptcy - shows that correlated jumps in credit risk (and defaults) are indeed possible. Just before the weekend of the 13th and 14th of September 2008, many institutions were considered at risk, but neither the credit ratings nor the CDS spreads indicated an extremely high likelihood of immediate default. For example, the Lehman 5 year CDS was trading at around 700bp per year, Merrill's at 400bp, and the credit ratings of their debt were still as high as 4 months before, with an implied default probability of less than 0.25% per annum. A buyer who bought a Lehman or a Merrill CDS at 350bp per year a month before the default would have seen the value of the contract (the present discounted value of the difference in spreads) grow to 15 cents and 5 cents on the dollar respectively on Friday September 12th. Therefore, even if the buyers had called for enough collateral to cover the current value of such contracts, they would have improved their recovery rate by only 5% to 15%.

For the reasons explained above, buyers were generally aware that the collateral agreements in place (if any) would have left them exposed to the risk of double default. In fact, several sources document that in early 2008 buyers of CDS contracts were buying additional CDS contracts *against their counterparties* to hedge the residual counterparty risk. For example, from the documents on the AIG bailout (Maiden Lane III) from the Financial Crisis Inquiry Commission, we see that starting November 2007, Goldman Sachs - which had bought $22bn of CDS on a super-senior tranche of a CDO from AIG - was adjusting the amount of CDS protection *against* AIG together with their margin calls to AIG (which were caused by increases in the default probability of the underlying asset). Up to June 2008, the nominal amount of protection bought against AIG was of the same order of magnitude as the total

[3]Bloomberg, "Goldman Sachs Demands Collateral It Won't Dish Out", March 15, 2010.

amount of collateral called by Goldman.

In a document issued by Goldman Sachs in 2009 regarding the AIG bailout[4], the firm declares: *"In mid-September 2008, prior to the government's action to save AIG, a majority of Goldman Sachs' exposure* [current market value] *to AIG was collateralized and the rest was covered through various risk mitigants. Our total exposure on the securities on which we bought protection was roughly $10 billion. Against this, we held roughly $7.5 billion in collateral. The remainder was fully covered through hedges we purchased, primarily through CDS for which we received collateral from our market counterparties. Thus, if AIG had failed, we would have had the collateral from AIG and the proceeds from the CDS protection we purchased."*. Similarly, in an interview with ABN Amro, Reuter reports[5]: *When counterparties* [to OTC derivatives] *are large corporations, which do not usually put up collateral, ABN buys protection in the CDS market against the default of the counterparty itself. ABN's trading desk must go into the market constantly to rebalance those CDS holdings so that its protection equals its counterparty risk profile."*.

This evidence indicates that buyers were understanding the direct and indirect costs of the residual counterparty risk. Note that the fact that collateral was not enough to eliminate counterparty risk does not mean that buyers were making a bad deal on their contracts. Simply, they would have been compensated by paying a lower spread for the contracts when the counterparty was at higher risk of double default. In fact, the 2008 Barclays report titles a section: *How much should I pay for a higher-rated counterparty?* (The analysis then quantifies this number for generic corporate reference entities of different credit rating).

Appendix B - Implementation of the Linear Programming Problem

This appendix describes in detail the algorithm employed to transform the probability bounds problem into a linear programming problem.

B.1 - Linear programming representation in the general case

This section describes the algorithm used to transform the probability problem

$$max\, P_r$$

s.t.

$$P(A_i) = a_i$$

[4] Goldman Sachs, "Overview of Goldman Sachs' Interaction with AIG and Goldman Sachs' Approach to Risk Management", March 2009.

[5] Reuters, "Banks move to guard against counterparty failures", Jan 24, 2008.

$$...$$

$$P(A_i \cap A_j) = a_{ij}$$

into the LP representation

$$max_p \, c_r' p$$

s.t.

$$p \geq 0$$

$$i'p = 1$$

$$Ap = b$$

for the general case of N banks.

Start with a matrix B of size $(2^N, N)$ whose rows contain the binary representation of all numbers between 0 and $2^N - 1$. For example, with N=4:

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ & ... & & \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Each row of this matrix corresponds to a particular element of the partition of the sample space: the event

$$A_1^* \cap A_2^* \cap ... \cap A_N^*$$

where $A_j^* = A_j$ if element j of the row is 1, and $A_j^* = \overline{A}_j$ if element j of the row is 0. The probability system p will then be determined as a vector of 2^N elements containing the probability of each of the elements of the partition represented by the 2^N rows of the matrix B. For example p_1 will be the probability that none of the A_i events occur, p_2 will represent the probability that event A_N occurs but none of the other events does, and so on. Finally, the element p_{2^N} will represent the probability that all events occur.

The maximization problem presented above tries to find the vector p that maximizes the probability of systemic event of degree r (P_r) while satisfying constraints on marginal and pairwise default probabilities, as well as the constraints implied by the consistency of the probability measure. The latter are immediate: because the events represented by the rows of B are a partition of the sample space, and p is a probability measure on these events, all elements of p need to be nonnegative and sum to one:

$$p \geq 0$$

$$p'i = 1$$

To obtain in LP form the inequalities and equalities that involve marginal and pairwise default probabilities, note first that because the elements of the partition are disjoint events, the probability of any union of them is equal to the sum of their probabilities. Therefore, to find the probability of an event A_i, $P(A_i)$, in terms of p, one needs to sum the probabilities of all the elements of the partition in which event A_i occurs. But this is immediate given the representation in B:

$$P(A_i) = \sum_{j:B(j,i)=1} p_j$$

or:

$$P(A_i) = a^{i'} p$$

for a vector a_i of size $(2^N, 1)$ s.t.:

$$a^i_j = B(j, i)$$

In other words, to find which elementary events form event A_i one needs to find all the rows of B in which element i is equal to 1. The union of these events will coincide with A_i, and therefore the sum of their probabilities will be $P(A_i)$. Given the linearity, this sum is equivalent to the product of the vector p with a vector a_i, whose elements are ones whenever the corresponding elementary event is a subset of A_i.

Similarly, the probability of a joint event:

$$P(A_i \cap A_k) = \sum_{j:B(j,i)=1 \, and \, B(j,k)=1} p_j$$

or:

$$P(A_i \cap A_k) = b^{ik'} p$$

for a vector b_{ik} of size $(2^N, 1)$ s.t.:

$$b^{ik}_j = B(j, i)B(j, k)$$

i.e., the probability of the joint default is obtained summing the elements of p s.t. the corresponding element of the partition involves *both* the occurrence of A_j and of A_k. All these constraint can then be collected in the matrix form $Ap = b$.

Finally, the probability that at least r events occur can be found as follows:

$$P_r = \sum_{j:\left(\sum_{h=1}^{N} B(j,h)\right)\geq r} p_j$$

or:

$$P_r = c^{r\prime} p$$

for a vector c^r of size $(2^N, 1)$ s.t.:

$$c_j^r = I\left[\sum_{h=1}^{N} B(j,h) \geq r\right]$$

where $I[]$ is the indicator function.

Given this decomposition, the LP representation follows immediately.

B.2 - Symmetry of the probability system: Proposition 2

This section introduces the relevant definitions for a formal exposition of Proposition 2, and proves that Proposition. Consider the vector $p \in \mathbb{R}^{2^N}$ representing a probability system on the σ-algebra generated by the basic events $A_1, ..., A_N$, as in Proposition 1. Consider a permutation J of the indices of the basic events: $A_{J_1}, ..., A_{J_N}$, and call $\mathcal{M}$ the set of permutations. Call $p_J \in \mathbb{R}^{2^N}$ the vector representing the probability system generated by $A_{J_1}, ..., A_{J_N}$ that corresponds to p, constructed as described in the text.

For example, take two events A_1 and A_2. The vector p would have four elements: $p_1 = P(\overline{A}_1 \cap \overline{A}_2)$, $p_2 = P(\overline{A}_1 \cap A_2)$, $p_3 = P(A_1 \cap \overline{A}_2)$ and $p_4 = P(A_1 \cap A_2)$. In this case, only one additional permutation of the generating events is possible, $J = \{2, 1\}$, with $p_{J1} = p_1$, $p_{J2} = p_3$, $p_{J3} = p_2$, and $p_{J4} = p4$.

Definition. A linear combination of the elements of p defined by the vector c is *symmetric* with respect to the generating events $A_1, ..., A_N$ if $c'p = c'p_J \;\; \forall J \in \mathcal{M}$. A linear programming problem, $max\, c'p \;\, s.t. \;\, Ap \leq b$, is *symmetric* if c and all rows of A are *symmetric* with respect to the generating events $A_1, ..., A_N$.

An example of a symmetric weighting vector c is the one corresponding to the probability of the union of the events, $c = [1\,1\,0\,1]'$, since $c'p = c'p_J = P(A_1 \cup A_2)$.

Definition. A probability system p is *symmetric* if every event in V, the finest partition of the sample space generated by the basic events, has the same probability in all permutations of the generating events.

For example, with three generating events ($N = 3$), the probability system is symmetric if $P(A_1) = P(A_2) = P(A_3)$ and $P(A_1 \cap A_2) = P(A_1 \cap A_3) = P(A_1 \cap A_3)$. I can now prove the following proposition:

Proposition. *Suppose that the probability bounds correspond to a* symmetric *LP problem. Then, the bounds are attained by a* symmetric *probability system.*

Proof. Start from a symmetric LP problem

$$max\, c'p$$

$$s.t.\, Ap \leq b$$

Suppose that p^* is a solution to the problem. Given the definition of symmetry presented in the text, it is clear that p^*_J is also a solution to the problem: $c'p^* = c'p^*_J$ and similarly hold for every row of the constraints, for every J.

Now, construct p^{**} as follows:

$$p^{**} = \frac{1}{\#J} \sum_J p^*_J$$

where the first J correspond to no permutation, and J cycles across all permutations of indices $A_1, ..., A_N$.

Note that it is also possible to construct p^{**} in the following way, considering the binary representation introduced in the text. Every b_i vector has O_i ones and $N - O_i$ zeros. Call H_i the set of all vectors of size N that have O_i ones and $N - O_i$ zeroes in different positions. Call b_{ih} the vector corresponding to element h from H_i. Then, for every i, construct p^{**} as:

$$p_i^{**} = \begin{pmatrix} O_i \\ N \end{pmatrix}^{-1} \sum_{h \in H_i} b_{ih}$$

From the first construction, it is clear why p^{**} is a solution to the maximization problem, being just an average of solutions. Additionally, p^{**} is symmetric, which proves the statement of the Proposition.

An example with $N = 3$. We can construct the probability system p^* as follows:

$$p^*_1 = Pr\{\overline{A}_1 \cap \overline{A}_2 \cap \overline{A}_3\}$$

$$p^*_2 = Pr\{\overline{A}_1 \cap \overline{A}_2 \cap A_3\}$$

$$p^*_3 = Pr\{\overline{A}_1 \cap A_2 \cap \overline{A}_3\}$$

$$p_4^* = Pr\{\overline{A}_1 \cap A_2 \cap A_3\}$$

$$p_5^* = Pr\{A_1 \cap \overline{A}_2 \cap \overline{A}_3\}$$

$$p_6^* = Pr\{A_1 \cap \overline{A}_2 \cap A_3\}$$

$$p_7^* = Pr\{A_1 \cap A_2 \cap \overline{A}_3\}$$

$$p_8^* = Pr\{A_1 \cap A_2 \cap A_3\}$$

Suppose p^* solves the maximization problem, and construct p^{**} as:

$$p_1^{**} = p_1^*$$

$$p_2^{**} = p_3^{**} = p_5^{**} = \frac{p_2^* + p_3^* + p_5^*}{3}$$

$$p_4^{**} = p_6^{**} = p_7^{**} = \frac{p_4^* + p_6^* + p_7^*}{3}$$

$$p_8^{**} = p_8^*$$

p^{**} solves the maximization problem and is symmetric.

□

Corollary. *(Proposition 2) The bounds on systemic events of the type "at least r institutions default" given a symmetric constraint set (for example, constraints on the* average *marginal and pairwise default probabilities) are attained by a symmetric probability system.*

The bounds obtained in a symmetric network in which we observe all marginal and pairwise probabilities will always be at least as wide as those obtained in an asymmetric network with the same averages of the low-order probabilities. The difference between the bounds obtained in the two cases captures precisely the extent to which asymmetry in the shape of the network affects the probability of systemic events.

Appendix C: additional pricing details

C.1 - Bond pricing model

The bond pricing model used corresponds to the one in Longstaff et al. (2005), with the exception that the process for h_t^i has no drift term. Following their notation, the closed-form solution for the bond price at time 0 is:

$$P(c,R,T) = c\int_0^T A(t)e^{B(t)h_0}C(t)D(t)e^{-\gamma_0 t}dt$$

$$+A(T)e^{B(T)h_0}C(T)D(T)e^{-\gamma_0 T}dt$$

$$+R\int_0^T e^{B(t)h_0}C(t)D(t)(G(t)+H(t)h_0)e^{-\gamma_0 t}dt \tag{1}$$

with:

$$A(t) = 1$$

$$B(t) = \frac{-\sqrt{2\sigma^2}}{\sigma^2} + \frac{2\sqrt{2\sigma^2}}{\sigma^2(1+e^{\sqrt{2\sigma^2}t})} = \frac{\sqrt{2}}{\sigma}\left(-1+\frac{2}{1+e^{\sqrt{2\sigma^2}t}}\right)$$

$$C(t) = exp(\frac{\eta^2 t^3}{6})$$

$$D_t = E[exp(\int_0^t r_s ds)]$$

$$G(t) = 0$$

$$H(t) = exp(\sqrt{2\sigma^2}t)(\frac{2}{1+e^{\sqrt{2\sigma^2}t}})^2$$

$$\phi = \sqrt{2\sigma^2}$$

C.2 - Additional details of CDS contracts

Besides those considered explicitly in this paper, there are other elements of CDS contracts that potentially affect their spreads.

First, liquidity of the CDS market could influence the CDS spreads, just as bond liquidity is known to affect bond prices. In this paper, I explicitly take into account liquidity premia in bond prices, but not in CDS spreads. For the case of CDSs, liquidity is much less likely to be an issue, especially because they require much less capital at origination and they are not in fixed supply.[6]

Also, I abstract from restructuring clauses and the cheapest-to-deliver option sometimes present in CDS contracts. A restructuring clause (under which payment is triggered for simple debt restructuring, in addition to bankruptcy) is more frequent for European bonds, and this

[6]For an additional discussion of this and on the supporting evidence, see Blanco, Brennan and Marsh (2003,2005).

results in the contract being triggered in cases close to the Chapter 11 for the US. Berndt, Jarrow and Kang (2007) estimate that the presence of such clause increases the value of the CDSs by 6-8%, and all the results in this paper are robust to an adjustment of CDS spreads of that magnitude. The value of the cheapest-to-deliver option (which allows the buyer to deliver to the seller the cheapest of the defaulted bonds of the same seniority as the reference bond) will be small relative to the CDS spread as long as in default all senior unsecured bonds have similar recovery rates. Additionally, as observed in the Delphi and Calpine defaults in 2005, the high demand for the cheapest bonds might determine shortages of such securities and therefore, anticipating this, a reduction in the ex-ante value of the option (De Wit 2006).

C.3 - CDS pricing approximation

This section shows how to derive the linearized version of the CDS pricing equation. For notational simplicity, we compute the CDS price at time 0. I obtain the derivation in continuous time, though the same derivation can be obtained in discrete time as well. Start from the pricing equation:

$$E\left[z_{ji}\int_0^T e^{-\int_0^t r_s+(h_s^i+h_s^j-h_s^{ij})ds}dt\right]$$

$$= E\left[\int_0^T \left\{\left[h_t^i-h_t^{ij}\right](1-R)+h_t^{ij}S(1-R)\right\}e^{-\int_0^t r_s+(h_s^i+h_s^j-h_s^{ij})ds}dt\right]$$

Assuming independence of the interest rate process, we can write

$$z_{ji}\int_0^T \delta(t)E\left[e^{-\int_0^t (h_s^i+h_s^j-h_s^{ij})ds}\right]dt$$

$$= \int_0^T \delta(t)E\left[\left\{\left[h_t^i-h_t^{ij}\right](1-R)+h_t^{ij}S(1-R)\right\}e^{-\int_0^t (h_s^i+h_s^j-h_s^{ij})ds}\right]dt$$

or:

$$\frac{z_{ji}}{1-R} = \frac{\int_0^T \delta(t)E\left[\left\{\left[h_t^i-h_t^{ij}\right]+h_t^{ij}S\right\}e^{-\int_0^t (h_s^i+h_s^j-h_s^{ij})ds}\right]dt}{\int_0^T \delta(t)E\left[e^{-\int_0^t (h_s^i+h_s^j-h_s^{ij})ds}\right]dt}$$

where $\delta(t)$ is the time-0 risk-free discount factor.

The expectations above are conditional on information at time 0, and they are a nonlinear function of the information at time 0 (h_0^i, h_0^j, h_0^{ij}). We now make two simple approximations to the formula above that will hold for small values of the intensities $\{h_s\}$, and we verify numerically that the approximation works extremely well for a wide range of calibrated values of the probabilities (h_0^i, h_0^j, h_0^{ij}) that span the values we observe in the data.

First, we consider a linear approximation of the functions that appear within the expectations, for all h^i_s, h^j_s and h^{ij}_s close to 0. Note that these functions are effectively deterministic because they condition on the realized path of the h's.

Consider first the denominator. Calling $X(t) = \int_0^t (h^i_s + h^j_s - h^{ij}_s) ds$, a first-order expansion around $X(t) = 0$ gives:

$$e^{-\int_0^t (h^i_s + h^j_s - h^{ij}_s) ds} = e^{-X(t)} \simeq 1 - X(t) = 1 - \int_0^t (h^i_s + h^j_s - h^{ij}_s) ds$$

For the numerator, consider the function

$$F = \left\{ h^i_t - h^{ij}_t + h^{ij}_t S \right\} e^{-\int_0^t (h^i_s + h^j_s - h^{ij}_s) ds} = \left\{ h^i_t - h^{ij}_t + h^{ij}_t S \right\} e^{-X(t)}$$

Taking a Taylor expansion of F for both $\left\{ h^i_t - h^{ij}_t + h^{ij}_t S \right\}$ and $X(t)$ close to 0, we obtain:

$$F \simeq h^i_t - h^{ij}_t + h^{ij}_t S$$

We can now express the right-hand side of the pricing equation as a function of h^i_0, h^j_0 and h^{ij}_0, exploiting the fact that h are martingales:

$$\frac{z_{ji}}{1-R} \simeq \frac{\int_0^T \delta(t) E\left[h^i_t - h^{ij}_t + h^{ij}_t S \right] dt}{\int_0^T \delta(t) E\left[1 - \int_0^t (h^i_s + h^j_s - h^{ij}_s) ds \right] dt} = \frac{\int_0^T \delta(t) \left[h^i_0 - h^{ij}_0 + h^{ij}_0 S \right] dt}{\int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt}$$

Finally, in order to have a linear constraint in the optimization, we approximate the resulting function linearly around $h^i_0 = h^j_0 = h^{ij}_0 = 0$. Call the function

$$G = \frac{\int_0^T \delta(t) \left[h^i_0 - h^{ij}_0 + h^{ij}_0 S \right] dt}{\int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt}$$

Notice first that $G(0) = 0$. Then compute:

$$G_{h^i_0}|_0 = \frac{\left[\frac{d}{dh^i_0} \int_0^T \delta(t) \left[h^i_0 - h^{ij}_0 + h^{ij}_0 S \right] dt \right]_0 \left[\int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt \right]_0}{\left(\int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt \right)^2} |_0$$

$$- \frac{\left[\int_0^T \delta(t) \left[h^i_0 - h^{ij}_0 + h^{ij}_0 S \right] dt \right]_0 \left[\frac{d}{dh^i_0} \int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt \right]_0}{\left(\int_0^T \delta(t) \left[1 - (h^i_0 + h^j_0 - h^{ij}_0) t \right] dt \right)^2}$$

$$= \frac{\left[\frac{d}{dh_0^i} \int_0^T \delta(t) \left[h_0^i - h_0^{ij} + h_0^{ij} S\right] dt\right]_0 \left[\int_0^T \delta(t) dt\right]}{\left[\int_0^T \delta(t) dt\right]^2} |_0 = \frac{\left[\int_0^T \delta(t) dt\right]^2}{\left[\int_0^T \delta(t) dt\right]^2} = 1$$

Similarly, one can compute

$$G_{h_0^j}|_0 = 0$$

and

$$G_{h_0^{ij}}|_0 = (S - 1)$$

Therefore, putting all together we can write

$$z_{ji} \simeq (h_0^i - h_0^{ij})(1 - R) + h_0^{ij} S(1 - R) = (1 - R)(h_0^i - (1 - S)h_0^{ij}) \tag{2}$$

The pricing formula can also be obtained directly in discrete-time, by writing down the pricing equation for CDSs in discrete time (again, assuming a martingale for h^i and h^{ij} and ignoring the convexity effects due to future variability of hazard rates), and employing first-order Taylor expansions analogous to the ones used in continuous time above. In particular, the approximation one obtains in discrete time is:

$$z_{ij} \simeq (1 - R) \frac{\left[\sum_{s=1}^T \delta(0, s)\right]}{\left[\sum_{s=1}^T \delta(0, s - 1)\right]} (h_0^i - (1 - S)h_0^{ij})$$

which is almost identical to eq. (2), apart of a (negligible) adjustment term that depends on the risk free rate and comes from the fact that when working in discrete time I assume that the payment of CDS premia happens at the beginning of each period while payouts happen at the end of each period.

Note also that the same pricing formula would hold if agents priced the CDS using a constant-hazard rate model, treating the hazard rates as if they will be constant at level h_t^i and h_t^{ij} for all the life of the CDS, when pricing a CDS at time t. This is a result of the first-order approximation that ignores the volatility adjustment together with the fact that hazard rates are assumed to be martingales.

It is important to check the accuracy of the approximation for a realistic range of parameters. For several different points in time (every 50 days) between 1/1/2007 and 3/31/2009, I compare the correct spread and the approximated spread, computed using the US yield curve at that time. I simulate the CDS spreads using the discretized version of the model, and compare it with the approximated one. I consider a large range of parameters:

- different values of $h_j = P(A_j)$: between 0 and the maximum probability implied by bond data under no liquidity assumptions ($max_j\{h_j(0)\}$).
- different values of $h^{ij} = P(A_i \cap A_j)$: between 0 and $P(A_j)$
- different values of R and S: between 0.1 and 0.4.

In all these simulations, the approximation error is always below 0.3% of the true value of the CDS spread. The approximation therefore is extremely accurate.

Appendix D - Robustness Tests

In this section I study the robustness of the main results of the paper to different assumptions. For each robustness test (all performed under the calibration of the liquidity to the basis of nonfinancial institutions, $\underline{\gamma}_t^i = \alpha_i \lambda_t^*$), I report in Appendix Table 1 the average value of the bounds, in basis points per month, during different periods: January to December 2007, January 2008 to March 15 2008 (the run-up to Bear Stearns' collapse), from Bear's episode to Lehman's default (on September 15th 2008), the month after Lehman's default (in which CDS spreads and bond yields spiked), the period between September 2008 and April 2009 (the latest peak of the crisis, just before the stress test results were released) and finally from May 2009 to June 2010. The subperiods were chosen to reflect the main events identified in the Figures. In the three panels, I show values for the lower and upper bound on P_1, and the upper bound on P_4 (the lower bound on P_4 is always 0).

Besides showing the level and the time series of the bounds, this Table allows us to check that the main results reported in the paper hold under different assumptions. The bold line in each panel of Appendix Table 1 reports the baseline case presented in the paper. We can confirm the result, presented in Figure 4, that systemic risk was low in the months preceding Bear Stearns' collapse, while idiosyncratic risk was already high. Besides, we can see that during the month after Lehman's default, idiosyncratic risk (P_1) spiked: it increased sharply and then decreased as sharply. The upper bound on systemic risk, P_4, increases as well, but not as much: in most cases, it does not even reach the level observed in the following 6 months. Only for two specifications (only US banks and alternative bond model) P_4 is actually higher in September 2008 than in the next 6 months, but in those cases P_1 is *much* higher in September 2008 than in the following 6 months: idiosyncratic risk still spikes *more* than systemic risk. The results, as explained below, only stop holding for the case of full recovery of CDSs in case of double default, $S = 100\%$, and in that case systemic and idiosyncratic risk are essentially indistinguishable.

D.1 - Assumptions on S

Let us start with robustness with respect to the assumed recovery rate of CDSs when double default occurs, $S \in [R, 1]$. The effect of changes in this assumption depends crucially on the liquidity-adjusted bond/CDS basis of each bank. For some banks, the basis is small enough that it can be completely explained by counterparty risk. For these banks, an increase in S means that the same basis can account for higher counterparty risk. For other banks, instead, the basis is large enough that, due to internal constraints of the probability system, it cannot be completely explained by counterparty risk: even at the upper bound for systemic risk, a part of the basis must be explained by liquidity. For these banks, an increase in S means that the same amount of counterparty risk - which was already at the maximum possible - will explain an even smaller fraction of the basis. This means that the marginal probability of default, $P(A_i)$, has to decrease. In turn, this directly reduces the maximum possible amount of counterparty risk for contracts written by i against other banks, since for each j we must have $P(A_i \cap A_j) \leq P(A_i)$.

An increase in S then has a different effect on banks with a relatively small basis and banks with a large basis. The two effects are also at play for each bank individually, for different starting levels of S: when S is low enough counterparty risk has a large effect on CDS spreads, and therefore the basis will be relatively small - it can be completely explained by counterparty risk. When S is large enough, not all the basis can be explained by counterparty risk, and the second mechanism operates. Typically, because of asymmetry in the basis across banks, for most values of S the two effects described above will operate for some banks in one direction and for other banks in the opposite direction. This explains why we see the bounds on systemic risk being very robust to changes in S (at least up to a recovery rate of 90%), as shown in Appendix Table 1.

To see formally the effect of S on the implied estimate of systemic risk, it is useful to look at a symmetric network. Remember that the upper bound on systemic risk is attained by the most correlated probability system that satisfies the constraints:

$$P(A_i) \leq h_i(\underline{\gamma}_t^i)$$

$$P(A_i) - (1 - S)\left(\frac{1}{N-1}\sum_{j \neq i} P(A_i \cap A_j)\right) = b_i$$

where $b_i = \frac{\bar{z}_i}{(1-R)}$ [7].

Intuitively, for a given S, one can obtain the most correlated probability system by setting $P(A_i)$ as high as possible (up to the constraint $h_i(\underline{\gamma}^i_t)$) for all banks and then increasing the term $\frac{1}{N-1}\sum_{i \neq j} P(A_i \cap A_j)$ to match the CDS spreads (b_i). Counterparty risk would explain the whole bond/CDS basis, and a higher recovery rate S would imply that a higher joint default probability is needed to match it, increasing the upper bound on systemic risk. This intuitive reasoning, however, does not take into account the internal restrictions of consistency of the probability system. For a symmetric network, call the marginal probability of default of each bank q_1 and the pairwise joint probabilities of default of each pair q_2. The previous constraints become:

$$q_1 \leq h$$

$$q_1 - (1-S)q_2 = b$$

where h is the (common) upper bound on the marginal probability of default and b is the (common) b_i.

To maximize systemic risk, we would intuitively set $q_1 = h$, and then q_2 will be set to match CDS spreads:

$$q_2 = \frac{q_1 - b}{1 - S} \tag{3}$$

For given q_1, q_2 is increasing in S, as is systemic risk. This captures the intuition that a higher recovery rate of CDSs implies that higher counterparty risk is needed to explain the same bond/CDS basis.

In fact, this effect is at play only when S is small enough. As S grows, q_2 keeps increasing, and at some point it will reach the level $q_2 = q_1$. At that point, the internal consistency of the probability system kicks in, preventing further increases: it would violate the implicit constraint that $q_2 \leq q_1$.

What happens then if S increases further? The only way to satisfy the constraints is to *lower* q_1 *below* h: for $q_1 = h$ there might exist no probability systems able to satisfy both constraints: matching the CDS spread and satisfying internal consistency. Instead, with a

[7] I focus on the upper bound for the probability of at least $r > 1$ events occurring. Following the analysis reported in section 3, the same argument holds for the *lower* bound for the probability that at least 1 institution defaults, since that is achieved for a very correlated system. It is easy to see why the results for the lower bound for $r > 1$ and the upper bound for $r = 1$ do not depend on S: these bounds look for the least correlated system, which can always be obtained by setting the marginal default probabilities at the levels implied by the CDS spreads and attributing the bond/CDS basis entirely to liquidity.

lower q_1, it is possible to set q_2 to be equal to q_1 and satisfy the CDS constraint, so that:

$$q_2 = q_1 = \frac{b}{S}$$

which is *decreasing* in S. This means that for large enough values of S, the bond/CDS basis is too large to be explained by counterparty risk. Even at the upper bound on systemic risk, liquidity has to explain part of the basis. In a symmetric system, then, the bounds on systemic risk first increase and then decrease with S.

These forces play out in similar but very nonlinear ways for asymmetric networks. The asymmetry in the bond/CDS basis across banks means that the upper bounds on marginal probabilities (that are obtained from bond prices) will bind for some banks and not for others. So, as S increases, the maximum counterparty risk increases for some banks and decreases for others. Empirically, these opposite effects tend to cancel out for S as large as 90%, as shown by Appendix Table 1. As a result, the upper bound on systemic risk does not depend much on S as long as S is not too large. As S approaches 100%, all the banks that have a basis of 0 or very close to zero, and thus are essentially disconnected from the financial system for all but very high S, can become more correlated to the rest of the financial system. When $S = 100\%$ the CDS spread is not affected by counterparty risk at all, so all the banks can be maximally correlated with each other, with no restriction coming from the basis. Even a bank with a basis of 0, which necessarily has zero counterparty risk as long as the basis contains *some* information about counterparty risk ($S < 100\%$), can become highly correlated with the other banks when $S = 100\%$, since then there is no information about counterparty risk in the basis. This explains why at $S = 100\%$ we find a higher level of systemic risk, and why in that case systemic risk is driven entirely by movements in the level of CDS spreads, not by the basis: the basis is completely uninformative about counterparty risk if recovery is always full.

D.2 - Assumptions on R

The case for the recovery rate of bonds R is different. R affects the prices of *both* bonds and CDSs. A higher expected recovery rate in case of default increases the value of a bond, and at the same time decreases the value of CDS insurance written on that bond, since the payment from the CDS seller covers only the amount of bond value not recovered in default. Because this recovery rate multiplies the marginal and joint default probabilities in the pricing formulas, when R changes all probabilities implied in bonds and CDSs are scaled up or down by approximately the same amount.[8] Therefore, the bounds on systemic risk will scale in a

[8] The difference between the two comes from differences in the cash flow timing of bonds and CDSs. They

similar way. However, the main results on the time series of the bounds will not change, as shown by Appendix Table 1.

D.3 - Time varying recovery rates

Above I have studied robustness to different assumptions about S and R, when these are assumed to be constant during the whole sample period. In theory, it is possible that these recovery rates vary over time in a way that affects the results on the time-series of systemic risk presented in section 4. Suppose that at every time t bonds and CDSs are still priced assuming that at all future periods $t+s$ the recovery rates are constant and equal to S_t and R_t; however, let now S_t and R_t vary over time. How will this affect the bounds on systemic risk?

The tests presented above show that the bounds on systemic risk scale in the same direction as R. If we believe that, during peak episodes like the one following Lehman's default, recovery rates R might have dropped, this would in fact strengthen the result that the spike in systemic risk was then relatively low, because it would further reduce the bound on P_4 during that month.

Another possibility is a reduction in the recovery rate of CDSs, S, in times when systemic risk increases. However, it is easy to see that this case actually reinforces the main empirical results. If the recovery rate S becomes smaller during the key episodes of the crisis, then joint default risk has to be smaller as well. This stems once more from the fact that during these episodes the bond/CDS basis is small relative to CDS and yield spreads. When S is reduced, the probability of joint default has a *greater* effect on the basis. To still match the basis even if S is higher, joint default risk has to decrease. Therefore, the main results in the paper will be robust to a decrease in the recovery rate S in times of crisis.

D.4 - Stochastic recovery rate on bonds R

Another possibility is that when pricing bonds and CDSs, agents incorporate the possibility that recovery rates might be stochastic and correlated with the default events in the financial sector. In particular, one could think that recovery rates of both bonds and CDSs might deteriorate the more defaults happen in the financial system.

Because of the limited data available, it is difficult to solve explicitly for the case of stochastic recovery rates. However, it is possible to gain some intuition on the effect of this assumption under simple modeling assumptions. Suppose that the recovery rate on bonds is R_H whenever one bank defaults alone, and $R_L < R_H$ whenever two or more banks default. Below I show that we can decompose as follows the change in the bounds on systemic risk,

are scaled by exactly the same amount in the simple two-period example of section 2.

going from a non-stochastic recovery rate R to the stochastic recovery process described above. First, we can shift the (constant) recovery rate downwards for both bonds and CDSs to R_L. This component scales down the bond-implied and the CDS-implied probabilities by a similar amount, as discussed above. This would scale the bounds on systemic risk downwards. Second, we increase the present value of bonds by an amount Y_{bond}, and decrease the present value of payments of the CDS contract by an amount $Y_{CDS} \approx Y_{bond}$ (in a first-order approximation with small probabilities of default). This second effect shifts the CDS spread and the yield spread in the same direction by a similar amount, with minimal effect on the basis and hence on counterparty risk. We then expect the bounds on systemic risk to become lower if we introduce a stochastic recovery rate with $R_L < R$. The reason is that for the purpose of systemic risk, the relevant recovery rate is the one that obtains in states of multiple defaults, or R_L in this case. However, as long as the recovery rates R_L and R_H themselves do not vary over time, the time series of the bounds should still look as in Figure 4.

Define $X_i \equiv \cup_{k \neq i} A_k$ the event of at least one default among the banks different from i, and similarly $X_{ij} \equiv \cup_{k \neq i,j} A_k$. Call h_{iX} the joint intensity of default of i and at least one other bank (X_i) during dt (the event $A_i \cap X_i$). Call $h^{i\overline{X}}$ the intensity of the event $A_i \cap \overline{X}_i$, $h^{i\overline{j}\overline{X}_{ij}}$ the intensity of the event $A_i \cap \overline{A}_j \cap \overline{X}_{ij}$, and $h^{i\overline{j}X_{ij}}$ the intensity of the event $A_i \cap \overline{A}_j \cap X_{ij}$. Call $B_R(0,T)$ the price of a bond under the assumption of constant recovery rate R and $B_{R_L,R_H}(0,T)$ the price of a bond with stochastic recovery rate described above, and similarly for the CDS spreads. Then the bond pricing equation – ignoring liquidity – is:

$$B_{R_L,R_H}(0,T) = E\left[c\int_0^T exp(-\int_0^t r_s + h_s ds)dt\right]$$

$$+E\left[exp(-\int_0^T r_s + h_s ds)\right] + E\left[R_H\int_0^T h_t^{i\overline{X}_i} exp(-\int_0^t r_s + h_s ds)dt\right]$$

$$+E\left[R_L\int_0^T h_t^{iX_i} exp(-\int_0^t r_s + h_s ds)dt\right]$$

while the CDS pricing equation is:

$$E[z_{ji}^{R_L,R_H}\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)dt]$$

$$= E[\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)\,\{h_t^{i\bar{j}\overline{X}_{ij}}(1 - R_H) + h_t^{i\bar{j}X_{ij}}(1 - R_L) + h_t^{ij} S(1 - R_L)\}dt]$$

Now, rewrite bond prices as:

$$B_{R_L,R_H}(0,T) = E\left[c\int_0^T exp(-\int_0^t r_s + h_s ds)dt\right]$$

$$+E\left[exp(-\int_0^T r_s + h_s ds)\right] + (R_H - R_L)E\left[\int_0^T h_t^{i\overline{X}_i} exp(-\int_0^t r_s + h_s ds)dt\right]$$

$$+R_L E\left[\int_0^T (h_t^{iX_i} + h_t^{i\overline{X}_i}) exp(-\int_0^t r_s + h_s ds)dt\right]$$

Noting that $h_t^{iX_i}dt + h_t^{i\overline{X}_i}dt = h_t^i dt$, we can rewrite

$$B_{R_L,R_H}(0,T) = E\left[c\int_0^T exp(-\int_0^t r_s + h_s ds)dt\right]$$

$$+E\left[exp(-\int_0^T r_s + h_s ds)\right]$$

$$+R_L E\left[\int_0^T h_t^i exp(-\int_0^t r_s + h_s ds)dt\right] + Y_{bond}$$

Or:

$$B_{R_L,R_H}(0,T) = B_{R_L}(0,T) + Y_{bond}$$

where

$$Y_{bond} = (R_H - R_L)E\left[\int_0^T h_t^{i\overline{X}_i} exp(-\int_0^t r_s + h_s ds)dt\right]$$

The price of the bond is now equal to the price of the bond in case that the recovery rate is constant and equal to R_L plus the last term, which is the present value of the additional recoveries in case i defaults alone.

A similar formula holds for CDS spreads. Since $h^{i\bar{j}\overline{X}_{ij}} = h^{i\overline{X}_i}$, we can rewrite the CDS spread as

$$E[z_{ji}^{R_L,R_H}\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)dt]$$

$$= E[\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)\,\{h_t^{i\overline{jX}_{ij}}(1-R_H) - h_t^{i\overline{jX}_{ij}}(1-R_L)$$

$$+h_t^{i\overline{jX}_{ij}}(1-R_L) + h_t^{i\bar{j}X_{ij}}(1-R_L) + h_t^{ij}S(1-R_L)\}dt]$$

$$= E[\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)\,\{h_t^{i\overline{jX}_{ij}}(R_L - R_H)$$

$$+(1-R_L)(h_t^{i\overline{jX}_{ij}} + h_t^{i\bar{j}X_{ij}} + h_t^{ij}S\}dt]$$

Since

$$h_t^{i\overline{jX}_{ij}}dt + h_t^{i\bar{j}X_{ij}}dt + h_t^{ij}Sdt = h^{i\bar{j}}dt + h_t^{ij}Sdt = h_t^i dt - (1-S)h_t^{ij}dt$$

we can write:

$$E[z_{ji}^{R_L,R_H}\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)dt]$$

$$= E[\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)\,\{h_t^{i\overline{jX}_{ij}}(R_L - R_H)$$

$$+(1-R_L)(h_t^i dt - (1-S)h_t^{ij})\}dt]$$

or:

$$E[z_{ji}^{R_L,R_H}\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)dt]$$

$$= E[z_{ji}^{R_L}\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)dt] - Y_{CDS}$$

where

$$Y_{CDS} = (R_H - R_L)E[\int_0^T exp(-\int_0^t r_s + (h_s^i + h_s^j - h_s^{ij})ds)\,h_t^{i\overline{X}_i}\}dt$$

A first-order approximation of both Y_{bond} and Y_{CDS} around $h_s^i = h_s^j = h_s^{ij} = 0$ just as in Appendix C immediately shows that to the first order $Y_{bond} = Y_{CDS}$. As discussed above the addition to the *same* term to both bond yields and CDS spread does not affect the bounds.

D.5 - Assumptions about the hazard rate process

In this section I allow for a more flexible form for the hazard rate process. Assume that for each institution i, from the perspective of an agent pricing bonds at time t, the hazard rate will evolve from the current level, h_t, following the deterministic but time-varying process:

$$dh_s = \rho(\overline{h} - h_s)ds \quad s > t \tag{4}$$

Here, h_t (the current level of the intensity process), $\overline{h}$ (the long-run default intensity level) and ρ (the speed of convergence to the long-run intensity level $\overline{h}$) are known at time t when the bond is priced. For simplicity, I consider here a deterministic path for the hazard rate, from the perspective of time t, ignoring the effect of the variability in future hazard rates on bond prices. I also assume that agents potentially revise their estimates of ρ and $\overline{h}$ at each time t. Therefore, the cross-section of bond prices at time t contains information about the current intensity h_t as well as the time-t beliefs of the agents about the parameters ρ and $\overline{h}$: these are used by agents to price bonds of all maturities at time t.

This simple specification allows me to use the cross-section of bond prices to estimate h, ρ and $\overline{h}$ separately at each t, at the same time allowing for time variation in the slope of the term structure of default intensities (determined by ρ and $\overline{h}$). After estimating h_t and the implied time-t values of parameters ρ and $\overline{h}$, I construct the bounds on systemic risk using h_t as an upper bound to the marginal probabilities (linear inequality from bond prices).

As discussed in section 3, the main problem with this approach is that while it is easy to estimate a more flexible function for the marginal hazard rate of default using bond prices, CDS data do not contain enough information to estimate a similarly flexible process for joint default risk process h^{ij} (because at each time t we only observe consistently the spread of one CDS, with maturity of 5 years). To tackle this limitation, I assume that the joint hazard process replicates the shape of the marginal hazard process of the reference entity: the process decays at the same rate (ρ) and displays the same ratio between short-term and long-term default hazards ($h/\overline{h}$) as that estimated from bond prices. Therefore, I can use the 5-year CDS spread to back out h_t^{ij} at each time t. Appendix Table 1 shows the results obtained using these assumptions. Note that, just as for the baseline case, I employ a first-order approximation of CDS spreads analogous to that presented in appendix C. I also discretize all pricing formulas at the monthly horizon, just like in the baseline case. Because we are trying to extract the time series of the one-month-ahead default probabilities using bonds of maturities much longer than one month, the estimated hazard rates will be quite noisy. However, Appendix Table

1 shows that the main empirical results are still valid under these assumptions. While the level of the short-term joint default risk increases, the broad time-series behavior of the upper bounds on P_1 and P_4 is the same as in the baseline case.

D.6 - Using interest rate swaps as the risk-free rate

While swap rates may not be the appropriate rate to discount cash flows under risk neutral probabilities (because they are indexed to a risky reference, LIBOR, and because they contain counterparty risk), it is interesting to check how the results change if we use them in place of Treasury rates.[9] Because these rates are higher than the Treasury rates, and therefore result in a lower basis for all banks, we would expect the upper bounds on systemic risk to decrease noticeably. At the same time, one must remember that we are calibrating the time variation in the liquidity process to the basis of non-financial firms, and the level of the liquidity process to the basis of each bank in 2004. Therefore, the change in the risk-free rate will be offset by a corresponding decrease in the liquidity process (even though the offset is not exactly one to one). Appendix Table 1 shows that the change in the bounds is very small.

D.7 - Assumptions about the weighting of contributors in CDS contracts

As discussed in section 3, the bounds are computed under the assumption that the CDS spreads are obtained by averaging quotes obtained from all the other dealers in the sample. If some dealers do not post quotes at all times, the average spread observed will, in expectation, overweight dealers which send quotes in more frequently. In turn, this is most likely related to how active the dealer is in the CDS market.

While we cannot obtain directly estimates of the activity of the dealers (in terms of number of contracts written and volume of CDS protection sold), Fitch Ratings[10] reports a ranking of the top 5 counterparties by trade count (which in turn is very correlated with gross positions sold), for each year between 2006 and 2010. We might then think that because these dealers are more active, quotes are more likely to be obtained from them, and therefore the average CDS spread observed will in expectation reflect more their contribution. Given this, as a robustness test I compute bounds that overweight the top 5 institutions in the formula for CDS contracts. I consider two relatively extreme weighting schemes. In all of them, institutions ranked below 5 have the same weight (I do not have information about the relative ranking of these dealers). In the first weighting scheme, I compute the bounds assuming that the top 5 institutions are 5 times more likely than the other 10 to contribute quotes, and therefore their contribution

[9] I bootstrap the zero-coupon yield curve from the par swap rate curve of the different currencies using linear interpolation.

[10] Fitch Ratings, 2008, Global Credit Derivatives Survey.

is weighted 5 times more than the other institutions in the sample. The second weighting scheme again assumes that all institutions ranked 6-15 have the same weight, and the top dealer has 10 times their weight, the second dealer 8 times, and so on up to the 5th largest dealer (with a weight twice that of the smaller dealers).

The effect of this overweighting on the bounds of systemic risk is not immediate. Suppose, for example, that in the bounds computed under equal weighting, systemic risk comes from the joint default risk among top-5 banks. Then, increasing the weight on these banks will have the effect, everything else constant, of lowering the weighted CDS spreads. But this is not possible because the CDS spreads were chosen to match the observed ones. Therefore, the joint default risk among these banks will have to decrease. At the same time, joint default risk with smaller banks can increase. But if these smaller banks were contributing little to default risk before the change in weights, an increase in the possibility of joint default risk with them might not make up for the reduction in maximum systemic risk coming from the top-5 dealers. In this example, systemic risk will likely decrease when we overweight top-5 dealers. It is easy to see that the opposite is true if systemic risk mainly comes from non top-5 dealers.

If instead in both groups (top-5 and non-top-5 dealers) we find dealers with large contribution to systemic risk as well as dealers with small contribution to systemic risk, under equal weighting, the bounds will be relatively robust to changes in the weights. In fact, this is the case. The top 5 banks include both banks with high contribution to systemic risk as well as banks with low contribution to systemic risk, such as one or two European banks. Appendix Table 1 shows that under both weighting schemes the main results still hold.

This robustness test also allows us to say something about heterogeneity in collateral agreements across counterparties. All the results in the paper have been derived assuming that the recovery rate in case of double default, S, is the same for all banks. How do the main results change if instead (because of different collateral agreements and exposure to other shocks) the recovery rate is different across institutions? While we have no direct information about the expected recovery rates of each counterparty, it is easy to show that if the recovery rates S_j are different across counterparties j, the average quote reflects not an equally weighted average across j's of the joint default probabilities $P(A_i \cap A_j)$, but rather a weighted average $\sum w_j P(A_i \cap A_j)$, where $w_j = \frac{(1-S_j)}{(N-1)(1-\overline{S})}$ and $\overline{S} = \frac{1}{N-1}\sum_j S_j$. Therefore, given a certain average recovery rate $\overline{S}$, the joint default risk with counterparty j will be weighted more in the observed quote if j's recovery rate is lower. Now, it is reasonable to assume that more important counterparties (that have a larger volume of the business) are also the counterparties that are able to obtain less stringent collateral agreements – and therefore buyers of CDSs from them might obtain a lower recovery rate in case of double default. As a consequence, the robustness test presented in this section can also be interpreted as robustness to this case of heterogeneity in recovery rates.

D.8 - Assumptions about the exchange rate

The construction of the bounds on systemic risk involves the estimation of risk-neutral probabilities from bond prices and of joint default probabilities from CDS spreads. Using probabilities obtained from different securities to obtain risk-neutral probabilities of joint default requires additional assumptions if the securities are denominated in different currencies. In particular, while most bonds issued by American firms and the CDSs written on them are denominated in dollars, European firms issue several bonds in Euros and in other currencies, and the CDSs written on them are denominated in Euros.

To simplify the discussion, consider one-period bonds and CDSs written by banks i and j. Call m_{se} the stochastic discount factor of a US investor in state (s, e). Here, s indicates the default state of the banks i and j, so that it can take values i (only i defaults), j (only j defaults), ij (both default), and 0 (none defaults). e indicates the exchange rate with a foreign currency. Call π_s the probability of s occurring, and note that $\pi_s E[m_{se}|s]$ is the price of a security that pays 1 if default state s happens. The price of a state-contingent security that pays a unit of foreign currency if default state s happens is then $\pi_s E[em_{se}|s]$.

It is easy to see that a sufficient condition for correctly estimating risk-neutral default probabilities using bonds and CDSs denominated in different currencies (using the risk-free rates denominated in the respective currencies to discount cash flows) is that for each s:

$$\frac{E[e \cdot m_{se}|s]}{E[m_{se}|s]} = \frac{E[e \cdot m_{se}]}{E[m_{se}]} \tag{5}$$

which requires that the relative price of domestic and foreign risk-free securities is the same as the relative price of domestic and foreign state-contingent securities that pay off in the various default states. Of course, it is reasonable to assume that the relative price of dollar-denominated and foreign currency-denominated default-contingent securities might be different depending on the default state (think for example of a flight-to-quality to US securities if several banks default). As a robustness test for the validity of the bounds in case these conditions are violated, I perform the estimation exercise including only American firms, for which all bonds and CDSs are dollar-denominated. Appendix Table 1 shows that the results still hold for this subset of banks.

To derive equation (5), we can start by considering the price of a bond issued by i denominated in different currencies. Consider the variable d_i which indicates i's default, and is 1 if $s = i$ or $s = ij$, and 0 otherwise. Cal $\pi_{d_i=1}$ the probability that i defaults.

Note that

$$\pi_{d_i=1}E[m_{se}|d_i = 1] + \pi_{d_i=0}E[m_{se}|d_i = 0] = E[m_{se}]$$

For a dollar-denominated risky bond (R is the recovery rate), the dollar price is:

$$p_i^{\$} = \pi_{d_i=0}E[m_{se}|d_i = 0] + R\pi_{d_i=1}E[m_{se}|d_i = 1] = E[m_{se}] - (1 - R)\pi_{d_i=1}E[m_{se}|d_i = 1]$$

Now consider a euro-denominated bond issued by the same firm, and of equal seniority. Calling e_0 the time-0 exchange rate, we obtain:

$$p_i^E e_0 = \pi_{d_i=0}E[e \cdot m_{se}|d_i = 0] + R\pi_{d_i=1}E[e \cdot m_{se}|d_i = 1]$$

$$= E[e \cdot m_{se}] - (1 - R)\pi_{d_i=1}E[e \cdot m_{se}|d_i = 1]$$

The prices of the respective risk-free securities are:

$$t^{\$} = E[m_{se}]$$

$$t^E e_0 = E[e \cdot m_{se}]$$

Combining defaultable and risk-free bonds we get:

$$p_i^{\$} = t^{\$}\left(1 - (1 - R)\pi_{d_i=1}\frac{E[m_{se}|d_i = 1]}{E[m]}\right)$$

$$p_i^E e_0 = t^E e_0\left(1 - (1 - R)\pi_{d_i=1}\frac{E[e \cdot m_{se}|d_i = 1]}{E[e \cdot m_{se}]}\right)$$

We can then use either bond to estimate the risk-neutral probability of default of firm i

$$P(A_i) = \pi_{d_i=1}\frac{E[m_{se}|d_i = 1]}{E[m_{se}]}$$

discounting cash flows by the appropriate risk-free rate as long as the following condition holds:

$$\frac{E[e \cdot m_{se}|d_i]}{E[m_{se}|d_i]} = \frac{E[e \cdot m_{se}]}{E[m_{se}]} \tag{6}$$

which requires that the relative price of domestic and foreign risk-free securities is the same as the relative price of domestic and foreign state-contingent securities that pay off if i defaults.

Now, consider the case of a CDS written by one bank on the default of another bank. The CDS is written on a european bank (i) but the counterparty (j) is american. The contract is denominated in euros.

The CDS contract costs z_{ji} euros. So we must have

$$z_{ji}e_0 = E[e \cdot m_{se}] \left((1-R)\pi_i \frac{E\left[e \cdot m_{se}|s=i\right]}{E[e \cdot m_{se}]} + (1-R)S\pi_{ij} \frac{E\left[e \cdot m_{se}|s=ij\right]}{E[e \cdot m_{se}]} \right)$$

Therefore, as long as the european yield curve is used to discount cash flows for euro-denominated CDSs the sufficient condition is:

$$\frac{E[e \cdot m_{se}|s]}{E[m_{se}|s]} = \frac{E[e \cdot m_{se}]}{E[m_{se}]} \tag{7}$$

for every default event in s. Note that equation (7) implies (6) so it is a sufficient condition for both bonds and CDSs of different currencies.

D.9 - Using only larger transactions from TRACE

A concern with using bond prices from Bloomberg is that they might incorporate stale information (for European bonds, for which I use quoted prices), or they might depend on very small trades, which might be less reflective of credit risk (see for example Dick-Nielsen et al. (2010)). To make sure results are robust to these problems, I compute the bounds for the subset of US firms using only transaction data from TRACE, and ignoring all trades with nominal amounts of less than \$100,000. Of course, this will exclude several bonds for several days. Appendix Table 1 reports that the bounds change very little.

Appendix Tables

Appendix Table 1a: max P1

Model		2007	Jan 2008 to Bear	Bear to Lehman	Month after Lehman	Oct 2008 to April 2009	After April 2009
		1/1/07	1/1/08	3/16/08	9/15/08	10/16/08	5/1/09
R	S	12/31/07	3/15/08	9/12/08	10/15/08	4/30/09	6/10/10
0.10	0.10	50.5	178.8	168.5	295.5	221.6	132.8
0.10	0.30	50.5	178.8	168.5	295.5	221.6	132.8
0.10	0.40	50.5	178.8	168.5	295.5	221.6	132.8
0.10	0.70	50.5	178.8	168.5	295.5	221.6	132.8
0.10	0.90	50.5	178.8	168.5	295.5	221.6	132.8
0.10	1.00	50.5	178.8	168.5	295.5	221.6	132.8
0.30	**0.30**	**64.9**	**229.9**	**216.7**	**380.0**	**284.9**	**170.7**
0.30	0.40	64.9	229.9	216.7	380.0	284.9	170.7
0.30	0.70	64.9	229.9	216.7	380.0	284.9	170.7
0.30	0.90	64.9	229.9	216.7	380.0	284.9	170.7
0.30	1.00	64.9	229.9	216.7	380.0	284.9	170.7
0.40	0.40	75.7	268.3	252.8	443.3	332.4	199.2
0.40	0.70	75.7	268.3	252.8	443.3	332.4	199.2
0.40	0.90	75.7	268.3	252.8	443.3	332.4	199.2
0.40	1.00	75.7	268.3	252.8	443.3	332.4	199.2
Using swap rates		64.9	230.1	216.8	380.3	285.1	170.8
US banks		49.5	167.3	156.7	274.9	182.8	101.9
US banks, larger trans		46.4	167.4	156.2	270.5	182.6	96.3
Reweight top 5 banks		65.1	229.9	216.7	380.0	285.3	171.2
Reweight, decreasing		65.1	229.9	216.7	380.0	285.3	171.2
Alternative bond model		34.7	120.2	164.7	665.4	242.7	53.5

Appendix Table 1b: max P4

Model		2007	Jan 2008 to Bear	Bear to Lehman	Month after Lehman	Oct 2008 to April 2009	After April 2009
		1/1/07	1/1/08	3/16/08	9/15/08	10/16/08	5/1/09
R	S	12/31/07	3/15/08	9/12/08	10/15/08	4/30/09	6/10/10
0.10	0.10	2.3	1.8	16.2	23.8	49.3	31.1
0.10	0.30	2.4	2.0	18.0	24.7	49.1	31.4
0.10	0.40	2.5	2.2	18.8	25.0	48.9	31.3
0.10	0.70	2.7	2.8	21.1	25.3	47.7	30.7
0.10	0.90	3.2	3.2	22.3	25.9	46.5	29.9
0.10	1.00	12.6	44.7	42.1	69.6	55.4	33.2
0.30	**0.30**	**3.2**	**2.8**	**25.2**	**37.8**	**65.3**	**40.6**
0.30	0.40	3.3	3.0	26.3	38.2	64.8	40.5
0.30	0.70	3.5	3.7	28.8	39.5	62.6	39.6
0.30	0.90	3.9	4.2	30.2	41.7	61.0	38.4
0.30	1.00	16.2	57.5	54.2	89.5	71.2	42.7
0.40	0.40	3.9	3.7	32.7	48.6	77.2	47.3
0.40	0.70	4.2	4.4	35.5	49.5	74.2	46.1
0.40	0.90	4.7	5.3	36.5	50.4	72.0	44.8
0.40	1.00	18.9	67.1	63.2	104.4	83.1	49.8
Using swap rates		2.0	1.9	19.2	38.6	58.4	28.1
US banks		1.2	0.5	11.7	31.9	34.7	17.1
US banks, larger trans		1.7	0.6	16.7	38.1	43.4	18.3
Reweight top 5 banks		1.5	1.1	24.3	36.9	69.2	42.3
Reweight, decreasing		1.6	1.3	24.8	36.0	70.3	42.6
Alternative bond model		4.6	7.0	25.1	83.9	62.7	8.0

Appendix Table 1c: min P1

Model			Average level of the bounds (bp per month)					
			2007	Jan 2008 to Bear	Bear to Lehman	Month after Lehman	Oct 2008 to April 2009	After April 2009
		Start	1/1/07	1/1/08	3/16/08	9/15/08	10/16/08	5/1/09
R	S	End	12/31/07	3/15/08	9/12/08	10/15/08	4/30/09	6/10/10
0.10	0.10		42.6	164.9	121.7	221.9	109.8	62.3
0.10	0.30		42.0	164.3	117.4	217.3	101.3	56.1
0.10	0.40		41.8	163.8	115.0	214.7	97.0	53.0
0.10	0.70		40.9	162.2	107.2	205.8	85.8	44.4
0.10	0.90		39.1	160.6	101.1	202.8	80.6	40.3
0.10	1.00		7.4	27.3	25.1	84.7	41.3	22.6
0.30	**0.30**		**53.9**	**210.3**	**146.4**	**252.2**	**125.9**	**71.3**
0.30	0.40		53.6	209.5	143.1	248.0	120.4	67.3
0.30	0.70		52.5	206.6	132.9	237.4	106.5	56.6
0.30	0.90		50.9	204.8	125.3	232.8	99.4	51.6
0.30	1.00		9.5	35.1	32.2	108.9	53.1	29.0
0.40	0.40		62.4	243.7	161.9	281.8	135.8	77.9
0.40	0.70		61.2	241.3	149.8	273.5	120.1	65.6
0.40	0.90		59.7	238.9	142.3	268.4	111.1	59.7
0.40	1.00		11.1	40.9	37.6	127.0	61.9	33.8
Using swap rates			56.6	217.0	156.1	258.2	141.5	97.8
US banks			43.3	157.9	117.0	187.0	110.6	62.3
US banks, larger trans			39.8	154.1	109.8	168.8	100.6	57.4
Reweight top 5 banks			50.4	204.8	128.2	243.5	121.5	67.5
Reweight, decreasing			50.9	204.4	128.2	249.7	121.1	64.6
Alternative bond model			25.2	92.4	107.5	397.8	121.3	26.6

Note: Table reports the average value of the bounds on monthly P1 (probability that at least one bank defaults) and P4 (probability that at least four banks default) for different nonoverlapping periods, under different assumptions discussed in the text. The lower bound for P4 is 0 throughout.