



UNIVERSITY OF AMSTERDAM

## UvA-DARE (Digital Academic Repository)

### Information processing in complex networks

Quax, R.

**Publication date**  
2013

[Link to publication](#)

#### **Citation for published version (APA):**

Quax, R. (2013). *Information processing in complex networks*. [Thesis, fully internal, Universiteit van Amsterdam].

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## **Appendix B**

### **Appendix to Chapter 3**

#### **B1 How IDL differs from other leading indicators**

The primary difference between IDL and other the alternative leading indicators analyzed here is that it filters out external correlation at each time point. The IRS prices across maturities are financial indicators which not only correlate amongst themselves, but may also correlate with external indicators such as the house-price index (HPI). A standard correlation computed between IRS prices therefore consists of two parts: their interdependence (cause-and-effect) and the external correlation they have in common. This is the case in the previously introduced spatial leading indicators (145, 176, 178, 181, 221).

As the causal relation between two IRS prices becomes more indirect (larger difference in maturity), the interdependence vanishes to zero, and the external correlation that they all have in common remains. IDL estimates the magnitude of this common external correlation at each time point, and computes the characteristic length scale of the decay of interdependence on top of it. This idea is illustrated in Figure 1c. It is analogous to the classical concept of correlation length, where instead the mutual information is used as correlation measure.

The second difference is the use of the mutual information measure to compute correlations between time series. Typically, Pearson-like and other linear correlation measures are used to construct leading indicators. In contrast, the mutual information measure is capable of detecting various non-linear forms of correlation as well (222).

## **B2 IDL in IRS data in the EUR market**

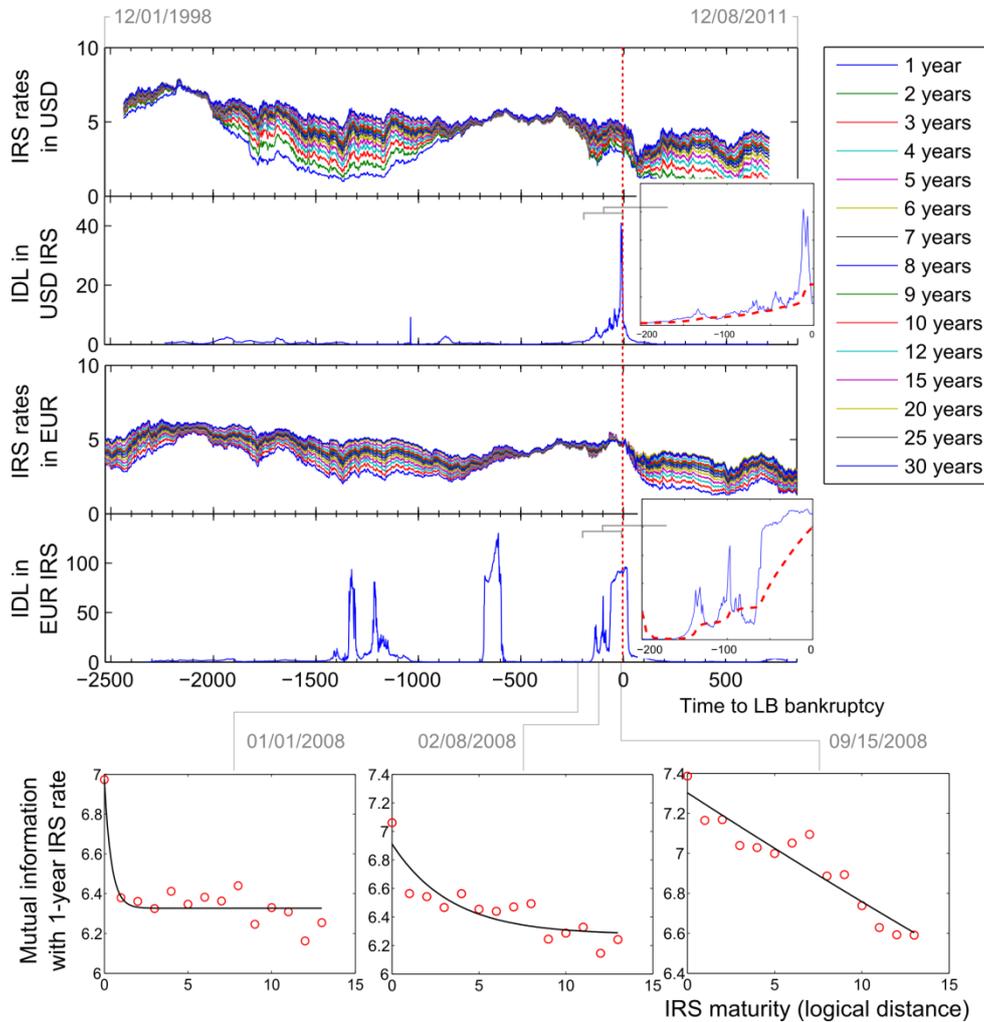
### **B2.1 The EUR data**

We have also analyzed time series of IRS rates of different maturities for the EUR currency. The EUR data range from 12/01/1998 to 12/08/2011, which encompasses the USD data which ranges from 04/29/1999 to 06/06/2011. The EUR data corresponds to IRSs with yearly fixed payments in exchange of semi-annual variable payments, which is different from the USD data where the variable payments are quarterly. The maturities in the EUR data are 2, ..., 10, 12, 15, 20, 25, and 30 years, i.e., they are the same as in the USD data except that it misses the 1-year maturity.

### **B2.2 Results in the EUR market**

In Figure B17 we show the original time series of IRS rates with the corresponding values of IDL, for the USD and EUR markets. The upper panel is equal to Figure 7 in the main text for comparison. In both cases the day of the Lehman Brothers bankruptcy is preceded by a significant increase of IDL of one and two orders of magnitude, respectively, and drops abruptly after the bankruptcy.

We also find that both IDL curves could have been used as an early warning signal. As in the main text we set the warning threshold at three standard deviations above the mean of a sliding window of 400 trade days. The earliest warning in the EUR market is more pronounced than in the USD market and starts at 161 trade days in advance, lasting for 34 trade days, followed by two additional extended warning periods.



**Figure B17:** The original time series in both the USD and EUR currency of the IRS rates for different maturities and the corresponding IDL indicator. The USD panel is the same as Figure 7 in the main text for comparison. Inset: the IDL indicator and a warning threshold during the 200 trade days leading up to the LB bankruptcy. The warning threshold is three standard deviations above the mean of a sliding window of 400 trade days. Bottom: the mutual information between the rates of the 1-year maturity IRS and all other maturities at three different dates in 2008, for the EUR data. The fitted exponential decay is used to estimate the IDL value for each trade day.

### **B2.3 Additional peaks in the EUR market**

The USD data is highly specific to the Lehman Brothers bankruptcy because its IDL peaks only once in twelve years. The EUR data contains three peaks in its IDL in the same period. Because of the exceptional magnitude and impact of the Lehman Brothers bankruptcy it seems reasonable to assume that it explains the two peaks that coincide with it.

This leaves two unexplained peaks of IDL in the EUR market. Although finding the underlying causes is highly speculative, it is important to evaluate the possibility that the IDL detected an increased instability of the financial market. Therefore we present three major events that coincided with the two peaks.

The first critical period starts in April 2003 and ends in July 2004. The first major event at this time was the largest simultaneous expansion of the E.U. by ten countries. The treaty was signed on April 16, 2003 and the expansion was completed on May 1, 2004. The critical period started in the same month and ended three months later. The second event was the uncovering of the largest corporate fraud in Europe's history by Parmalat, which filed for bankruptcy on December 24, 2003. This corresponds to trade day -1233 in Figure B17 and marks the transition of a medium IDL (~10) to a rapid increase to a high IDL (~80) 15 trade days later.

The second critical period starts in January 2006 and ends in June 2006. At this time the house price bubble emerged (223) because many homeowners became unable to pay their mortgage debts. The U.S. season-adjusted house-price index had been growing at an increasing rate from 1991 to 2006 (224), which stimulated the sale of low-rate mortgages based on the premise that the house prices would keep growing. Around the year-end of 2005, however, the growth stopped increasing and in March 2006 the growth had its largest drop to below zero. The house-price indices in Europe follow the U.S. trend quite closely (225, 226). If the house price bubble caused the IDL peak in the EUR data then there is possibly a false negative in the USD data, but it may also be explained by other factors such as a difference of

effectiveness of the policies and responses of the corresponding central banks.

Note that a critical transition may not have occurred at these times: while a high IDL indicates the potential for a system-wide abrupt change, it does not guarantee that it will actually happen. The built up stress in the financial markets may also have been detected through other indicators and relieved through measures such as artificially decreasing the interest rates by central banks.

In the pessimistic scenario that both additional peaks are false positives, i.e., the financial markets were in fact not capable of a system-wide critical transition at that time, the IDL indicator would still be accurate in half of its warnings.

### **B3 Robustness of IDL as a leading indicator**

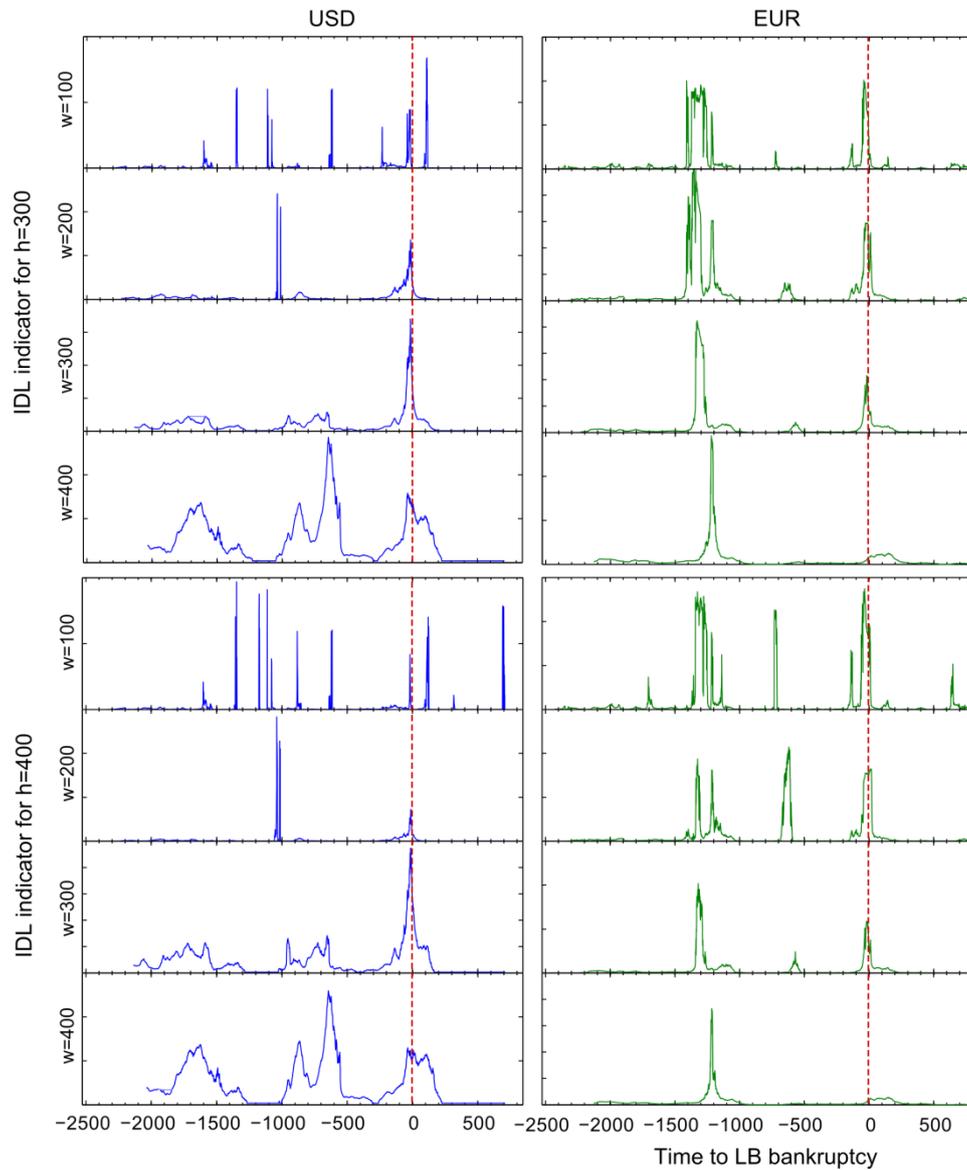
Estimating the IDL requires two parameters: the size of the sliding window  $w$  and the size of a bin  $h$  in the contingency table. At each time point we use the preceding  $w$  price values in two time series in order to estimate their mutual information. The bin size determines the price equivalence relation, i.e., which price values are considered equal in each sliding window. This is necessary to calculate the mutual information using the discrete version of Eq. (1), i.e., to estimate the joint probability distribution of the two price values using a finite set of observations. In financial terms, we partition the data histogram into ranges of  $h$  ‘basis points’.

The higher the sliding window size  $w$ , the more accurate can mutual information be estimated but the less sensitive it is to detecting short-term events or sudden changes. Therefore,  $w$  should be as low as permitted by the accuracy of calculating the mutual information.

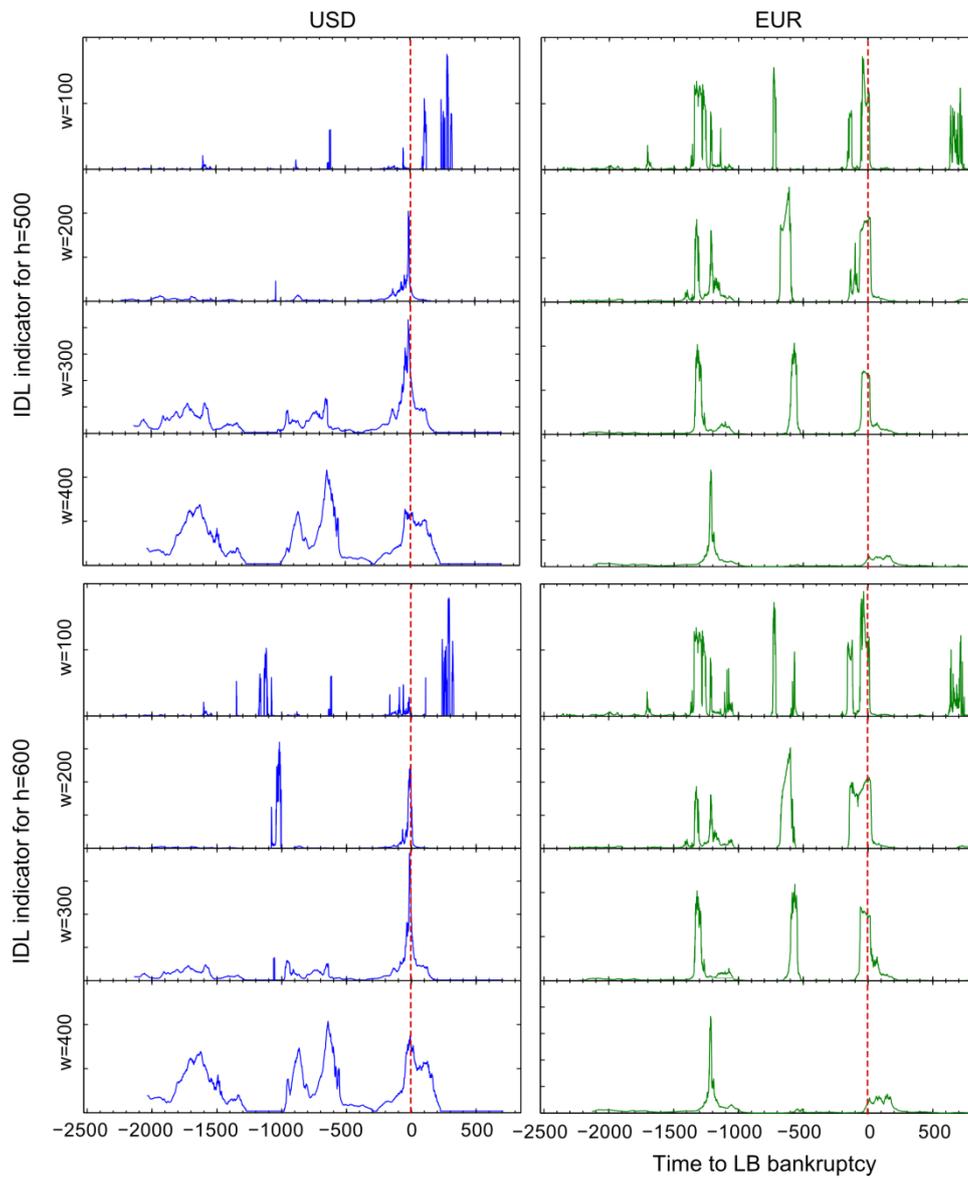
If the bin size  $h$  is too small then no price values will be considered equal, which means that each observed pair of prices is unique and the mutual information is invariably maximum. Increasing the bin size implies a lower sensitivity to small correlations of price fluctuations, so the bin size determines the magnitude of price changes that are correlated. Here too is a trade-off between accuracy and sensitivity. In the case of computing mutual information,  $h$  can be taken relatively small since the bins with zero occurrences do not change the mutual information as the  $p \log p$  for  $p \rightarrow 0$  is taken to be zero.

We show the IDL for a wide range of parameter values for  $w$  and  $h$  in Figure B18 through Figure B20. We show one value (100) for  $w$  which is ‘too small’, i.e., we observe many narrow spikes, and we also show one value (400) which is ‘too large’, i.e., the IDL curve becomes too gradual so that the threshold is no longer crossed. This shows in what range the window size should be chosen; since the smaller the better, we chose  $w = 200$  for Figure 7 in the main text.

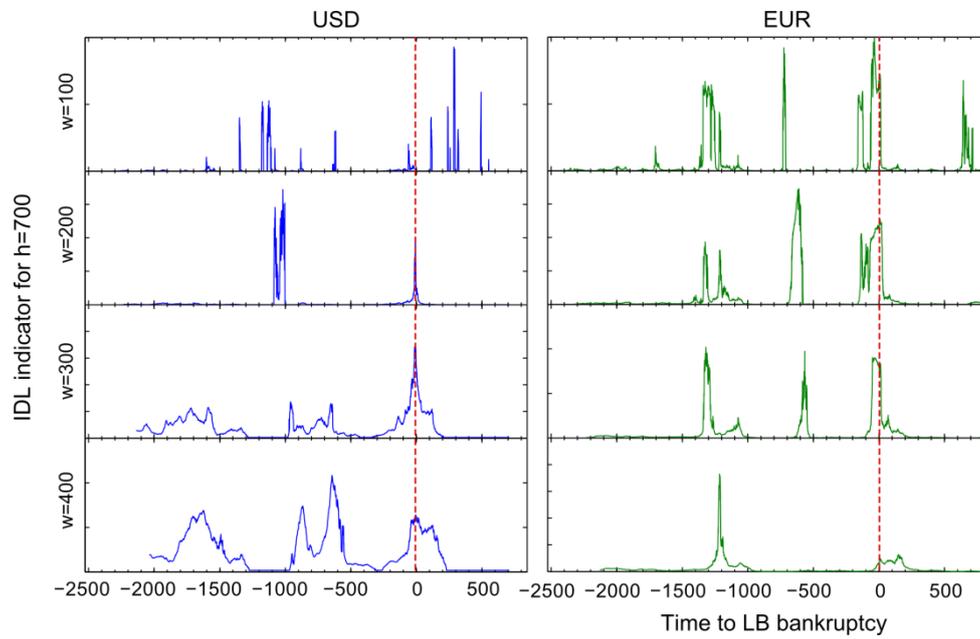
We also observe that the bin size  $h$  has a modest effect on the IDL curves, except that for some combinations of  $w$  and  $h$  a sudden spike of IDL occurs in the time range -1100 to -1000. This peak is not as consistent across the parameter values as the peak around Lehman Brothers, so we investigated further what could be the cause. We show in the next subsection that during this time period the IRS rates were suddenly recorded at the resolution of 0.05 basis points, whereas around that time it is recorded in 0.5 basis points. (During the Lehman Brothers bankruptcy it is recorded in 0.1 basis points.) It is clear that the manner of recording the IRS prices had changed in the time range -1100 to -1000, so we should be careful when interpreting the results during this period. This observation, combined with the erratic shape of the peak and its reduced consistency, leads us to choose the parameter value  $h=1/500$  for Figure 7 in the main text, which is also close to the recording resolution of  $1/1000$  around the Lehman Brothers bankruptcy.



**Figure B18:** The IDL indicator for IRS rates in the USD and EUR currency, for bin sizes of 1/300 and 1/400 percentage points and sliding window sizes 100, 200, 300, and 400. The IDL peak in the USD data for bin size 1/400 and sliding window size  $w=200$  still peaks significantly (to more than  $IDL=40$ , as in the main text): it appears smaller due to the erratic peak at time -1050 which is discussed in the text.



**Figure B19: The IDL indicator for IRS rates in the USD and EUR currency, for bin sizes of 1/500 and 1/600 percentage points and sliding window sizes 100, 200, 300, and 400.**



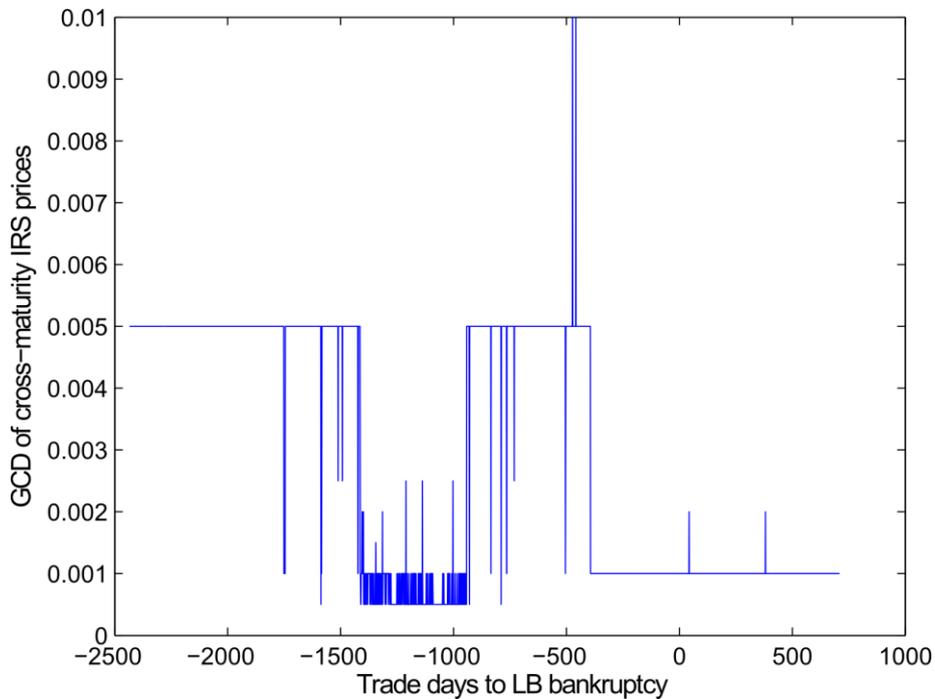
**Figure B20: The IDL indicator for IRS rates in the USD and EUR currency, for the bin size of 1/700 percentage points and sliding window sizes 100, 200, 300, and 400. The IDL peak in the USD data for bin size 1/400 and sliding window size  $w=200$  peaks to even more than 100: it appears smaller due to the erratic peak at time -1050 which is discussed in the text.**

### **B3.1 Choosing the bin size based on the data**

A bin size of, e.g.,  $h=1/500$  means that the IRS prices are divided into ranges of 0.002 percentage points, or 0.2 basis points. The choice for the bin size should be consistent with the data, that is, the bin size should be of the order of the size of the fluctuations. Also it is pointless to choose the bin size to be smaller than the significance of the data: if the data is accurate up to 0.5 basis points, then choosing  $h$  smaller than  $1/200$  would yield the same result as for  $h=1/200$ . This is because a smaller bin size would only create additional bins that are empty in the contingency table, and due to the common convention  $0\log 0=0$  these bins do not change the calculated mutual information of two vectors.

In Figure B21 we show the minimum greatest common divisor (GCD) between the 1-year IRS price and the other 14 IRS prices on each trade day in the USD data. From this figure we expect that  $h$  should be at least  $1/1000$  and at most  $1/100$ . Indeed we find experimentally in the above sensitivity analysis that  $h$  should roughly be in the range  $1/200$  to  $1/700$ . Also we see that in the first 500 or so trade days the choice of  $h$  does not change the IDL curves, consistent with the observation that the initially reported USD IRS data is only significant up to 0.5 basis points.

We also observe that during the trade days  $-1400$  through  $-900$  the USD IRS data is suddenly recorded at a resolution of 0.05 basis points. During this period we observe a contingent and erratic additional IDL peak in the USD IRS data for certain combinations of  $h$  and  $w$ . It is clear that the manner of recording the IRS prices had changed in this time range, so we should be careful when interpreting the results during this period, as we discussed above.



**Figure B21: The minimum greatest common divisor (GCD) between the 1-year IRS price and the other 14 IRS prices on each trade day in the USD data. This provides a rough guide to choose the parameter value  $h$ : if  $h$  is smaller than these GCDs then it has the same effect as choosing it as the smallest GCD; if  $h$  is larger than these GCDs then the contingency table may start to lose its ability to correlate fluctuations between the 1-year IRS price and the other 14 IRS prices in any sliding window. In this figure we calculated the minimum GCD of the cross-maturity USD IRS prices per 1 day; for a sliding window of e.g. 200 trade days, the corresponding GCD would be the minimum of the 200 preceding per-day GCDs.**

## **B4 Comparison with previously introduced leading indicators**

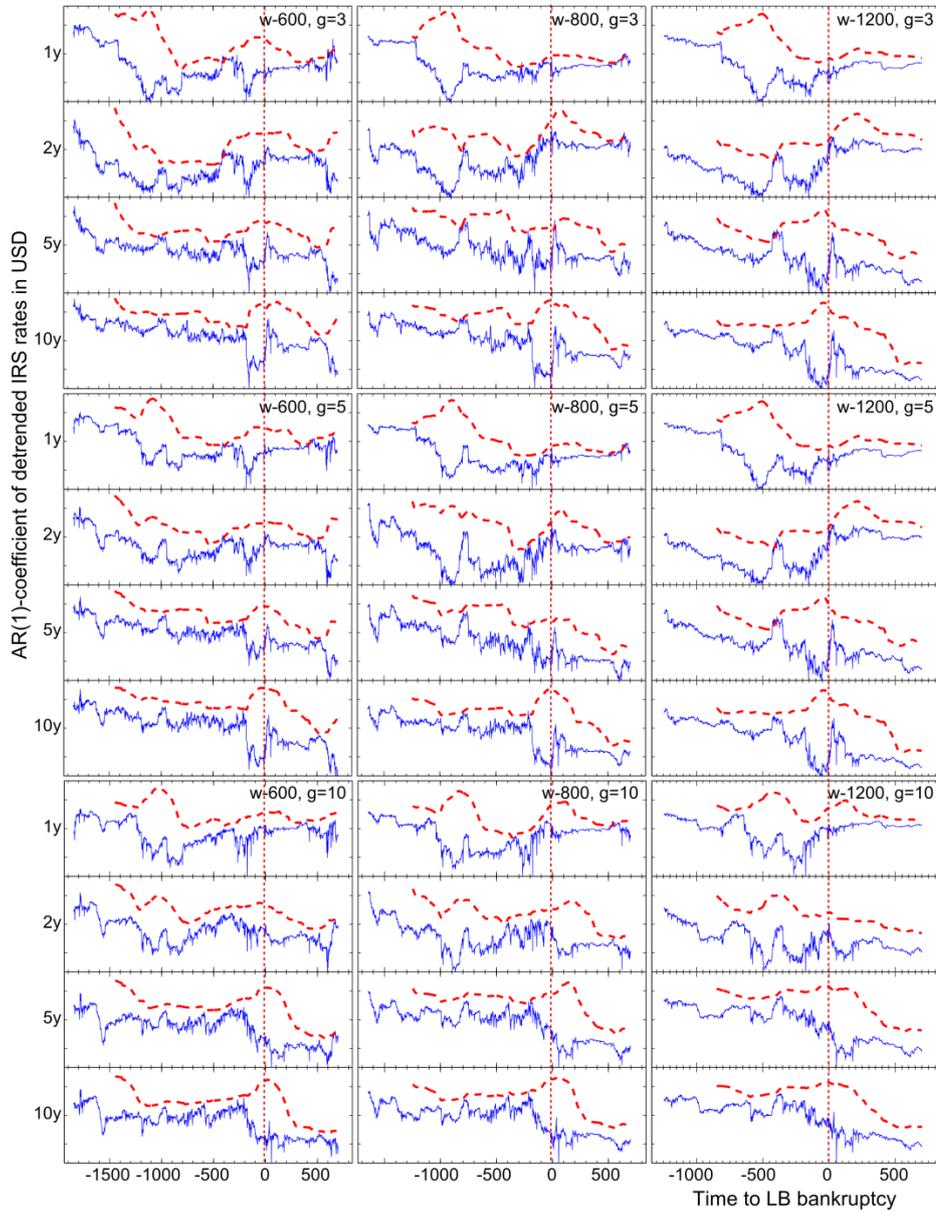
### **B4.1 Critical slowing down in IRS rates**

The effect of critical slowing down (173) can be measured by the coefficient of a first-order autoregression of the fluctuations of a signal (144, 145, 174, 176, 179, 182, 221). Calculating this coefficient requires two parameters: the size of the smoothing kernel, which de-trends the signal, and the size of the sliding window, which is used to compute the autoregression. Here we investigate whether there is a set of parameter values for which the critical slowing down can provide a clear leading indicator of the Lehman Brothers bankruptcy. In Figure 8 in the main text we show a representative set of results, where the smoothing kernel has a standard deviation of 5 trade days and the sliding window to compute the autoregression was 1000 trade days. Here we show the results for a wide range of parameter values. We do not find a clear warning for the Lehman Brothers bankruptcy for any combination of sliding window size and Gaussian smoothing kernel width.

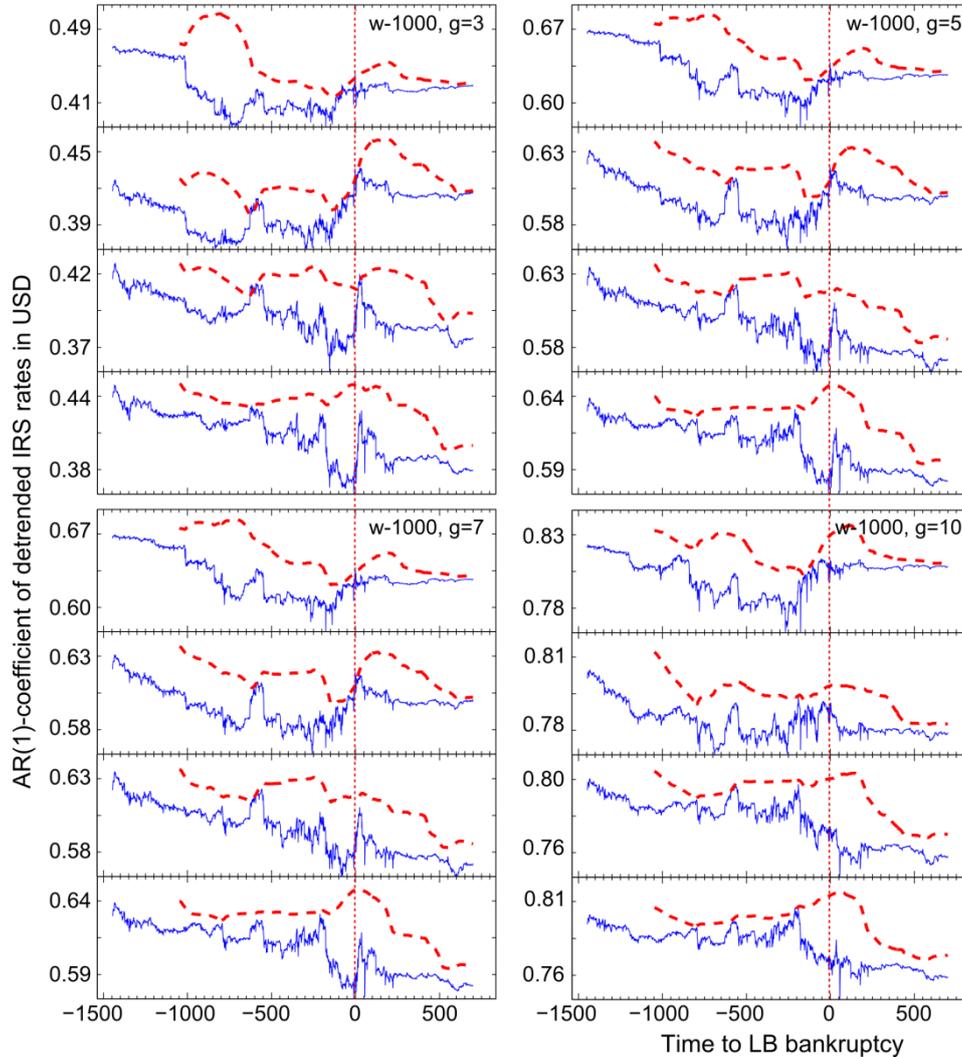
The smoothing kernel is used to filter long-term price trends from the time series. We use a Gaussian smoothing kernel, following e.g. Dakos et al. (174). The smoothing kernel is used to remove the long-term trend from a signal, because the effect of critical slowing down is detected in the short-term fluctuations of the time series: it is the time it takes for the price value to return to its long-term trend after a small perturbation. In effect we compute a running weighted average of each time series, where each price value becomes the weighted average of its neighbors, and subtract it from the original time series to obtain the de-trended signal. The weights are Gaussian distributed and the width of the distribution is the free parameter. Figure 8 in the main text was created using a Gaussian kernel with a standard deviation of 5 trade days. Here we show the first-order autoregression coefficient for the parameter values 3, 5, and 10.

At each time point the autoregression coefficient is calculated using the preceding  $w$  price values, where  $w$  is the size of the sliding window. The

higher the value of  $w$ , the more accurate can the coefficient be calculated but the less sensitive it is to short-term effects. The first drawback that we find is that the calculation of the coefficient requires a considerably larger sliding window than for calculating the IDL. Where the IDL indicator starts to be meaningful at a size of about 150 trade days, the autoregression coefficient requires a minimum window size of approximately 600 trade days. This problem has already been recognized by others (145, 176, 182, 221). Figure 8 in the main text was created using a sliding window of 1000 trade days and a Gaussian smoothing kernel with a standard deviation of 5 trade days. In Figure B22 and Figure B23 we show the first-order autoregression coefficient of the de-trended IRS rates for the sliding window sizes 600, 800, and 1200 trade days, a Gaussian smoothing kernel with standard deviations 3, 5, and 10 trade days, for a representative sample of maturities: 1 year, 2 years, 5 years, and 10 years.



**Figure B22: The critical slowing down indicator for the USD data for the sliding window sizes  $w=\{600,800,1200\}$  trade days (in Figure 8 we used  $w=1000$ ) and a Gaussian smoothing kernel with a standard deviation of  $g=\{3,5,10\}$  trade days. The red dashed curve is the warning threshold, computed at each time point as three times the standard deviation of the preceding 400 values above its mean.**



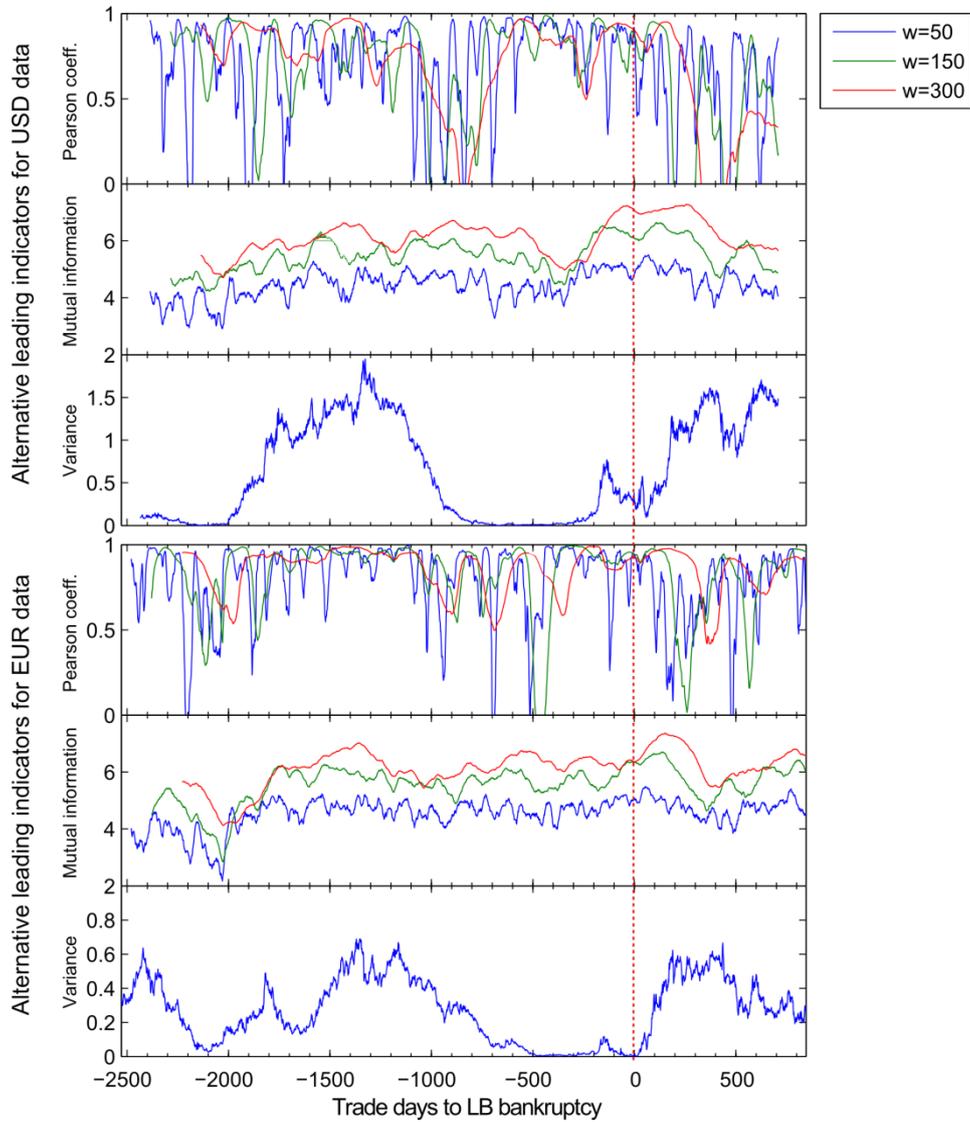
**Figure B23:** The critical slowing down indicator for the USD data for the sliding window size  $w=1000$  trade days, which was used in Figure 8 in the main text, and Gaussian smoothing kernels with standard deviations  $g=\{3,5,7,10\}$ , where  $g=5$  in Figure 8 in the main text. The red dashed curve is the warning threshold, computed at each time point as three times the standard deviation of the preceding 400 values above its mean. Time point 0 on the horizontal axis corresponds to the date of the Lehmann Brothers bankruptcy.

## B4.2 Spatial leading indicators

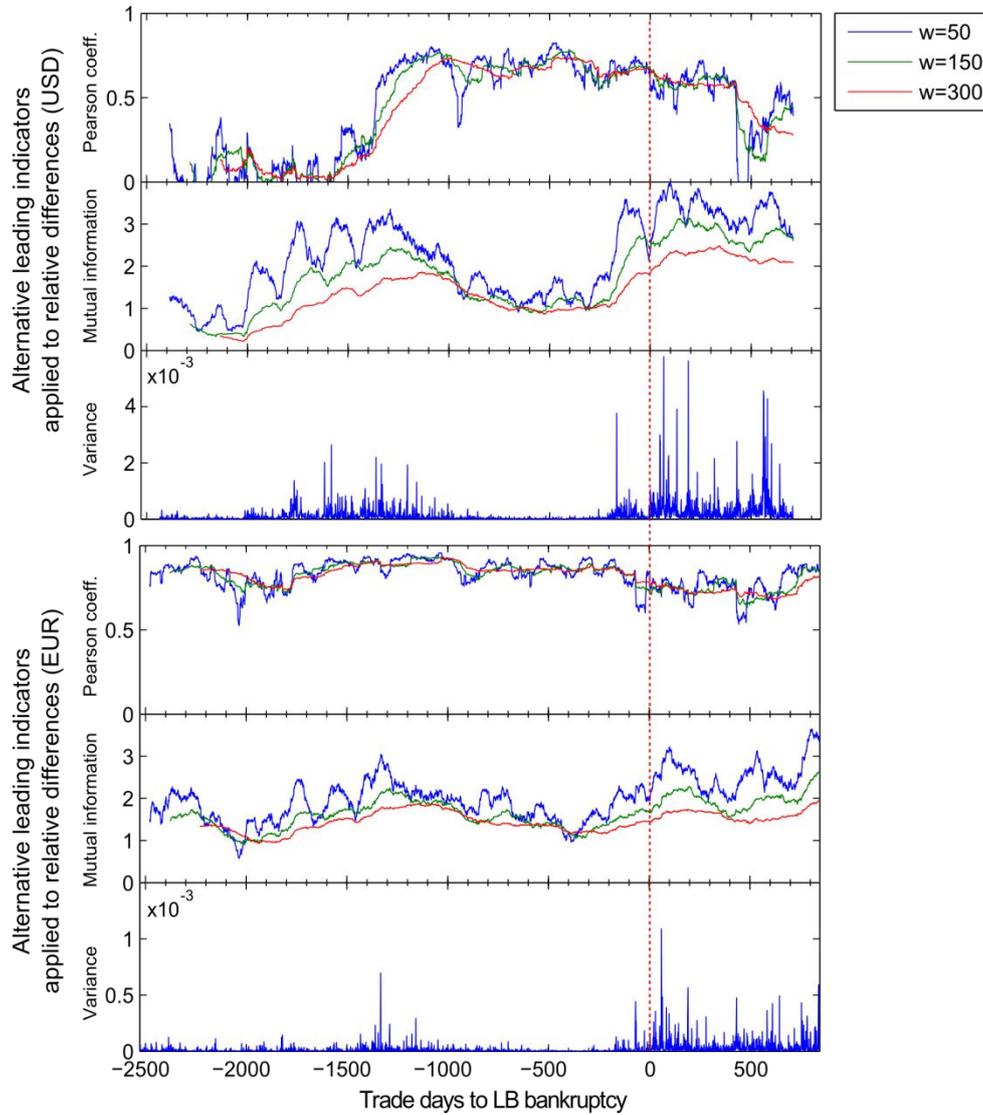
One of the prominent leading indicators reported on in the literature is the increasing spatial correlation among the units of a system; another is the spatial variance of the signals of the units (145, 178, 181, 221). Since there is a linear ‘spatial’ component in the prices of IRSs of increasing maturities, we investigate whether a spatial leading indicator could be used as an early-warning signal. In the following we qualify the linear dimension of maturities as ‘spatial’.

We calculate the spatial correlation and variance in our system for different sizes of the sliding window  $w$  and using Pearson’s linear correlation coefficient and mutual information. The sliding window sizes that we used are 50, 150, and 300 trade days; lower values yield erratic curves whereas higher values yield nearly constant curves. At each time point we calculate the correlation coefficient  $C_t = \langle F_{\text{corr}}(s_1^{t-w}, \dots, s_1^t; s_i^{t-w}, \dots, s_i^t) \rangle_i$  using the preceding  $w$  IRS rates of maturity 1 and maturity  $i$ ,  $i = 1, 2, \dots, 15$ , where  $F_{\text{corr}}$  is one of the correlation measures and  $s_i^t$  is the price of an IRS of maturity  $i$  at time  $t$ . In words, we use the average spatial correlation of all maturity prices with the price of a 1-year IRS, due to the logical ordering of the prices of the maturities as described in the main text. The variance  $\sigma^2(t)$  at time  $t$  is calculated at each time step as  $\sigma^2(t) = \sum_i (s_i^t - \langle s^t \rangle)^2$ , i.e., it is computed only of the IRS rates of the 15 maturities at time  $t$ . We show the results in Figure B24.

In finance, correlations are often calculated from the relative differences of time series instead of the absolute values (171). To investigate whether this has an effect on the ability of the alternative indicators to provide an advance warning, we first replace each original rate  $s_i^t$  by its relative difference (or ‘return’)  $(s_i^t - s_i^{t-1})/s_i^{t-1}$ . Then we calculate the same measures as in Figure B24. The results are shown in Figure B25.



**Figure B24: Alternative leading indicators for the IRS time series in both the USD and EUR data. For the Pearson correlation and mutual information we computed the correlations for sliding window sizes  $w=50$  (blue line),  $w=150$  (green line), and  $w=300$  (red line). Time point 0 on the horizontal axis corresponds to the date of the Lehmann Brothers bankruptcy.**



**Figure B25: Alternative leading indicators for the daily relative changes of IRS rates in both the USD and EUR data. Here, each original rate  $s_i^t$  is replaced by the relative difference  $(s_i^t - s_i^{t-1}) / s_i^{t-1}$  before the indicators are computed. For the Pearson correlation and mutual information we computed the correlations for sliding window sizes  $w=50$  (blue line),  $w=150$  (green line), and  $w=300$  (red line). Time point 0 on the horizontal axis corresponds to the date of the Lehmann Brothers bankruptcy.**



We conclude that none of the spatial measures provides an unambiguous leading indicator for the Lehmann Brothers bankruptcy. One possible explanation is that all IRS prices correlate strongly with external financial indices (such as the home-price index), which may dominate the observed correlations in the IRS prices across the maturities. In this scenario the IDL can still be a leading indicator because the calculation of IDL ignores the correlation that is shared among all IRS prices ('baseline information'). That is, the information in the IRS prices of different maturities decays as  $a + b \cdot (f^t)^{1-i}$  where  $a$  is the information (or correlation) shared among all IRS prices, and the estimated rate of decay  $f^t$  is independent of  $a$ .

#### **B4.3 The onset of a LIBOR-OIS spread as leading indicator**

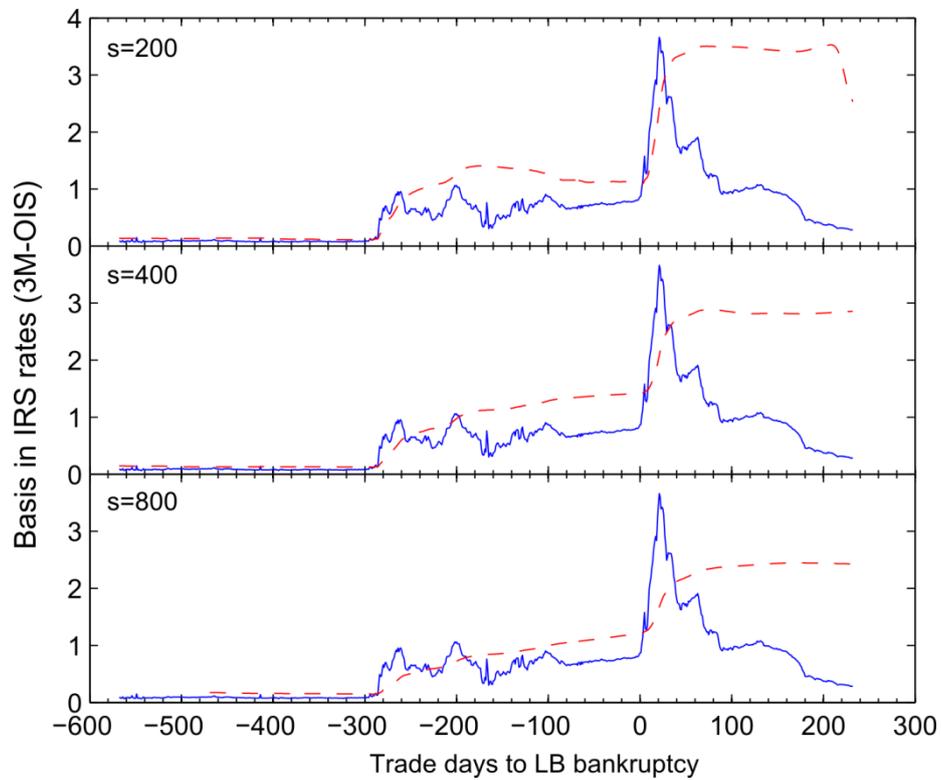
During the build-up of the recent crisis in 2007 a LIBOR-OIS spread, the so-called 'basis' in swap contracts emerged. The prices of the same swap but with different frequencies of variable payments had always been roughly equal, but around August 2007 the prices of swaps with less frequent payments started to increase. This is a very significant event that had never occurred before. The phenomenon is a symptom of calculating the risk of default of a financial institute, or in other words, a lack of trust in the stability of financial institutes (227). The price difference is essentially an insurance premium to compensate the risk that the variable interest payer would default during the swap contract's duration. Such a default had hardly been considered before 2007.

The question is whether the onset of a basis in interest rates can be used to anticipate the bankruptcy of Lehman Brothers. Therefore we interpret the basis as a leading indicator and test whether it could be used as an early warning for the bankruptcy. As warning threshold we use the same definition as for all other leading indicators, namely three times the standard deviation above the mean of a sliding window of  $s=400$  trade days. In addition, we test the case for a sliding window of  $s=200$  and  $s=800$  trade days. As a basis we use the daily differences between the 3-months (3M) LIBOR rates, which is based on a financial contract with one single

payment, and the 3-months overnight indexed swaps (OIS), which have daily payments. The first reason for taking the difference with the 3M swaps is that it is the most liquidly traded; the second reason is that it is the same frequency as the USD data analyzed in the main text and here.

A 3M-OIS swap is a different type of swap from an IRS which party A and B can negotiate. Party A pays  $x$  USD to B immediately after signing the contract, and B pays back  $x + \text{LIBOR} \cdot 3/12 \cdot x$  to A after three months. In such a swap there is no notional exchange; only one fixed rate is exchanged with one floating rate. This floating rate is the average daily interest rate cumulated over the period of three months. There is also more credit risk in such swaps: if B defaults then A loses its  $x$  USD.

See Figure B26 for the 3M-OIS interest rate and the three different warning thresholds. Although in each case there are two periods where a warning is issued, neither warning can be used to anticipate the Lehman Brothers bankruptcy. The first warning that lasts at least two days are at -291 trade days ( $s=200$ , lasting 33 days), -285 ( $s=400$ , lasting 29 days), and -285 ( $s=800$ , lasting 47 days). The basis may arguably be warning for a significant financial event around this time (227), however it is too early and too short-lasting to be used to anticipate the Lehman Brothers bankruptcy. The second warning is consistently too late (starting 3, 9, and 4 trade days after the bankruptcy), so it could be interpreted more as a consequence of the bankruptcy rather than anticipating it.



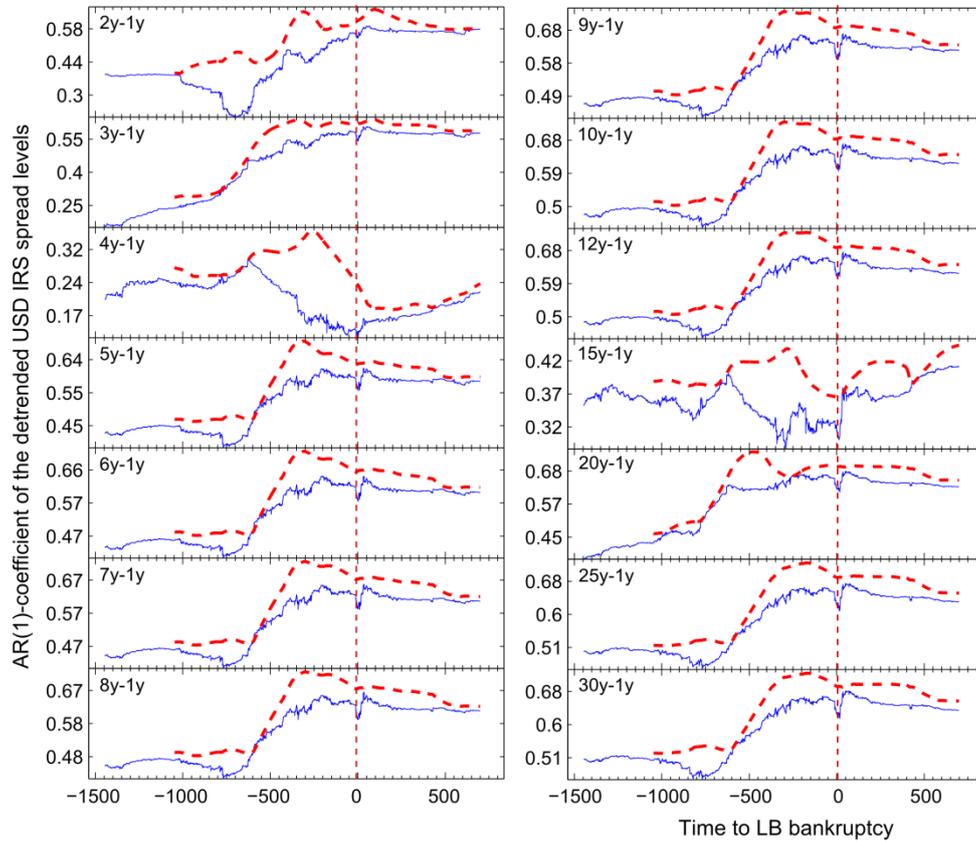
**Figure B26: The basis in the USD IRS rates in the time span March 2006 through November 2009. The red dashed curve is the warning threshold, computed as three times the standard deviation above the mean of a sliding window of  $s=200$ ,  $s=400$ , and  $s=800$  trade days respectively. As a basis we use the daily differences between the swap rates with a variable payments frequency of three months (3M) and overnight indexed swaps (OIS), which have a daily frequency. Time 0 on the x-axis corresponds to the day of the Lehman Brothers bankruptcy.**

#### **B4.4 Critical slowing down in IRS spread levels**

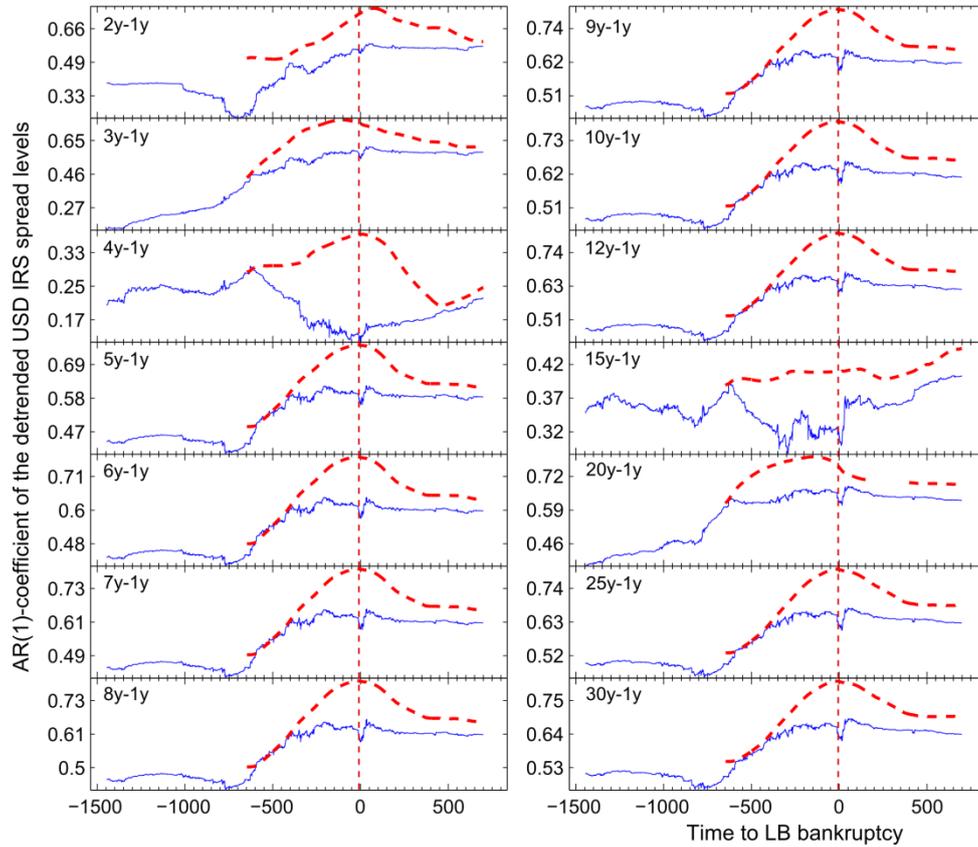
A traditional financial indicator is the ‘spread’ of (in this case) IRSs across maturities (183). It is already evident from the IRS prices plot in Figure B17 and the cross-maturity variance plot in Figure B24 that the spread levels themselves do not provide an early warning for the Lehman Brothers bankruptcy. Nonetheless, since they are often used as underlying indices in complex interest rate derivatives (‘spread options’ (171)), it is possible that the critical slowing down (CSD) indicator applied to the spread levels provides an early warning signal. We already showed in Section B4.1 that the CSD does not anticipate the bankruptcy when it is applied to the original IRS levels.

We compute the spread levels as the daily differences of IRSs of all maturities compared to IRSs with a 1-year maturity, all in USD. In other words, the smallest spread is the 2-year IRS rate minus the 1-year IRS rate, then the 3-year IRS rate minus the 1-year IRS rate, etc. Next we calculate the first-order autoregression coefficient in the same manner as in Figure 8 in the main text, that is, with a sliding window of 1000 trade days and a Gaussian smoothing kernel with a standard deviation of 5 trade days. To compute the warning threshold we take two sizes for the sliding window: 400 trade days and 800 trade days. 400 trade days is consistent with all other warning thresholds computed in the main text and here; 800 trade days was tested because the AR(1) signal turns out to grow quite gradually, so possibly it required a threshold that moves more gradual as well. The results are shown in Figure B27 and Figure B28.

Interestingly, the effect of critical slowing down is more apparent in the spread levels than in the original IRS rates, which were shown in Section B4.1. However, in the two years preceding the Lehman Brothers bankruptcy we find no warning for any combination of spread level and threshold window size.



**Figure B27: The first-order autoregression coefficient of all de-trended USD IRS spread levels (sliding window size 1000) and a warning threshold computed over a sliding window size 400. A caption such as ‘2y-1y’ denotes the spread level between a 2-year IRS and a 1-year IRS, i.e., the daily difference of their prices. Time 0 on the x-axis corresponds to the day of the Lehman Brothers bankruptcy.**



**Figure B28: The first-order autoregression coefficient of all de-trended USD IRS spread levels (sliding window size 1000) and a warning threshold computed over a sliding window size 800. A caption such as ‘2y-1y’ denotes the spread level between a 2-year IRS and a 1-year IRS, i.e., the daily difference of their prices. Time 0 on the x-axis corresponds to the day of the Lehman Brothers bankruptcy.**

## **B5 Verification of IDL using generated time series**

Calculating the IDL of generated time series allows us to address two questions. Firstly, does the IDL curve contain peaks even if the time series are uncorrelated (false alarms)? Secondly, does the IDL curve contain a peak at the time where we let the time series correlate?

### **B5.1 Generating the time series**

The generated data consists of 15 time series which consist of 3145 real-valued elements. These dimensions are equal to that of the USD IRS data so that the verification in this section is as comparable as possible to the real data used in the main text.

The first time series, i.e. the first ‘maturity’, is a copy of the 1-year USD IRS rates. Each subsequent  $i^{\text{th}}$  time series is a vector of random values, with a mean and standard deviation that equals that of the corresponding  $i^{\text{th}}$  maturity IRS data in the USD currency. The exception to this rule is a range of 500 elements starting at the 2000<sup>th</sup> element, where an artificial correlation is introduced as follows. The values of the elements 2000, ..., 2499 in the  $i^{\text{th}}$  time series, excluding  $i=1$ , is a randomized copy of the values of the  $(i-1)^{\text{th}}$  time series. The randomization is an added noise factor that has a zero mean and a standard deviation  $s$ , which is the free parameter.

We obtain a rough estimate of the range of values for  $s$  from the IRS data as follows. We denote the vector of USD IRS rates for the  $i^{\text{th}}$  maturity as  $r_i$ , and the average rate over time as  $\langle r_i \rangle$ . Since the calculation of mutual information ignores additive constants we consider the zero-mean time series  $r_i - \langle r_i \rangle$ . The best-fit model that assumes that the rates of maturity  $i+1$  are the rates of maturity  $i$  plus independently normally-distributed noise uses the standard deviation of the residuals  $r_{i+1} - \langle r_{i+1} \rangle - (r_i - \langle r_i \rangle)$  in order to predict the rates of subsequent maturities. For the USD IRS data, the expected standard deviation of residuals between subsequent maturities

is 0.1189, and its standard deviation is 0.0857. The values of  $s$  that we use in this section should be roughly in this range.

## **B5.2 Results**

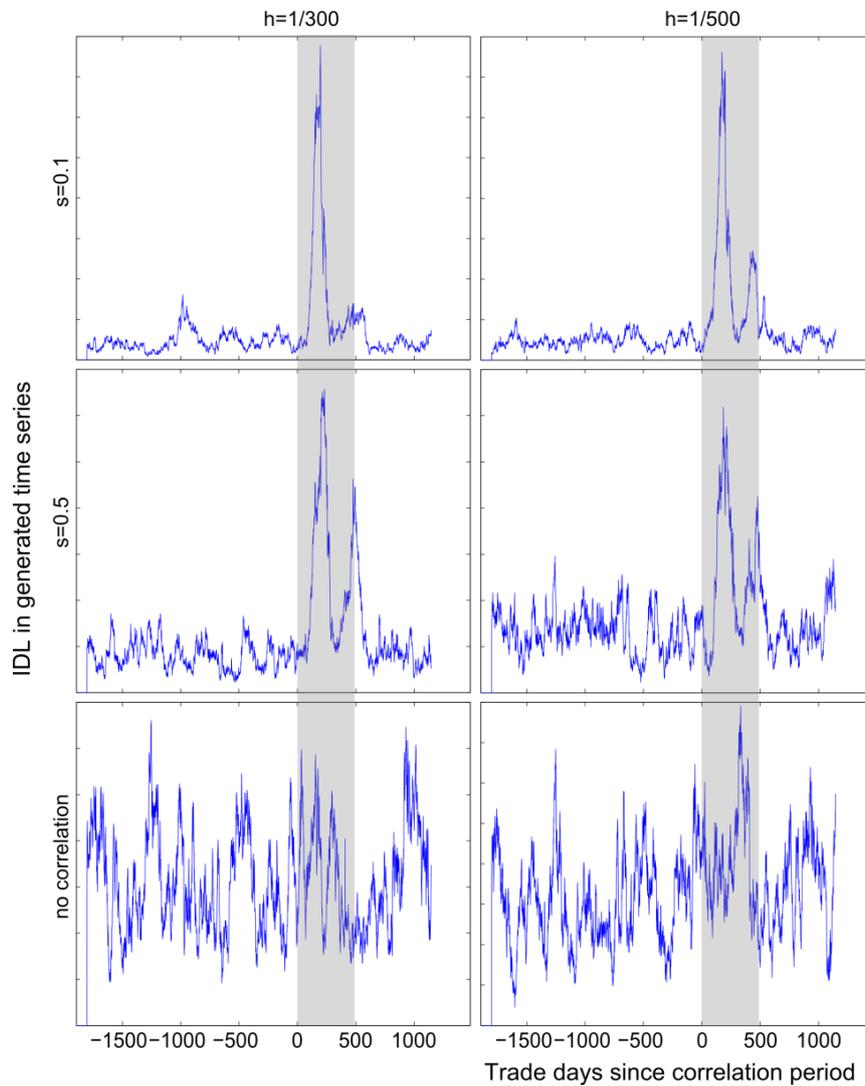
The lower the value of  $s$ , the higher the correlation between subsequent time series, and expectedly the higher the IDL. In other words, the higher the value of  $s$ , the less information about the first time series propagates through subsequent time series. Therefore we expect that the IDL peaks during the 500 correlated ‘trade days’ with a magnitude that decays for increasing  $s$ . In the absence of correlations, we expect that the IDL curve contains no discernible peaks.

In Figure B29 we show the IDL of the generated data for  $s = \{0.1, 0.5\}$  and for the case of uncorrelated time series. Here,  $s = 0.1$  represents a ‘reasonable’ correlation value and  $s = 0.5$  represents a ‘very low’ correlation value. The IDL is calculated using a sliding window size of  $w = 200$  and bin sizes  $h = \{1/300, 1/500\}$ , matching the parameters used in the main text. The x-axis is now renumbered so that time point 0 indicates the start of the 500 correlated elements (grey area).

Indeed we observe that the IDL peaks during the period of correlated elements among subsequent time series, confirming our hypothesis that the IDL is capable of detecting correlations between subsequent maturities. The delay of the peak compared to time point 0 is expected because a sliding window is used, and no gradual onset of correlation is generated. However, we do observe that the IDL indicator may decrease significantly within the correlation period, even toward its long-term average, suggesting that the IDL indicator may be prone to ‘false negatives’, i.e., the absence of a warning even though correlations arise between the time series.

Further we observe that the IDL does not contain significant peaks in the absence of correlations, suggesting that the IDL curve does not tend to

generate 'false positive' warnings, i.e., warnings in the absence of correlated time series.



**Figure B29: The IDL indicator computed of generated time series. The time series are uncorrelated except during the time points  $0, \dots, 499$ ; the higher  $s$ , the weaker the correlation. The sliding window size is  $w = 200$  and the bin sizes  $h = \{1/300, 1/500\}$ , corresponding to the parameter values used in the main text. The x-axis is now renumbered so that time point 0 indicates the start of the 500 correlated elements (grey area).**