Where is the Market?
Three Econometric Approaches to Measure Contributions to Price Discovery

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Wirtschafts- und Sozialwissenschaftlichen Fakultät
der Eberhard-Karls-Universität Tübingen

vorgelegt von
Franziska Julia Peter
aus Stuttgart

2011

Dekan:	Professor Dr. rer. soc. Josef Schmid
Erstberichterstatter:	Professor Dr. rer. pol. Joachim Grammig
Zweitberichterstatter:	Professor Dr. rer. pol. Martin Biewen
Tag der mündlichen Prüfung:	01. April 2011

To whom, if not to you.

Acknowledgements

Thanks to my parents for their never-ending support and trust,
To my little sister for teaching me how to surf the waves,
To my brothers, for finally remembering my birthday,
And to my grandparents for being wise and open-minded.

Thanks to my supervisor for picking me up from the econometric gutter,
Shoving me into conferences all around the world,
And making me realize that "So what?" is not an appropriate answer to a stupid question.

Thanks to my co-authors for enduring to work with me,
To Jo, for making up new words and teaching me the importance of being persistent,
To Kerstin, my kindred spirit, for the scarves and always calling back.

Thanks to my colleagues,
To Tommy for his heartily hugs,
To Stephan for his glass being always half full,
To Luis for finally attributing some humanity to me,
To Ramona for our joint analysis of the complexity of the past,
To Tobi for uttering pieces of wisdom,
To Oli, for pointing out that sometimes there is nothing more important than the next soccer game.

To all the student assistants for their delicious cakes and for breathing fresh air into the daily routines of a working day.

Franziska Julia Peter, Tübingen 03. April 2011

Contents

List of Figures

List of Tables

1 Introduction

The question "Where is the market?" has been of interest to researchers for over a decade. It was first addressed in the early 1990s, when a number of studies examined stocks that were listed on several regional U.S. exchanges. The simultaneous trading of one stock on several exchanges gave rise to competition among the different trading venues and measuring the importance of each market for the price discovery process of the common stock became the subject of research (see Hasbrouck 1995, Harris et al. 1995). In the following years the competition intensified due to the rapidly growing number of internationally cross-listings and the question "Where is the market?" went global. In particular for smaller stock exchanges, the large U.S. markets posed a threat, since they might take over the price discovery process of the dually listed stocks and thereby diminish the importance of the respective home market. In recent years the focus of price discovery studies has shifted away from stock markets and spread to various fields in empirical finance. Whether in the case of commodity markets (see Figuerola-Ferrett and Gonzalo 2010), the treasury market (see Mizrach and Neely 2008) or newly developed derivative markets such as the one for credit default swaps (see Blanco et al. 2005), the question which market leads price discovery is always one of the first to be asked.

However, the huge and diversified amount of empirical research into measuring contributions to price discovery is not mirrored by an adequate number of studies concerning the methodological aspects. The two prevalent measures both date back to 1995, Hasbrouck's (1995) information shares and the Gonzalo and Granger (1995) approach. In 2002, the *Journal of Financial Markets* devoted an issue to measuring contributions to price discovery, in which both approaches are compared and critically evaluated (Journal of Financial Markets, 2002, Vol. 5, Issue 3). Since then only very few advances have been made with regard to improving the methodologies used or proposing innovative approaches to measure contributions to price discovery.

This thesis presents three methods that either resolve drawbacks of the standard methodologies or offer new approaches to quantify contributions to price discovery. In Chapter 2, the standard approaches are summarized and discussed. Since a number of studies point out that the method developed by Hasbrouck's (1995) has more economic appeal than the Gonzalo and Granger (1995) approach and it further is the method that is more frequently

applied, emphasize is put on the former. The main drawback of the Hasbrouck (1995) information shares is that it does not deliver a unique measure. The method is based on a Vector Error Correction Model that involves an identification problem concerning the contemporaneous effects of price innovations. The solution proposed by Hasbrouck (1995) results in upper and lower bounds for a market's contribution to price discovery. In the case of tight bounds the non-uniqueness of the measure is less severe. However, in many application the bounds diverge substantially and the price leadership becomes blurred.
Chapter 3 offers a solution to this drawback. It is based on a joint research paper with Joachim Grammig, *A new approach to estimate unique market information shares*. We propose to use distributional assumptions in order to resolve the indeterminacy problem of the Hasbrouck (1995) approach. These assumptions match the stylized facts detected for financial return data, namely fat tails and tail dependencies. Fat tails imply that we observe extreme values in the empirical distributions more often than predicted by a normal distribution. Tail dependence refers to the phenomenon that the correlation of innovations in various markets in the center of their distribution differs from the correlation in the tails. These features can be described by modeling price innovations as a mixture of two normal distributions, so that the tail and center observations are generated from two regimes associated with different variances. The idea to use mixture normals to identify idiosyncratic innovations was first brought forward by Lanne and Lütkepohl (2010). In order to identify unique information shares, further restrictions are necessary, which we provide from the economics behind the "one security-multiple markets" setting of the price discovery analysis. Our method is illustrated in an empirical application to credit default swap and corporate bond markets data, for which we measure contributions to the pricing of credit risk. The results emphasize the informational leadership of the more liquid credit derivatives market during the pre-crisis period by the gain in accuracy achieved by being able to deliver a unique measure rather than upper and lower bounds.
The standard approaches to measuring contributions to price discovery are based on a Vector Error Correction Model and thereby model the evolution of prices and price changes. As mentioned by Hasbrouck (1995), the main issue when measuring the importance of different markets to the pricing of one common asset is to determine which market moves first. Since all markets are assumed to be linked by an arbitrage relation all prices will reflect

the same information in the long run. Yet, which market incorporates new information first? If the underling data is available in equally spaced intervals, such as daily frequency, the chronology of information processing cannot be fully observed. However, in the case of high frequency data, for instance in form of intraday quote or transaction data, the sequence of events can deliver useful information concerning the consecutive incorporation of information in different markets. Chapter 4 presents an approach that analyzes price discovery from this perspective. It is based on a joint research paper with Kerstin Kehrle, *International Price Discovery in Stock Markets - A Unique Intensity Based Information Share*. Rather than sticking to the standard approach of modeling price changes, we propose an innovative method that is based on a model for the arrival rates (intensities) of the price processes in a bivariate intensity model. The intensity roughly gives the probability of a transaction event within the next instant. We use Russell's (1999) ACI model that allows for a flexible interaction between the two markets' conditional intensities. When information is incorporated into one market's prices, arbitrage makes sure that the other market's prices subsequently adjust to retain the equilibrium. An increased intensity in one market arising from the incorporation of new information should subsequently increase the intensity in the other markets. We suggest a new information share measure which is based on the relative cross effects of shocks to the intensities in the different markets. Our model is applied to high frequency data of a large sample of Canadian stocks, which are cross-listed on the New York Stock Exchange. The results are in favor of the home market as the leading market, although the contribution of the U.S. market is not negligible. In a cross sectional analysis we also examine potential determinants of contributions to price discovery. We find no evidence for stock related factors being of any importance, but that relative liquidity is the main determinant of information shares.

While the approaches in Chapters 3 and 4 rely on rather heavily parametrized models and distributional assumptions, Chapter 5 applies the model free concept of transfer entropy to measure information flows between financial markets. It is based on the research paper *Using transfer entropy to measure information flows from and to the CDS market.* Transfer entropy was introduced within the context of information theory, whose general aim was to optimally encode messages such that they can be transmitted more quickly. For that purpose it was necessary to quantify the information that can be gained from a specific sequence of transmitted symbols. Transfer entropy is designed as the Kullback-Leibler distance of transition probabilities. The assumptions concerning the data are minimal. It does not require the existence of a cointegration relation between different time series, but quantifies information transfer without being restricted to linear dynamics. We propose an information share measure constructed from this method. Furthermore, we derive standard errors based on a block bootstrap that enable inference, in particular to examine the statistical significance of the estimated information flow - an issue that has not been addressed so far. The method is applies to examine the information flow between a sample of European credit default swaps and the corporate bond market. In addition, the information transfer between the iTraxx as a measure for credit risk and the VIX as a proxy for market risk is analyzed. The results indicate that the information flow from the credit default swap market to the bond market is larger than vice verse. Concerning the dynamic relation between market and credit risk, we find uni-directional information flow from the VIX to the iTraxx.

2 Standard Measures for Contributions to Price Discovery

There exist currently two standard measures for contributions to price discovery in a multiple market setting. Hasbrouck's (1995) information shares and the Gonzalo and Granger (1995) permanent-transitory decomposition. Both approaches rely on the law of one price. Prices in different trading venues which refer to the same underlying asset and are linked by an arbitrage relation cannot deviate from each other in the long run. In econometric terms these prices are cointegrated. While prices in each market are I(1) variables, a linear combination of them yields a stationary process. This means that there exists one stochastic trend common to all price series. According to Hasbrouck (1995) this common trend can be interpreted as the efficient price of the asset underlying all markets. Generally, the Hasbrouck information shares as well as the Gonzalo and Granger (1995) approach rely on a Vector Error Correction Model (VECM) to describe the dynamics of the price processes:

$$\Delta \boldsymbol{p}_t = \boldsymbol{\alpha}\boldsymbol{\beta}'\boldsymbol{p}_{t-1} + \boldsymbol{\Gamma}_1 \Delta \boldsymbol{p}_{t-1} + \ldots + \boldsymbol{\Gamma}_{q-1} \Delta \boldsymbol{p}_{t-q+1} + \boldsymbol{u_t} \quad , \tag{2.1}$$

where $\boldsymbol{p}_t = (p_t^1, ..., p_t^n)'$, $\boldsymbol{\Gamma}_1$ to $\boldsymbol{\Gamma}_{q-1}$ are $n \times n$ parameter matrices. $\boldsymbol{u}_t = (u_t^1, ..., u_t^n)'$ is a white noise vector with zero means and covariance matrix $\boldsymbol{\Sigma}_u$. $\boldsymbol{\beta}$ denotes the $n \times 1$ cointegration vector. The vector $\boldsymbol{\alpha} = (\alpha^1, ..., \alpha^n)'$ contains the coefficients associated with the speed of adjustment of each price series to deviations from the equilibrium. There exists a Vector Moving Average Representation of the model given by

$$\boldsymbol{p}_t = \boldsymbol{u}_t + \boldsymbol{\Xi}_1 \boldsymbol{u}_{t-1} + \boldsymbol{\Xi}_2 \boldsymbol{u}_{t-2} + ... = \boldsymbol{\Xi}(L)\boldsymbol{u}_t \quad , \tag{2.2}$$

from which the Stock and Watson (1988) common trends representation can be derived:

$$\boldsymbol{p}_t = \boldsymbol{\Xi}(1) \sum_{s=0}^{t} \boldsymbol{u}_s + \boldsymbol{\Xi}^*(L)\boldsymbol{u}_t \quad . \tag{2.3}$$

The first term on the right hand side captures the common random walk component. The second term includes the transitory effects. $\boldsymbol{\Xi}(1)$ is the sum of the moving average coefficients and gives the permanent impact of an innovation in each series. The special case of a "one security-multiple market " setting with cointegrated prices implies that the permanent impact of innovations in each price series is equal for all markets and the row vectors in $\boldsymbol{\Xi}(1)$ are identical. In the following we denote this common row vector by $\boldsymbol{\xi}$, which gives

$$\boldsymbol{p}_t = \boldsymbol{\iota\xi} \sum\nolimits_{s=0}^{t} \boldsymbol{u}_s + \boldsymbol{\Xi}^*(L)\boldsymbol{u}_t \quad , \tag{2.4}$$

where $\boldsymbol{\iota}$ is a column vector of ones. According to Hasbrouck (1995) $\boldsymbol{\iota\xi}$ is the component of the price change that is due to new information as it is permanently impounded into the prices. He interprets this component as the common factor or underlying efficient price (compare also Lehmann (2002)and Hasbrouck (2002)).

In Gonzalo and Granger (1995) the common factor is defined by a linear combination of current prices in different markets

$$\boldsymbol{p}_t = \boldsymbol{f}_t + \boldsymbol{g}_t \quad , \tag{2.5}$$

where $\boldsymbol{f}_t = \boldsymbol{\Gamma p}_t$. $\boldsymbol{\Gamma}$ gives the common factor coefficient vector and $\boldsymbol{g}_t$ includes the transitory components. Identification of the common factor is achieved by assuming that these transitory components do not Granger-cause prices in the long run. They also show that $\boldsymbol{\Gamma}$ is orthogonal to the vector of adjustment coefficients in Equation (2.1). Harris et al. (2002) suggest that these factor loadings in the permanent-transitory decomposition of the price processes can be used to measure contribution to price discovery. The measure, which in the following we will refer to as the adjustment coefficient ratio, is given by

$$Adj^j = \frac{|\alpha^j|}{\sum_{j=1}^{n} |\alpha^j|} \quad . \tag{2.6}$$

A straightforward advantage of the adjustment coefficient ratio is that it is easy to compute and a unique measure. However, Baillie et al. (2002) and Hasbrouck (2002) point out that the common trend $\boldsymbol{f}_t$ in Equation 2.5 does not have to be a random walk and therefore might not be a martingale, which renders its economic relevance questionable. Apart from that the approach neglects further dynamics of the price process, such as their variances and contemporaneous correlation in the innovations (see Hasbrouck 1995, Hasbrouck 2002, De Jong 2002, Lehmann 2002, Baillie et al. 2002).

The Hasbrouck information shares are based on a decomposition of the variance of the efficient price innovations (v_t). These are given from Equation (2.4) by $v_t = \boldsymbol{\xi u}_t$ and their variance by $\boldsymbol{\Sigma}_v = \boldsymbol{\xi \Sigma}_u \boldsymbol{\xi'}$. Decomposing the variance into contributions from each market is straightforward if $\boldsymbol{\Sigma}_u$ is a diagonal matrix, which means that the residuals $\boldsymbol{u}_t$ are not contemporaneously correlated. Yet, this is not the case in most applications. As a solution to this identification problem, Hasbrouck (1995) models the contemporaneously correlated innovations $\boldsymbol{u}_t$ as a linear combination of uncorrelated innovations $\boldsymbol{\varepsilon}_t$:

$$\boldsymbol{u_t} = \boldsymbol{B\varepsilon}_t \quad , \tag{2.7}$$

where $\boldsymbol{\varepsilon_t}$ is a vector of zero mean and unit variance variables and the $n \times n$ coefficient matrix $\boldsymbol{B}$ is restricted to be the lower triangular matrix resulting from a Cholesky decomposition of $\boldsymbol{\Sigma}_u$, so that $\boldsymbol{\Sigma}_u = \boldsymbol{CC'}$ and

$$\boldsymbol{u}_t = \boldsymbol{C\varepsilon}_t \quad . \tag{2.8}$$

For $\boldsymbol{C}$ being triangular implies that prices in the market ordered first are not contemporaneously affected by innovation in the remaining markets, while vice verse contemporaneous effects of its own innovations are not restricted. This hierarchy goes down to the market ordered last, which can be contemporaneously affected by innovations in all other markets, but its own cross effects are restricted to zero. Hasbrouck information shares are then computed by decomposing the efficient price innovation variance, $\boldsymbol{\Sigma}_v = \boldsymbol{\xi CC'\xi'}$ into contributions from the different markets

$$IS^j = \frac{([\boldsymbol{\xi' C}]_j)^2}{\boldsymbol{\xi' CC' \xi}} \quad , \tag{2.9}$$

where $[\boldsymbol{\xi' C}]_j$ denotes the j^{th} element of the vector $\boldsymbol{\xi' C}$. According to Johansen (1991) the common row vector $\boldsymbol{\xi}$ in the matrix $\boldsymbol{\Xi}$ can be derived from the VECM parameters as follows

$$\boldsymbol{\Xi} = \boldsymbol{\beta}_\perp [\boldsymbol{\alpha}'_\perp (I_n - \sum_{i=1}^{q-1} \boldsymbol{\Gamma}_i) \boldsymbol{\beta}_\perp]^{-1} \boldsymbol{\alpha}'_\perp \quad . \tag{2.10}$$

Concerning the information shares the Cholesky decomposition implies that the contribution of the market ordered first is maximized and that of the market ordered last is minimized. Since there is no theoretical justification for such a hierarchy, the common solution is to permutate the ordering of the markets. This yields upper and lower bounds of information shares.

Several studies point out that the Hasbrouck information shares are based on a definition of price discovery that has more economic appeal compared to the Gonzalo and Granger (1995) approach (see Hasbrouck 2002, De Jong 2002, Lehmann 2002, Baillie et al. 2002). Yet, the fact that it does not deliver a unique measure but upper and lower bounds for information shares is its crucial drawback. If the bounds are tight, their midpoint might well be taken as a proxy for the true information share. If the contemporaneous correlation is high, these bounds can diverge considerably. Particularly, when dealing with high frequency data, the non-uniqueness of the Hasbrouck shares constitute a severe problem. Figure 2.1 illustrates the link between the sampling frequency and the divergence of the Hasbrouck bounds.

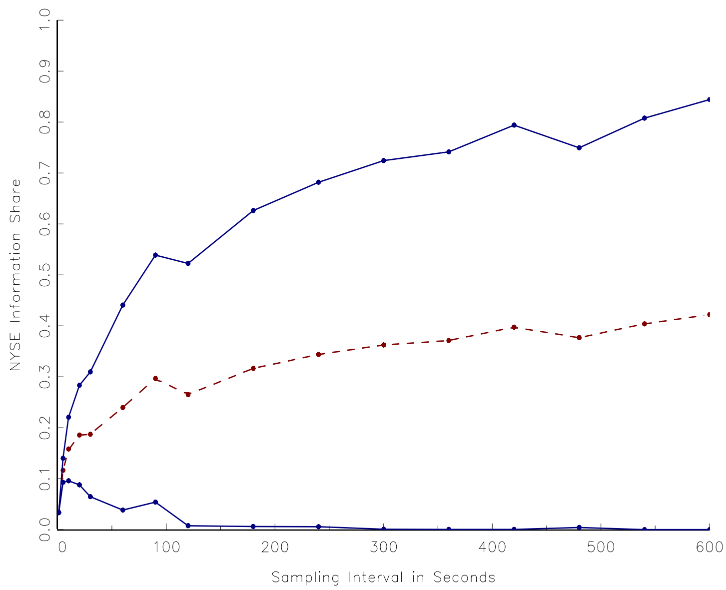


Figure 2.1: Information share estimates at different frequencies. The graph shows the dependence of Hasbrouck information shares on the sampling frequency. It displays the upper and lower bound (solid lines) of the NYSE information share as well as the associated midpoint (dotted line) for the Canadian NYSE interlisted stock *Abitibi Consolidated Inc.* (ABY), estimated at different frequencies. The estimates are calculated over 62 trading days (January first 2004 to March 31st 2004) using the first two hours of trading.

It shows the New York Stock Exchange (NYSE) information share for ABY, a Canadian stock, whose home market is the Toronto Stock Exchange. It is obvious that at a sampling frequency of two minutes, the bounds have already become extremely wide. This is the result of the contemporaneous correlation increase when lowering the sampling frequency, since the lead-lag relationship gets blurred at lower frequencies. Lien and Shrestha (2009) propose a solution to that problem based on another decomposition of Σ_u. However, while the Cholesky decomposition used by Hasbrouck can be interpreted as lower and upper bounds for the true information share, the decomposition of Lien and Shrestha (2009) lacks any economic intuition and presents merely one of many ways to decompose a covariance matrix. The issue of identification of unique shares therefore remains unresolved in current literature.

Estimation of the measures outlined above is usually preceded by a cointegration analysis of the data. This includes unit root tests of the variables as well as testing for the existence and the number of cointegration relations. Commonly, the Dickey-Fuller test statistic is used to examine whether the variables are I(1) and the Johansen trace and maximum eigenvalue statistics are applied to test for cointegration (see Dickey and Fuller 1981, Johansen 1988). Estimation of the VECM in Equation (2.1) can be done via Least Squares (see Lütkepohl 2005) or Maximum Likelihood according to Johansen (1991). The number of lags in the VECM can be conveniently determined by the Schwarz of Akaike information criterion (see Schwarz 1978, Akaike 1969, Akaike 1971). Standard errors for the price discovery measures can be derived via the delta method in the case of the Gonzalo and Granger (1995) approach. Concerning the Hasbrouck information shares Grammig et al. (2005) propose a bootstrap along the lines of MacKinnon (2002). This procedure works as follows: First VECM parameters in Equation (2.1) are estimated and then employed to simulate price series using observations from the original price series as starting values. In the case of a parametric bootstrap, the residuals u_t in Equation (2.1) are independent draws from a specific distribution. In the case of a non-parametric bootstrap, the empirical distribution of the VECM innovation is used. Next, the VECM parameters and price discovery measures are estimated from the simulated series. This procedure is repeated and finally standard errors for parameter and information share estimates are computed from the empirical distribution of the bootstrap estimates.[1]

[1] Davidson and MacKinnon (2000) recommend choosing the number of bootstrap replications B such that $\alpha(B+1)$ is an integer. Testing one-sided at 5% significance, $B = 399$ implies that the 20th largest bootstrap estimate is the critical value at $\alpha = 0.05$.

3 A Data Driven Approach to Estimate Unique Information Shares

The trading of securities on multiple markets raises the question of each market's share in the discovery of the informationally efficient price. We exploit salient distributional features of multivariate financial price processes to uniquely determine these contributions. Thereby we resolve the main drawback of the widely used Hasbrouck (1995) methodology which merely delivers upper and lower bounds of a market's information share. When these bounds diverge, as is the case in many applications, informational leadership becomes blurred. We show how fat tails and tail dependence of price changes, which emerge as a result of differences in market design, can be exploited to estimate unique information shares. The empirical application of our methodology emphasizes the leading role of the credit derivatives market compared to the corporate bond market in pricing credit risk during the pre-crisis period.

This chapter is based on the article *Tell-Tale Tails - A new approach to estimate unique market information shares* by Joachim Grammig and Franziska J. Peter (2010).

3.1 Introduction

One of the most frequently asked questions in empirical finance is “Where is the market?” Whether in the case of cross-listings of stocks or newly developed derivative markets, this question has stirred up an enormous amount of research. Booth et al. (2002), for instance, examine the role of upstairs and downstairs markets in the price discovery process at the Helsinki Stock Exchange, while Huang (2002) estimates the contributions of market makers and electronic crossing networks to the price formation of NASDAQ stocks. Hasbrouck (2003) analyzes the importance of different trading venues for the price discovery process of U.S. equity indices. The share of the futures market in U.S. treasury price discovery is the focus of a study by Mizrach and Neely (2008), and Blanco et al. (2005) estimate the share of the bond market and the market of credit default swaps (CDS) in the process of pricing credit risk. While dealing with the same question in different trading environments, all these studies report Hasbrouck (1995) information shares which is the most prevalent approach to measure contributions to price discovery.

In this chapter we resolve the main drawback of Hasbrouck’s (1995) methodology which does not deliver a unique measure, but merely information share upper and lower bounds. These bounds can diverge considerably and hinder a clear detection of the market that leads price discovery. Our approach identifies unique information shares by exploiting distributional properties of financial data, namely fat tails and tail dependence. Thereby we deliver a more accurate measure which can be applied to study price discovery in various fields of financial research.

Within Hasbrouck’s methodology, information shares are defined as each market’s contribution to the variance of the efficient price innovations. However, within a vector equilibrium correction framework the efficient price variance can generally not be decomposed without further restrictions. For that purpose Hasbrouck (1995) uses the Cholesky factorization of the innovation covariance matrix which implies a hierarchical ordering in terms of the contemporaneous information flow. Permuting the ordering of markets results in upper and lower information share bounds. When these bounds diverge, they measure contributions to price discovery very inaccurately.

Our approach towards estimating unique information shares is related to the identification of structural shocks through heteroskedasticity (see Rigobon 2003) and non-normal innovations (see Lanne and Lütkepohl 2010). These papers show that structural innovations within a multiple time series framework can be identified if the data exhibit heteroskedasticity that can be described by a multi-regime process associated with different innovation variances. We connect this insight with two salient facts of financial price processes: fat tailed return distributions combined with tail dependence. We show how these features, which may result from differences in market liquidity, can be exploited to disentangle the contemporaneous correlations of the price innovations across markets. In particular, the occurrence of a large price movement in one market can either represent an informative event or a transitory liquidity shock. Contemporaneous price movements of the other markets reveal the informational content of the large price change, and thereby identify market idiosyncratic innovations. Those *tell-tale* tail observations are the key to deliver unique information shares.

Drawing on the approach put forth by Lanne and Lütkepohl (2010), we assume that market idiosyncratic price innovations come from mixture distributions, and that the observed price innovations emerge as a linear combination of these structural shocks. We show that the resulting multivariate mixture distribution can account for fat tails and tail dependence, which we exploit for the computation of unique information shares. The basic data requirement to achieve this goal is that the correlations of the market price innovations in the tails and in the center of their joint distribution are sufficiently different.

Since there are no identifying restrictions suggested by finance theory, the possibility to disentangle the contemporaneous correlation structure of price innovations based on distributional properties of financial data is quite appealing. However, Lanne and Lütkepohl's method delivers ambiguous results if applied without further restrictions. In particular, we show that one would merely identify sets of information shares, while not being able to allocate them uniquely to the markets. We offer a solution by proposing identifying restrictions which naturally arise from the *one security-multiple markets* framework.

We use the new methodology to measure the contribution of the CDS and the corporate bond market to the pricing of credit risk. The results emphasize the informational leadership of the more liquid credit derivatives market during the pre-crisis period. They also corroborate the conclusions of previous studies that identify relative market liquidity as the most important variable for explaining market information shares (see e.g. Yan and

Zivot 2010). Liquidity, as a result of market design, attracts trading volume and promotes a market's leadership in price discovery. The methodology proposed in this chapter systematically exploits the informational content of those market design effects to deliver a unique measure of a market's information share.

3.2 Fat Tails, Tail Dependence, and Unique Information Shares

3.2.1 Motivation and Econometric Specification

The identification of variance shares and idiosyncratic innovations as defined in Equations (2.9) and (2.7) is a prevalent problem in various fields of economics. As an alternative to the Cholesky decomposition, macroeconomic VAR analyses exploit theoretically motivated restrictions on long run effects, by imposing constraints on $\mathbf{\Xi B}$, and/or short run effects, by imposing restrictions on $\mathbf{B}$ (see Lütkepohl 2008). However, finance theory does not suggest such restrictions concerning the one security-multiple markets framework. As a result, the indeterminacy of Hasbrouck's information share measure remained a caveat for 15 years.

Our proposed solution exploits two stylized facts of financial price processes: fat tails and tail dependence. Fat tails mean that large negative or positive price changes occur more frequently than predicted by a normal distribution (see e.g. Haas et al. 2004). By tail dependence we refer to the phenomenon that the correlation of price changes in the tails of the distribution is different from that in the center (see e.g. Longin 2001). While these empirical facts are not at odds with finance theory, there are no first principles explanations for their existence.

Before we outline the mathematical details of our methodology, let us first illustrate how fat tails and tail dependence can help disentangle the contemporaneous correlation of the price innovations. For that purpose we follow Rigobon (2003) who uses scatter plots to visualize identification through heteroskedasticity. Our illustration focuses on the case of $n = 2$ markets.

The three panels in Figure 3.1 depict scatter plots of composite price innovations u_1 and u_2. The upward sloping regression lines indicate the positive contemporaneous correlation of the price innovations on the two markets. All three panels show price innovations clustering

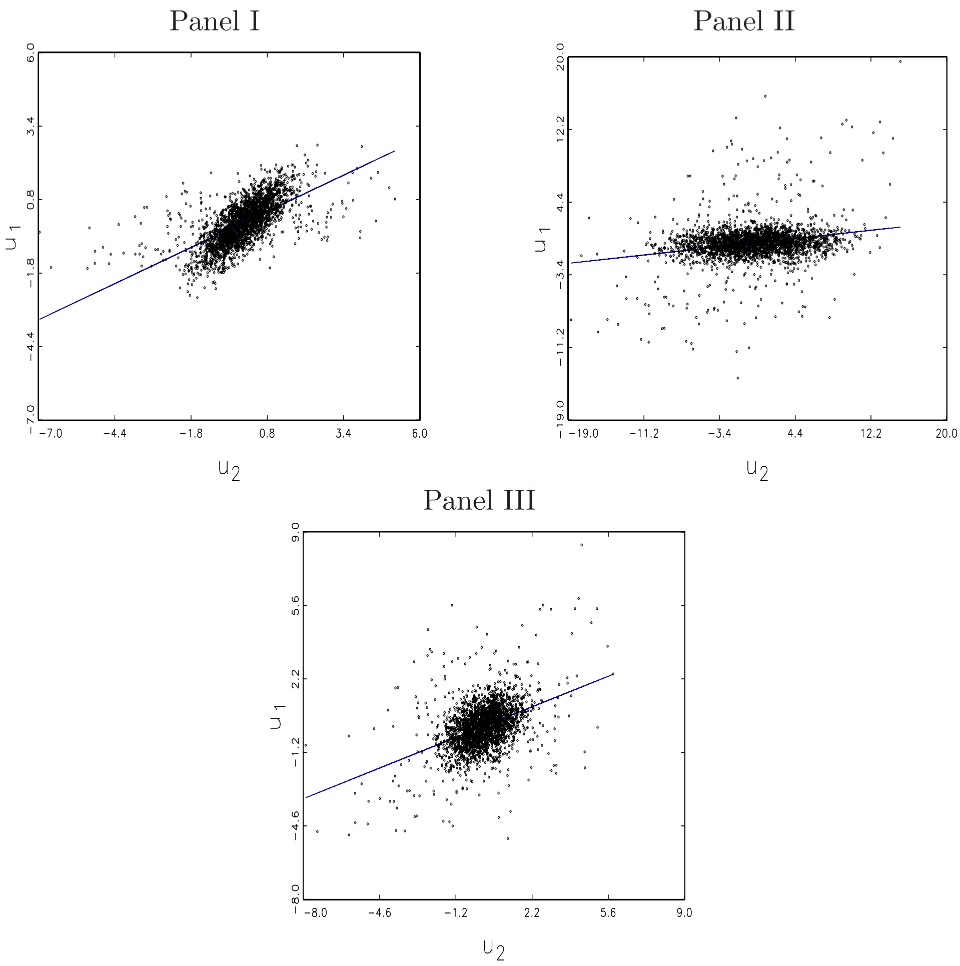


Figure 3.1: Scatter plots of composite price innovations. The lines result from a regression of u_1 on u_2.

in the dense center of the bivariate distributions. In Panels I and II the correlations of the innovations in the center and in the tails of the bivariate distribution are distinctly different. Price innovations in the dense center of the Panel I distribution are positively correlated. However, tail observations in market two do not tend to be accompanied by particularly large absolute values of u_1. The Panel II data also exhibit tail dependence, but the correlation in the dense center is smaller than in the tails. Here the marginal distribution of u_1 is more leptokurtic, with price innovations that are mostly small in absolute value, but with occasional large positive or negative shocks. If, however, there is a large and positive (negative) innovation in market one, then the market two innovation tends to be large and positive (negative), too. The converse does not hold true: The horizontally flattened dense center of the Panel II scatter plot implies that extreme market two price innovations do not tend to be accompanied by u_1 observations that are large in absolute value.

An economic explanation for such observations is that the design of the trading process on market two may entail temporary shortages of liquidity, which cause large absolute price changes. These liquidity shocks on market two do not affect the common efficient price, and thus do not contemporaneously spill over to market one. As it turns out, such market imperfections are very useful for our quest. They reveal the contemporaneous dependence structure which is the key to identify unique information shares. In detail, given the factor structure

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} b_{1,1} & b_{1,2} \\ b_{2,1} & b_{2,2} \end{pmatrix} \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \quad , \tag{3.1}$$

the Panel I and II scatter plots suggest that the weight $b_{1,2}$, which transfers an idiosyncratic price shock occurring on market two into the price innovation of market one, is small, while $b_{2,1}$ is large.

Let us now set up a statistical model that accounts for fat tails and tail dependence. For that purpose, we draw on Lanne and Lütkepohl's (2010) idea to identify structural shocks in a VAR framework by assuming mixture distributions for the residuals. Such an assumption may not be obvious or sensible in a macroeconomic analysis involving variables like GDP, money supply, unemployment and interest rates. In the present application, however, it perfectly matches the stylized facts observed in financial data.

We retain the factor structure $\mathbf{u}_t = \mathbf{B}\boldsymbol{\varepsilon}_t = \mathbf{W}\mathbf{e}_t$, where $\mathbf{W}$ denotes a non-singular matrix, and $\mathbf{e}_t$ is an n-dimensional vector of contemporaneously and serially uncorrelated innovations. It results from a mixture of two serially independent Gaussian random vectors,

$$\mathbf{e}_t = \begin{cases} \mathbf{e}_{1,t} \sim \mathscr{N}(\mathbf{0}, \mathbf{I}_n) & \text{with probability } \gamma \\ \mathbf{e}_{2,t} \sim \mathscr{N}(\mathbf{0}, \boldsymbol{\Psi}) & \text{with probability } 1-\gamma \quad , \end{cases} \tag{3.2}$$

where $0 < \gamma < 1$ and $\boldsymbol{\Psi}$ is a diagonal matrix with positive elements $\psi_1, .., \psi_n$.

As shown by Rigobon (2003), the identification of structural shocks through heteroskedasticity relies on the existence of regimes with different innovation variances. Unlike Rigobon (2003), who assumes exogenously defined variance regimes, Equation (3.2) specifies only a regime probability. This entails the necessity to deal with and deliver identifying restrictions. We will address this issue in the next section and for now assume that the set of mixture parameters $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$ can be uniquely identified.

It follows from (3.2) that the covariance matrix of the idiosyncratic innovations $\mathbf{e}_t$ is given by

$$\mathbf{\Sigma}_e = \gamma \mathbf{I}_n + (1-\gamma)\mathbf{\Psi} \quad , \tag{3.3}$$

such that

$$\mathbf{\Sigma}_u = \mathbf{B}\mathbf{B}' = \mathbf{W}\mathbf{\Sigma}_e\mathbf{W}' \quad , \tag{3.4}$$

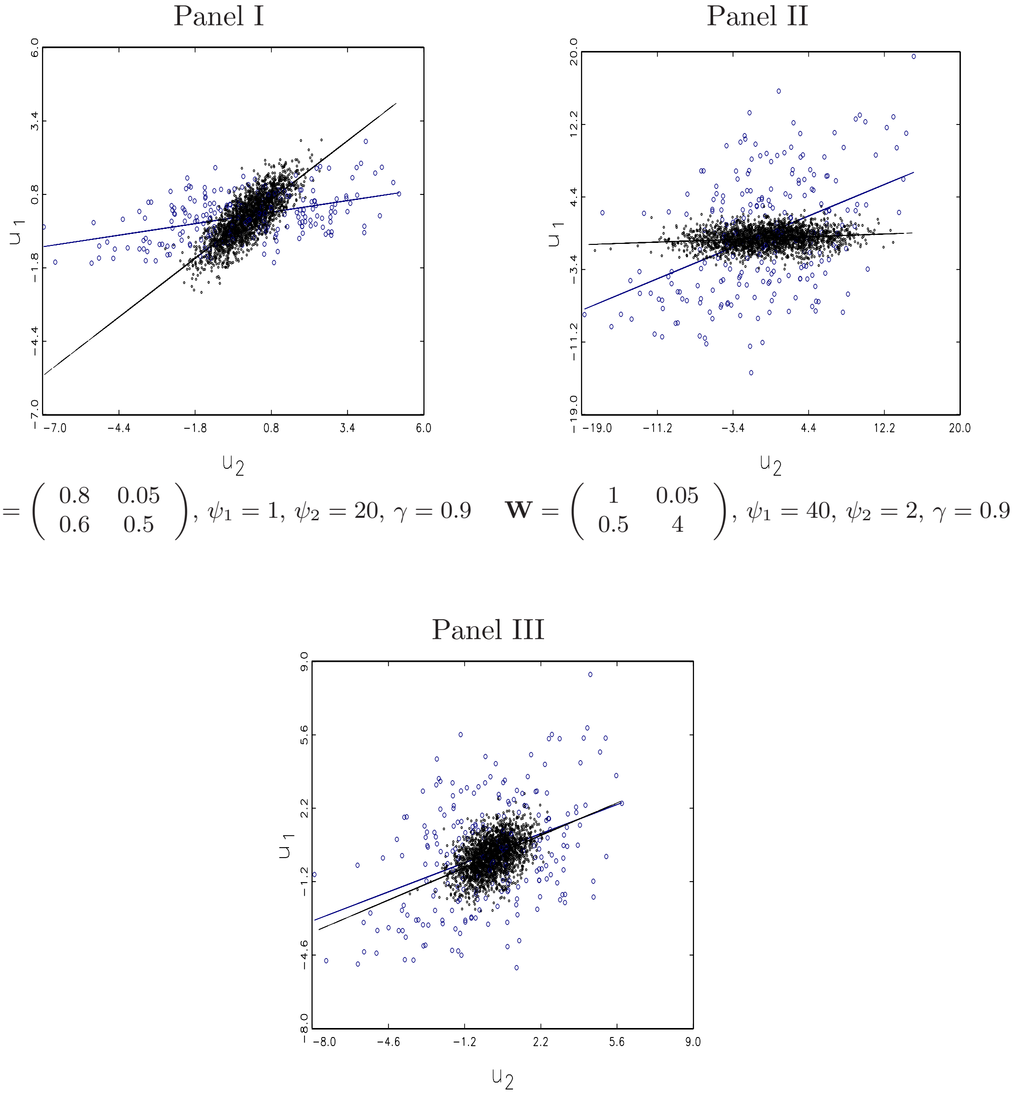
which implies that $\mathbf{B} = \mathbf{W}\mathbf{\Sigma}_e^{0.5}$. Information shares which are independent of the ordering of markets, can then be computed replacing the Cholesky matrix $\mathbf{C}$ in Equation (2.8) by $\mathbf{W}\mathbf{\Sigma}_e^{0.5}$, viz

$$\mathbf{IS}_X(\theta_m, \theta_v) = \frac{\left[\xi'\mathbf{W}\mathbf{\Sigma}_e^{0.5}\right]^{(2)}}{\xi'\mathbf{W}\mathbf{\Sigma}_e\mathbf{W}'\xi} \quad , \tag{3.5}$$

where $\theta_v = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_q\}$ collects the VECM parameters. The X subscript indicates that the identification of information shares exploits the informational content of extreme (tail) observations.

Figure 3.2 illustrates how mixture of normal distributions can produce fat tails and tail dependence. The three Panels reveal that the innovations displayed in Figure 3.1 were drawn from bivariate normal mixtures with a low and a high variance regime. Tail dependence prevails in Panels I and II, since here $\psi_1 \neq \psi_2$, while in Panel III $\psi_1 = \psi_2$. Identical regime variances imply that the correlation of the innovations is the same in the low and the high variance regime. In other words, the dependence of innovations in the tails of the distribution is not different from that in the center when $\psi_1 = \psi_2$.

Figure 3.2 also reveals that the off-diagonal elements of the weight matrix $\mathbf{W}$ are as suspected by eyeballing the Panel I and II scatter plots in Figure 3.1. The parameter $w_{1,2}$ − the weight with which the idiosyncratic market two innovation e_2 contemporaneously affects the price on market one − is smaller than $w_{2,1}$, the weight with which the market one idiosyncratic innovation e_1 contemporaneously affects the price on market two. Fat tails along with tail dependence represent the basic data features to successfully apply our methodology. Using mixtures of normal distributions, with regime variances that are dif-



$\mathbf{W} = \begin{pmatrix} 0.8 & 0.05 \\ 0.6 & 0.5 \end{pmatrix}$, $\psi_1 = 1$, $\psi_2 = 20$, $\gamma = 0.9$

$\mathbf{W} = \begin{pmatrix} 1 & 0.05 \\ 0.5 & 4 \end{pmatrix}$, $\psi_1 = 40$, $\psi_2 = 2$, $\gamma = 0.9$

$\mathbf{W} = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}$, $\psi_1 = 10$, $\psi_2 = 10$, $\gamma = 0.9$

Figure 3.2: Scatter plots of composite price innovations with DGPs revealed. Data are generated by bivariate mixture distributions. The small dots represent observations from regime 1, the circles represent observations from regime 2. The lines result from regressions of u_1 on u_2 using data from the respective regimes.

ferent across markets, one can account for these features in a statistical model. However, as we will outline in the next section, additional restrictions are required to identify the vector of IS_X information shares according to Equation (3.5).

3.2.2 Identification

The identification of unique information shares involves two aspects, namely to determine the *set* of information shares and to allocate them to the n markets. As shown by Lanne and Lütkepohl (2010), the identification of the weighting matrix $\mathbf{W}$ requires that the diagonal elements of $\mathbf{\Psi}$ (the idiosyncratic innovation variances) are all different. This result corresponds to Rigobon's (2003) finding that in order to permit identification through heteroskedasticity the regime variances have to be different.

In particular, Lanne and Lütkepohl (2010) show that if $\mathbf{\Psi}$ contains different elements on its main diagonal, then the columns of $\mathbf{W}$ are identified up to a multiplication of one or many of its columns by -1. However, being able to identify the columns of $\mathbf{W}$ only up to a sign shift does not affect the information shares computed according to Equation (3.5). Furthermore, the sign indeterminacy can be easily resolved by restricting the main diagonal elements of $\mathbf{W}$ to be greater than zero. This is a sensible restriction in almost any application. In the present context it implies that an idiosyncratic price innovation on market i, $e_{i,t}$, contemporaneously impacts on the composite innovation $u_{i,t}$ with the same sign and a nonzero weight.

However, distinct main diagonal elements of $\mathbf{\Psi}$ ensure the identification of the *columns* of $\mathbf{W}$, but not their *ordering*. The consequences are severe, as it is only possible to identify the set of information shares, but not to assign them uniquely to the n markets. As we prove in Appendix A, there exist $n!$ possibilities to allocate information shares to the n markets. These information share vectors result from alternative parametrization which are observationally equivalent to $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$. They imply the same joint density of the random vector $\mathbf{u}_t$ which, resulting from Equation (3.2) and $\mathbf{u}_t = \mathbf{W}\mathbf{e}_t$, is given by

$$\begin{aligned} f(\mathbf{u}_t;\theta_m) &= \gamma \times (2\pi)^{-\frac{n}{2}} \det(\mathbf{W})^{-1} \exp\left\{-\frac{\mathbf{u}_t'(\mathbf{W}\mathbf{W}')^{-1}\mathbf{u}_t)}{2}\right\} \\ &+ (1-\gamma) \times (2\pi)^{-\frac{n}{2}} \det(\mathbf{\Psi})^{-0.5} \det(\mathbf{W})^{-1} \exp\left\{-\frac{\mathbf{u}_t'(\mathbf{W}\mathbf{\Psi}\mathbf{W}')^{-1}\mathbf{u}_t}{2}\right\}. \end{aligned} \tag{3.6}$$

We refer the reader to Appendix A for a formal proof. The key insight is that distinct diagonal elements of $\mathbf{\Psi}$ identify $\mathbf{W}$ uniquely only if the ordering of the columns of $\mathbf{W}$ cannot be altered. However, the re-parametrization $\theta_m^* = \{\gamma, \mathbf{\Psi}^*, \mathbf{W}^*\}$, where $\mathbf{W}^* = \mathbf{W}\mathbf{P}$ and $\mathbf{\Psi}^* = \mathbf{P}'\mathbf{\Psi}\mathbf{P}$, with $\mathbf{P}$ a permutation matrix of order n, is observationally equivalent to the original parametrization $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$, such that $f(\mathbf{u}_t; \theta_m^*) = f(\mathbf{u}_t; \theta_m)$.[2] This implies that there exist $n! - 1$ sets of mixture parameters which are observationally equivalent to the original parametrization. Furthermore, there exist $n!$ additional parametrization $\tilde{\theta}_m = \{\tilde{\gamma}, \widetilde{\mathbf{\Psi}}, \widetilde{\mathbf{W}}\}$ where $\widetilde{\mathbf{W}} = \mathbf{W}\mathbf{\Psi}^{0.5}\mathbf{P}$, $\widetilde{\mathbf{\Psi}} = \mathbf{P}'\mathbf{\Psi}^{-1}\mathbf{P}$ and $\tilde{\gamma} = 1 - \gamma$. These parametrization are also observationally equivalent to θ_m.

As we show in Appendix A, these alternative parametrization permute the original information shares according to

$$\mathbf{IS}_X(\theta_m^*, \theta_v) = \mathbf{IS}_X(\tilde{\theta}_m, \theta_v) = \mathbf{IS}_X(\theta_m, \theta_v) \times \mathbf{P} \quad , \tag{3.7}$$

such that there exist $n!$ different, but observationally equivalent information share vectors. In other words, it is impossible to determine which information share belongs to a single market.

Equation (3.7) implies that in order to ensure identification we need additional restrictions that prevent the permutation of the columns of $\mathbf{W}$ and the diagonal elements of $\mathbf{\Psi}$. Fortunately, the one security-multiple markets application framework suggests the following constraints:

$$\begin{aligned} w_{i,i} &> 0 && \forall \quad i \\ w_{i,i} &> |w_{j,i}| && \forall \quad j \neq i \quad , \end{aligned} \tag{3.8}$$

where $w_{i,j}$ is the row i, column j element $\mathbf{W}$. The restriction that the diagonal elements of $\mathbf{W}$ are larger than the remaining elements in the same column is economically plausible, since we expect the weight with which the idiosyncratic shock originating in market i, $e_{i,t}$, contemporaneously affects the own market composite price innovation $u_{i,t}$ to be larger in absolute value than the weights with which it contemporaneously affects the composite price innovations of all other markets.

The restrictions in (3.8) leave $\mathbf{P} = \mathbf{I}_n$ as the only eligible permutation matrix. The two remaining parametrization θ_m and $\overline{\theta}_m = \{1-\gamma, \mathbf{\Psi}^{-1}, \mathbf{W}\mathbf{\Psi}^{0.5}\}$ imply the same allocation of information shares to the n markets. Restricting one of the regime variances to be greater

[2] A permutation matrix $\mathbf{P}$ results from permuting the rows of an identity matrix. Every row and column therefore contains one element that equals one and the remaining elements are zero. Consequently, there exist $n!$ distinct permutation matrices of order n, one of which is the identity matrix. Post-(pre-) multiplication by a permutation matrix results in a matrix where the columns (rows) of a matrix are interchanged according to the permutation implied $\mathbf{P}$. The operation $\mathbf{WP}$ thus permutes the columns of $\mathbf{W}$. The operation $\mathbf{P}'\mathbf{\Psi}\mathbf{P}$ permutes the diagonal elements of $\mathbf{\Psi}$ accordingly.

than one leaves θ_m as the only eligible parametrization. Together with the restriction that all elements of $\boldsymbol{\Psi}$ are distinct, the constraints in (3.8) suffice to identify the set of information shares and allocate them uniquely.

3.2.3 Estimation

Maximum Likelihood presents the natural method to estimate the model parameters. Using $\mathbf{u}_t = A(L)\mathbf{p}_t$, where

$$A(L) = 1 - L - \boldsymbol{\alpha}\boldsymbol{\beta}'L - \boldsymbol{\Gamma}_1\Delta L - \ldots - \boldsymbol{\Gamma}_{q-1}\Delta L^{q-1} \tag{3.9}$$

and Equation (3.6), the conditional log-likelihood function reads

$$\begin{aligned}\mathscr{L}(\theta_m, \theta_v) = &\sum_{t=1}^{T} \ln\Bigg(\gamma \times (2\pi)^{-\frac{n}{2}} \det(\mathbf{W})^{-1} \exp\left\{-\frac{\mathbf{u}_t'(\mathbf{W}\mathbf{W}')^{-1}\mathbf{u}_t}{2}\right\} \\ &+ (1-\gamma) \times (2\pi)^{-\frac{n}{2}} \det(\boldsymbol{\Psi})^{-0.5} \det(\mathbf{W})^{-1} \exp\left\{-\frac{\mathbf{u}_t'(\mathbf{W}\boldsymbol{\Psi}\mathbf{W}')^{-1}\mathbf{u}_t}{2}\right\}\Bigg).\end{aligned} \tag{3.10}$$

Estimation of the VECM parameters θ_v and the mixture parameters θ_m in a single step is computationally burdensome. We therefore adopt the two-step estimation strategy outlined by Lütkepohl (2005) and Vlaar (2004). The first step either estimates the cointegrating vectors, or uses those suggested by theory (i.e. $\boldsymbol{\beta}'=[\boldsymbol{\iota}_{n-1} \quad -\mathbf{I}_{n-1}]$). Equation by equation OLS of the VECM in (2.1) then delivers consistent estimates of θ_v which can be used to compute an estimate of the long run impacts vector $\boldsymbol{\xi}$ from Equation (2.10). The second estimation step maximizes the concentrated log-likelihood which results from replacing the VECM parameters in (3.10) by their first step estimates, i.e. $\mathbf{u}_t$ is replaced by

$$\widehat{\mathbf{u}}_t = (1 - L - \widehat{\boldsymbol{\alpha}}\widehat{\boldsymbol{\beta}}'L - \widehat{\boldsymbol{\Gamma}}_1\Delta L - \ldots - \widehat{\boldsymbol{\Gamma}}_{q-1}\Delta L^{q-1})\mathbf{p}_t \quad , \tag{3.11}$$

to obtain estimates of θ_m. Maximization of the concentrated log-likelihood imposes the identifying constraints (3.8). Plugging in the first step estimates $\hat{\theta}_v$ and the second step estimates $\hat{\theta}_m$ in (3.5) delivers IS_X information share estimates. Standard errors for the estimates resulting from this two-step procedure can be delivered by a parametric bootstrap along the lines of MacKinnon (2002). We conduct a parametric bootstrap to provide standard errors and confidence intervals for parameter and information share estimates resulting from the two-step estimation procedure outlined in Section 3.2.3. The procedure works as follows. We first draw an iid sequence of random variables from a normal mixture distribution. This distribution is generated using the mixture parameters which are estimated in the second (Maximum Likelihood) step of the estimation procedure. Next, we generate simulated price series according to Equation (2.1) using observations from the original price series as starting values, the estimated or pre-specified cointegrating vectors, the first step OLS estimates of the VECM parameters, and the simulated mixture residuals. The length of the simulated series equals the number of observations in the original data set plus 100. We discard the first 100 data points in order to reduce the dependence on the starting values. The two-step estimation procedure described in Section 3.2.3 is then applied to the simulated data. We store the resulting parameter estimates and compute estimates of $\boldsymbol{\xi}$ using Equation (2.10), upper and lower bounds of Hasbrouck information shares according to Equation (2.9), and IS_X information shares according to (3.5). This procedure is repeated $B = 399$ times, as suggested by Davidson and MacKinnon (2000). They recommend choosing the number of bootstrap replications B such that $\alpha(B + 1)$ is an integer. $B = 399$ implies that the 20th largest bootstrap estimate is the critical value at $\alpha = 0.05$. Standard errors for parameter and information share estimates are computed from the empirical distribution of the bootstrap estimates.

3.3 Empirical Application

3.3.1 Credit Default Swaps, Credit Spread, and the Price of Credit Risk

To illustrate the benefit of our methodology we revisit a research question addressed by Blanco et al. (2005) who quantify the information share of the corporate bond market and the market for credit derivatives in pricing credit risk. Given the importance of credit securitization and the controversial role played by credit derivatives during the recent financial crisis, research on this topic is more relevant than ever.

Both corporate bonds and credit derivatives, of which CDSs are the most important instruments, are traded on over-the-counter markets. The corporate bond market determines credit spreads (p_{CS}), the difference between risky bond yields and the risk-free rate. A

CDS is a contract between two counterparties trading credit risk. The protection buyer transfers default risk by paying a fee to the protection seller who is willing to assume the risk. In return, the buyer receives a payoff if the underlying financial instrument defaults. The economic effect of a CDS is thus similar to that of an insurance contract, but the buyer of credit protection via a CDS does not necessary have to hold the insured security. The annualized fee, expressed in basis points of the notional volume, is referred to as the CDS price (p_{CDS}). Since credit spread and CDS price are linked by an approximate arbitrage relation (see Duffie 1999, Hull and White 2000a, Hull and White 2000b), Blanco et al. (2005) assume cointegration between the two $I(1)$ price series such that $p_{CDS,t} - p_{CS,t}$ is $I(0)$.[3]

Blanco et al.'s (2005) study is an exemplary application of Hasbrouck's (1995) methodology. They set up the VECM in Equation (2.1) with $p_{1,t} = p_{CDS,t}$ and $p_{2,t} = p_{CS,t}$. Here the common stochastic trend can be interpreted as the price of credit risk. This research question is especially interesting for the application of our methodology, since liquidity matters on markets for credit risk. Collin-Dufresne et al. (2001) point out that movements in liquidity premia explain a large proportion of the total variation in credit spreads. As outlined in Section 3.2, differences in market liquidity are the key to identify unique information shares.

3.3.2 Data

We make use of the data on CDS prices and credit spreads collected by Blanco et al. (2005).[4] The time series of CDS prices are midpoints of daily close- of-business indicative quotes supplied by the CDS broker CreditTrade and J.P. Morgan Securities. The CDS

[3] The arbitrage relation can be explained as follows. Suppose an investor buys a T-year par bond with yield to maturity of y issued by the reference entity. The investor also buys credit protection on that entity for T years at p_{CDS}. The net annual return is $y - p_{CDS}$ which, by arbitrage, and because default risk is eliminated, should be equal to the T-year risk-free rate denoted by x. If $y - p_{CDS} < x$, then shorting the risky bond, writing protection on the CDS market, and buying the risk free rate would present an arbitrage opportunity. If $y - p_{CDS} > x$, then buying the risky bond and protection, and shorting the risk-free bond becomes profitable. Accordingly, the price of the CDS should equal the credit spread, $p_{CDS} = p_{CS} = y - x$. However, with market imperfections such as liquidity premia, not exactly matching maturity dates, and cheapest to delivery options in case of default, the arbitrage relation is not perfect. Assuming cointegration accounts for the approximate nature of the arbitrage relation between CDS price and credit spread.

[4] We are grateful to R. Blanco for making these data available.

	Country	Sector	Rating	Mean		Std. Dev.		Kurtosis		Corr
				p_{CDS}	p_{CS}	Δp_{CDS}	Δp_{CS}	Δp_{CDS}	Δp_{CS}	
AOL	United States	Internet	BBB	93.20	80.17	5.48	7.46	17.04	14.54	0.02
Bank of Am.	United States	Banking	A	36.14	39.69	2.58	4.41	12.24	4.73	0.07
Bank One	United States	Banking	A	45.17	50.78	2.71	5.90	5.79	0.51	-0.02
Bear Stearns	United States	Banking	A	71.40	80.93	3.84	6.76	9.83	1.66	0.02
Citigroup	United States	Banking	AA	32.17	26.43	2.72	4.94	9.88	0.70	-0.02
Fleet Boston	United States	Banking	A	49.32	44.08	2.05	4.99	12.31	17.77	0.19
Ford	United States	Automobile/finance	BBB	143.47	140.89	7.57	6.70	7.31	6.57	0.26
GE Capital	United States	Finance	AAA	30.40	7.18	2.27	5.56	71.63	0.63	0.07
General Mot.	United States	Automobile/finance	BBB	119.04	108.39	5.72	6.6	4.22	0.915	0.15
Goldman Sachs	United States	Banking	A	51.91	55.72	2.98	5.40	7.62	0.80	-0.01
JPMorgan	United States	Banking	AA	44.52	42.02	2.67	3.8	7.82	2.648	-0.10
Morgan St.	United States	Banking	AA	47.67	47.98	3.21	4.97	18.69	4.82	0.03
Lehman Bros.	United States	Banking	A	69.86	77.61	3.82	7.28	6.80	1.80	-0.03
Merrill Lynch	United States	Banking	AA	50.24	43.56	2.98	5.58	18.08	0.67	-0.02
Wal Mart	United States	Retail	AA	19.77	-0.85	0.99	4.59	43.99	9.30	0.04
Wells Fargo	United States	Banking	A	26.32	30.17	2.38	5.37	15.60	3.78	-0.07
British Tel.	United Kingdom	Telecom.	A	103.02	113.04	4.12	4.69	2.67	0.60	0.27
Commerzbank	Germany	Banking	A	27.31	14.70	1.11	3.64	32.73	0.44	-0.07
Daimler	Germany	Automobile	BBB	128.50	120.65	4.92	6.01	3.65	1.15	0.28
Deutsche Tel.	Germany	Telecom.	BBB	144.64	121.46	7.70	4.67	5.34	4.36	0.47
Fiat	Italy	Automobi	A	106.30	100.52	4.48	3.27	5.95	3.60	0.31
Iberdrola	Spain	Utilities	A	32.54	49.25	1.01	3.00	40.57	17.80	0.04
Metro	Germany	Retail	BBB	62.94	80.29	1.99	3.55	42.29	6.70	0.06
Siemens	Germany	Telecom.	AA	44.69	33.68	2.04	3.47	9.43	33.48	0.12
Telefonica	Spain	Telecom.	A	85.65	73.31	4.06	2.75	10.13	0.77	0.22
Volvo	Sweden	Automobile	A	72.50	79.83	3.95	2.86	19.19	4.18	0.13

Table 3.1: Data descriptives. The table lists the reference entities and basic descriptives of CDS prices and corporate bond spreads. We report the mean of the CDS price and credit spreads (in basis points) as well as the standard deviation, kurtosis and correlation of their first differences. The sampling period is January 2, 2001 to June 20, 2002 (383 trading days).

prices are for single-name standard ISDA benchmark contracts for physical settlement, a notional volume of $ ten million, and five years maturity, the most liquid maturity in the CDS market.[5] Risky bond yields are from Bloomberg. By linearly interpolating yields between three and five years and yields with more than six and a half years to maturity at the start of the sample, a five-year yield to maturity is estimated to match the CDS maturity. Euro and Dollar five year swap rates, respectively, are used as proxies for the risk-free rate. The resulting time series of CDS prices and credit spreads for 33 reference entities (16 U.S. and 17 European companies) run from January 2, 2001 to June 20, 2002 (383 trading days). In the following, we focus on those 26 reference entities for which the data support the existence of the hypothesized cointegrating relation (see Table III in Blanco et al. 2005). Descriptive statistics are reported in Table 3.1.

3.4 Estimation Results and Discussion

Two-step estimation of IS_X information shares is performed as described in Section 3.2.3. Estimation results are reported in Tables 3.2 and 3.3. For the first step estimation, we assume the theoretical cointegrating vector $\beta = (1, -1)'$ and $q = 2$ in Equation (2.1). The first step estimates are used to compute upper and lower bounds of Hasbrouck information share estimates and alternative measures of contributions to price discovery.

Table 3.2 reports the mixture parameter estimates and the Wald test results for the null hypothesis of identical regime variances, $\psi_1 = \psi_2$. For all reference entities the null is rejected at conventional significance levels. As outlined above, this is a necessary condition for the identification of unique information shares according to our methodology.

Along with IS_X estimates, Table 3.3 contains lower and upper bounds of the Hasbrouck information share estimates of the CDS market. We further include the estimates of the long run impact coefficients $\boldsymbol{\xi} = (\xi_{CDS}, \xi_{CS})'$, and the ratio of adjustment coefficients $\lambda_{CS} = \frac{|\alpha_{CS}|}{|\alpha_{CDS}|+|\alpha_{CS}|}$.[6] Standard errors for these estimates as well as for Hasbrouck information shares are obtained applying the non-parametric bootstrap procedure outlined

[5] The International Swaps and Derivatives Association (ISDA) contracts define default events and ways of settlement in case of default (cash or physical delivery, i.e. delivery of a reference asset).

[6] Baillie et al. (2002) show that with $\beta = (1, -1)'$ it follows from Equation (2.10) that $\frac{|\alpha_{CS}|}{|\alpha_{CDS}|+|\alpha_{CS}|} = \frac{|\xi_{CDS}|}{|\xi_{CDS}|+|\xi_{CS}|}$.

Reference Entity	ψ_1	ψ_2	γ	w_{11}	w_{12}	w_{21}	w_{22}	Wald test
AOL	14954.108	4.851	0.729	0.086	-0.020	0.004	4.711	892
	(1279.240)	(0.813)	(0.038)	(0.197))	(0.029)	(0.048)	(0.207)	< 0.01
Bank of Am.	363.193	1.423	0.573	0.203	0.014	0.032	3.865	375
	(73.767)	(0.257)	(0.030)	(0.061)	(0.022)	(0.032)	(0.211)	< 0.01
Bank One	32.856	1.365	0.640	0.737	-0.053	0.102	4.971	119
	(5.939)	(0.282)	(0.039)	(0.064)	(0.065)	(0.099)	(0.262)	< 0.01
Bear Stearns	119.039	1.900	0.661	0.587	0.000	0.014	5.122	184
	(21.405)	(0.345)	(0.031)	(0.076)	(0.054)	(0.074)	(0.265)	< 0.01
British Tel.	725.767	1.168	0.433	0.201	0.031	0.059	4.157	505
	(149.080)	(0.240)	(0.035)	(0.119)	(0.037)	(0.042)	(0.277)	< 0.01
Citigroup	66.632	0.835	0.649	0.535	0.010	0.045	4.652	290
	(11.593)	(0.146)	(0.034)	(0.049)	(0.044)	(0.054)	(0.212)	< 0.01
Commerzbank	488.851	0.663	0.893	0.151	-0.003	-0.038	3.229	233
	(112.523)	(0.237)	(0.023)	(0.032)	(0.011)	(0.033)	(0.153)	< 0.01
Daimler	39.521	2.142	0.494	1.043	0.013	0.388	4.277	83
	(8.166)	(0.442)	(0.043)	(0.123)	(0.139)	(0.101)	(0.334)	< 0.01
Deutsche Tel.	173.237	3.656	0.369	0.730	-0.064	0.227	2.298	115
	(38.506)	(0.721)	(0.033)	(0.206)	(0.137)	(0.073)	(0.210)	< 0.01
Fiat	310.047	1.192	0.500	0.352	-0.001	0.074	2.713	126
	(64.739)	(0.296)	(0.043)	(0.149)	(0.062)	(0.039)	(0.209)	< 0.01
FleetBoston	708.814	1.715	0.551	0.110	-0.008	0.061	3.834	329
	(132.471)	(0.359)	(0.035)	(0.057)	(0.018)	(0.044)	(0.257)	< 0.01
Ford	22.850	6.439	0.681	2.654	-0.199	0.765	3.655	12
	(4.188)	(1.171)	(0.037)	(0.194)	(0.401)	(0.324)	(0.284)	< 0.01
GE Capital	1588.611	1.629	0.866	0.155	0.002	0.036	4.734	438
	(383.333)	(0.353)	(0.023)	(0.058)	(0.011)	(0.036)	(0.208)	< 0.01
General Mot.	28.993	2.095	0.550	1.523	0.132	0.342	4.883	65
	(5.498)	(0.433)	(0.045)	(151)	(0.182)	(0.137)	(0.357)	< 0.01
Goldman Sachs	123.067	1.148	0.644	0.428	-0.056	0.015	4.693	325
	(21.717)	(0.215)	(0.031)	(0.055)	(0.037)	(0.047)	(0.247)	< 0.01
Iberdrola	632.948	5.582	0.885	0.117	-0.007	0.020	2.138	188
	(139.851)	(1.284)	(0.020)	(0.019)	(0.008)	(0.040)	(0.088)	< 0.01
JPMorgan	1229.124	4.515	0.481	0.105	-0.023	-0.011	2.171	367
	(208.009)	(0.852)	(0.033)	(0.081)	(0.031)	(0.029)	(0.169)	< 0.01
Lehman Bros.	82.699	1.067	0.560	0.612	-0.007	-0.027	6.210	208
	(14.513)	(0.197)	(0.034)	(0.076)	(0.060)	(0.072)	(0.344)	< 0.01
Merrill Lynch	861.935	0.922	0.662	0.172	-0.014	-0.008	4.995	882
	(164.308)	(0.149)	(0.031)	(0.080)	(0.018)	(0.027)	(0.231)	< 0.01
Metro	485.798	8.201	0.804	0.199	-0.043	0.022	2.115	166
	(103.403)	(1.821)	(0.028)	(0.046)	(0.023)	(0.056)	(0.112)	< 0.01
Morgan St.	345.376	2.220	0.679	0.293	-0.021	0.025	4.029	269
	(65.204)	(0.402)	(0.028)	(0.068)	(0.027)	(0.040)	(0.210)	< 0.01
Siemens	1393.361	4.135	0.695	0.098	-0.010	0.022	2.223	395
	(250.220)	(0.817)	(0.036)	(0.075)	(0.021)	(0.039)	(0.130)	< 0.01
Telefonica	567.298	1.463	0.623	0.275	0.004	0.044	2.291	600
	(117.300)	(0.248)	(0.030)	(0.105)	(0.031)	(0.023)	(0.118)	< 0.01
Volvo	225.962	2.786	0.689	0.460	0.075	0.052	2.124	261
	(40.362)	(0.481)	(0.029)	(0.076)	(0.042)	(0.030)	(0.109)	< 0.01
Wal Mart	408.636	2.157	0.882	0.138	-0.007	0.054	4.087	213
	(84.480)	(0.526)	(020)	(0.018)	(0.009)	(0.055)	(0.174)	< 0.01
Wells Fargo	82.850	1.418	0.749	0.496	0.000	-0.022	4.529	110
	(14.654)	(0.305)	(0.030)	(0.044)	(0.038)	(0.073)	(0.223)	< 0.01

Table 3.2: Mixture model estimation results. The table shows second step ML estimates of the mixture parameters using the first step VECM residuals as input. The CDS price is the first series, the bond spread the second. In parentheses we report standard errors from a parametric bootstrap (see Chapter 2). The last column gives the values of the Wald statistic for a test of $\psi_1 = \psi_2$ along with the corresponding p-values.

in Chapter 2. Table 3.3 shows that the mean of $\hat{\lambda}_{CS}$, averaged across reference entities, amounts to 0.84. This indicates a strong (weak) adjustment of the credit spread (CDS price) to previous day price. The Hasbrouck information share estimates also indicate a larger contribution of the CDS market to price discovery. While for some reference entities the bounds of the Hasbrouck information shares are narrow, they are quite wide for others. For instance, the lower bound of the CDS market Hasbrouck information share estimate for *Ford* amounts to 52.3 %, (s.e. = 20.3), the upper bound is 80.0 % (s.e. = 16.9).

The last column in Table 3.3 reports the estimates of the CDS market IS_X information shares. For the reference entity *Ford* the IS_X estimate amounts to 83.4 % (s.e. = 16.8), a value above the Hasbrouck information share upper bound estimate.

Table 3.3 shows that the more pronounced leadership of the CDS market indicated by our unique information share measure is a general result. For those reference entities with wide bounds, the IS_X information shares tend to be close to the Hasbrouck information share upper bounds. The CDS market IS_X estimate averaged across entities amounts to 86.1 % which is close to the mean upper bound of the Hasbrouck share.

This result of a distinct informational leadership of the more liquid CDS market corroborates the conclusions of Grammig et al. (2008) who study price discovery for internationally cross-listed stocks and identify relative market liquidity as the most important variable for explaining the information shares of home and foreign market. Liquidity, as a result of market design, attracts trading volume and promotes a market's leadership price discovery (see also Yan and Zivot 2010). Our findings suggest that this conclusion also holds for markets trading credit risk.

The scatter plots of the VECM residuals depicted in Figure 3.3 match and illustrate the liquidity story. The four panels show horizontally flattened dense centers of the bivariate distributions, which imply that tail observations for the credit spread residuals do not tend to be accompanied by extreme CDS residuals. However, when the CDS residual is large and positive (negative), the credit spread residual tends to be large and positive (negative), too. This pattern complies with the notion of a corporate bond market where transitory price changes may occur only due to a lack of liquidity. Price innovations in the more liquid CDS market, on the other hand, tend to convey information with respect to the price of credit risk which spills over contemporaneously to the credit spreads.

Reference Entity	λ_{CS}	ξ_{CDS}	ξ_{CS}	Hasbrouck IS(CDS)			IS_X(CDS)
				low	up	mid	
Ford	0.58	0.31	0.22	52.3	80.0	66.2	83.4
	(0.21)	(0.09)	(0.08)	(20.3)	(16.9)	(18.3)	(16.8)
Daimler	0.81	0.48	0.12	71.3	94.1	82.7	93.9
	(0.23)	(0.12)	(0.11)	(21.8)	(14.2)	(17.6)	(13.8)
Telefonica	0.59	0.29	0.20	65.9	87.0	76.4	86.9
	(0.25)	(0.12)	(0.10)	(23.3)	(20.0)	(21.4)	(18.0)
Fiat	0.79	0.45	0.11	79.4	97.6	88.5	97.6
	(0.20)	(0.16)	(0.14)	(20.2)	(12.5)	(15.8)	(9.1)
General Mot.	0.75	0.40	0.14	72.1	90.1	81.1	88.3
	(0.23)	(0.07)	(0.06)	(16.0)	(11.0)	(13.1)	(11.0)
Volvo	0.52	0.26	0.24	59.0	76.5	67.8	74.5
	(0.22)	(0.08)	(0.07)	(20.4)	(17.8)	(18.8)	(18.0)
British Tel.	0.87	0.50	0.07	82.9	97.8	90.3	97.5
	(0.14)	(0.15)	(0.12)	(20.9)	(13.5)	(16.7)	(13.1)
Fleet Boston	0.93	0.48	0.04	84.2	97.1	90.7	97.3
	(0.05)	(0.07)	(0.05)	(19.4)	(10.7)	(14.3)	(9.6)
Commerzbank	0.82	0.42	0.09	69.0	76.9	72.9	69.2
	(0.24)	(0.07)	(0.04)	(20.9)	(22.7)	(21.6)	(23.0)
Wal Mart	0.85	0.41	0.07	58.1	65.8	62.0	66.5
	(0.03)	(0.08)	(0.03)	(21.2)	(20.5)	(20.7)	(19.6)
Siemens	0.88	0.45	0.06	88.7	95.8	92.2	96.1
	(0.23)	(0.10)	(0.08)	(18.0)	(14.0)	(15.8)	(13.3)
Deutsche Tel.	0.72	0.83	-0.33	89.6	95.2	92.4	94.6
	(0.16)	(0.99)	(0.89)	(23.7)	(9.5)	(14.5)	(8.9)
Iberdrola	0.77	0.40	0.12	58.9	64.5	61.7	65.3
	(0.23)	(0.07)	(0.05)	(22.2)	(21.8)	(21.9)	(20.5)
Citigroup	0.72	0.33	0.13	65.8	70.7	68.3	70.4
	(0.15)	(0.04)	(0.03)	(14.9)	(13.7)	(14.0)	(13.6)
Bank One	0.69	0.30	0.14	51.5	56.0	53.8	58.2
	(0.17)	(0.03)	(0.03)	(16.7)	(16.2)	(16.2)	(14.8)
Bank of Am.	0.93	0.43	0.03	95.5	98.7	97.1	98.5
	(0.06)	(0.09)	(0.09)	(17.0)	(13.5)	(15.0)	(11.5)
Morgan St.	0.83	0.42	0.09	88.4	91.3	89.8	91.7
	(0.17)	(0.06)	(0.06)	(16.1)	(14.2)	(15.0)	(12.8)
Wells Fargo	0.75	0.33	0.11	67.1	69.1	68.1	67.1
	(0.11)	(0.04)	(0.04)	(18.4)	(17.2)	(17.2)	(15.6)
Lehman Bros.	0.80	0.39	0.10	84.0	86.0	85.0	84.2
	(0.11)	(0.04)	(0.03)	(10.9)	(11.1)	(10.9)	(10.1)
GE Capital	0.98	0.53	0.01	97.8	99.8	98.8	99.8
	(0.16)	(0.08)	(0.05)	(10.1)	(7.5)	(8.6)	(7.9)
Metro	0.88	0.43	0.06	94.0	95.4	94.7	96.8
	(0.19)	(0.07)	(0.06)	(15.0)	(13.8)	(14.2)	(12.4)
Bear Stearns	0.78	0.36	0.10	82.5	83.7	83.1	83.7
	(0.16)	(0.04)	(0.04)	(14.1)	(13.1)	(13.4)	(12.1)
Merrill Lynch	0.90	0.43	0.05	96.5	97.7	97.1	96.7
	(0.26)	(0.05)	(0.04)	(8.7)	(9.3)	(8.9)	(9.7)
JP Morgan	0.94	0.47	0.03	99.2	100.0	99.6	99.4
	(0.12)	(0.09)	(0.08)	(11.6)	(13.7)	(12.4)	(10.5)
Aol	0.97	0.48	0.01	99.6	99.9	99.7	99.9
	(0.23)	(0.08)	(0.07)	(6.4)	(5.2)	(5.7)	(6.9)
Goldman Sachs	0.80	0.37	0.10	84.1	84.1	84.1	85.6
	(0.21)	(0.04)	(0.04)	(12.9)	(12.6)	(12.5)	(11.5)
Mean	0.84	0.42	0.08	78.4	86.6	82.5	86.3
Std. Dev.	0.11	0.10	0.10	15.2	12.9	13.5	13.0

Table 3.3: Alternative measures for contributions to price discovery. The table reports the adjustment coefficient ratio ($\lambda_{CS} = \frac{|\alpha_{CS}|}{|\alpha_{CS}|+|\alpha_{CDS}|}$), long run impact coefficients (ξ_{CDS} and ξ_{CS}), Hasbrouck information shares for the CDS price (lower bound, upper bound, midpoint) and modified information shares for the CDS price (IS_X(CDS)). The values in parentheses are bootstrap standard errors.

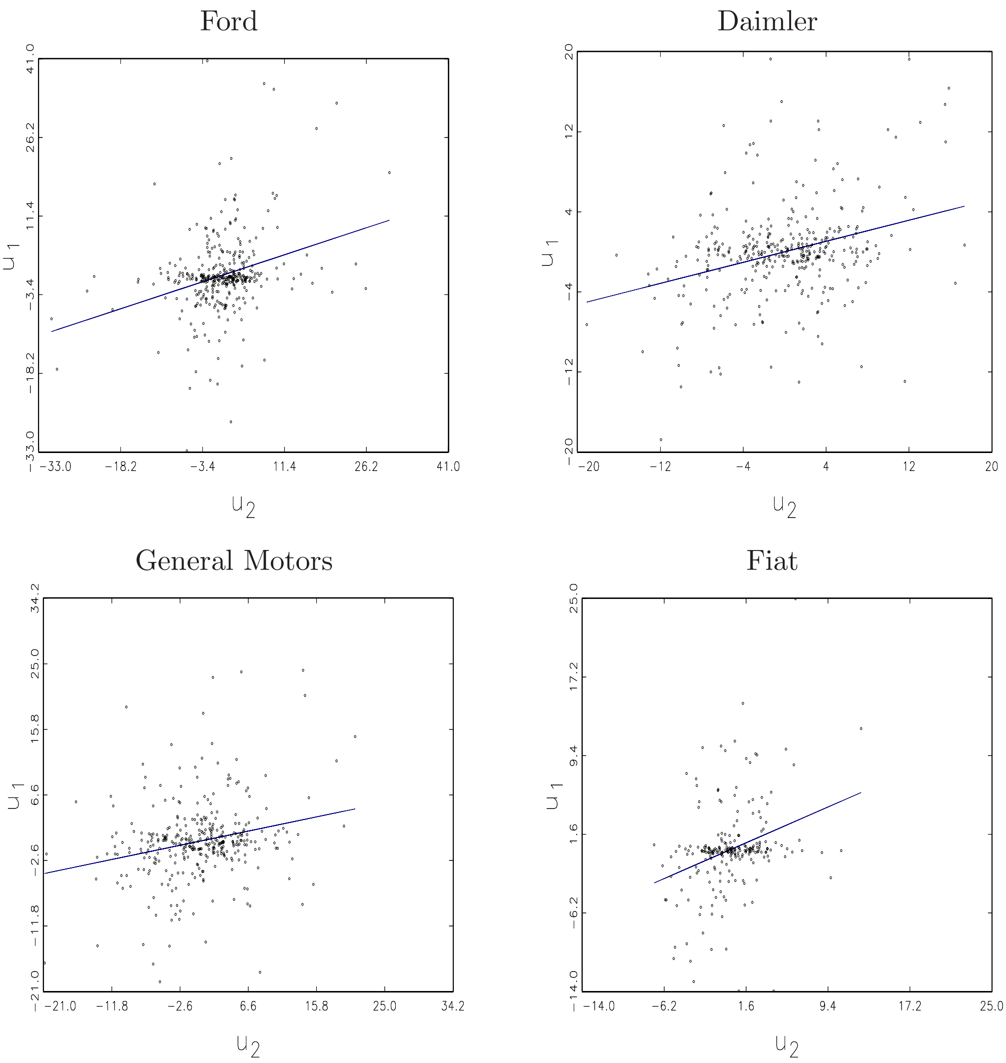


Figure 3.3: Scatter plots of VECM residuals. The four panels show scatter plots of residuals from the first step VECM estimation for four reference entities. u_1 are CDS residuals, u_2 credit spread residuals. The lines result from a regression of u_1 on u_2.

The estimates of the weight matrix $\mathbf{W}$ reported in Table 3.2 are in line with these scatter plots. The estimate of $w_{2,1}$, the weight with which an idiosyncratic CDS innovation contemporaneously affects the credit spread, tends to be larger than the estimate of $w_{1,2}$, the weight with which an idiosyncratic credit spread innovation contemporaneously affects the CDS price. The estimate of $w_{1,2}$ is in most cases not significantly different from zero. The relative illiquidity of the corporate bond market thus helps to identify contemporaneous effects and facilitates the estimation of unique information shares.

For some of the reference entities, the bounds of the Hasbrouck information shares are narrow because the contemporaneous correlations of the credit spread and CDS price residuals are small. In these cases, the estimates of the off-diagonal elements of **W** are small and not significantly different from zero, and the IS_X estimates are close to the Hasbrouck information share midpoints. We take it as a sign of robustness that both the standard identification method and the one proposed in this chapter deliver very similar results when no ambiguity in terms of wide bounds prevails. Furthermore, the estimation precision in terms of standard errors is comparable for Hasbrouck and IS_X information shares. Hence, the increase in precision offered by our methodology is unambiguous.

3.5 Concluding Remarks

"Where does price discovery take place?" is one of the key questions in empirical finance. It is raised when studying the competition for order flow between traditional and alternative trading platforms, national and international exchanges, and parallel markets for traditional and innovative financial instruments. Hasbrouck's (1995) methodology is the standard approach to address this research question empirically. He proposes to estimate the information share for each of the parallel markets on which financial instruments linked by the law of one price are traded. Information shares result from a variance decomposition of the innovations of the markets prices' common stochastic trend which is associated with the notion of the efficient price of the underlying security.

The competitive edge of Hasbrouck's information shares over alternative methodologies to measure contributions to price discovery is widely accepted (see the synopsis by Lehmann 2002). However, most applications suffer from a lack of identification since the contemporaneous dependence structure of price innovations across markets cannot be disentangled without further restrictions. As a solution, Hasbrouck (1995) performs a Cholesky decomposition of the covariance matrix of the price innovations. Thereby a hierarchical ordering of markets is assumed that is hardly ever justifiable. In empirical work researchers often resort to permuting the ordering of the markets, which yields upper and lower bounds of information shares rather than a unique measure. These bounds can become so wide that it is impossible to determine even the leading market.

This chapter resolves the problem of undetermined information shares by exploiting the informational content of distributional properties of financial prices. We show that different dependencies of contemporaneous price innovations in the tails and in the center of the distributions deliver the necessary information to determine unique information shares. Such tail dependence can be caused by the design of the trading process which may induce market specific liquidity effects. Since in most applications of the Hasbrouck methodology the market structures are clearly different - this is why alternative trading platforms emerge in the first place - our methodology presents an appealing solution. Regarding the pricing of credit risk, it is the relatively higher liquidity of the CDS market compared to the corporate bond market which sharpens the finding of the informational leadership of the credit derivatives market during the pre-crisis period.

The relation between market liquidity and contributions to price discovery has recently been emphasized by Yan and Zivot (2010). Our methodology systematically exploits the informational content of those market design effects and thereby delivers a unique measure for a market's information share. Researchers concerned with quantifying contributions to price discovery have a new tool to sharpen their conclusions.

Appendix A: Propositions and Proofs

Proposition 1. *Denote by* $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$ *the set of mixture parameters that yields the density of* $f(\mathbf{u}_t;\theta_m)$ *given in Equation (3.6), and by* $\theta_v = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_q\}$ *a set of VECM parameters. Suppose the main diagonal elements of* $\mathbf{W}$ *are all greater than zero, and that the elements of the diagonal matrix* $\mathbf{\Psi}$ *are distinct. Furthermore, let* $\mathbf{IS}_X(\theta_m, \theta_v)$ *denote the vector of information shares given by Equation (3.5). Then, holding the mixture probability* γ *fixed, there exist* $n! - 1$ *further sets of mixture parameters* $\theta_m^* = \{\gamma, \mathbf{\Psi}^*, \mathbf{W}^*\}$ *given by* $n! - 1$ *distinct permutations of the columns in* $\mathbf{W}$ *and the corresponding elements in* $\mathbf{\Psi}$*,*

$$\mathbf{\Psi}^* = \mathbf{P}'\mathbf{\Psi}\mathbf{P} \tag{3.12}$$

$$\mathbf{W}^* = \mathbf{W}\mathbf{P} \tag{3.13}$$

where $\mathbf{P}$ *is a permutation matrix of order* n*. The parametrization* θ_m^* *are observationally equivalent to* θ_m *in that*

$$f(\mathbf{u}_t;\theta_m) = f(\mathbf{u}_t;\theta_m^*) \quad . \tag{3.14}$$

and permute the original vector of information shares according to

$$\mathbf{IS}_X(\theta_m^*, \theta_v) = \mathbf{IS}_X(\theta_m, \theta_v) \times \mathbf{P} \quad . \tag{3.15}$$

Proof: To prove the first part of Proposition 1 note that the observational equivalence of two mixture parametrization $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$ and $\theta_m^* = \{\gamma, \mathbf{\Psi}^*, \mathbf{W}^*\}$ entails identity of the variance covariance matrices $Var(\mathbf{u}_t) = \gamma\mathbf{W}\mathbf{W}' + (1-\gamma)\mathbf{W}\mathbf{\Psi}\mathbf{W}' = \gamma\mathbf{W}^*\mathbf{W}^{*\prime} + (1-\gamma)\mathbf{W}^*\mathbf{\Psi}^*\mathbf{W}^{*\prime}$. Hence, let $\mathbf{Q}$ be a matrix, such that $\mathbf{W}^* = \mathbf{W}\mathbf{Q}$, and $\mathbf{\Psi}^*$ be a diagonal matrix with distinct positive elements. Then $f(\mathbf{u}_t;\theta_m) = f(\mathbf{u}_t;\theta_m^*)$ implies

$$\mathbf{W}[\gamma\mathbf{I}_n + (1-\gamma)\mathbf{\Psi}]\mathbf{W}' = \mathbf{W}\mathbf{Q}[\gamma\mathbf{I}_n + (1-\gamma)\mathbf{\Psi}^*]\mathbf{Q}'\mathbf{W}' \quad . \tag{3.16}$$

Multiplication of 3.16 from the left with $\mathbf{W}^{-1}$ and from the right with its transpose and rearranging terms yields

$$\gamma(\mathbf{I}_n - \mathbf{Q}\mathbf{Q}') = (1-\gamma)(\mathbf{Q}\mathbf{\Psi}^*\mathbf{Q}' - \mathbf{\Psi}) \quad . \tag{3.17}$$

This holds for $0 < \gamma < 1$ only if both sides of Equation (3.17) are zero which implies that

$$\mathbf{\Psi} = \mathbf{Q}\mathbf{\Psi}^*\mathbf{Q}' \tag{3.18}$$

and

$$\mathbf{Q}\mathbf{Q}' = \mathbf{I}_n \quad . \tag{3.19}$$

It follows from Equation (3.19) that $\mathbf{Q}$ has to be orthogonal, i.e. $\mathbf{Q}' = \mathbf{Q}^{\mathbf{-1}}$. Hence Equation (3.18) can be regarded as a spectral decomposition of $\mathbf{\Psi}$, where $\mathbf{\Psi}^*$ contains the eigenvalues of $\mathbf{\Psi}$ on its diagonal, and the columns of $\mathbf{Q}$ are the corresponding eigenvectors. As all elements of $\mathbf{\Psi}$ are assumed to be distinct, the columns of $\mathbf{Q}$ are linearly independent, unit length vectors. Consequently, all possible solutions for $\mathbf{Q}$ are given by $\mathbf{Q} = \mathbf{PS}$, where $\mathbf{P}$ is an n-dimensional permutation matrix and $\mathbf{S}$ an n-dimensional diagonal matrix, whose diagonal elements are either 1 or -1. Therefore $\mathbf{W}^* = \mathbf{WPS}$, which implies that the columns of $\mathbf{W}$ are identified up to multiplication by -1. However, as the main diagonal elements of $\mathbf{W}$ are restricted to be greater than zero, only $\mathbf{S} = \mathbf{I}_n$ is eligible which yields (3.13). This implies that there exit $n!$ permutations of the columns in $\mathbf{W}$ of which $n! - 1$ yield a matrix $\mathbf{W}^*$ which is distinct from $\mathbf{W}$. The only permutation matrix that leaves the ordering of the columns in $\mathbf{W}$ unchanged is $\mathbf{P} = \mathbf{I}_n$. Regarding Equation (3.18) it follows that

$$\mathbf{\Psi} = \mathbf{PS}\mathbf{\Psi}^*\mathbf{S}'\mathbf{P}' = \mathbf{P}\mathbf{\Psi}^*\mathbf{P}' \quad . \tag{3.20}$$

Solving for $\mathbf{\Psi}^*$ yields (3.12). $\mathbf{\Psi}^* = \mathbf{P}'\mathbf{\Psi}\mathbf{P}$ is a diagonal matrix, which results from a permutation of the diagonal elements of $\mathbf{\Psi}$. This proves the first part of Proposition 1.

To prove (3.15), start from Equation (3.5), which written in detail reads

$$\mathbf{IS}_X(\theta_m, \theta_v) = \frac{\left[\boldsymbol{\xi}'\mathbf{W}(\gamma\mathbf{I}_n + (1-\gamma)\mathbf{\Psi})^{0.5}\right]^{(2)}}{\boldsymbol{\xi}'\mathbf{W}(\gamma\mathbf{I}_n + (1-\gamma)\mathbf{\Psi})\mathbf{W}'\boldsymbol{\xi}} \quad . \tag{3.21}$$

Since (3.16) holds, θ_m and θ_m^* imply the same covariance matrix of $\mathbf{u}_t$, the denominator in Equation (3.21) is not affected by the permutation of elements in $\mathbf{W}$ and $\mathbf{\Psi}$ according to Equations (3.12) and (3.13). Therefore $\mathbf{IS}_X(\theta_m^*, \theta_v)$ can differ from $\mathbf{IS}_X(\theta_m, \theta_v)$ only by their numerators, which relate to each other by

$$
\begin{aligned}
[\boldsymbol{\xi}'\mathbf{W}^*(\gamma\mathbf{I}_n + (1-\gamma)\boldsymbol{\Psi}^*)^{0.5}]^{(2)} &= [\boldsymbol{\xi}'\mathbf{W}\mathbf{P}(\gamma\mathbf{I}_n + (1-\gamma)\mathbf{P}'\boldsymbol{\Psi}\mathbf{P})^{0.5}]^{(2)} \\
&= [\boldsymbol{\xi}'\mathbf{W}\mathbf{P}\mathbf{P}'\boldsymbol{\Sigma}_e^{0.5}\mathbf{P}]^{(2)} \\
&= [\boldsymbol{\xi}'\mathbf{W}'\boldsymbol{\Sigma}_e^{0.5}]^{(2)}\mathbf{P} \quad .
\end{aligned}
$$

Thus, $\mathbf{IS}_X(\theta_m^*, \theta_v)$=$\mathbf{IS}_X(\theta_m, \theta_v)\mathbf{P}$, such that $\mathbf{IS}_X(\theta_m^*, \theta_v) \neq \mathbf{IS}_X(\theta_m, \theta_v) \; \forall \; \mathbf{P} \neq \mathbf{I}_n$. This leaves $n! - 1$ distinct permutation matrices $\mathbf{P}$ associated with $n! - 1$ different sets of mixture parameters which are observationally equivalent, but imply different information share vectors. Thereby the proposition is proven. □

Proposition 2. *Denote by* $\theta_m = \{\gamma, \mathbf{\Psi}, \mathbf{W}\}$ *the set of mixture parameters that yields the density of* $f(\mathbf{u}_t; \theta_m)$ *given in Equation (3.6), and by* $\theta_v = \{\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{\Gamma}_1, \ldots, \mathbf{\Gamma}_q\}$ *a set of VECM parameters. Suppose that the elements of the diagonal matrix* $\mathbf{\Psi}$ *are distinct. Furthermore, let* $\mathbf{IS}_X(\theta_m, \theta_v)$ *denote the vector of information shares given by Equation (3.5). If it holds for the elements of* $\mathbf{W}$ *that*

$$\begin{aligned} w_{i,i} &> 0 \qquad \forall \quad i \\ w_{i,i} &> |w_{j,i}| \qquad \forall \quad j \neq i \end{aligned} \tag{3.22}$$

then there exists only one set of mixture parameters $\overline{\theta} = \{\overline{\gamma}, \overline{\mathbf{\Psi}}, \overline{\mathbf{W}}\}$*, given by*

$$\begin{aligned} \overline{\gamma} &= 1 - \gamma \\ \overline{\mathbf{\Psi}} &= \mathbf{\Psi}^{-1} \\ \overline{\mathbf{W}} &= \mathbf{W}\mathbf{\Psi}^{0.5} \quad , \end{aligned} \tag{3.23}$$

that is observationally equivalent to θ_m *in that* $f(\mathbf{u}_t; \theta_m) = f(\mathbf{u}_t; \overline{\theta}_m)$*. Furthermore, the parametrization* $\overline{\theta}_m$ *implies*

$$\mathbf{IS}_X(\overline{\theta}_m, \theta_v) = \mathbf{IS}_X(\theta_m, \theta_v) \quad . \tag{3.24}$$

Proof: Let $\mathbf{Q}$ be a matrix, such that $\overline{\mathbf{W}} = \mathbf{W}\mathbf{Q}$, then it has to hold that

$$\mathbf{W}[\gamma \mathbf{I}_n + (1-\gamma)\mathbf{\Psi}]\mathbf{W}' = \mathbf{W}\mathbf{Q}[(1-\gamma)\mathbf{I}_n + \gamma\overline{\mathbf{\Psi}}]\mathbf{Q}'\mathbf{W}' \quad . \tag{3.25}$$

By multiplying (3.25) from the left with $\mathbf{W}^{-1}$ and from the right with its transpose and rearranging terms yields

$$\gamma(\mathbf{I}_n - \mathbf{Q}\overline{\mathbf{\Psi}}\mathbf{Q}') = (1-\gamma)(\mathbf{Q}\mathbf{Q}' - \mathbf{\Psi}) \quad . \tag{3.26}$$

This holds only if both sides of Equation (3.26) are zero, which implies that $\mathbf{Q}\overline{\mathbf{\Psi}}\mathbf{Q}' = \mathbf{I}_n$ and that $\mathbf{Q}\mathbf{Q}' = \mathbf{\Psi}$. The latter equation gives $\mathbf{Q} = \mathbf{\Psi}^{0.5}$, and using this result for the first yields $\overline{\mathbf{\Psi}} = \mathbf{\Psi}^{-1}$ and $\overline{\mathbf{W}} = \mathbf{W}\mathbf{\Psi}^{0.5}$, which shows Equation (3.23). The restrictions in (3.22) rule out permuting the columns of $\overline{\mathbf{W}}$ and the diagonal elements of $\overline{\mathbf{\Psi}}$. Without these restrictions $\widetilde{\mathbf{W}} = \overline{\mathbf{W}}\mathbf{P}$ and $\widetilde{\mathbf{\Psi}} = \mathbf{P}\overline{\mathbf{\Psi}}\mathbf{P}'$ with $\mathbf{P} \neq \mathbf{I}_n$ would yield $n!-1$ observationally equivalent parametrization. Thereby the first part of Proposition 2 is proven.

Since (3.25) holds, the vector $\mathbf{IS}_X(\overline{\theta}_m)$ can only differ from $\mathbf{IS}_X(\theta_m)$ by the vectors in the numerators, but

$$\begin{aligned}
[\boldsymbol{\xi}'\overline{\mathbf{W}}(\overline{\gamma}\mathbf{I}_n + (1-\overline{\gamma})\overline{\boldsymbol{\Psi}})^{0.5}]^{(2)} &= [\boldsymbol{\xi}'\mathbf{W}\boldsymbol{\Psi}^{0.5}((1-\gamma)\mathbf{I}_n + \gamma\boldsymbol{\Psi}^{-1})^{0.5}]^{(2)} \\
&= [\boldsymbol{\xi}'\mathbf{W}((1-\gamma)\boldsymbol{\Psi}\mathbf{I}_n + \gamma\boldsymbol{\Psi}\boldsymbol{\Psi}^{-1})^{0.5}]^{(2)} \\
&= [\boldsymbol{\xi}'\mathbf{W}(\gamma\mathbf{I}_n + (1-\gamma)\boldsymbol{\Psi})^{0.5}]^{(2)} \quad .
\end{aligned} \tag{3.27}$$

As the right hand side of (3.27) is the numerator of $\mathbf{IS}_X(\theta_m)$ it follows that $\mathbf{IS}_X(\overline{\theta}_m) = \mathbf{IS}_X(\theta_m)$ which proves the second part of Proposition 2. □

4 An Intensity Based Information Share

In this chapter we propose a new measure for contributions to price discovery based on Russell's (1999) autoregressive conditional intensity model. While previous studies rely on equally spaced high frequency data, we use the information conveyed by quote revision intensities to determine a market's information share. Thereby, we account for the irregular nature of the data. An empirical application to U.S.listed Canadian stocks supports previous evidence for the home market leadership in price discovery. Based on a cross sectional analysis we confirm the positive link between liquidity and contributions to price discovery.

This chapter is based on the article *A new approach to estimate unique market information shares* by Kerstin Kehrle and Franziska J. Peter (2010).

4.1 Introduction

According to Coffee (2002), increasing globalization and improved technology will lead to a decay in the number of securities exchanges around the world. Small national exchanges will lose their share in trading to large international exchanges, which provide a more efficient trading environment. Carpentier et al. (2007) examine this development for the Canadian stock exchanges with respect to the U.S. markets. They report a rapidly growing share of U.S. markets in trades of Canadian stocks up to the point where interlisted stocks are absorbed by the foreign market and delisted on the home market. These developments foreshadow small national stock exchanges as markets for illiquid stocks that fail to attract investors on the large markets (see Gaa et al. 2002). Thus, within the context of international cross-listed stocks, it is of paramount interest to national stock exchanges that they remain the dominant market with regard to price discovery.[7] The competition among smaller national and the large U.S. markets for the leadership in price discovery has therefore grown immensely and has stirred up an increasing amount of research. The main contribution of this chapter is summarized as follows. We develop a new information share, i.e. a measure for the home and foreign market contributions to the price discovery process, by applying Russell's (1999) autoregressive conditional intensity model (ACI). The bivariate intensity approach accounts for the informational content of time between consecutive quoted price changes within a market and the timing interdependencies between the price processes on both markets. In contrast to the commonly applied Hasbrouck (1995) methodology we exploit the irregular occurrence of price changes and deliver a unique information share rather than lower and upper bounds. In an empirical application we analyze the price discovery process of Canadian stocks that are traded on the Toronto Stock Exchange (TSX) and cross-listed on the New York Stock Exchange (NYSE). Furthermore, we examine potential determinants of information shares in a cross sectional analysis.

Evidence from previous studies suggests that the main part of price discovery for cross-listed stocks takes place in the home market. Eun and Sabherwal (2003) examine a sample of U.S. listed Canadian stocks based on the relative adjustment of prices in a market to deviations from the equilibrium price. They conclude that the contribution of the

[7] For a comprehensive study concerned with cross-listings in stock markets see Karolyi (2006).

U.S. market cannot be neglected, while the home market clearly leads price discovery. Adjustment coefficients as a measure for price discovery, however, have been criticized, since they do not account for the contemporaneous correlations and variances of market's price innovation (see De Jong 2002, Baillie et al. 2002). The major part of empirical studies applies Hasbrouck's (1995) method, who defines the information share as the contribution of a market's price innovation to the variance of the efficient underlying price innovations. Grammig et al. (2005), Hupperets and Menkveld (2002) and Korczak and Phylaktis (2010) use the Hasbrouck (1995) methodology to estimate the home and foreign market share in price discovery for U.S. listed stocks from various countries. They conclude that trading on the home market stock exchanges contributes most to price discovery, while trading on the NYSE primarily takes place to offset arbitrage opportunities.

The main drawback of the Hasbrouck (1995) approach is that it merely delivers upper and lower bounds for an information share. The method requires equidistant sampled quotes and depending on the chosen sampling frequency the information share bounds can diverge considerably. Consequently, the conclusions concerning the leading market are rather vague, see Hupperets and Menkveld (2002) and Korczak and Phylaktis (2010). We revisit the question of how to measure the contribution to price discovery in a multiple market setting based on the following considerations. First, as it is acknowledged in the financial markets literature (see e.g. Dufour and Engle 2000, Engle 2000, Engle and Lunde 2003, Frijns and Schotman 2009), the irregular occurrence of trades and quotes and the time between consecutive financial market events reveal the dynamics of price responses to new information in the market. Hence, arbitrary sampling schemes used to obtain regular spaced data neglect this part of the price dynamics and induce an undesirable loss of information. Second, as pointed out by Hasbrouck (1995), "...the information share measures *who moves first* in the process of price adjustment." If price discovery is indeed understood as which market "moves first", an information share measure that is derived from an approach that directly embeds the irregular sequence and timing of the price process seems to be a straightforward consequence. We take this irregularity of the data into account and derive an information share by modeling the arrival rates (intensities) of the price processes using Russell's (1999) ACI model. An intensity roughly gives the probability of a cumulated quote change within the next instant.

The dynamics of the intensity functions are driven by innovations that allow for a flexible interaction and simultaneously affect the conditional intensities on both markets. Since the arbitrage relation between prices in parallel markets force an immediate incorporation of new information arising in one market in the second market's price, we expect market spill over effects due to the innovations. We argue that the larger the effect of an innovation in one market on the other markets intensity, the more the former contributes to the price discovery process. Our proposed method therefore uses these cross effects to derive a new unique measure for contributions to price discovery which does not suffer from an identification problem inherent in the Hasbrouck (1995) approach. We empirically analyze the price discovery process of Canadian stocks, which are traded on the TSX and cross-listed on the NYSE. Our results show a clear leadership of the TSX in the price discovery process. With an average information share of 73%, the contribution of the TSX is more pronounced than indicated by Eun and Sabherwal (2003). Furthermore, we show that the intensity based information share is able to detect the leading market for the majority of the sample stocks. We examine potential determinants of a market's contribution to price discovery by conducting a cross sectional regression of the intensity based information share on stock specific factors and liquidity related variables. Our results show that only liquidity proxies as the relative spread, medium trades and trading volume contribute to a market's price determination. This implies that providing an efficient and liquid trading environment, is of special interest for small national stock exchanges that seek to maintain their dominance in the price discovery process of cross-listed stocks.

4.2 A New Measuring for Price Discovery in International Stock Markets

4.2.1 The Autoregressive Conditional Intensity Model

In the empirical application we consider stocks that are simultaneously traded on the TSX and the NYSE. This section introduces a bivariate autoregressive conditional intensity (ACI) model that is applied to price events[8] of a stock listed on both markets. We start

[8] We define price events as cumulated absolute midquote changes that we refer to as informative price events. A detailed description follows in Section 4.2.2.

by defining the point process $\{t_i^s\}_{i=1}^{n^s}$ as the stochastic sequence of price changes on the market s in calendar time t, where $s = 1$ corresponds to a price event on the TSX and $s = 2$ refers to a price event on the NYSE. The associated counting functions that count the number of s-type events through t are indexed by $N^s(t)$. Pooling and ordering of the arrival times, t_i^1 and t_i^2 yields a simple point process $\{t_i\}_{i=1}^{n}$ with counting function $N(t)$. We assume that the arrival times are strictly distinct, $0 < t_1 < t_2 ... < t_n$. Due to this assumption the individual point processes are strictly orderly, too. Figure 4.1 gives an illustration of a pooled point process $N(t)$ consisting of two individual processes $N^1(t)$ and $N^2(t)$.

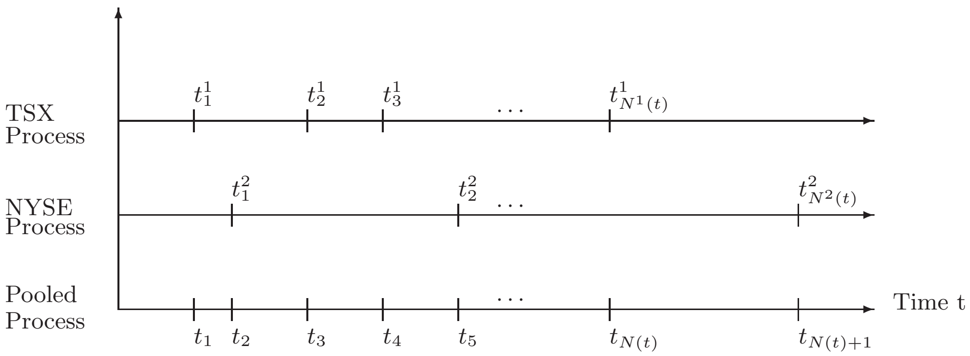


Figure 4.1: Pooled point process illustration. The figure gives an illustration of a simple point process $N(t)$ that consists of two individual counting processes $N^1(t)$ and $N^2(t)$. $\{t_i^1\}_{i=1}^{n^1}$ denotes the arrival times of events on the TSX and $\{t_i^2\}_{i=1}^{n^2}$ corresponds to price event times on the NYSE. A time sequence $\{t_i\}_{i=1}^{n}$ containing both event time series is obtained by pooling and ordering the individual event times. As a consequence, an event occurring on the TSX does not depend only on its own history but is allowed to depend on the history of the NYSE process, as well, and vice verse.

The internal filtration denoted by $\Im_t$ consists of the complete information path of the left continuous counting process $N(t)$. The $\Im_t$-intensity process that characterizes the evolution of $N^s(t)$ is then

$$\lambda^s(t;\Im_t) = \lim_{\Delta\to 0}\frac{1}{\Delta}\mathbb{P}\left[N^s(t+\Delta) - N^s(t) > 0, N^{s'}(t+\Delta) - N^{s'}(t) = 0|\Im_t\right] \tag{4.1}$$

$\forall s \neq s'$, where $s' = 1, 2$. Equation (4.1) gives an instantaneous probability of observing an s-type price event conditional on the information set available at t.

Russell's (1999) ACI model defines the s-type conditional intensity function as,

$$\lambda^s(t;\Im_t) = \lambda_o^s\psi^s(t)\phi^s(t) \quad , \tag{4.2}$$

with a baseline intensity function given by $\lambda_o^s = \exp(w^s)$ and $\psi^s(t)$ which captures the dynamic structure of the conditional intensity. $\phi^s(t)$ accounts for diurnal patterns which are common when dealing with high frequency financial data. The 2×1 vector, $\boldsymbol{\psi}_i = (\psi_i^1, \psi_i^2)'$, is parametrized in terms of a vector autoregressive moving average (VARMA(1,1)) process,

$$\tilde{\boldsymbol{\psi}}_i = \sum_{s=1}^{2} (\mathbf{a}^s \varepsilon_{i-1}^s + \mathbf{B}\tilde{\boldsymbol{\psi}}_{i-1}) y_{i-1}^s \quad , \tag{4.3}$$

where $\mathbf{a^s}$ is a 2×1 coefficient vector and $\mathbf{B}$ is a 2×2 autoregressive coefficient matrix. In order to ensure positivity of $\lambda^s(t; \Im_t)$ in Equation (4.2), we define $\psi^s(t) = \exp\left(\tilde{\psi}_{N(t)}^s\right)$. y_i^s denotes an indicator variable that takes on the value one if the i^{th} price event of the pooled process comes from the TSX ($s = 1$) or the NYSE ($s = 2$) process and zero otherwise. Henceforth, we denote an innovation originating in the home market (TSX) with superscript 1 and the corresponding coefficient vector with $\mathbf{a}^1$. The first element of $\mathbf{a}^1$ (a_1^1) then measures the impact of a TSX innovation on the TSX conditional intensity. The second element (a_2^1) gives the cross effect of an innovation in the TSX process on the NYSE conditional intensity. Analogously, we denote NYSE associated shocks with a superscript 2. Following earlier studies (see Russell 1999, Bauwens and Hautsch 2006, Hall and Hautsch 2006), we restrict the autoregressive coefficient matrix $\mathbf{B}$ to be diagonal. Then, the diagonal elements of $\mathbf{B}$ determine the long run impact of a shock and stationarity of the process is ensured if the eigenvalues of $\mathbf{B}$ (i.e. its diagonal elements) lie inside the unit circle.

According to Russell (1999) the specification of the innovation in Equation (4.3) is based on the integrated intensity which is computed by piecewise integration of $\lambda^s(t; \Im_t)$,

$$\Lambda^s(t_{i-1}^s, t_i^s) = \int_{t_{i-1}^s}^{t_i^s} \lambda^s(u; \Im_u) du = \sum_j \int_{\tilde{t}_j}^{\tilde{t}_{j+1}} \lambda^s(u; \Im_u) du \quad , \tag{4.4}$$

for j denoting all points with $t_{i-1}^s < \tilde{t}_j < \tilde{t}_{j+1} \leq t_i^s$.

Using the random time change theorem any non-Poisson process can be transformed into a standard Poisson process which implies an iid standard exponential distributed integrated intensity, i.e. $\Lambda^s(t_{i-1}^s, t_i^s) \sim$ iid Exp(1) (see Hautsch 2003, Brémaud 1981, Bowsher 2007). Following Bauwens and Hautsch (2006), we then define the innovation in Equation (4.3) as logarithm of an iid exponential variate centered by its unconditional expectation,[9]

$$\varepsilon_i^s = -0.5772 - \ln \Lambda^s(t_{i-1}^s, t_i^s) \quad . \tag{4.5}$$

Hence, an innovation in the ACI model has the interpretation of the deviation between the realized number of events and the expected number of events within the interval $(t_{i-1}^s, t_i^s]$. This implies that positive values of ε_i indicate an underprediction of arrival rates and negative values an overprediction.

The model parameters are estimated by the method of maximum likelihood. The log-likelihood function of the two-dimensional ACI process can be expressed as

$$\ln \mathscr{L}(\boldsymbol{\theta}) = \sum_{s=1}^{2} \sum_{i=1}^{n} \{-\Lambda^s(t_{i-1}, t_i) + y_i^s \ln \lambda^s(t_i; \Im_{t_i})\} \quad , \tag{4.6}$$

with $\boldsymbol{\theta} = (\omega^s, \mathbf{a}^s, \mathbf{B})$ collecting the parameters of interest for $s = 1, 2$. The first term on the right hand side of Equation (4.6) corresponds to the s-type intensity integrated over $(t_{i-1}, t_i]$ and the second to the probability of the arrival times in the pooled process. The log-likelihood can be maximized by standard nonlinear optimization algorithms.

If the model is specified correctly, the resulting s-type residuals, $\tilde{\varepsilon}_i^s = \Lambda^s(t_{i-1}^s, t_i^s)$, should be iid unit exponentially distributed. Hence, the dynamic and distributional properties of the estimated residuals can be evaluated by an overdispersion test suggested by Engle and Russell (1998). Their test statistic against excess dispersion, $\sqrt{\frac{n^s}{8}}(\sigma^2_{\tilde{\varepsilon}^s} - 1)$, follows asymptotically a normal distribution, where $\sigma^2_{\tilde{\varepsilon}^s}$ denotes the variance of the s-type residual series and n^s denotes the number of price events in the process s. The amount of autocorrelation in the residuals not explained by the specified model is assessed by a Ljung-Box test. However, if the assumption of exponentially distributed innovations is not supported by the data, we achieve consistent quasi-maximum likelihood estimators of $\boldsymbol{\theta}$ by specifying only the conditional mean of the innovation distribution (see White 1982, Bollerslev and Wooldridge 1992).

[9] As indicated by Hautsch (2003), the logarithm of an iid exponential variate yields a minimum Gumbel variate, $\ln \Lambda^s(t_{i-1}^s, t_i^s)$, with mean -0.5772 and variance $\sigma_\varepsilon^2 = \pi^2/6$.

4.2.2 The Data

We empirically examine Canadian stocks traded on the TSX, which are cross-listed on the NYSE. These cross-listings are particularly well suited for price discovery analysis, since the trading times of the TSX and NYSE coincidence and the whole trading period can be examined. Apart from that the large number of Canadian NYSE listed stocks allows for further cross sectional analysis concerning the determinants of information shares. In detail we use quote data for 83 Canadian NYSE listed stocks. The NYSE data are extracted from the Trade and Quote (TAQ) DVDs supplied by the NYSE. Toronto quote and trade data were taken from the Equity Trades and Quotes data set provided by the TSX. The sample period covers 62 trading days from 1^{st} of January 2004 to 31^{st} of March 2004. Continuous trading on both exchanges takes place from 9:30am to 4:00pm. Table 4.1 gives the stock ticker and company names.

Following Engle and Russell (1997) and Bauwens and Hautsch (2006), the quote data were thinned based on a function of price marks. We first compute bid and ask midquotes and construct cumulative absolute price changes that are retained whenever they exceed a specific threshold. We set this threshold to 0.025 Canadian Dollars for the TSX returns, which after accounting for the exchange rate and the minimum tick size corresponds to 0.02 U.S. Dollars for the NYSE return series. The thinning algorithm is applied to reduce the amount of noise due to microstructure effects that should not be considered as a movement in the fundamental price. Thereby, we obtain what we refer to as informative price events.[10] The ACI model introduced in the previous section assigns zero probability to the simultaneous occurrence of two events and therefore quote revisions with the same time stamp within one market are treated as one. Finally, events with the same time stamp in both markets are deleted.[11] Table 4.2 gives summary statistics for the filtered data. Across 83 stocks the average daily number of quote revisions is 158 on TSX and 181 on NYSE. Quote revisions on TSX and NYSE occur on average every 4.5 min and 3.5 min, respectively. On the TSX the most frequent stock has a midquote change every 23 s and the least quoted stock updates its midquotes every 21 min on average. Accordingly,

[10] The filtering process is subject to an arbitrary threshold and for robustness checks we conduct all subsequent analyses for the next higher and next lower thresholds. The results are qualitatively similar and are available upon request.

[11] On average across all sample stocks only 5% of the data are affected by this selection rule.

Ticker	Company Name	Industry
ABX	Barrick Gold	Gold Mining
ABY	Abitibi Consolidated Inc.	Paper
AEM	Agnico Eagle Mines Ltd.	Gold Mining
AGU	Agrium Inc.	Chemicals (Specialty)
AL	Alcan Inc.	Metals and Mining
BCE	BCE Inc.	Foreign Telecom.
BCM	Canadian Imp. Bank of Commerce	Bank
BEI	Boardwalk Equities	Real Estate Holding
BGM	General Motors Corp.	Automobile
BMO	Bank of Montreal	Bank
BNN	Brascan Corp.	Real Estate Holding
BNS	Bank of Nova Scotia	Bank
BPO	Brookfield Properties Corporation	Real Estate Holding
BR	Broadridge Inc.	IT Services and Consulting
BVF	Biovail Corp.	Pharmaceuticals
CCJ	Cameco Corp.	Nonferrous Metals
CGT	CAE Inc.	Aerospace
CJR	Chorus Entertainment Inc.	Broadcasting and Entertainment
CLS	Celestica Inc.	Electronics
CNI	Canadian National Railway	Transport
CNQ	Canadian Natural Ressources	Petroleum (Producing)
COT	Cott Corp.	Soft Drinks
CP	Canadian Pacific Railway	Transport
CWG	CanWest Global Communications Corp.	Broadcasting and Entertainment
DTC	Domtar Corp.	Paper
ECA	EnCana Corp.	Energy
ENB	Enbridge Inc.	Gas Distribution
ERF	Enerplus Resource Fund	Exploration and Production
EXEA	Extendicare Inc.	Health Services
FDG	Fording Canadian Coal Trust	Mining (Other Mines)
FFH	Fairfax Financial Holdings Ltd.	Property and Casualty Insurance
FHR	Fairmont Hotels Resorts Inc.	Hotels
FS	Four Seasons Hotels Inc.	Hotels
GG	Goldcorp Inc.	Gold Mining
GIB	CGI Group Inc.	Computer Services
GIL	Gildan Activewear Inc.	Clothing and Accessories
GLG	Glamis Golds Ltd.	Gold Mining
HBG	Hub International Ltd.	Insurance
IDR	Intrawest Corp..	Hotels
IPS	IPSCO Inc.	Metals and Mining
IQW	Quebecor World	Publishing
ITN	Intertan Inc.	Electronics
ITP	Intertape Polymer Group Inc.	Containers and Packaging
KFS	Kingsway Financial Services Inc.	Insurance
KGC	Kinross Gold Corp.	Gold Mining
LAF	Lafarge North America Inc.	Construction Materials
MDG	Meridian Gold Inc.	Gold Mining
MDZ	MDS Inc.	Medical Equipment
MFC	Manulife Financial Corp.	Insurance
MGA	Magna International Inc.	Auto Parts
MHM	Masonite International Corp.	Building Products
MIM	MI Developments Inc.	Gambling
MWI	Moore Wallace	Computer Services
N	Inco Ltd.	Metals and Mining
NCX	Nova Chemicals Corp.	Commodity Chemicals
NRD	Noranda Inc.	Metals and Mining
NT	Nortel Networks	Foreign Telecom.
NXY	Nexen Inc.	Energy
OPY	Oppenheimer Holdings Inc.	Investment Services
PCZ	Petro-Canadian Com.	Integrated Oil and Gas
PDG	Placer Dome	Precious Metals
PDS	Precision Drilling Corp	Oil Equipment and Services
PGH	Pengrowth Energy	Exploration and Production
PKZ	PetroKazakhstan Inc.	Petroleum
POT	Potash Corp.	Chemical
PWI	Primewest Energy Trust	Energy
RBA	Ritchie Bros Auctioneers	Industrial Equipment
RCN	Radiant Communications	Telecommunications
RG	Rogers Publishing Limited	Publishing
RY	Royal Bank of Canada	Bank
RYG	Royal Group Technologies Ltd.	Building Products
SLF	Sun Life Financial Serv.	Insurance
SU	Suncor Energy	Petroleum
TAC	TransAlta Corp.	Conventional Electricity
TD	Toronto-Dominion	Bank
TEU	CP Ships Ltd.	Maritime
TLM	Talisman Energy	Energy
TOC	Thomson Corp.	Information Services
TRA	Terra Industries	Chemicals
TRP	TransCanada Corp.	Energy
TU	Telus Corp.	Telecommunications
VTS	Veritas DGC Inc.	Energy
ZL	Zarlink Semiconductor Inc.	Semiconductors

Table 4.1: Sample stocks. The table shows the ticker symbols of the 83 Canadian sample stocks together with the full company name and their industry.

the NYSE duration averages range between 46 s and 20 min.

Descriptive	$\#Q^1$	$\#Q^2$	τ^1	τ^2
Min	14.43	16.39	22.79	46.31
$Q25$	73.65	92.67	110.33	91.21
M	159.79	181.61	263.47	205.45
$Q75$	197.30	237.85	288.65	242.50
Max	999.89	499.82	1241.81	1217.31

Table 4.2: Descriptive statistics across 83 sample stocks. For TSX the superscript s equals 1 and for NYSE $s = 2$. The columns labeled $\#Q^s$ give the average of the number of observations per day. Columns labeled τ^s contain the daily average of transaction durations in seconds. The table displays the mean (M), the first ($Q25$) and third quartile ($Q75$), and the minimum (Min) and maximum (Max) of $\#Q^s$ and τ^s across 83 sample stocks. All statistics are calculated over the sample period from January 1st to 31st of March 2004.

Several authors (see e.g. Engle and Russell 1997) point out that price durations exhibit an intraday pattern in the rate of arrival. To account for that we diurnally adjust the data prior to estimation. Assuming the separability of the time function and stochastic function in Equation (4.2), the elimination of the time-of-day effects proceeds in two steps. First, the typical intraday pattern (ϕ_i) is estimated by regressing the transaction durations ($\tau_i = t_i - t_{i-1}$) of the pooled process on polynomial and trigonometric time functions (see Eubank and Speckman 1990 and Appendix B1). Second, dividing the durations by their estimated typical shape gives intraday adjusted durations, i.e. $\tilde{\tau}_i = \frac{\tau_i}{\phi_i}$. Finally, a diurnally adjusted arrival times series of the pooled process is achieved by setting the first arrival time of the day to zero and cumulating adjusted durations for each day.

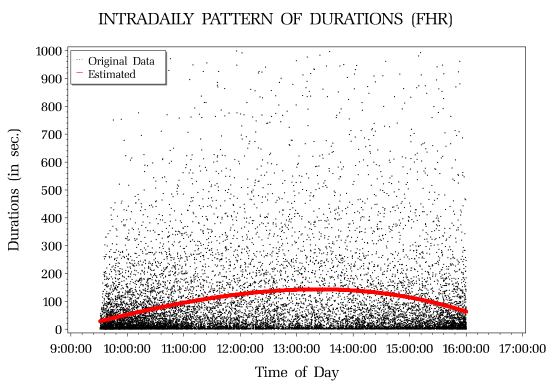


Figure 4.2: Intraday pattern of durations. The figure shows the transaction durations (Original Data: black dots) of the pooled process for one of our sample stocks (FHR). As visible from the figure, transaction durations exhibit a ∩-shape. The typical intraday pattern is captured by the estimated seasonal figure (Estimated: solid line).

Figure 4.2 shows the quote durations on of our sample stocks before removing the intraday pattern. As clearly visible from the figure, transaction durations exhibit the typical ∩-shape. The intraday pattern is captured by the estimated intraday effects.

4.2.3 Intensity Based Unique Information Shares

In order to measure contributions to price discovery, we draw our attention to the cross effects of the conditional intensities associated with quote changes on the home and foreign market. As prices in both markets refer to the same underlying asset, the law of one price holds. This implies that informative price events occurring in one market will subsequently be incorporated into the prices of the other market to offset arbitrage opportunities. Consequently, surprising quote adjustments in one market trigger price events in the other market and increase the other market's conditional intensity. Analogously, the unexpected absence of price changes in one market reduces the other market's intensity. This corresponds to positive cross effects. We propose to use the size of the markets' cross effects to measure the relative importance of a market in the price discovery process.

We start by examining the long run impacts of an innovation shock on the conditional intensities by analyzing the full dynamics of the markets' interdependencies. For this purpose, we compute impulse response functions and derive the cumulated effects of a shock in period i on $\tilde{\boldsymbol{\psi}}_{i+h}$. Iterating the ARMA specification in Equation (4.3) h periods forward yields

$$\tilde{\boldsymbol{\psi}}_{i+h} = \mathbf{a}^s\varepsilon^s_{i+h-1} + \mathbf{B}\mathbf{a}^s\varepsilon^s_{i+h-2} + ... + \mathbf{B}^{h-2}\mathbf{a}^s\varepsilon^s_{i+1} + \mathbf{B}^{h-1}\mathbf{a}^s\varepsilon^s_i + \mathbf{B}^h\tilde{\boldsymbol{\psi}}_i \quad .$$

To isolate the standard deviation shock $\sigma_\varepsilon = \sqrt{\pi^2/6}$ in period i, all subsequent shocks are set to their unconditional mean, $\mathbb{E}\left[\varepsilon^s_i\right] = 0$. Furthermore, the unconditional mean $\mathbb{E}\left[\tilde{\boldsymbol{\psi}}_i\right] = 0$ is used as a starting value for $\tilde{\boldsymbol{\psi}}$. The impulse response functions are then given by,

$$IR^1(h) = \mathbf{B}^{h-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad IR^2(h) = \mathbf{B}^{h-1}\mathbf{a}^2\sigma_\varepsilon \quad , \tag{4.7}$$

where IR^1 denotes the bivariate impulse response function associated with an innovation in the TSX price process and analogously IR^2 gives the bivariate impulse response function associated with a shock in the NYSE price process. Summing up the effects in each period delivers the cumulative impulse response functions,

$$CIR^1(h) = \sum_{j=1}^{h} \mathbf{B}^{j-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad CIR^2(h) = \sum_{j=1}^{h} \mathbf{B}^{j-1}\mathbf{a}^2\sigma_\varepsilon \quad . \tag{4.8}$$

For a stationary process the effects of a shock die out in the long run. Thus, the cumulative impulse response functions in Equation (4.8) for $h \to \infty$ converge to a finite vector given below:

$$\lim_{h\to\infty} CIR^1 = [1-\mathbf{B}]^{-1}\mathbf{a}^1\sigma_\varepsilon \qquad \text{and} \qquad \lim_{h\to\infty} CIR^2 = [1-\mathbf{B}]^{-1}\mathbf{a}^2\sigma_\varepsilon \quad . \tag{4.9}$$

Figure 4.3 shows cumulated impulse response functions as derived in Equation (4.8) and their convergence to the terms in equation Equation (4.9) averaged over 73 sample stocks. The left panel depicts the impact of a standard deviation shock in the TSX intensities and its impacts on TSX and NYSE processes. Analogously, the right panel illustrates the impacts of a NYSE standard deviation shock.

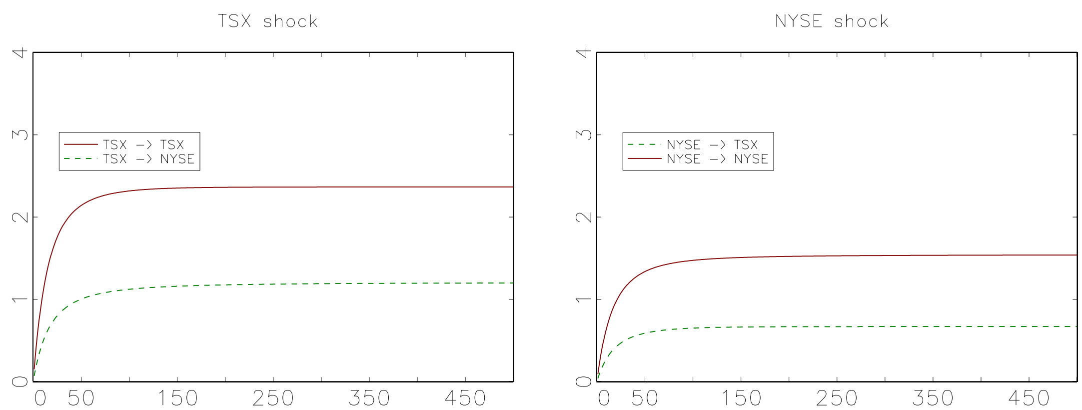


Figure 4.3: Cumulated impulse response function of a standard deviation innovation shock. The figure shows cumulated impulse response functions for the recursive process $\tilde{\psi}$ in Equation (4.3) averaged across our sample stocks (see Table 4.1). The left panel depicts the impact of a standard deviation shock on the TSX and its impact on the TSX process (solid line) and on the NYSE (dashed line). Analogously, the right panel illustrates the impact of a standard deviation shock on the NYSE on the NYSE process (solid line) and on the TSX (dashed line).

Due to the arbitrage relation between prices in both markets, we find positive spillover effects, i.e. positive intensity shocks have significant positive impacts on the other market's conditional intensity. More precisely, the other market experiences an increase in the frequency of quote adjustments which we associate with an information flow into this market. The extreme case of no cumulated cross impacts would indicate that the market's conditional intensity does not react to unexpected events in the other market. As expected, own market's shocks, i.e. TSX $\rightarrow$ TSX and NYSE $\rightarrow$ NYSE, have larger impacts than shocks of the other market, i.e. TSX $\rightarrow$ NYSE and NYSE $\rightarrow$ TSX. The long run effects of a TSX shock on the NYSE intensity tend to be larger than the effects of a NYSE shock on the TSX intensity. These differences in the markets' reaction might be due to differences in the market characteristics on the TSX and NYSE and should be accounted for when comparing effects across markets.

Hence, in order to determine the contributions to price discovery, we focus on the magnitude of cumulated cross effects in either market (CIR_2^1 and CIR_1^2) and account for differences in the markets specific adjustment to intensity shocks. We therefore suggest to *standardize* the cumulative cross effects by the cumulative impact of the own market's shock. Consequently, $\frac{CIR_1^2}{CIR_1^1}$ denotes the cross effect of a NYSE intensity shock on the TSX conditional intensity, standardized by the impact of a TSX shock on TSX's intensity. $\frac{CIR_2^1}{CIR_2^2}$ gives the analog ratio for the NYSE. Considering the equations in (4.9), this ratio straightforward reduces to $\frac{a_2^1}{a_2^2}$ for TSX and $\frac{a_1^2}{a_1^1}$ for NYSE. Finally, in order to simplify the interpretation of our measure, we confine the information shares to lie between zero and one by taking the standardized cross effects of each market relative to the sum of standardized cross effects:

$$IIS^1 = \frac{\frac{a_2^1}{a_2^2}}{\frac{a_1^2}{a_1^1} + \frac{a_2^1}{a_2^2}} \qquad \text{and} \qquad IIS^2 = \frac{\frac{a_1^2}{a_1^1}}{\frac{a_1^2}{a_1^1} + \frac{a_2^1}{a_2^2}} \quad . \tag{4.10}$$

Equation (4.10) gives the unique intensity based information shares, where IIS^1 denotes the TSX and IIS^2 the NYSE contribution to price discovery. Standard errors of the IIS can be computed via the delta method.

4.3 Results and Discussion

As outlined in Section 4.2.1 estimation of the model parameters is done via maximizing the likelihood function in Equation (4.6). We do not allow previous day shocks to affect the next day's intensity, hence, the likelihood function is re-initialized each day and becomes the sum of independent day-likelihoods. Table 4.3 contains descriptive statistics for the estimated ACI coefficients over 73 stocks which have positive $\mathbf{a}^s$ estimates.[12] Stock specific results can be found in Table 4.7 in Appendix B2.

Estimates	M	Std	$Q25$	$Q75$	Min	Max	$M(SE)$	$\#sig$
$\hat{\omega}^1$	-1.285	0.329	-1.499	-1.079	-2.031	-0.169	0.0351	73
$\hat{\omega}^2$	-0.936	0.286	-1.150	-0.737	-1.630	-0.278	0.0317	73
$\hat{a}_1^1$	0.116	0.030	0.095	0.131	0.051	0.213	0.0074	73
$\hat{a}_2^1$	0.052	0.023	0.039	0.062	0.005	0.171	0.0050	72
$\hat{a}_1^2$	0.030	0.017	0.020	0.035	0.004	0.132	0.0044	70
$\hat{a}_2^2$	0.070	0.031	0.046	0.088	0.020	0.178	0.0051	73
$\hat{b}^1$	0.927	0.032	0.905	0.947	0.834	0.981	0.0076	73
$\hat{b}^2$	0.929	0.038	0.905	0.955	0.801	0.993	0.0082	73

Table 4.3: Estimation summary results. The table contains descriptive statistics for the estimated parameters of the ACI model in equations (4.2) and (4.3). The table displays the mean (M), the standard deviation (Std), the first ($Q25$) and third quartile ($Q75$), and the minimum (Min) and maximum (Max) of the estimated parameters over all sample stocks. $M(SE)$ is the average standard error of the estimates and $\#sig$ gives the number of significant estimates on a 5% significance level over the sample stock. The descriptive statistics are computed over 73 stocks that have positive $\mathbf{a}^s$ estimates.

The estimated baseline intensities exhibit small standard errors and are statistically significant at the 1% level of confidence. The cross sectional correlation between the ratio of baseline coefficients, $\frac{\exp(\hat{\omega}^1)}{\exp(\hat{\omega}^1)+\exp(\hat{\omega}^2)}$, and the number of quotes on TSX relative to the total number of quotes on TSX and NYSE is 0.9. Therefore, the baseline function captures very well the differences in the intensity levels in the two markets.

According to Table 4.3 the short run impacts of TSX innovations on the NYSE process are on average about twice as large as the effects of NYSE innovations on the TSX intensity. For 70 stocks the innovation spill over effects from the TSX on NYSE and the effects from NYSE on TSX are significant at the 5% level. From the duration modeling literature (see e.g. Engle and Russell 1998), we expect and find strong persistence of innovation shocks

[12] The stocks which are excluded according to these criteria are BEI, BR, CJR, FFH, and OPY. BGM and ITN were excluded prior to estimation due to erroneous data. VTS, EXEA, and LAF do not achieve proper convergence.

Ticker	TSX $\bar{\tilde{\varepsilon}}^1$	$\sigma_{\tilde{\varepsilon}^1}$	OD^1	AC^1	LB^1	NYSE $\bar{\tilde{\varepsilon}}^2$	$\sigma_{\tilde{\varepsilon}^2}$	OD^2	AC^2	LB^2
ABX	1.00	1.32	26.93	0.05	22.14	1.00	1.17	14.41	0.04	20.40
ABY	1.04	1.46	16.85	0.02	0.87	1.02	1.43	18.97	0.03	2.11
AEM	1.00	1.40	27.84	0.01	0.67	1.02	1.35	26.10	0.03	6.32
AGU	1.01	1.62	41.36	0.02	2.52	1.01	1.44	28.01	0.03	6.31
AL	1.00	1.38	51.42	0.02	6.85	1.00	1.30	43.45	0.04	45.18
BCE	1.00	1.41	21.11	0.03	2.98	1.01	1.37	23.72	0.03	5.22
BCM	0.99	1.50	48.43	0.01	2.09	1.01	1.54	70.42	0.03	25.14
BMO	1.00	1.50	42.89	0.04	18.30	1.00	1.54	63.86	0.03	12.29
BNN	1.01	1.73	61.05	0.01	0.72	1.01	1.75	85.60	0.05	41.15
BNS	0.99	1.52	39.97	0.01	0.38	1.01	1.56	63.06	0.04	26.77
BPO	1.03	1.92	63.40	0.04	6.59	1.00	1.58	48.88	0.05	18.12
BVF	1.01	1.54	60.50	0.06	55.63	1.01	1.45	47.39	0.07	80.92
CCJ	0.99	1.77	92.18	0.05	32.65	1.01	1.73	113.79	0.05	76.00
CGT	1.05	1.50	11.77	-0.02	0.29	1.04	1.39	10.25	0.08	5.91
CLS	1.01	1.46	44.92	0.06	50.52	1.00	1.44	42.75	0.06	44.53
CNI	0.99	1.62	71.95	0.04	23.46	1.00	1.47	56.05	0.02	8.44
CNQ	0.98	1.54	59.77	0.03	10.66	1.00	1.65	106.15	0.02	14.03
COT	1.01	1.77	57.50	-0.01	0.47	1.02	1.60	50.79	0.03	7.22
CP	1.00	1.72	57.51	-0.01	1.39	1.01	1.42	36.68	0.05	25.18
CWG	1.01	2.27	83.99	0.00	0.06	1.03	1.90	47.16	0.10	24.96
DTC	1.02	1.59	28.59	0.00	0.01	1.03	1.42	22.92	0.03	2.68
ECA	0.99	1.46	49.55	0.03	16.26	1.01	1.30	32.17	0.06	59.40
ENB	1.06	2.07	79.71	0.01	0.17	1.01	1.75	84.97	0.00	0.24
ERF	1.01	1.70	62.99	0.06	30.93	1.02	1.55	53.21	0.06	36.68
FDG	1.01	1.92	112.31	0.07	62.91	1.01	2.01	151.04	0.07	103.29
FHR	1.02	1.71	56.17	0.02	3.43	1.00	1.69	57.72	0.01	0.99
FS	1.01	1.66	80.52	0.06	57.52	1.00	1.74	92.32	0.03	12.16
GG	1.02	1.37	27.16	0.05	16.20	1.01	1.21	15.37	0.08	61.91
GIB	1.04	1.53	15.16	-0.03	1.04	1.02	1.56	20.09	0.04	2.37
GIL	1.02	2.14	92.95	0.04	8.83	1.01	2.13	130.89	0.05	25.49
GLG	1.01	1.52	43.99	0.04	12.29	1.01	1.37	31.83	0.06	39.80
HBG	1.03	1.94	81.02	0.06	21.74	1.01	1.93	61.17	0.08	28.71
IDR	1.01	1.78	56.06	0.03	4.83	1.01	1.66	50.42	0.02	2.39
IPS	1.03	1.97	59.37	0.03	2.30	1.02	1.90	67.38	0.07	27.57
IQW	1.02	1.82	55.23	0.01	0.48	1.00	1.55	42.79	0.07	33.52
ITP	1.03	1.91	52.96	0.04	4.50	1.02	1.98	76.35	0.03	6.53
KFS	1.07	1.97	44.44	0.03	2.33	1.02	1.69	42.80	0.04	5.88
KGC	1.04	1.37	16.68	0.04	3.76	1.02	1.19	8.62	0.07	16.18
MDG	1.01	1.49	34.65	0.03	4.34	1.01	1.33	25.06	0.03	9.98
MDZ	1.02	1.65	33.02	0.02	0.88	1.03	1.58	34.88	0.05	10.18
MFC	1.00	1.47	36.84	0.04	10.20	1.00	1.52	48.22	0.02	6.42
MGA	0.99	1.66	89.45	0.04	28.85	1.00	1.48	64.74	0.02	10.84
MHM	1.00	2.10	83.03	0.00	0.00	1.01	1.97	108.71	0.04	17.56
MIM	1.02	1.96	82.26	0.03	6.64	1.00	1.95	83.27	0.02	4.04
MWI	1.03	1.93	70.35	0.08	32.93	1.01	1.73	43.38	0.03	4.53
N	1.00	1.42	57.20	0.04	38.40	1.00	1.40	53.39	0.05	51.21
NCX	1.02	1.65	46.38	0.03	3.69	1.00	1.49	39.98	0.02	3.52
NRD	0.99	1.48	34.65	0.02	1.76	1.00	1.48	41.62	0.03	6.41
NT	0.97	1.09	4.86	0.05	10.53	0.97	1.13	6.07	0.01	0.12
NXY	0.99	1.67	67.34	0.02	4.59	1.00	1.73	100.08	0.03	18.76
PCZ	0.99	1.60	62.38	0.04	19.34	1.02	1.58	75.86	0.04	26.52
PDG	1.01	1.30	25.25	0.04	16.42	1.00	1.21	17.58	0.03	8.90
PDS	0.99	1.86	111.20	0.02	4.82	1.00	1.55	64.08	0.04	22.69
PGH	1.06	1.92	64.56	0.07	20.22	1.05	1.69	51.12	0.07	29.07
PKZ	0.99	2.12	167.71	0.04	24.13	1.01	1.73	92.52	0.05	45.61
POT	0.99	1.86	146.13	0.05	68.52	0.99	1.67	95.75	0.04	31.37
PWI	1.03	1.59	38.14	0.03	5.82	1.03	1.49	32.48	0.08	37.45
RBA	1.01	2.21	100.67	-0.03	4.50	1.01	2.07	97.92	0.08	41.63
RCN	1.00	2.71	188.03	0.00	0.04	1.02	2.42	199.52	0.02	8.11
RG	1.00	1.84	59.01	-0.01	0.16	1.02	1.55	42.62	0.02	4.52
RY	0.99	1.46	38.76	0.02	4.46	1.00	1.43	47.91	0.04	22.19
RYG	0.99	1.95	66.46	0.03	5.55	1.02	1.76	59.05	0.07	27.84
SLF	0.99	1.53	44.03	0.03	6.99	1.01	1.39	36.96	0.03	9.01
SU	1.00	1.51	46.44	0.02	3.10	1.02	1.41	38.96	0.05	26.12
TAC	1.10	1.89	36.60	0.00	0.02	1.02	1.68	38.47	-0.01	0.43
TD	1.00	1.41	28.52	0.02	4.05	1.01	1.43	38.37	0.02	6.02
TEU	1.03	1.73	45.89	0.00	0.00	1.01	1.62	47.52	0.04	12.67
TLM	0.99	1.64	70.05	0.00	0.26	1.01	1.53	74.69	0.03	21.39
TOC	0.99	1.66	49.05	0.03	6.49	1.02	1.54	47.93	0.04	14.63
TRA	1.02	3.07	216.86	0.02	1.86	1.06	1.94	56.44	0.07	17.78
TRP	1.01	1.57	31.08	0.04	4.92	1.02	1.35	24.52	0.05	18.36
TU	1.01	1.77	58.69	0.02	2.66	1.02	1.82	92.45	0.01	2.61
ZL	1.04	1.66	22.98	0.04	2.73	1.02	1.46	15.41	0.03	1.08

Table 4.4: Residual diagnostics for the ACI model. The table presents residual diagnostics for the estimated residuals corresponding to TSX ($s = 1$) and NYSE ($s = 2$). AC^s denotes the value of the first order autocorrelation and columns labeled with LB^s contain the corresponding Ljung-Box statistic. $\bar{\tilde{\varepsilon}}^s$ and $\sigma_{\tilde{\varepsilon}^s}$ contain the mean and the standard deviation of the estimated residuals and OD^s gives the test statistic of the overdispersion test of Engle and Russell (1998). This statistics has a limiting normal distribution under the null with a 5% critical value of 1.645. The statistics are computed for 73 stocks that have positive $\mathbf{a}^s$ estimates. For full company names see Table 4.1.

which is reflected by large autoregressive coefficients. Across our sample stocks $\hat{b}^1$ is on average 0.927 and $\hat{b}^2$ is 0.929. Turning to the test statistics concerning the model specification, the results for the Ljung-Box test in Table 4.4 for the first autocorrelation of the estimated residuals are mixed. For some stocks the null hypothesis of no autocorrelation cannot be rejected. Furthermore, the table shows that the means of the estimated residuals are on average close to one for both markets. Overall, the results indicate that some excess dispersion is still present, nevertheless we obtain consistent estimators based on the quasi maximum likelihood method.

	TSX					NYSE				
DESCRIPTIVE	IIS^1	HIS^1_{low}	HIS^1_{up}	HIS^1_{mid}	Adj^1	IIS^2	HIS^2_{low}	HIS^2_{up}	HIS^2_{mid}	Adj^2
M	72.6	39.4	86.2	62.8	67.6	27.4	13.8	60.6	37.2	32.4
$M(SE)$	6.2	2.9	1.9	2.3	2.6	6.2	1.9	2.9	2.3	2.6
Std	17.6	24.1	19.7	20.3	26.0	17.6	19.7	24.1	20.3	26.0
$Q25$	63.2	15.9	81.3	48.2	50.0	15.1	1.0	39.7	20.7	9.2
$Q75$	84.8	59.5	98.7	78.3	90.6	36.2	18.6	83.9	49.8	49.7
Min	17.1	0.0	1.1	0.6	0.2	1.3	0.0	17.6	8.8	0.0
Max	98.7	82.4	100.0	91.2	100.0	82.9	98.9	100.0	99.4	99.8

Table 4.5: Intensity based information shares – descriptives. The table presents descriptives computed over the information shares using an intensity based and the standard Hasbrouck approach in percent. The descriptives are the mean (M), the standard deviation (Std), the mean of the information share standard error ($M(SE)$), the 25% quantile ($Q25$), the 75% quantile ($Q75$), the minimum (Min) and maximum (Max) over the cross sectional information shares. The midpoint and the lower and upper bounds of Hasbrouck are denoted by HIS^s_{mid}, HIS^s_{low} and HIS^s_{up}, respectively. Columns labeled IIS^s give the unique intensity based information share. For TSX $s = 1$ and NYSE $s = 2$. The descriptives are computed over 73 stocks that have positive $\mathbf{a}^s$ estimates.

Table 4.5 displays summary results for the unique intensity based information share according to Equation (4.10). Stock specific results can be found in Table 4.8 in Appendix B2. As the Hasbrouck methodology is a well established measure, we also report Hasbrouck information shares. Additionally, we present results for the adjustment coefficient ratios.

The average home market intensity based information share (IIS) amounts to 73%, which implies a clear leadership of the TSX in the price discovery process. The adjustment coefficient ratios (Adj) also support the TSX as the leading market and match results of Eun and Sabherwal (2003), who examine Canadian stock data from a period in 1999. They conclude that about two thirds of adjustment is done by NYSE prices. The Hasbrouck information share midpoints (HIS) deliver similar results with an average TSX midpoint of 63%. Stock specific findings are illustrated in Figure 4.4, which shows the intensity based information shares along with their 0.95 confidence bounds and the corresponding Hasbrouck information shares midpoints with the upper 0.95 confidence interval for the upper HIS share and the lower 0.95 confidence interval for the lower HIS share.

On average Hasbrouck information shares differ by 47 percentage points. These results correspond to previous findings by Korczak and Phylaktis (2010), who examine Candian data using the Hasbrouck methodology. Note that the Hasbrouck and our intensity approach model different aspects of the price processes. We therefore do not a priori expect them to deliver exactly the same results. However, as the Hasbrouck methodology is a well established measure, we take the qualitatively similar results as evidence that information can be gained from modeling bivariate intensities. Regarding the extremely wide bounds of the Hasbrouck information share estimates, the advantage that can be derived from our unique intensity based information share is straightforward.

Since the HIS bounds differ by 47 percentage points on average the leading market can only be determined for 28 out of 73 stocks. For the remaining stocks the HIS bounds do not allow for any conclusion concerning the importance of the trading venues for the formation of the fundamental price. Using the IIS, we are able to answer the question "which market leads?" for 54 out of 73 stocks. We find 50 stocks for which the TSX significantly dominates price discovery and in only 4 cases we detect a higher NYSE information share. These findings emphasize the advantage of the intensity based information share which is a quite accurate measure to determine a market's contribution to the price discovery process.

4.4 Cross Sectional Analysis

As visible from the descriptive statistics in Table 4.5 and Figure 4.4, the variation of the information shares among the sample stocks is considerably high. An interesting task is therefore to analyze the factors that influence a market's intensity based information share by running cross sectional regressions.

For this purpose, we use the TSX information share as dependent variable and perform a logistic transformation, $\ln\left(\frac{IIS^1}{1-IIS^1}\right)$, to ensure that predicted values lie within zero and one. We follow previous studies by Korczak and Phylaktis (2010), Eun and Sabherwal (2003) and Grammig et al. (2005) and select a set of liquidity related variables and firm

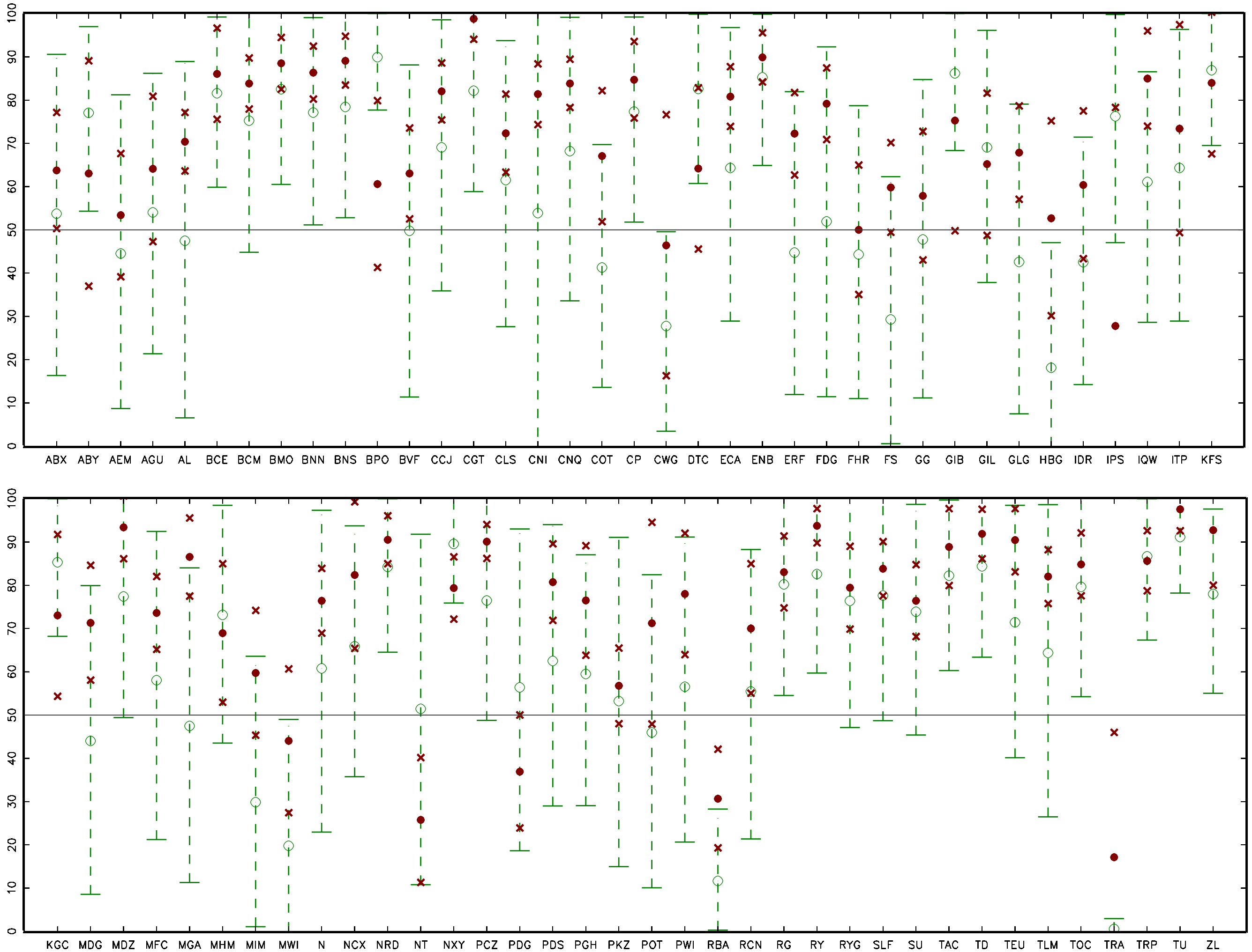


Figure 4.4: Stock specific of intensity based and Hasbrouck information shares on TSX. The figure shows the estimated intensity based information shares (solid circle) and their 0.95 confidence bounds (cross) as well as the Hasbrouck information share midpoints (circle) along with the upper 0.95 confidence interval for the upper HIS share and the lower 0.95 confidence interval for the lower HIS share (vertical dashed lines).

specific factors.

As liquidity related variables we consider the relative spread ($Spread$), measured by the quoted spread on the TSX relative to the spread on the NYSE, and traded volume ($TsxVol$) given by the number of shares traded on TSX relative to the total number of shares traded on TSX and NYSE.

Hasbrouck (1995) classifies trades into volume categories and finds that a relative higher number of medium size trades are positively correlated with the relative information share. We compute medium trades $MedTrad$ as the number of the medium shares to total shares traded on TSX relative to the ratio of the medium shares to total shares traded on NYSE.[13]

We also control for firm and sector specific properties. We use sector dummies as explanatory variables. Furthermore, we include the number of years a firm has been listed on the NYSE through 2004 ($YearListed$) and $FSale$

Additionally, the regression contains the firm size measured by the log of a firm's market capitalization, $LMktCap$.[14]

As $Spread$, $MedTrad$ and $TsxVol$ are highly correlated, we include them into the regressions one by one. Thus, regressions (1)-(3) in Table 4.6 control for the firm size and each contains a proxy for liquidity. Concerning the $Spread$, the results reveal a negative and statistically significant coefficient at the 1% level, which implies that a lower relative spread on the home market is linked to a higher TSX information share. The coefficient of the relative proportion of medium trades ($MedTrad$) and the coefficient for $TsxVol$ are positive and significant on the 1% level. As regressions (4) and (5) show, the firm specific variables are not statistically significant. We find evidence neither for a relation between a firm's size and a market's information share, nor do the duration of listing on the NYSE, the proportion of a company's foreign sales or industry characteristics have any significant effects. Using the intensity based information share we confirm a positive link between the relative liquidity on a market and its contribution to price discovery (see Eun and Sabherwal 2003, Korczak and Phylaktis 2010). As acknowledged by Grammig et al. (2005), these findings indicate a link between those two variables rather than a causal relation. Nevertheless we argue that liquidity related variables rather than stock specific factors determine a market's dominance and thus suggest it is the market's design

[13] A medium trade is based on a trade category of 2,501-10,000 shares.

[14] Market capitalization as reported on 31 December 2003 in the TSX Factbook (2003).

Explanatory Variable	(1)	(2)	(3)	(4)	(5)
Constant	1.120	0.540	-1.175	-1.515	-1.323
	(1.767	(2.300)	(1.663)	(1.629)	(1.513)
LMktCap	0.574	-0.070	0.283	0.416	0.170
	(0.786)	(1.133)	(0.855)	(0.855)	(0.808)
Spread	-0.925 ***				
	(0.162)				
MedTrad		0.507 ***			
		(0.139)			
TsxVol			2.551 ***	2.522 ***	2.643 ***
			(0.412)	(0.451)	(0.459)
YearListed				-0.004	-0.002
				(0.003)	(0.004)
FSale				0.002	0.004
				(0.003)	(0.003)
Mining					0.126
					(0.312)
Manufacturing					-0.075
					(0.361)
Transport					0.370
					(0.302)
Finance					0.353
					(0.313)
R^2_{adj}	0.423	0.150	0.432	0.423	0.413

Table 4.6: Cross sectional regression results. The table reports cross sectional OLS estimates. The dependent variable is the logistic transformation of the TSX information share, $\ln\left(\frac{IIS^1}{1-IIS^1}\right)$. The logistic transformation ensures predicted regression values within 0 and 1. The regression includes 73 stocks that have positive $\mathbf{a}^s$ estimates. $LMktCap$ is the log market capitalization as reported on 31 December 2003 in the TSX Factbook (2003). $Spread$ denotes the ratio between the percentage spread on the TSX and percentage spread on the NYSE. $MedTrad$ gives the ratio of proportions of shares traded in Canada and on the NYSE in medium-sized lots of 2,501to 10,000 shares. $TsxVol$ denotes the ratio between trade volume on the TSX and NYSE denoted in CAD. $YearListed$ denotes the number of years a company has been listed on the NYSE and $FSale$ is the ratio of foreign sales to total sales. The remaining variables are sector dummies. We classify five industry groups. Mining, Manufacturing, Finance, and Transport/Utility are dummies corresponding to four of these groups. The fifth industry group serves as benchmark sector and includes service and retail firms. R^2_{adj} give the regression's adjusted R squared. The numbers in parentheses below the estimates are standard errors computed from the heteroskedasticity-consistent covariance matrix following White (1980). ***, **, and * indicate the corresponding statistical two-tailed significance at $\alpha =$ 1%, 5%, and 10%, respectively.

that matters for price discovery. These results imply that providing an efficient trading environment and attracting investors in order to increase liquidity, is of special interest for small national stock exchanges that seek to maintain their dominance in the price discovery process of cross-listed stocks.

4.5 Concluding Remarks

Investors' decision to invest and companies' intention to list their stocks on a stock exchange depends on the ability of an exchange to provide a prospering trading environment. As a result of an increasing globalization and improved technology, small national exchanges fear to lose their attractiveness for investors and companies. In particular, within the context of international cross-listed stocks, it is of paramount concern for a national stock exchange to remain the dominant market with regard to price discovery.

We propose a new approach to measure the contribution of trading venues to the price discovery process of internationally cross-listed stocks. We use a bivariate intensity approach as an alternative to the commonly applied vector error correction model in order to take the irregularity of the data into account. Based on the autoregressive conditional intensity model of Russell (1999), contributions to price discovery are measured by modeling the interdependencies of the trading processes in both markets. In contrast to the Hasbrouck (1995) approach, the new intensity based information shares deliver unique results rather than upper and lower bounds. Furthermore, we show that the suggested approach provides a more accurate analysis to determine the leading market in the price discovery process.

In our empirical application we examine Canadian stocks which are listed on the TSX as well as on the NYSE. We find that despite the concern of the TSX to lose its share in price discovery to the NYSE, trading on the TSX still plays the most important role. This confirms previous results by Korczak and Phylaktis (2010), Eun and Sabherwal (2003), and Grammig et al. (2005), who also analyze Canadian stocks. We show that the leadership of the TSX is even more pronounced than indicated by previous studies. The average TSX information share amounts to 73%.

Concerning potential determinants of a market's information share, we conduct cross sectional regressions and find as Grammig et al. (2005), Eun and Sabherwal (2003), and

Korczak and Phylaktis (2010) that liquidity contributes positively to a market's price determination. These results imply that providing an efficient trading environment and attracting investors in order to increase liquidity, is of special interest for small national stock exchanges that seek to maintain their dominance in the price discovery process.

Appendix B1: Adjustment of Intraday Effects

In order to diurnally adjust the quote data, we follow Eubank and Speckman (1990) and regress the quote durations of the pooled process on polynomial and trigonometric time functions. The regression equation reads for some integers $d \geq 0$ and $\delta \geq 0$ as follows,

$$\tau_i = \beta_0 + \sum_{j=1}^{d} \beta_j^p t_i^j + \sum_{j=1}^{\delta} [\beta_j^c \cos(jt_i) + \beta_j^s \sin(jt_i)] + \epsilon_i \quad , \tag{4.11}$$

where the duration is $\tau_i = t_i - t_{i-1}$. The number of polynomial and trigonometric terms are selected by a generalized cross-validation measure defined as,

$$GCV = \frac{nRSS}{(n - 2\delta - d - 1)^2} \quad , \tag{4.12}$$

where RSS denotes the residual sum of squares and n the number of observations. In order to avoid overfitting we restrict in the selection d and δ to be smaller than three. To compute a typical time-of-day function we select the specification of Equation (4.11) that minimizes the GCV in Equation (4.12).

Appendix B2: Additional Tables

Ticker	$\hat{\omega}^1$	$\hat{\omega}^2$	$\hat{a}_1^1$	$\hat{a}_2^1$	$\hat{a}_1^2$	$\hat{a}_2^2$	$\hat{b}^1$	$\hat{b}^2$
ABX	-0.860	-0.678	0.082	0.050	0.029	0.078	0.932	0.909
	(0.027)	(0.020)	(0.015)	(0.007)	(0.008)	(0.011)	(0.024)	(0.020)
ABY	-0.952	-0.560	0.126	0.055	0.048	0.084	0.853	0.880
	(0.049)	(0.056)	(0.031)	(0.021)	(0.015)	(0.021)	(0.047)	(0.050)
AEM	-0.925	-0.702	0.110	0.058	0.056	0.100	0.834	0.852
	(0.030)	(0.029)	(0.016)	(0.010)	(0.010)	(0.013)	(0.038)	(0.035)
AGU	-1.074	-0.909	0.098	0.050	0.036	0.076	0.916	0.901
	(0.053)	(0.055)	(0.013)	(0.009)	(0.008)	(0.014)	(0.017)	(0.029)
AL	-0.940	-0.738	0.106	0.056	0.032	0.077	0.935	0.930
	(0.019)	(0.022)	(0.007)	(0.006)	(0.004)	(0.008)	(0.008)	(0.014)
BCE	-1.079	-0.655	0.177	0.051	0.025	0.060	0.875	0.930
	(0.035)	(0.039)	(0.021)	(0.009)	(0.008)	(0.010)	(0.025)	(0.018)
BCM	-1.268	-0.737	0.117	0.055	0.029	0.043	0.915	0.938
	(0.031)	(0.045)	(0.014)	(0.010)	(0.005)	(0.006)	(0.016)	(0.016)
BEI	-0.936	0.011	0.243	-0.024	0.021	0.044	0.702	0.976
	(0.145)	(0.128)	(0.202)	(0.010)	(0.038)	(0.016)	(0.359)	(0.016)
BMO	-1.378	-0.840	0.105	0.040	0.015	0.036	0.948	0.960
	(0.041)	(0.049)	(0.014)	(0.007)	(0.004)	(0.006)	(0.011)	(0.009)
BNN	-1.604	-0.998	0.114	0.052	0.020	0.045	0.943	0.948
	(0.049)	(0.052)	(0.010)	(0.007)	(0.004)	(0.007)	(0.008)	(0.010)
BNS	-1.435	-0.645	0.133	0.067	0.024	0.045	0.919	0.910
	(0.039)	(0.052)	(0.017)	(0.016)	(0.005)	(0.007)	(0.017)	(0.029)
BPO	-1.500	-0.752	0.113	0.034	0.025	0.102	0.914	0.889
	(0.056)	(0.048)	(0.020)	(0.009)	(0.008)	(0.011)	(0.023)	(0.021)
BR	-2.304	-1.678	0.079	0.020	-0.023	0.095	0.987	0.960
	(0.067)	(0.033)	(0.004)	(0.004)	(0.003)	(0.012)	(0.001)	(0.010)
BVF	-1.250	-1.143	0.081	0.050	0.035	0.069	0.976	0.970
	(0.043)	(0.051)	(0.006)	(0.005)	(0.005)	(0.009)	(0.003)	(0.005)
CCJ	-1.747	-0.925	0.107	0.063	0.016	0.093	0.968	0.924
	(0.047)	(0.057)	(0.010)	(0.006)	(0.004)	(0.009)	(0.004)	(0.013)
CGT	-0.658	-0.278	0.213	0.171	0.011	0.042	0.860	0.833
	(0.087)	(0.080)	(0.042)	(0.047)	(0.017)	(0.021)	(0.048)	(0.091)
CJR	-2.135	-1.535	0.101	-0.034	-0.006	0.076	0.897	0.972
	(0.076)	(0.113)	(0.033)	(0.006)	(0.014)	(0.009)	(0.049)	(0.004)
CLS	-1.031	-1.057	0.091	0.049	0.032	0.054	0.961	0.966
	(0.040)	(0.048)	(0.009)	(0.007)	(0.005)	(0.009)	(0.006)	(0.008)
CNI	-1.200	-0.833	0.121	0.065	0.020	0.089	0.937	0.887
	(0.033)	(0.035)	(0.009)	(0.007)	(0.004)	(0.010)	(0.007)	(0.023)
CNQ	-1.552	-0.779	0.137	0.062	0.023	0.071	0.941	0.926
	(0.042)	(0.061)	(0.014)	(0.010)	(0.003)	(0.010)	(0.010)	(0.020)
COT	-1.462	-1.039	0.070	0.028	0.018	0.052	0.960	0.962
	(0.050)	(0.047)	(0.008)	(0.006)	(0.005)	(0.008)	(0.006)	(0.007)
CP	-1.223	-0.774	0.117	0.057	0.025	0.049	0.918	0.925
	(0.040)	(0.080)	(0.014)	(0.016)	(0.006)	(0.016)	(0.016)	(0.042)
CWG	-1.291	-1.266	0.054	0.010	0.025	0.026	0.937	0.973
	(0.075)	(0.082)	(0.017)	(0.005)	(0.011)	(0.011)	(0.024)	(0.014)
DTC	-1.180	-0.715	0.095	0.055	0.034	0.084	0.918	0.877
	(0.054)	(0.048)	(0.019)	(0.015)	(0.009)	(0.012)	(0.025)	(0.032)
ECA	-1.076	-0.871	0.105	0.051	0.032	0.040	0.951	0.962
	(0.033)	(0.040)	(0.009)	(0.007)	(0.005)	(0.006)	(0.007)	(0.009)
ENB	-1.825	-0.855	0.137	0.059	0.016	0.057	0.940	0.932
	(0.051)	(0.051)	(0.016)	(0.009)	(0.005)	(0.006)	(0.010)	(0.010)
ERF	-1.584	-1.282	0.087	0.043	0.031	0.045	0.969	0.977
	(0.050)	(0.049)	(0.008)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)
FDG	-1.819	-1.630	0.093	0.049	0.022	0.056	0.967	0.970
	(0.058)	(0.070)	(0.009)	(0.005)	(0.004)	(0.008)	(0.004)	(0.005)
FFH	-1.317	-1.533	0.145	0.015	-0.014	0.147	0.948	0.918
	(0.045)	(0.030)	(0.006)	(0.004)	(0.004)	(0.008)	(0.004)	(0.007)
FHR	-1.155	-1.118	0.085	0.045	0.042	0.091	0.917	0.929
	(0.053)	(0.053)	(0.012)	(0.007)	(0.007)	(0.010)	(0.016)	(0.012)
FS	-1.099	-1.088	0.090	0.071	0.027	0.159	0.931	0.845
	(0.040)	(0.047)	(0.008)	(0.006)	(0.006)	(0.018)	(0.010)	(0.029)
GG	-0.909	-0.771	0.066	0.034	0.035	0.046	0.950	0.970
	(0.030)	(0.035)	(0.009)	(0.005)	(0.006)	(0.008)	(0.010)	(0.007)
GIB	-0.954	-0.550	0.122	0.052	0.038	0.054	0.873	0.927
	(0.069)	(0.085)	(0.024)	(0.022)	(0.015)	(0.021)	(0.029)	(0.044)
GIL	-1.825	-1.286	0.141	0.030	0.023	0.098	0.925	0.946
	(0.060)	(0.066)	(0.014)	(0.006)	(0.007)	(0.007)	(0.009)	(0.006)
GLG	-1.111	-0.903	0.086	0.060	0.034	0.072	0.938	0.932
	(0.041)	(0.049)	(0.009)	(0.008)	(0.006)	(0.012)	(0.011)	(0.018)
HBG	-1.107	-1.520	0.080	0.041	0.026	0.114	0.925	0.900
	(0.059)	(0.092)	(0.007)	(0.010)	(0.009)	(0.033)	(0.009)	(0.041)
IDR	-1.170	-0.968	0.138	0.039	0.034	0.104	0.850	0.905
	(0.046)	(0.050)	(0.014)	(0.009)	(0.009)	(0.011)	(0.023)	(0.018)
IPS	-1.389	-1.219	0.112	0.005	0.025	0.059	0.901	0.970
	(0.081)	(0.074)	(0.029)	(0.007)	(0.010)	(0.009)	(0.040)	(0.006)
IQW	-1.521	-0.911	0.115	0.045	0.024	0.039	0.932	0.952
	(0.053)	(0.060)	(0.015)	(0.009)	(0.009)	(0.011)	(0.013)	(0.015)

continued

Ticker	$\hat{\omega}^1$	$\hat{\omega}^2$	$\hat{a}_1^1$	$\hat{a}_2^1$	$\hat{a}_1^2$	$\hat{a}_2^2$	$\hat{b}^1$	$\hat{b}^2$
ITP	-1.386	-1.039	0.111	0.029	0.016	0.073	0.902	0.932
	(0.061)	(0.066)	(0.012)	(0.008)	(0.008)	(0.009)	(0.014)	(0.010)
KFS	-1.486	-0.745	0.123	0.044	0.015	0.070	0.944	0.938
	(0.064)	(0.058)	(0.016)	(0.010)	(0.007)	(0.009)	(0.009)	(0.010)
KGC	-0.786	-0.600	0.138	0.079	0.044	0.093	0.878	0.868
	(0.036)	(0.033)	(0.022)	(0.015)	(0.013)	(0.024)	(0.033)	(0.051)
MDG	-1.063	-0.644	0.114	0.078	0.033	0.107	0.910	0.801
	(0.036)	(0.028)	(0.014)	(0.013)	(0.008)	(0.017)	(0.018)	(0.057)
MDZ	-1.224	-0.766	0.099	0.076	0.013	0.041	0.914	0.877
	(0.055)	(0.057)	(0.016)	(0.017)	(0.006)	(0.009)	(0.020)	(0.036)
MFC	-1.080	-0.839	0.134	0.086	0.037	0.113	0.902	0.882
	(0.033)	(0.038)	(0.012)	(0.008)	(0.006)	(0.012)	(0.014)	(0.019)
MGA	-1.216	-0.908	0.116	0.051	0.010	0.097	0.935	0.890
	(0.031)	(0.041)	(0.007)	(0.005)	(0.003)	(0.013)	(0.006)	(0.025)
MHM	-2.031	-1.270	0.134	0.023	0.017	0.081	0.935	0.952
	(0.070)	(0.069)	(0.018)	(0.006)	(0.006)	(0.007)	(0.012)	(0.005)
MIM	-1.263	-1.434	0.114	0.050	0.033	0.114	0.903	0.926
	(0.073)	(0.068)	(0.013)	(0.008)	(0.007)	(0.011)	(0.017)	(0.010)
MWI	-1.369	-1.386	0.089	0.055	0.052	0.120	0.961	0.931
	(0.082)	(0.085)	(0.013)	(0.011)	(0.008)	(0.019)	(0.007)	(0.016)
N	-0.899	-0.934	0.097	0.048	0.023	0.061	0.940	0.951
	(0.025)	(0.033)	(0.008)	(0.005)	(0.004)	(0.007)	(0.009)	(0.009)
NCX	-1.137	-0.818	0.098	0.036	0.012	0.065	0.902	0.911
	(0.041)	(0.043)	(0.014)	(0.007)	(0.006)	(0.008)	(0.020)	(0.017)
NRD	-1.213	-0.849	0.141	0.062	0.028	0.033	0.926	0.954
	(0.042)	(0.063)	(0.015)	(0.012)	(0.006)	(0.010)	(0.013)	(0.018)
NT	-0.169	-0.486	0.154	0.053	0.132	0.178	0.970	0.963
	(0.071)	(0.061)	(0.017)	(0.013)	(0.016)	(0.018)	(0.005)	(0.007)
NXY	-1.537	-1.032	0.113	0.046	0.022	0.063	0.945	0.947
	(0.041)	(0.068)	(0.009)	(0.006)	(0.003)	(0.008)	(0.006)	(0.011)
OPY	-1.196	-1.552	0.126	0.023	-0.053	0.143	0.934	0.834
	(0.097)	(0.076)	(0.014)	(0.021)	(0.010)	(0.034)	(0.009)	(0.061)
PCZ	-1.692	-1.152	0.129	0.060	0.023	0.036	0.968	0.975
	(0.047)	(0.042)	(0.009)	(0.006)	(0.004)	(0.005)	(0.003)	(0.003)
PDG	-0.834	-0.710	0.072	0.038	0.054	0.086	0.947	0.943
	(0.027)	(0.030)	(0.008)	(0.006)	(0.007)	(0.015)	(0.010)	(0.017)
PDS	-1.115	-1.041	0.145	0.048	0.024	0.070	0.896	0.931
	(0.036)	(0.062)	(0.012)	(0.007)	(0.005)	(0.011)	(0.013)	(0.020)
PGH	-1.664	-1.270	0.094	0.054	0.035	0.045	0.981	0.984
	(0.062)	(0.054)	(0.009)	(0.007)	(0.007)	(0.010)	(0.002)	(0.003)
PKZ	-1.435	-1.341	0.147	0.038	0.053	0.079	0.918	0.952
	(0.054)	(0.065)	(0.010)	(0.005)	(0.006)	(0.009)	(0.009)	(0.008)
POT	-1.023	-1.156	0.131	0.022	0.009	0.132	0.897	0.887
	(0.037)	(0.036)	(0.010)	(0.004)	(0.004)	(0.013)	(0.015)	(0.020)
PWI	-1.499	-1.366	0.084	0.056	0.034	0.040	0.979	0.984
	(0.056)	(0.053)	(0.011)	(0.008)	(0.009)	(0.011)	(0.004)	(0.004)
RBA	-2.008	-1.459	0.103	0.027	0.051	0.124	0.957	0.955
	(0.083)	(0.090)	(0.008)	(0.005)	(0.008)	(0.013)	(0.004)	(0.007)
RCN	-1.912	-1.467	0.188	0.031	0.025	0.100	0.899	0.939
	(0.068)	(0.076)	(0.026)	(0.006)	(0.008)	(0.007)	(0.014)	(0.006)
RG	-1.368	-0.840	0.135	0.070	0.037	0.052	0.895	0.907
	(0.049)	(0.053)	(0.014)	(0.011)	(0.008)	(0.007)	(0.017)	(0.018)
RY	-1.339	-0.706	0.118	0.054	0.013	0.033	0.944	0.951
	(0.033)	(0.039)	(0.014)	(0.008)	(0.004)	(0.005)	(0.010)	(0.011)
RYG	-1.838	-1.401	0.131	0.066	0.043	0.052	0.952	0.964
	(0.075)	(0.065)	(0.017)	(0.010)	(0.009)	(0.010)	(0.008)	(0.007)
SLF	-1.176	-0.751	0.127	0.089	0.033	0.066	0.914	0.892
	(0.033)	(0.036)	(0.011)	(0.012)	(0.005)	(0.009)	(0.012)	(0.024)
SU	-1.074	-0.809	0.112	0.064	0.037	0.061	0.939	0.931
	(0.035)	(0.041)	(0.010)	(0.008)	(0.006)	(0.010)	(0.010)	(0.017)
TAC	-1.308	-0.484	0.122	0.048	0.031	0.023	0.916	0.953
	(0.061)	(0.055)	(0.026)	(0.014)	(0.009)	(0.006)	(0.026)	(0.015)
TD	-1.180	-0.698	0.120	0.066	0.015	0.046	0.931	0.925
	(0.033)	(0.050)	(0.016)	(0.015)	(0.005)	(0.012)	(0.016)	(0.031)
TEU	-1.357	-0.785	0.130	0.083	0.018	0.063	0.923	0.894
	(0.049)	(0.064)	(0.013)	(0.015)	(0.007)	(0.012)	(0.011)	(0.031)
TLM	-1.358	-0.732	0.128	0.064	0.025	0.070	0.917	0.915
	(0.031)	(0.051)	(0.011)	(0.010)	(0.004)	(0.012)	(0.011)	(0.025)
TOC	-1.364	-0.824	0.138	0.071	0.036	0.049	0.900	0.921
	(0.066)	(0.060)	(0.031)	(0.014)	(0.007)	(0.009)	(0.035)	(0.024)
TRA	-1.502	-1.251	0.051	0.006	0.020	0.079	0.973	0.964
	(0.118)	(0.082)	(0.018)	(0.005)	(0.009)	(0.015)	(0.009)	(0.009)
TRP	-1.405	-0.611	0.204	0.078	0.054	0.050	0.894	0.938
	(0.051)	(0.048)	(0.039)	(0.016)	(0.010)	(0.011)	(0.033)	(0.021)
TU	-1.515	-1.075	0.115	0.035	0.004	0.025	0.939	0.971
	(0.053)	(0.050)	(0.014)	(0.005)	(0.004)	(0.003)	(0.011)	(0.003)
ZL	-0.838	-0.673	0.111	0.028	0.012	0.020	0.968	0.993
	(0.080)	(0.063)	(0.023)	(0.009)	(0.009)	(0.005)	(0.010)	(0.004)

Table 4.7: Stock specific estimation results. The table contains estimated parameters of the ACI model. Standard errors are reported in parentheses. For full company names see Table 4.1

				TSX			
TICKER	IIS^1	$SE(IIS)$	HIS^1_{low}	HIS^1_{up}	HIS^1_{mid}	$SE(HIS_{mid})$	Adj^1	$SE(Adj^1)$
ABX	63.7	6.7	19.6	87.9	53.7	1.5	59.2	2.3
ABY	63.0	13.0	58.9	95.2	77.0	1.5	81.2	1.7
AEM	53.4	7.1	11.7	77.3	44.5	1.7	44.3	2.6
AGU	64.1	8.4	26.4	81.6	54.0	2.3	54.2	2.7
AL	70.4	3.4	9.2	85.8	47.5	1.4	46.3	3.1
BCE	86.1	5.3	66.3	96.9	81.6	2.2	84.8	2.7
BCM	83.8	3.0	50.9	99.5	75.2	1.7	91.9	3.4
BMO	88.5	3.0	65.4	99.7	82.5	1.3	94.6	2.1
BNN	86.3	3.1	56.8	97.5	77.1	1.7	84.0	2.4
BNS	89.1	2.8	57.1	99.7	78.4	1.1	94.7	1.6
BPO	60.6	9.7	79.8	100.0	89.9	0.5	100.0	0.3
BVF	63.0	5.2	14.6	84.9	49.8	1.6	51.1	2.6
CCJ	82.0	3.3	41.5	96.7	69.1	1.8	79.1	2.8
CGT	98.7	2.4	65.5	98.8	82.2	2.0	89.6	3.0
CLS	72.3	4.5	32.7	90.4	61.5	1.8	68.1	2.3
CNI	81.4	3.5	15.6	92.1	53.9	27.7	52.7	17.0
CNQ	83.8	2.8	38.4	98.0	68.2	1.4	82.8	2.4
COT	67.0	7.6	16.4	66.1	41.3	1.5	41.4	1.4
CP	84.7	4.4	56.6	98.1	77.4	1.4	86.4	2.0
CWG	46.4	15.1	17.8	37.7	27.7	6.4	10.0	0.9
DTC	64.2	9.3	67.3	97.9	82.6	2.1	86.9	2.9
ECA	80.8	3.5	34.6	94.0	64.3	2.1	70.9	3.2
ENB	89.8	2.8	72.2	98.2	85.2	1.7	91.9	1.2
ERF	72.2	4.8	16.0	73.5	44.7	2.7	49.8	2.6
FDG	79.1	4.1	14.9	89.0	52.0	1.5	55.4	3.0
FHR	50.0	7.5	14.9	73.8	44.3	2.1	44.2	2.6
FS	59.8	5.2	1.3	57.3	29.3	1.3	13.6	1.7
GG	57.9	7.4	14.1	81.4	47.7	1.6	50.0	2.4
GIB	75.3	12.7	73.3	99.1	86.2	1.4	91.8	2.0
GIL	65.2	8.2	46.4	91.8	69.1	3.2	69.7	3.6
GLG	67.8	5.4	10.0	75.3	42.6	1.5	39.4	2.4
HBG	52.7	11.3	3.0	33.2	18.1	4.4	12.1	4.6
IDR	60.4	8.5	17.9	67.2	42.5	1.8	41.4	1.8
IPS	27.8	25.2	54.3	98.3	76.3	2.1	85.1	3.2
IQW	85.0	5.5	41.2	81.0	61.1	4.1	60.8	2.2
ITP	73.4	12.0	36.6	92.1	64.3	2.9	71.4	3.7
KFS	83.9	8.2	75.3	98.5	86.9	1.8	90.6	2.3
KGC	73.0	9.4	70.7	100.0	85.3	0.6	100.0	0.5
MDG	71.3	6.6	11.0	77.1	44.0	1.2	43.3	1.8
MDZ	93.4	3.6	55.5	99.4	77.4	1.7	90.9	3.1
MFC	73.6	4.2	31.5	84.5	58.0	4.5	60.8	5.0
MGA	86.5	4.5	15.9	79.1	47.5	2.3	46.1	3.2
MHM	68.9	8.0	50.5	95.9	73.2	2.3	79.4	3.1
MIM	59.7	7.2	2.9	56.8	29.9	1.8	19.0	2.3
MWI	44.0	8.3	1.5	38.1	19.8	2.6	9.2	1.7
N	76.4	3.7	25.3	96.3	60.8	0.8	73.1	1.7
NCX	82.3	8.5	40.9	91.0	65.9	1.9	69.0	2.2
NRD	90.5	2.8	69.3	99.2	84.3	1.4	91.3	2.0
NT	25.7	7.2	14.3	88.7	51.5	1.6	50.0	3.2
NXY	79.3	3.6	79.1	100.0	89.6	0.8	99.9	0.7
PCZ	90.1	2.0	53.1	99.9	76.5	1.1	96.8	1.9
PDG	36.9	6.5	22.8	90.1	56.4	1.7	62.9	2.7
PDS	80.7	4.4	33.9	91.1	62.5	1.9	67.6	2.5
PGH	76.5	6.3	37.8	81.2	59.5	3.0	61.9	2.2
PKZ	56.7	4.4	19.2	87.3	53.3	1.8	56.7	2.8
POT	71.2	11.7	12.6	79.3	46.0	1.3	44.3	2.0
PWI	78.0	7.0	26.0	87.1	56.5	2.3	62.8	2.7
RBA	30.7	5.7	1.7	21.7	11.7	1.8	12.0	2.2
RCN	70.0	7.5	28.5	82.4	55.5	3.1	56.7	3.4
RG	83.0	4.2	60.4	100.0	80.2	1.5	98.8	2.0
RY	93.7	2.0	65.7	99.6	82.6	1.7	93.5	2.8
RYG	79.4	4.8	53.5	99.3	76.4	1.8	90.4	3.0
SLF	83.8	3.1	55.8	99.5	77.6	2.0	91.9	3.7
SU	76.4	4.2	50.7	97.1	73.9	1.7	81.1	2.4
TAC	88.8	4.4	66.3	98.2	82.3	1.8	88.0	2.4
TD	91.8	2.9	68.9	99.9	84.4	1.5	96.5	2.3
TEU	90.4	3.7	46.7	96.2	71.5	2.1	79.5	2.9
TLM	82.0	3.1	33.0	95.7	64.4	2.4	75.3	4.0
TOC	84.8	3.6	59.8	99.6	79.7	1.5	92.8	2.7
TRA	17.1	14.4	0.0	1.1	0.6	0.5	0.2	0.4
TRP	85.6	3.5	74.4	99.0	86.7	1.8	92.4	1.8
TU	97.6	2.5	82.4	99.9	91.2	1.0	98.2	1.2
ZL	92.7	6.3	60.5	95.4	78.0	1.9	80.3	2.2

Table 4.8: Stock specific intensity based information shares. The table presents TSX stock specific information shares using an intensity based (IIS^1) and the standard Hasbrouck approach in percent. The midpoint and the lower and upper bounds of TSX Hasbrouck are denoted by HIS^1_{mid}, HIS^1_{low} and HIS^1_{up}, respectively. The standard error for IIS is reported in the column labeled $SE(IIS)$ and for HIS_{mid} it is $SE(HIS_{mid})$. NYSE information shares can be calculated by $IIS^2 = 100 - IIS^1$, $HIS^2_{low} = 100 - HIS^1_{up}$, $HIS^2_{up} = 100 - HIS^1_{low}$ and $HIS^2_{mid} = 100 - HIS^1_{mid}$. The information shares are computed for 73 stocks that have positive $\mathbf{a}^s$ estimates. For full company names see Table 4.1.

5 Using Transfer Entropy to Measure Information Flows

Using the concept of transfer entropy we quantify the information flow between the credit default swap (CDS) and bond market in order to determine their importance in the process of pricing credit risk. The results show that overall information flows almost to an equal amount into both directions with a slight informational dominance of the CDS market. Furthermore, the dynamic relation between market risk and credit risk is examined by measuring the information transmission between the iTraxx and the VIX. Transfer entropy estimates indicate uni-directional information flow from the VIX to the iTraxx. We also conduct block bootstraps to allow for statistical inference, an issue that has not been addressed so far.

This chapter is based on the article *Using transfer entropy to measure information flows from and to the CDS market* by Franziska J. Peter (2010).

5.1 Introduction

Detecting and measuring interactions between different time series has been the subject of research studies in various areas. In finance the informational link between financial markets is of particular interest. Yet, there exists only a small range of methods to empirically examine these linkages. The predominant concept is that of *Granger causality* (see Granger 1969), which is widely applied to detect causality in the sense of a lead-lag relationship between time series. However, the conclusions that can be gained from this method are limited to the mere existence of information flows rather than their quantification. A measure for information transfer between financial markets exists only for a particular setting of empirical applications: if the prices in different markets refer to the same underlying asset, price discovery measures such as the Hasbrouck (1995) information shares or the adjustment coefficient ratio (see Gonzalo and Granger 1995, Baillie et al. 2002) can be used to determine informational dominance in a multiple market framework. These methods require a cointegration relationship between the different time series and only provide a sensible interpretation of the results if cointegration is supported by the data as well as from an economic point of view. Furthermore, Granger causality and price discovery measures are based on a Vector Autoregressive (VAR) or Vector Error Correction Model (VECM) framework, which states a rather restrictive assumptions concerning the underlying (linear) dynamics.

As an alternative to these standard models we propose to apply the concept of *transfer entropy* to measure information flows between different financial time series. Transfer entropy is a model-free measure which is designed as the Kullback-Leibler distance of transition probabilities. With very little assumptions this approach allows to quantify information transfer without being restricted to linear dynamics.

There exist only few studies so far that apply transfer entropy within the context of financial markets. Kwon and Yang (2008a) analyze the information flow between the S&P 500, the Dow Jones index and selected individual companies on a daily basis. Baek et al. (2005) examine the information transfer between groups of NYSE listed stocks to determine market sensitive and market leading companies, and Kwon and Yang (2008b) investigate the strength and direction of information transfer between various stock indices using transfer entropy. The measurement of interactions between the Indian stock and commodity market is the subject of a study by Reddy and Sebastin (2009). Finally, Marschinski and Kantz (2002), who examine the strength of the coupling between Dow Jones and DAX, propose a modification of the standard measure, the *effective transfer entropy*. Effective transfer entropy corrects for the potential upward bias in the standard measure due to finite sample effects. Still, none of these studies allows for any assessment concerning the statistical significance of the detected information flows. We fill this gap by conducting block bootstraps in order to simulate the empirical distribution of the measure and enable statistical inference concerning the estimation results.

This chapter includes two empirical applications of transfer entropy. First it is used to examine the information flows between the CDS and bond market, analyzing data on 36 iTraxx Europe companies. Both markets reflect the price of credit risk for the same reference entity and as outlined in Blanco et al. (2005) assuming cointegration between the time series seems plausible. Since the data do not support this assumption for several reference entities, the standard price discovery measure cannot be applied to the whole sample (see Blanco et al. 2005, Doetz 2007). Transfer entropy does not rely on cointegration and using this approach we find that overall the information flow from the CDS to the bond market is slightly larger than vice verse, which is in line with previous findings (see Blanco et al. 2005, Doetz 2007, Grammig and Peter 2010).

Second, after determining the dominant market for pricing credit risk, further factors that might influence the CDS market are analyzed by examining the question of causality between market risk and credit risk. Thereby we follow Figuerola-Ferretti and Paraskevopoulos (2009) and consider the dynamic relation between iTraxx and VIX. We find that the transfer entropy estimates for the flow of information from the VIX to the iTraxx are statistically significant and exceed the information flow from the iTraxx to the VIX.

5.2 The Concept of Transfer Entropy

The concept of transfer entropy is best understood within the context of information theory. In the era of early telecommunications, where communication was based on morse code, Hartley (1928) introduced a measure for information based on the logarithm of the number of all possible symbol sequences that can occur. The general aim was to optimally encode messages such that they can be transmitted more quickly. For that purpose it was necessary to quantify the information that can be gained from a specific sequence of transmitted symbols.

Consider the following example: When flipping a fair coin, there are two equally likely outcomes, heads or tails. According to Hartley (1928) the information that can be gained from flipping a coin once is given by $\mathrm{H} = \log(\frac{1}{0.5}) = \log(2)$. If the base of the logarithm is 2, $\mathrm{H} = \log_2(2) = 1$ and the measurement unit will be *bits*. Consequently, n flips of the coin yield n bits of information ($\mathrm{H} = \log(2^n) = n \times \log_2(2) = n$) and we would need n binary digits to specify the resulting sequence (such as 1 for heads and 0 for tails).

In the case of symbols that are not equally likely, but occur with different probabilities, p_j, the amount of information gained from a specific symbol j is given by $\log(1/p_j)$. The average amount (per symbol) of information one can get from such a sequence is defined as $\mathrm{H} = \sum_{j=1}^{n} p_j \log(\frac{1}{p_j})$, where n is the number of distinct symbols. This results in the general formula of Shannon (1948): Assume that J is a discrete variable with probability distribution $p(j)$, where j labels the different values (or states) that J can take. Then the Shannon entropy

$$\mathrm{H_J} = -\sum_{j} \mathrm{p}(j) \times \log_2 \mathrm{p}(j) \tag{5.1}$$

gives the average number of bits needed to optimally encode independent draws from the distribution of J. In the following *log* denotes the base 2 logarithm and the summation runs over the distinct values of J.[15] Shannon's formula in equation (5.1) is a measure for uncertainty. The more bits are needed to optimally encode realizations of the process, the higher is its uncertainty. The largest amount of uncertainty will be given if all values of J are equally likely, i.e., if J is uniformly distributed and a random draw can produce any realization of J with the same probability.

The link between uncertainty and information follows from drawing on the Kullback-Leibler distance (see Kullback and Leibler 1951). It can be used to define the excess amount of bits needed for encoding when erroneously assuming a probability distribution $q(j)$ of J different from $p(j)$:

[15] This is just a question of units measurement. Base 2 logarithm indicates bits, base 10 gives digits and the base of the natural logarithm yields nats.

$$\mathrm{K_J} = \sum_j \mathrm{p}(j) \times \log \frac{\mathrm{p}(j)}{\mathrm{q}(j)} \quad . \tag{5.2}$$

Turning to the bivariate case, let there be two discrete variables I and J with marginal probability distributions $p(i)$ and $p(j)$ and joint probability $p_{IJ}(i,j)$. The *mutual information* of the two processes is given by the reduction in uncertainty compared to the case where both processes are independent, i.e. where the joint distribution is given by the product of the marginal distributions, $p_{IJ} = p(i)p(j)$. The corresponding Kullback entropy known as the formula for mutual information is given by

$$\mathrm{M_{IJ}} = \sum_{i,j} \mathrm{p}(i,j) \times \log \frac{\mathrm{p}(i,j)}{\mathrm{p}(i)\mathrm{p}(j)} \quad , \tag{5.3}$$

where the summation runs over all possible values i and j. Mutual information can detect any form of statistical dependencies between different variables. However, it is a symmetric measure and therefore does not deliver any evidence concerning the dynamics of information exchange.

Let us switch to a time series context. Here, dynamical structure can be introduced when transition probabilities are considered (see Schreiber 2000). Be I a stationary Markov process of order k then it holds for the probability to observe I at time $t+1$ in state i conditional on the k previous observations that $p(i_{t+1}|i_t,...,i_{t-k+1}) = p(i_{t+1}|i_t,...,i_{t-k})$. The average number of bits needed to encode one more time series observation if the previous values are known is given by

$$\mathrm{h_I}(k) = -\sum_i \mathrm{p}(i_{t+1}, i_t^{(k)}) \times \log \mathrm{p}(i_{t+1}|i_t^{(k)}) \quad , \tag{5.4}$$

where $i_t^{(k)} = (i_t,...,i_{t-k+1})$. Turning again to the bivariate case Schreiber (2000) proposes to measure information flow from process J to process I by quantifying the deviation from the generalized Markov property $p(i_{t+1}|i_t^{(k)}) = p(i_{t+1}|i_t^{(k)}, j_t^{(l)})$, where l gives the order of the assumed Markov process for J. In the case of no information flow from J to I the transition probabilities of I are not affected by previous observations of J. Schreiber (2000) then draws once more on the Kullback-Leibler distance to measure the deviation of the bivariate system from this assumption and derives the formula for transfer entropy as

$$\mathrm{T}_{\mathrm{J}\rightarrow\mathrm{I}}(k,l) = \sum \mathrm{p}(i_{t+1}, i_t^{(k)}, j_t^{(l)}) \times \log \frac{\mathrm{p}(i_{t+1}|i_t^{(k)}, j_t^{(l)})}{\mathrm{p}(i_{t+1}|i_t^{(k)})} \quad . \tag{5.5}$$

$\mathrm{T}_{\mathrm{J}\rightarrow\mathrm{I}}$ consequently measures the information flow from J to I. As in empirical applications the transition probabilities have to be estimated from a specific sample, Marschinski and Kantz (2002) show that this measure is likely to be biased due to small sample effects. They propose a modification, the *effective transfer entropy*

$$\mathrm{ET}_{\mathrm{J}\rightarrow\mathrm{I}}(k,l) := \mathrm{T}_{\mathrm{J}\rightarrow\mathrm{I}}(k,l) - \mathrm{T}_{\mathrm{J}_{\mathrm{shuffled}}\rightarrow\mathrm{I}}(k,l) \quad , \tag{5.6}$$

where $\mathrm{T}_{\mathrm{J}_{\mathrm{shuffled}}\rightarrow\mathrm{I}}(k,l)$ indicates the transfer entropy with series J shuffled. This is done by randomly drawing values from the distribution of J and realigning them to generate a new time series, which implies that all statistical dependencies between the two series have been destroyed. $\mathrm{T}_{\mathrm{J}_{\mathrm{shuffled}}\rightarrow\mathrm{I}}(k,l)$ consequently converges to zero with increasing sample size and any nonzero value of $\mathrm{T}_{\mathrm{J}_{\mathrm{shuffled}}\rightarrow\mathrm{I}}(k,l)$ is due to small sample effects representing the bias in the standard entropy measure. Commonly, the data are shuffled several times and the transfer entropy estimate averaged over the simulations is used to calculate the effective transfer entropy.

To measure relevant but small scale causal structure and to allow for straight forward conclusions concerning the dominant direction of information flow, some research studies use the *normalized directionality index* (NDI). It is given by

$$\mathrm{NDI(I,J)} = \frac{\mathrm{ET}_{\mathrm{J}\rightarrow\mathrm{I}}(k,l) - \mathrm{ET}_{\mathrm{I}\rightarrow\mathrm{J}}(k,l)}{\mathrm{ET}_{\mathrm{J}\rightarrow\mathrm{I}}(k,l) + \mathrm{ET}_{\mathrm{I}\rightarrow\mathrm{J}}(k,l)} \quad . \tag{5.7}$$

The index varies between -1 and 1, where negative values imply that the information flow from I to J dominates and positive values indicate a larger information transfer from J to I. If the index equals 1 (-1) then there is uni-directional causality from J to I (I to J).

In most empirical applications - even if already discrete - the data have to be further discretized in order to reduce the number of possible states. The difficulties of discretizing the data and determining the partitions is generally referred to as the *generating partitions problem*. There exist different methods for discretization of the data as well as for estimation of the joint and conditional probabilities in Equation (5.5) (see for instance Hlavackova-Schindler et al. 2007). In the following applications we will use a simple partitioning of the data into three disjoint bins (symbolic encoding), which is motivated by economic considerations. Joint and conditional probabilities are then approximated by the observed frequencies.

Furthermore, the choice of the block lengths, l and k is a crucial point when calculating transfer entropy measures. Generally, k and l have to be large enough to capture the information flow between two time series. On the one hand finite sample effects will become more severe when increasing the block length. On the other hand, if k in $\mathrm{T}_{\mathrm{J}\to\mathrm{I}}(k,l)$ is too low then information contained in the past values of I might erroneously be assigned to come from J. This will not happen if I is independent from itself with a delay of k. Consequently, the dynamics within a single time series have to be considered when choosing k. According to Reddy and Sebastin (2009) and Fraser and Swinney (1986) the selection of the appropriate block length can be based on the mutual information. This is done by calculation of mutual information between a time series and its own series lagged with delay k as a function of the increasing block length k. The value of k associated with the first local minimum of this function can be used as the optimal block length. The length l of J in $\mathrm{T}_{\mathrm{J}\to\mathrm{I}}(k,l)$ is then commonly set to 1 or $l = k$.

5.3 Empirical Applications

5.3.1 Pricing Credit Risk: Information Flows between the CDS market and the Corporate Bond Market

With the emergence of the CDS market, default risk has become directly tradable in an over-the-counter (OTC) market. A CDS is a contract between two counterparties which transfers credit risk from the protection buyer to the protection seller, who is willing to assume the risk for a pre-specified fee. In case of default of the underlying financial instrument the buyer receives a payoff. The CDS price is denoted in basis points and gives the annualized fee of the notional volume. Credit risk is also determined implicitly by the corporate bond market as the difference between risky bond yields and the risk-free rate. Consequently, as both markets price credit risk, they are linked by an approximate arbitrage relation (see Duffie 1999, Hull and White 2000a, Hull and White 2000b). The econometric translation of such a setting is a cointegration relation between credit spread and CDS price, which means that both series follow one common stochastic trend that according to Hasbrouck (1995) can be regarded as the efficient price of credit risk. Hasbrouck (1995) furthermore proposes a Vector Error Correction Model (VECM) (see Blanco et al. (2005), Grammig and Peter (2010) and Doetz (2007)) for such multiple market settings. Decomposing the variance of the efficient price innovations that can be derived from such a model into contributions from either market delivers the Hasbrouck (1995) information shares (for details see Chapter 2). However, due to market imperfections such as liquidity premia which not exactly match maturity dates and cheapest to delivery options in case of default the arbitrage relation is not perfect. If cointegration is not supported by the data, application of the Hasbrouck methodology might yield inconclusive results. In addition, the VECM is a rather restrictive model based on linear dynamics. Therefore we present the transfer entropy estimates as an alternative approach, which relies on minimal assumptions and does not require a cointegration relation between both time series.

We use CDS price and credit spread data collected by Doetz (2007).[16] The data comprise 36 iTraxx companies and ranges from 21 January 2004 to 31 October 2006. They are obtained from Bloomberg and Thomson Financial Datastream. Both series are closing prices. The CDS spreads are midpoints of indicative bid-ask prices for 5-year contracts. The 5-year bond spread was calculated by interpolation over to bonds with different maturities. The risk-free rate over which the bond spread is calculated is proxied by the swap rate. Table 5.1 gives information on the 36 reference entities used in this application.

[16] We are grateful to N. Doetz for making these data available.

Ticker	Company	Country	Sector
ALL	Allianz	Germany	Financial
ALT	Altadis	Spain	Consumers
ARC	Arcelor	France	Industrials
BAY	Bayer	Germany	Industrials
BBI	Bco Bilbao	Spain	Financial
BMW	BMW	Germany	Autos
BSA	Bco Santander Central Hispano	Spain	Financial
CAR	Carrefour	France	Consumers
CAS	Casino Guichard-Perrachon and Cie	France	Consumers
COM	Commerzbank	Germany	Financial
DAI	DaimlerChrysler	Germany	Autos
DBA	Deutsche Bank	Germany	Financial
DET	Deutsche Telekom	Germany	TMT
EDP	Energias de Portugal	Portugal	Energy
ELT	Electricite de France	France	Energy
ENB	Energie Baden Wuerttemberg	Germany	Energy
END	Endesa	Spain	Energy
FOR	Fortum Oyi	Finland	Energy
FRAT	France Telecom	France	TMT
LAF	Lafarge	France	Industrials
LOU	LVMH	France	Consumers
NAT	National Grid	UK	Energy
OTE	Organisation Societe Anonyme	Greece	TMT
PSA	Peugeot	France	Autos
REP	Repsol	Spain	Energy
RWE	RWE	Germany	Energy
STG	St Gobain	France	Industrials
TELE	Telefonica	Spain	TMT
TELI	Telecom Italia	Italy	TMT
THY	ThyssenKrupp	Germany	Industrials
TNOR	Telenor	Norway	TMT
VAT	Vattenfall	Sweden	Energy
VIV	Vivendi	France	TMT
VOD	Vodafone	UK	TMT
VW	VW	Germany	Autos
WOL	Wolters Kluver	Netherlands	TMT

Table 5.1: **Reference entities.** The table shows the ticker symbols, company names, country and industry sector of the sample reference entities.

We calculate first differences of CDS prices $r_t^{CDS} = p_t^{CDS} - p_{t-1}^{CDS}$ and the corresponding credit spreads $r_t^{CS} = p_t^{CS} - p_{t-1}^{CS}$ to ensure stationarity. The observations are then partitioned into discretized values:

$$\begin{aligned} S(t) &= 1 \text{ for } r(t) <= q_1 \\ S(t) &= 2 \text{ for } q_1 < r(t) < q_2 \\ S(t) &= 3 \text{ for } r(t) >= q_2 \end{aligned}$$

The symbolic encoding above replaces each value in the return series of either market by its symbol (1,2,3). We choose the 0.05 quantile of either series for q_1 and the 0.95 quantile for q_2.[17]. The first bin corresponds to extremely large negative changes, the second to intermediate and the third to extremely large changes in the CDS prices and the credit spread. This choice is motivated by the leptokurtic distribution of changes in CDS prices and credit spreads. As can be seen from Figure 5.1 they deviate from a normal distribution and reveal fat tails and a peaked center. Within the context of price discovery the main task is to determine which market moves first. Consequently, if a market dominates price discovery, extreme changes in this market should be incorporated subsequently into the other market's prices. The observations in the tails of the leptokurtic return distributions of CDS prices and bond spreads therefore are of major interest. In addition, since the time series are likely to contain a considerable amount of noise due to the illiquidity of the bond market and the over-the-counter trading of both assets, the intermediate bin is kept rather large. Thereby, we seek to identify extreme changes more clearly. Based on the conditional mutual information criteria, the block lengths are set to $k = l = 3$ for all reference entities (see Table 5.6 in Appendix C).

Table 5.2 shows results for the calculated effective transfer entropy along with the normalized directionality index (NDI). Standard errors are derived from a block bootstrap using 150 repetitions and a block length of 9 symbols.[18] For all but two reference entities

[17] Using the values of the 0.1 quantile and 0.9 quantile as cut-off points does not change the overall results. When the observations in the second bin are further reduced, however, transfer estimates become mostly insignificant.

[18] The block length is increased to contain 9 symbols as it is well documented that a non-overlapping block bootstrap with blocks too short destroys too much of the dependence structure and therefore yields a downward bias of the estimated measure as well as an upward bias in the resulting variance. The block length was chosen such that the bias is reduced, i.e. the distribution of the simulated transfer entropy measures is centered around the measure derived from the original data.

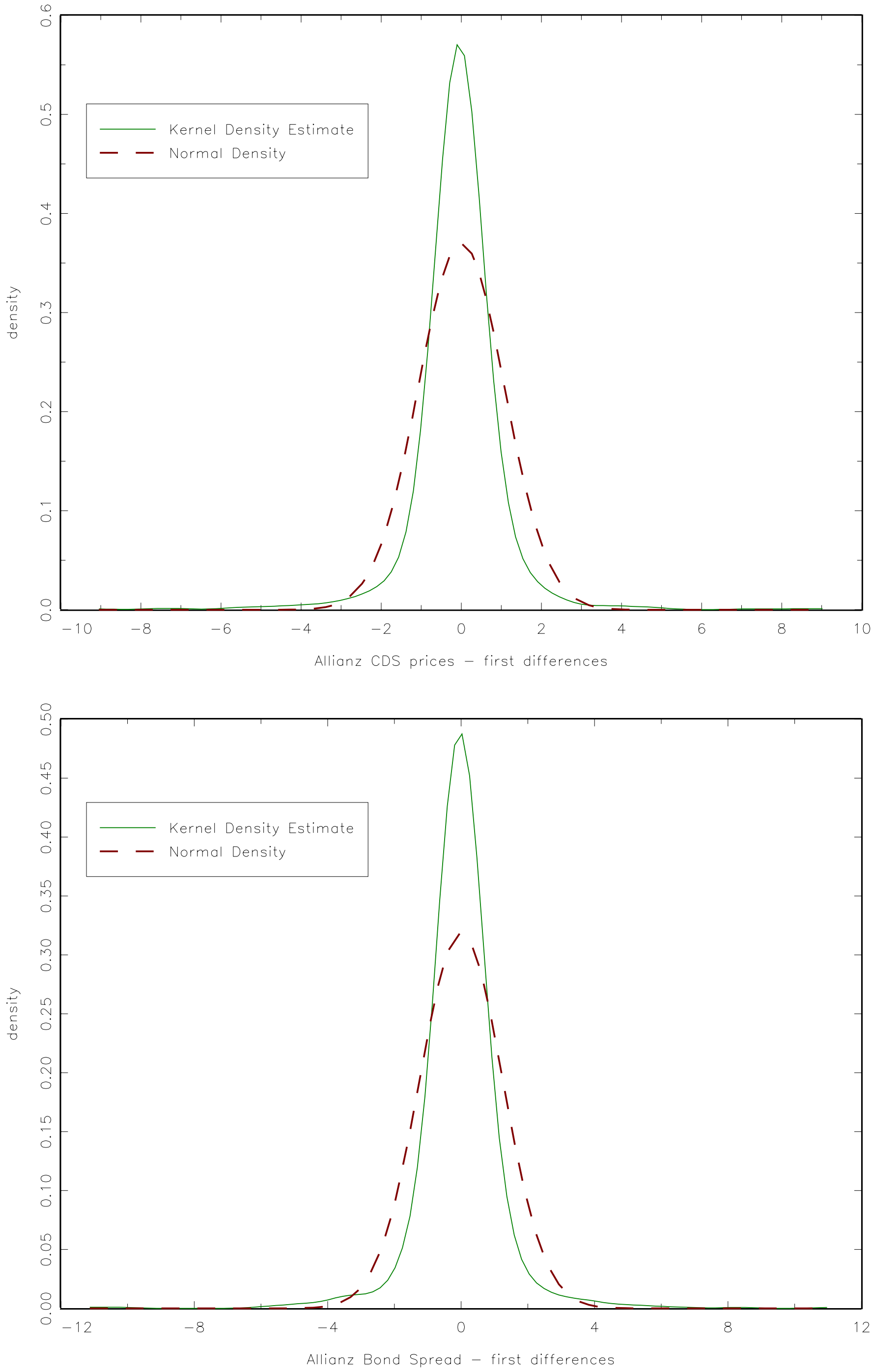


Figure 5.1: Kernel density plots of Allianz CDS and credit spread first differences. The figure shows kernel density plots of CDS and credit spread first differences for Allianz (solid line) together with the normal density (dashed line).

(BMW and VAT) we observe statistically significant bi-directional information flows. For ten reference entities the information flow from the bond to the CDS market is larger than vice verse. This is also shown by the positive values of the NDI. Consequently, for the majority of reference entities, the information flow from the CDS to the bond market dominates. Again this is mirrored by the average NDI which is negative, indicating larger transfer entropy estimates for the direction from the CDS to the bond market.
In order to render the results more easily comparable with the standard Hasbrouck information shares, we calculate the CDS market effective transfer entropy based share (ET share) as the information flow from CDS market to the bond market relative to the whole information flow between the markets:

$$\text{ET Share (CDS)} = \frac{\text{ET}(\text{CDS}\rightarrow\text{CS})}{\text{ET}(\text{CDS}\rightarrow\text{CS}) + \text{ET}(\text{CS}\rightarrow\text{CDS})} \quad . \tag{5.8}$$

Table 5.3 shows the estimated adjustment coefficients and CDS market Hasbrouck information share bounds and midpoints for those reference entities for which the data support a cointegration relation (see Table 5.5 in Appendix C). The standard errors are in parentheses and come from a non-parametric bootstrap. The last column shows the ET share of the CDS market with standard errors from the block bootstrap.
Overall both measures reveal that the contributions of the CDS market are slightly larger than the bond market contributions. The ET share shows a relative information flow of 59.0 % on average. This result is similar to the Hasbrouck information share midpoint of 58.0 %. However, there are only 17 reference entities left for which the Johansen tests indicate a cointegration relation and modeling the data using a VECM seems a plausible approach. Considering the adjustment coefficients, there are five more reference entities for an equilibrium relationship seems questionable. The error correction process implies that prices correct for deviations from the equilibrium price. This notion is not supported by adjustment coefficients which have the same sign. Furthermore, most of the estimated coefficients are not statistically significant and the bootstrapped standard errors of the information share estimates are also rather large. Apart from that, lower and upper information share bounds diverge considerably in most cases. These findings render the information share bound midpoint a questionable proxy for the contribution the price discovery process in this application. It is difficult to draw any clear conclusion by considering

Ticker	ET(CS→CDS)	ET(CDS→CS)	NDI
ALL	0.050***	0.081***	-0.242
	(0.015)	(0.018)	(0.160)
ALT	0.038***	0.038***	-0.007
	(0.014)	(0.017)	(0.301)
ARC	0.023**	0.031**	-0.145
	(0.014)	(0.016)	(0.409)
BAY	0.056***	0.060***	-0.034
	(0.014)	(0.015)	(0.193)
BBI	0.027**	0.025**	0.030
	(0.012)	(0.014)	(0.402)
BMW	0.011	0.019	-
	(0.013)	(0.014)	(0.537)
BSA	0.012**	0.044***	-0.560
	(0.014)	(0.015)	(0.353)
CAR	0.030**	0.056**	-0.294
	(0.015)	(0.019)	(0.306)
CAS	0.069***	0.051***	0.146
	(0.015)	(0.013)	(0.181)
COM	0.036***	0.034**	0.040
	(0.014)	(0.013)	(0.290)
DAI	0.029***	0.053***	-0.296
	(0.011)	(0.015)	(0.207)
DBA	0.074***	0.065**	0.066
	(0.020)	(0.019)	(0.179)
DET	0.090***	0.091***	-0.009
	(0.017)	(0.017)	(0.151)
EDP	0.045***	0.035***	0.132
	(0.013)	(0.013)	(0.245)
ELT	0.119***	0.080***	0.199
	(0.018)	(0.018)	(0.135)
ENB	0.036***	0.066***	-0.292
	(0.014)	(0.015)	(0.180)
END	0.046***	0.068***	-0.194
	(0.013)	(0.016)	(0.171)
FOR	0.045***	0.032***	0.175
	(0.013)	(0.013)	(0.261)
FRAT	0.063***	0.083***	-0.135
	(0.014)	(0.014)	(0.141)
LAF	0.015*	0.018**	-0.090
	(0.013)	(0.014)	(0.461)
LOU	0.024**	0.029***	-0.108
	(0.012)	(0.014)	(0.308)

continued

Ticker	ET(CS→CDS)	ET(CDS→CS)	NDI
NAT	0.057***	0.017*	0.544
	(0.018)	(0.013)	(0.292)
OTE	0.016**	0.051***	-0.529
	(0.014)	(0.019)	(0.275)
PSA	0.051***	0.077***	-0.200
	(0.013)	(0.017)	(0.172)
REP	0.092***	0.069***	0.144
	(0.019)	(0.016)	(0.128)
RWE	0.069***	0.117***	-0.258
	(0.015)	(0.018)	(0.137)
STG	0.066***	0.077***	-0.075
	(0.016)	(0.018)	(0.161)
TELE	0.056***	0.050***	0.059
	(0.013)	(0.013)	(0.196)
TELI	0.028**	0.026***	0.046
	(0.013)	(0.015)	(0.344)
THY	0.059***	0.068***	-0.072
	(0.016)	(0.015)	(0.151)
TNOR	0.007*	0.024**	-0.546
	(0.013)	(0.014)	(0.363)
VAT	-0.007	0.004	-
	(0.010)	(0.010)	(0.594)
VIV	0.056***	0.064***	-0.069
	(0.012)	(0.016)	(0.165)
VOD	0.033**	0.044***	-0.145
	(0.014)	(0.018)	(0.292)
VW	0.032***	0.067***	-0.346
	(0.014)	(0.015)	(0.236)
WOL	0.033**	0.043***	-0.136
	(0.012)	(0.015)	(0.297)
Mean	0.044	0.052	-0.094

Table 5.2: Effective transfer entropy estimates. The table shows effective transfer entropy estimates calculated using the mean over shuffled data from 100 repetitions. The last column gives the normalized directionality index. Negative (positive) values indicate a larger information flow from the CDS (bond) to the bond (CDS) market. Standard errors derived from a block bootstrap are in parenthesis. *,**,*** indicate significance at the 10, 5, 1 % significance level.

TICKER	ADJ. COEFF.		HASBROUCK			ET SHARE
	CS	CDS	LOW CDS	UP CDS	MID CDS	CDS
ALL	-	-	-	-	-	62.12
						(7.36)
ALT	-0.014	0.001	78.70	99.00	88.85	50.34
	(0.010)	(0.009)	(33.41)	(28.49)	(29.35)	(15.52)
ARC	-0.007	0.010	30.53	67.07	48.80	57.26
	(0.008)	(0.011)	(30.22)	(22.39)	(24.15)	(20.39)
BAY	-0.055	0.062	4.87	87.62	46.24	51.68
	(0.042)	(0.029)	(4.42)	(13.40)	(7.90)	(9.16)
BBI	-	-	-	-	-	48.52
						(16.55)
BMW	-	-	-	-	-	-
						(26.77)
BSA	-0.018	0.000	99.80	99.99	99.90	77.98
	(0.010)	(0.001)	(13.99)	(14.00)	(13.93)	(17.36)
CAR	-0.023	0.006	61.66	72.17	66.92	64.72
	(0.015)	(0.006)	(28.46)	(27.03)	(27.62)	(13.95)
CAS	-	-	-	-	-	42.71
						(10.28)
COM	-	-	-	-	-	48.00
						(14.89)
DAI	0.007	0.016	1.71	34.23	17.97	64.80
	(0.011)	(0.013)	(32.99)	(21.84)	(21.38)	(10.54)
DBA	-	-	-	-	-	46.68
						(8.27)
DET	-	-	-	-	-	50.44
						(6.59)
EDP	-	-	-	-	-	43.42
						(12.95)
ELT	-	-	-	-	-	40.06
	-	-	-	-	-	(6.60)
ENB	-	-	-	-	-	64.61
						(9.81)
END	-	-	-	-	-	59.69
						(8.51)
FOR	-0.004	0.020	4.20	9.19	6.70	41.23
	(0.011)	(0.010)	(22.94)	(25.45)	(24.00)	(13.04)
FRAT	-	-	-	-	-	56.75
						(8.00)
LAF	-0.015	0.008	37.40	72.08	54.74	54.52
	(0.012)	(0.008)	(32.38)	(26.62)	(28.55)	(22.56)
LOU	-	-	-	-	-	55.40
						(14.78)
continued						

TICKER	Adj. Coeff.		Hasbrouck			ET share
	CS	CDS	LOW CDS	UP CDS	MID CDS	CDS
NAT	-	-	-	-	-	22.80
						(16.11)
OTE	-0.011	0.010	38.53	54.13	46.33	76.46
	(0.008)	(0.007)	(27.78)	(27.99)	(27.73)	(13.47)
PSA	-0.009	0.011	19.99	33.56	26.77	59.99
	(0.011)	(0.008)	(32.74)	(32.61)	(32.44)	(10.23)
REP	-	-	-	-	-	42.82
						(7.11)
RWE	-	-	-	-	-	62.90
						(6.70)
STG	-	-	-	-	-	53.76
						(8.23)
TELE	-0.021	-0.002	97.85	98.44	98.15	47.03
	(0.006)	(0.005)	(8.82)	(14.69)	(10.87)	(13.00)
TELI	-0.002	0.006	5.20	69.41	37.31	47.68
	(0.009)	(0.011)	(31.89)	(32.78)	(25.26)	(19.12)
THY	0.001	0.051	0.05	22.90	11.47	53.58
	(0.015)	(0.018)	(8.66)	(22.87)	(13.85)	(8.58)
TNOR	-	-	-	-	-	77.31
						(22.02)
VAT	-0.018	-0.001	99.62	99.72	99.67	-
	(0.008)	(0.003)	(21.22)	(23.44)	(22.21)	(31.49)
VIV	-	-	-	-	-	53.44
						(8.24)
VOD	-0.008	-0.001	99.32	96.86	98.09	57.27
	(0.008)	(0.004)	(32.68)	(35.63)	(33.49)	(14.61)
VW	-0.015	0.017	19.54	72.34	45.94	67.28
	(0.011)	(0.009)	(31.11)	(31.12)	(28.88)	(11.79)
WOL	-0.120	0.008	91.79	95.23	93.51	56.82
	(0.020)	(0.006)	(12.89)	(11.14)	(11.89)	(14.85)
Mean			46.35	69.59	57.97	59.01

Table 5.3: Information share estimates. The tables shows the CDS market Hasbrouck information share bounds and midpoints (for details see Section 2) as well as the CDS market effective transfer entropy relative to the total information flow between the markets in percents. Hasbrouck shares are only calculated for those reference entities, for which the Johansen cointegration tests support the existence of a equilibrium price (see Table 5.5 in Appendix C). The transfer share is calculated only for those reference entities for which at least one of the estimates is statistically significant at the 10% level. Standard errors are in parentheses and derived from a non-parametric bootstrap for the Hasbrouck estimates and a blockbootstrap for the effective transfer entropy share.

Hasbrouck shares. In contrast the ET share standard errors are smaller and the transfer entropy measure is able to deliver a clearer picture. As it does not rely on a cointegration relation it is also applicable to all reference entities.[19]

5.3.2 The Information Transfer between Market Risk and Credit Risk

Concerning the process of pricing credit risk, there is evidence that the bond market constitutes the less important trading venue when compared to the CDS market. Yet, when agents price CDS, they need information concerning credit risk, i.e the default probability of the underlying reference entity. Some information might be gained from a rating agencies, but can information concerning credit risk also be extracted from the stock market? A theoretical link between stock market and credit risk can be found in the model by Merton (1974). Empirical studies also document this link. Byström (2005) examines the relation between the iTraxx indexes as a measure for credit risk and the stock price movements of the underlying entities. He detects a positive correlation between stock index return volatility and the iTraxx. Furthermore, Longstaff et al. (2007) as well as Pan and Singleton (2007) find a link between sovereign credit risk and the VIX index, which constitutes a measure for market risk. Overall, although the relationship between credit risk and market risk has been the subject of research recently, none of these empirical studies reaches beyond the point of detecting correlations.

To fill this gap the following section uses the concept of transfer entropy to determine dynamic link between market risk and credit risk by quantification of the information transfer between the iTraxx the VIX Index. The VIX is used as proxy for market risk. It is based on the implied volatilities of S&P 500 index options and measures the expectations of stock market volatility over the next 30 days.[20]

The data comprise daily closing prices of the VIX ranging from 21 January 2004 to 31

[19] The ET share was not calculated for VAT and BMW since the estimates in Table 5.2 do not show significant information flows between the time series.

[20] First constructed from the CBOE S&P 100 Index option prices the VIX was introduced in 1993. In 2003 the construction of the VIX was revised and the underlying index is changed the CBOE S&P 500 Index (see Whaley (2008)). As high VIX values are generally associated with a large amount of volatility as a result of investor uncertainty the VIX is often used as a measure of market risk.

October 2006 and the corresponding iTraxx Europe 5-year index data.[21] Figure 5.2 shows the time series of VIX and iTraxx over the pre-crisis sampling period.

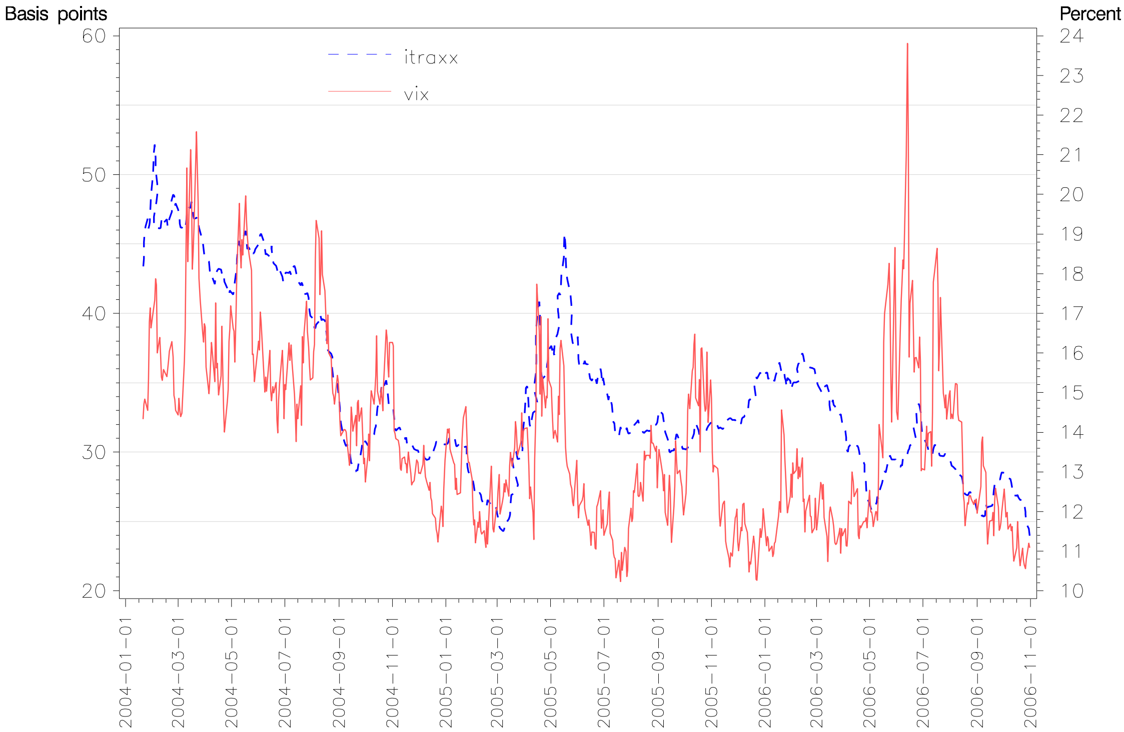


Figure 5.2: VIX and iTraxx. The figure shows the VIX and iTraxx time series. The left axis scale belongs to the iTraxx (dashed line) and is denoted in basis points. The right axis scale measures the VIX (solid line) given in percent.

Figuerola-Ferretti and Paraskevopoulos (2009), who also examine the dynamic link between market risk and credit risk. Using time series data before and throughout the recent crisis they find cointegration between CDS prices and the VIX and estimate price discovery measures based on a VECM. Yet, economic intuition does not readily provide an interpretation for a common stochastic trend in market risk and credit risk measures. We therefore use the concept of transfer entropy to examine information flows. After

[21] The iTraxx Europe is a CDS index which is composed of the most liquid 125 CDS referencing European investment grade credits. The index resulted from a 2004 merger of the two main CDS indices iBoxx and Trac-c and is traded over-the-counter mostly with 5 years maturity.

QUANTILES USED FOR q1	q2	ET(VIX→iTRAXX)	ET(iTRAXX→VIX)	NDI
0.05	0.95	0.013*	0.003	0.602
		(0.009)	(0.009)	
0.10	0.90	0.018**	0.011*	0.261
		(0.011)	(0.010)	
0.15	0.85	0.032***	0.012**	0.463
		(0.012)	(0.010)	
0.20	0.80	0.028***	0.001	0.901
		(0.012)	(0.008)	
0.25	0.75	0.031***	0.003*	0.812
		(0.011)	(0.009)	
0.30	0.70	0.039***	0.007**	0.692
		(0.010)	(0.008)	
0.35	0.65	0.041***	0.010*	0.606
		(0.011)	(0.009)	
0.40	0.60	0.050***	0.023**	0.364
		(0.013)	(0.011)	
MEAN		0.031	0.009	0.588
STD. DEV.		0.012	0.007	0.218

Table 5.4: Effective transfer entropy for VIX and iTraxx. The table shows the effective transfer entropy from VIX to iTraxx series and vice verse as well as the net directional index. All measures are calculated with the data separated into three bins according to the values associated with the quantiles given in the first two columns. Bootstrapped standard errors are in parenthesis. *,**, *** denote statistical significance on the 10, 5, 1% level of significance.

computing first differences in both series the observations are selected into three bins (see Equation 5.8). Effective transfer entropy estimates are then calculated for both directions according to Equation (5.6). The number of lags included is set to one to reduce finite sample effects due to the rather short time series. The shuffled value is the mean over 150 repetitions and the standard errors come from a non-overlapping block bootstrap with a block length of 5 and 200 repetitions. Furthermore, the thresholds for selection into the three bins are varied to examine the robustness of the results, which due to the large number of reference entities and time consuming calculation was not possible in the previous application. The net directional index (NDI) is given by Equation (5.7) and computed so that positive values denote a net information flow from the VIX to the iTraxx and negative values a net information flow from the iTraxx to the VIX.

Results are presented in Table 5.4. The information transfer from VIX to iTraxx is significant at least at the 10% independent of the different binning. Vice verse, effective transfer entropy estimates from iTraxx to the VIX are statistically significant at least at the 10% for only 6 out of the 8 different binning. Averaged over all estimations, the effective transfer entropy of VIX to iTraxx amounts to 0.031, vice verse the average is 0.009. Generally, the effective transfer entropy estimate for the flow from VIX to the iTraxx is larger than the corresponding estimate of the reverse information flow and these results are robust for different selection of observations into the three bins. This is also reflected in the positive values of the NDI, which has an average of 0.588. To sum up, these results show that there exists a dynamic relation between market risk and credit risk, with a predominant influence of the VIX to the iTraxx. However, as usual causality does not imply that the VIX itself causes the iTraxx in a sense of generating information concerning credit risk. In this particular setting, the different market designs should be taken into account, namely the OTC trading of the iTraxx versus electronic trading of the VIX. The information flow from VIX to the iTraxx could also imply that both react to other (macro) factors, which are incorporated into the VIX first, while the non-electronic OTC trading of the iTraxx might induce a delay of incorporation of information.

5.4 Concluding Remarks

This chapter uses the concept of transfer entropy to examine the information transfer between financial markets. Transfer entropy as defined by Schreiber (2000) quantifies information transmission based on the Kullback-Leibler distance. Its main advantages are that it is non-parametric and accounts for linear as well as nonlinear dynamics. Thereby, it constitutes an interesting alternative to standard measures such as Granger causality, which can detect causality, but not quantify the amount of information transfer. In the particular setting of measuring contributions to price discovery, it also states an appealing alternative to the standard Hasbrouck measure. In particular, if the data cannot be modeled within a VECM framework due to the theoretical or empirical lack of a cointegration relation or if the resulting information share bounds are extremely wide, transfer entropy with its minimal data requirements can be applied to derive results concerning the dominant direction of information flow. We apply the concept of transfer entropy to examine the information flows between the CDS and bond market using data on 36 iTraxx companies. The results show significant information flow into both directions, while the information transfer from the CDS market to the bond market is slightly higher. These results are in line with previous findings concerning the informational dominance of the CDS market over the bond market.

Furthermore, we examine information transmission between market risk and credit risk as proxied by the VIX and iTraxx Europe Index. We find that information flows mainly from the VIX to the iTraxx, which states an interesting result for the evaluation of credit risk. Finally, the block bootstrap proposed in this chapter allows to conduct inference of the estimated information flow measures, an issue that has not been addressed so far.

Appendix C: Additional Tables

Ticker	H0: at most one	H0: none
ALL	6.297**	9.081***
ALT	0.555	3.969**
ARC	2.295	6.819***
BAY	2.263	48.859***
BBI	4.852**	10.043***
BMW	3.248*	7.319***
BSA	2.341	23.458***
CAR	1.963	11.198***
CAS	3.594*	10.144***
COM	6.813***	10.897***
DAI	1.046	4.070**
DBA	6.050**	12.640***
DET	9.651***	17.708***
EDP	3.712*	14.102***
ELT	6.859***	8.932***
ENB	4.238**	20.931***
END	6.628***	10.954***
FOR	1.917	8.502***
FRAT	10.799***	78.809***
LAF	1.184	7.252***
LOU	10.815***	22.846***
NAT	2.880*	6.045**
OTE	0.428	9.466***
PSA	2.218	5.757**
REP	14.883***	47.748***
RWE	4.510**	29.520***
STG	19.939***	42.513***
TELE	1.780	28.305***
TELI	0.236	2.014
THY	1.587	13.677***
TNOR	3.206*	14.522***
VAT	2.737	14.252***
VIV	4.441**	6.945***
VOD	1.429	2.736
VW	1.575	4.385**
WOL	0.515	34.092***

Table 5.5: Johansen cointegration test statistics. The table shows the Johansen trace and maximum eigenvalue statistic. *,**, *** denote rejection of the null hypothesis on the 10, 5, 1% level of significance.

Ticker	CS		CDS	
	local minimum	k	local minimum	l
ALL	0.019	3	0.078	1
ALT	0.009	3	0.011	2
ARC	0.005	4	0.008	4
BAY	0.053	2	0.054	2
BBI	0.000	3	0.017	5
BMW	0.006	2	0.020	2
BSA	0.030	3	0.001	3
CAR	0.010	2	0.005	5
CAS	0.008	3	0.026	3
COM	0.008	2	0.007	5
DAI	0.009	3	0.014	3
DBA	0.027	3	0.011	3
DET	0.040	2	0.060	2
EDP	0.001	2	0.018	3
ELT	0.041	2	0.020	5
ENB	0.024	3	0.006	6
END	0.011	5	0.029	4
FOR	0.002	2	0.060	1
FRAT	0.026	4	0.058	1
LAF	0.003	2	0.009	3
LOU	0.006	2	0.016	3
NAT	0.023	2	0.010	4
OTE	0.028	1	0.001	2
PSA	0.021	2	0.008	5
REP	0.036	2	0.027	3
RWE	0.020	2	0.060	4
STG	0.007	2	0.030	2
TELE	0.008	2	0.019	2
TELI	0.016	3	0.032	2
THY	0.015	5	0.021	4
TNOR	0.007	6	0.008	2
VAT	0.002	5	0.013	2
VIV	0.011	4	0.013	3
VOD	0.005	2	0.023	3
VW	0.011	3	0.034	2
WOL	0.005	2	0.023	3

Table 5.6: Optimal block length selection. The table shows the first local minimum of the mutual information criteria for the optimal block length selection for each series.

6 Summary and Conclusion

Measuring the contribution of different markets to the price discovery process of a common asset has been the subject of many research studies in the last decade. In particular the newly developed derivatives markets have given rise to an increasing number of empirical studies examining the importance of each trading venue with respect to price discovery. The information shares developed by Hasbrouck (1995) are the most prevalent measure that has been applied in numerous empirical analyses. However, this approach suffers from a lack of identification which often renders it very difficult to draw clear conclusions. Chapter 3 proposes a data driven approach to resolve the identification problem inherent in the Hasbrouck (1995) approach. Based on a distributional assumption the contemporaneous effects of idiosyncratic price innovations can be determined, which results in unique information shares. The empirical application to the credit default swap and corporate bond market shows that the credit default swap market leads in price discovery for a set of European and U.S. reference entities.

The approach presented in Chapter 4 takes a different perspective in order to develop a new measure for contributions to price discovery. Price discovery is seen as the timing of informative events and we use a model that accounts for the irregularity of high frequency quote data. Estimates from a bivariate autoregressive conditional intensity model are used to develop a new intensity based information share. Applied to a large set of Canadian cross-listed stocks, it shows that the home market plays a more important role in price discovery than the foreign (U.S.) market.

Finally, Chapter 5 refrains from any restrictive assumption with respect to the price discovery process across markets, but estimates statistical dependencies between two time series. It outlines and applies the concept of transfer entropy to measure information flows from and to the credit default swap market. This non-parametric method relies on only a few assumptions and is able to capture linear as well as nonlinear information flows between time series. The empirical application provides evidence for bi-directional information transfer between the credit default swap and corporate bond market for a set of European reference entities, while there is evidence for credit default swap market being slightly more important with respect to information transfer.

The three new econometric approaches presented in this thesis all seek to answer the same question. However, a general evaluation in the sense of which approach is best, is difficult. The standard Hasbrouck (1995) information share and the method outlined in Chapter 3 are developed within a theoretical framework that models the efficient price of the common underlying asset. It requires cointegration between the different price series and defines contributions to price discovery by the share of a market in the variance of the innovations to the efficient price. If the return series considered show high contemporaneous correlation, leptokurtic distributions and tail dependence, the mixture normal assumption outlined in Chapter 3 could be used to deliver a measure that is more accurate than the standard Hasbrouck (1995) approach. These approaches need equally spaced data. In the case of irregular high frequency data, the intensity based information share model brought forward in Chapter 4 is able to capture the information in durations between price changes, which gets lost when sampling at specific intervals. It also relies on the idea of arbitrage, i.e. that informative price changes are subsequently incorporated into the prices in the different markets. However, it does not explicitly require the existence of a cointegration relation, nor does it make any assumptions concerning the equilibrium price. In particular, in the case of electronic trading platforms such as limit order books subsequent incorporation of information in several markets is often a question of a few seconds. Modeling this specific irregular structure of high frequency data in a multivariate framework in order to determine the leading market in price discovery therefore offers an appealing alternative to the standard methods, for which the data are sampled arbitrarily at a specific interval.

Transfer entropy as outlined in Chapter 5 is an overall less restrictive approach. It is not based on any assumptions or model, but measures statistical dependencies. One might argue that it lacks economic appeal, however, its advantage lies in the fact that it is not restricted to nonlinear dependencies and offers a very general framework to measure information flows.

The choice of the most suitable method therefore depends on the specific empirical application that is considered, the economic framework, availability and structure of the data. Still, results from all three approaches should be subject to extensive robustness checks. This requires testing for the existence of cointegration, the choice of sampling intervals, and the suitability of the distributional assumptions in the case of the unique information share model of Chapter 3. The choice of the appropriate thresholds in the data filtering process of the intensity based information share should be made with care and the desicion of how to discretize the data necessary to apply the concept of transfer entropy demands robustness checks. If these requirements are met any of the three methods presented in this thesis has the capability to provide empirical evidence in order to answer the question "Where is the market?" in a one security-multiple markets setting.

References

AKAIKE, H. (1969): "Fitting Autoregressive Models for Prediction," *Annals of the Institute of Statistical Mathematics*, 21, 243–247.

——— (1971): "Autoregressive Model Fitting for Control," *Annals of the Institute of Statistical Mathematics*, 23, 163–180.

BAEK, K. S., W.-S. JUNG, O. KWON, AND H.-T. MOON (2005): "Transfer Entropy Analysis of the Stock Market," *ArXiv Physics e-prints*.

BAILLIE, R. T., B. G. GEOFFREY, Y. TSE, AND T. ZABOTINA (2002): "Price Discovery and Common Factor Models," *Journal of Financial Markets*, 5, 309–321.

BAUWENS, L., AND N. HAUTSCH (2006): "Stochastic Conditional Intensity Processes," *Journal of Financial Econometrics*, 4, 450–493.

BLANCO, R., S. BRENNAN, AND I. W. MARSH (2005): "An Empirical Analysis of the Dynamic Relation Between Investment-grade Bonds and Credit Default Swaps," *Journal of Finance*, 60(5), 2255–2281.

BOLLERSLEV, T., AND J. WOOLDRIDGE (1992): "Quasi-maximum Likelihood Estimation and Inference in Dynamic Models with Time-varying Covariances," *Econometric Reviews*, 11(2), 143–172.

BOOTH, G., J.-C. LIN, T. MARTIKAINEN, AND Y. TSE (2002): "Trading and Pricing in Upstairs and Downstairs Stock Markets," *Review of Financial Studies*, 15(4), 1111–1135.

BOWSHER, C. G. (2007): "Modelling Security Market Events in Continuous Time: Intensity Based, Multivariate Point Process Models," *Journal of Econometrics*, 141, 876–912.

BRÉMAUD, P. (1981): *Point Processes and Queues, Martingale Dynamics*. Springer, New York.

BYSTRÖM, H. (2005): "Credit Defult Swaps and Equity Prices: the iTraxx CDS Index Market," Working Paper, Lund University.

CARPENTIER, C., J.-F. L'HER, AND J.-M. SURET (2007): "Competition and Survival of Stock Exchanges: Lessons from Canada," CIRANO Working Papers.

Coffee, J. C. (2002): "Racing Towards the Top?: The Impact of Cross-listings and Stock Market Competition on International Corporate Governance," *Columbia Law Review*, 102(7), 1757–1832.

Collin-Dufresne, P., R. S. Goldstein, and M. Spencer (2001): "The Determinants of Credit Spread Changes," *Journal of Finance*, 56, 2177–2207.

Davidson, R., and J. G. MacKinnon (2000): *Econometric Theory and Methods*. Oxford University Press, Oxford.

De Jong, F. (2002): "Measures on Contributions to Price Discovery: A Comparison," *Journal of Financial Markets*, 5, 323–327.

Dickey, D., and W. Fuller (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root," *Econometrica*, 49, 1057–1072.

Doetz, N. (2007): "Time-Varying Contributions by the Corporate Bond and CDS Markets to Credit Risk Price Discovery," Discussion Paper Series 2: Banking and Financial Studies.

Duffie, D. (1999): "Credit Swap Valuation," *Financial Analysts Journal*, 55, 73–87.

Dufour, A., and R. F. Engle (2000): "Time and the Price Impact of a Trade," *Journal of Finance*, 55, 2467–2498.

Engle, R. F. (2000): "The Econometrics of Ultra-high Frequency Data," *Econometrica*, 68, 1–22.

Engle, R. F., and A. Lunde (2003): "Trades and Quotes: A Bivariate Point Process," *Journal of Financial Econometrics*, 1, 159–188.

Engle, R. F., and J. R. Russell (1997): "Forecasting the Frequency of Changes in Quoted Foreign Exchange Prices with the Autoregressive Conditional Duration Model," *Journal of Empirical Finance*, 4, 187–212.

——— (1998): "Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data," *Econometrica*, 66, 1127,1162.

Eubank, R., and P. Speckman (1990): "Curve Fitting by Polynomial-trigonometric Regression," *Biometrika*, 77, 1–9.

Eun, C. S., and S. Sabherwal (2003): "Cross-border listings and Price Discovery: Evidence from U.S.-listed Canadian stocks," *Journal of Finance*, 58, 549–575.

Figuerola-Ferrett, I., and J. Gonzalo (2010): "Modelling and Measuring Price Discovery in Commodity Markets," *Journal of Econometrics*, 158, 95–107.

Figuerola-Ferretti, I., and I. Paraskevopoulos (2009): "The Dynamic Relation between CDS Markets and the VIX Index," Working Paper, University of Madrid.

Fraser, A. M., and H. L. Swinney (1986): "Independent Coordinates for Strange Attractors from Mutual Information," *Physical Review A*, 33(2), 1134–1140.

Frijns, B., and P. Schotman (2009): "Price Discovery in Tick Time," *Journal of Empirical Finance*, 16(5), 759–776.

Gaa, C., S. Lumkin, R. Ogrodnik, and P. Thurlow (2002): "The Future Prospects of National Markets and Trading Centers," Discussion paper, Bank of Canada Working Paper 2001-10.

Gonzalo, J., and C. W. J. Granger (1995): "Estimation of Common Long-memory Components in Co-integrated Systems," *Journal of Business and Economic Statistics*, 13, 27–36.

Grammig, J., M. Melvin, and C. Schlag (2005): "Internationally Cross-listed Stock Prices During Overlapping Trading Hours: Price Discovery and Exchange Rate Effects," *Journal of Empirical Finance*, 12, 139–164.

Grammig, J., M. Melvin, and C. Schlag (2008): "The Role of U.S. Trading in Pricing Internationally Cross-listed Stocks," Working Paper.

Grammig, J., and F. J. Peter (2010): "Tell-Tale Tails: A New Approach to Estimate Unique Market Information Shares," Working Paper, University of Tuebingen.

Granger, C. W. J. (1969): "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, 37(3), 424–438.

Haas, M., S. Mittnik, and M. Paolella (2004): "Mixed Normal Conditional Heteroskedasticity," *Journal of Financial Econometrics*, 2, 211–250.

Hall, A. D., and N. Hautsch (2006): "Order Aggressiveness and Order Book Dynamics," *Empirical Economics*, 30, 973–1005.

Harris, F. H. d., T. H. McInish, G. Shoesmith, and R. A. Wood (1995): "Cointegration, Error Correction, and Price Discovery on Informationally Linked Securitiy Markets," *Journal of Financial and Quantitative Analysis*, 4, 563–579.

Harris, F. H. d., T. H. McInish, and R. A. Wood (2002): "Common Factor Components Versus Information Shares: A Reply," *Journal of Financial Markets*, 5, 341–348.

Hartley, R. V. L. (1928): "Transmission of Information," *Bell System Technical Journal*, July, 535–563.

Hasbrouck, J. (1995): "One security, Many markets: Determining the Contributions to Price Discovery," *Journal of Finance*, 50(4), 1175–1199.

——— (2002): "Stalking the "Efficient Price" in Market Microstructure Specifications: An Overview," *Journal of Financial Markets*, 5, 329–339.

——— (2003): "Intraday Price Formation in U.S. Equity Markets," *Journal of Finance*, 58, 2375–2399.

Hautsch, N. (2003): *Modelling Irregularly Spaced Financial Data*. Springer, Berlin.

Hlavackova-Schindler, K., M. Palus, M. Vejmelka, and J. Bhattacharya (2007): "Causality Detection Based on Information-theoretic Approaches in Time Series Analysis," *Physics Report*, 441(1), 1–46.

Huang, R. D. (2002): "The Quality of ECN and Nasdaq Market Maker Quotes," *The Journal of Finance*, 57, 1540–6261.

Hull, J. C., and A. White (2000a): "Valuing Credit Default Swaps I: No Counterparty Default Risk," *Journal of Derivatives*, 8, 29–40.

——— (2000b): "Valuing Credit Default Swaps II: Modeling Default Correlations," *Journal of Derivatives*, 8, 12–22.

Hupperets, E. C., and A. J. Menkveld (2002): "Intraday Analysis of Market Integration: Dutch Blue Chips Traded in Amsterdam and New York," *Journal of Financial Markets*, 5, 57–82.

Johansen, S. (1988): "Statistical Analysis of Cointegration Vectors," *Journal of Economic Dynamis and Control*, 12, 231–254.

——— (1991): "Estimation and Hypothesis Testing of Cointegrated Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580.

Karolyi, G. A. (2006): "The World of Cross-listings and Cross-listings of the World: Challenging Conventional Wisdom," *Review of Finance*, 10(1), 99–152.

Korczak, P., and K. Phylaktis (2010): "RelatedSecurities and Price Discovery: Evidence from NYSE-listed Non-U.S. Stocks," *Journal of Empirical Finance*, 17(4), 566 – 584.

Kullback, S., and R. A. Leibler (1951): "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 1, 79–86.

Kwon, O., and J.-S. Yang (2008a): "Information Flow between Composite stock Index and Individual Stocks," *Physica A: Statistical Mechanics and its Applications*, 387(12).

——— (2008b): "Information Flow between Stock Indices," *Europhysics Letters*, 82(6).

Lanne, M., and H. Lütkepohl (2010): "Structural Vector Autoregression with Nonnormal Residuals," *Journal of Business and Economic Statistics*, 28(1), 159–168.

Lehmann, B. N. (2002): "Some Desiderata for the Measurement of Price Discovery Across Markets," *Journal of Financial Markets*, 5, 259–276.

Lien, D., and K. Shrestha (2009): "A New Information Share Measure," *Journal of Futures Markets*, 29, 377–395.

Longin, F. (2001): "Extreme Correlation of International Equity Markets," *Journal of Finance*, 56(2), 649–676.

Longstaff, F., J. Pan, P. Lasse, and K. Singleton (2007): "How Sovereign is Sovereign Credit Risk?," NBER Working Paper.

Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, New York.

——— (2008): "Problems Related to Over-identifying Restrictions for Structural Vector Error Correction Models," *Economics Letters*, 99, 512–515.

MacKinnon, J. (2002): "Bootstrap Inference in Econometrics," *Canadian Journal of Economics*, 29, 305–325.

Marschinski, R., and H. Kantz (2002): "Analysing the Information Flow between Financial Time Series," *The European Physical Journal B*, 30(2), 275–281.

Merton, R. (1974): "On the Pricing of Corporate Debt: the Risk Structure of Interest Rates," *Journal of Finance*, 29.

Mizrach, B., and C. Neely (2008): "Information Shares in the US Treasury Market," *Journal of Banking and Finance*, 32, 1221–1233.

Pan, J., and K. Singleton (2007): "Default and Recovery Implicit in the Term Structure of Sovereign CDS Spreads," Working Paper, MIT Sloan School of Management and Stanford Graduate School of Business.

Reddy, Y. V., and A. Sebastin (2009): "Are Commodity and Stock Markets Independent of Each Other? Cacase Study in India," *Journal of Alternative Investments*, 11(3), 85–99.

Rigobon, R. (2003): "Identification through heteroskedasticity," *Review of Economics and Statistics*, 85(4), 777–792.

Russell, J. R. (1999): "Econometric Modeling of Multivariate Irregularly-spaced High-frequency Data," Working Paper, University of Chicago.

Schreiber, T. (2000): "Measuring Information Transfer," *Physical Review Letters*, 85(2), 461–464.

Schwarz, G. (1978): "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461–464.

SHANNON, C. E. (1948): "A Mathematical Theory of Communication," *Bell System Technical Journal*, 27, 379–423.

STOCK, J. H., AND M. W. WATSON (1988): "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097–1107.

TORONTO STOCK EXCHANGE (2003): *TSX Factbook*. TSX Datalinx, Toronto.

VLAAR, P. (2004): "On the Asymptotic Distribution of Impulse Response Functions with Long-run Restrictions," *Review of Economics and Statistics*, 20, 891–903.

WHALEY, R. (2008): "Understanding VIX," Working Paper, Vanderbilt University.

WHITE, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrics: Journal of the Econometric Society*, 50(1), 1–25.

YAN, B., AND E. ZIVOT (2010): "A Structural Analysis of Price Discovery Measures," *Journal of Financial Markets*, 13, 1–19.