THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

Délivré par l'Université Toulouse Capitole

École doctorale : **Sciences Economiques-Toulouse School of Economics**

Présentée et soutenue par

NGUENANG KAPNANG Christian

le 25 juillet 2018

Essays in Financial Econometrics: Interlinked assets and High-Frequency Data

---

Discipline : **Sciences Economiques**

Unité de recherche : **TSE-R (UMR CNRS 5314 – INRA 1415)**

Directeur de thèse : Monsieur, Nour, MEDDAHI, Professeur, Université de Toulouse 1

JURY

Rapporteurs Monsieur, Serge, DAROLLES, Professeur, Université Paris Dauphine
Monsieur, Anders, RAHBEK, Professeur, Copenhagen university

Suffragants Madame, Sophie, MOINAS, Professeur, Université de Toulouse 1
Monsieur, Nour, MEDDAHI, Professeur, Université de Toulouse 1

Essays in Financial Econometrics: Interrelated assets and High-Frequency data

PhD Dissertation

Christian NGUENANG KAPNANG

Université de Toulouse 1 Capitole

Toulouse School of Economics

2018

Acknowledgments

This thesis is the outcome of a long journey full of enriching encounters that shaped my ideas, my thoughts, my skills, my behavior, and my fillings.

First, I would like to express my heartfelt gratitude to my Advisor Pr. Nour Meddahi. I had wished to deepen my knowledge of Financial Econometrics; he helped me get the best of this interesting field. I am grateful for the continuous support and guidance during the years of PhD research, for his patience and his priceless investment.

I would also like to thank Pr. Sophie Moinas. Many discussions with her edified me in the field of Market microstructure. Her insightful suggestions and comments helped me improve the quality of my work and highlight the importance of my research for the finance community.

My thanks also go to Fany Declerck. Many discussions with her helped me to better understand the markets data I have dealt with during my research.

I thank the thesis committee ....

I thank the members of the Econometrics group and of the Finance group at Toulouse School of Economics. Their remarks and questions during seminars and workshops certainly contribute to broaden my knowledge, improve my papers and my presentations skills.

I would like to thanks Aude Schloesing for her precious help during Out-of school administrative and housing difficulties. I also thank my fellow PhD for after-school activities that made those years very joyful and amazing.

I am grateful to my parents and to my brothers and sisters for the moral support throughout writing this thesis and in my life in general.

I would definitely like to thanks my wife Elza who is beside me since the beginning of this research, and my son Nathan for illuminating my life. Their contribution to my well-being during those stressful times is priceless.

Abstract

During the recent years, financial markets have known many institutional changes and new regulations. The improvements in ITC, the multiplication and fragmentation of markets increase the High Frequency trading activity, and the cross listing of assets in many towns or countries. The prices for a given security on those interrelated markets are strongly linked by arbitrage activities. A similar situation arises for one security and its derivatives: The Cash prices are related to futures prices, the CDS prices are related to the Credit spread, spot is related to options Markets. In those multiple market settings, it is interesting for regulators, investors and academia to understand how each market contributes to the dynamic of the common underlying fundamental value. My thesis develops new frameworks, with respect to the sampling frequency, to measure the contribution of each market to the formation of prices (Price discovery) and to the formation of volatility (Volatility discovery).

In the first chapter, I consider the problem of measuring price discovery using High-frequency data. I show that existing measures of price discovery lead to misleading conclusions when using High-frequency data, due to uninformative microstructure noises. I then propose robust-to-noise measures, good at detecting "which market incorporates quickly new information". Using the Dow Jones stocks traded on NYSE and NASDAQ on the period March 1st to May 30th 2011, I show that the data are in line with my theoretical conclusions. In addition, when the Information Share measure gives wide bounds making it unusable, my proposed robust IS has very close bounds. I later obtain that price discovery mostly happens on NYSE and Nasdaq is dominant for the four nasdaq-listed stocks. The contribution of NYSE is positively correlated with its liquidity and its market share in small size transactions. And, NASDAQ contribution to price discovery increases slightly the days with macroeconomic announcements.

In the second Chapter, I provide a new way to evaluate price adjustment across linked markets by building an Impulse Response measuring the permanent impact of market's innovation and I give its asymptotic distribution. The framework innovates in providing testable results for price discovery measures based on Hasbrouck (1995) innovation variance and gives a rationale to the Information Share Upper bound. I later present an equilibrium model of different maturities futures

markets with convenience yield and show that it supports my measure: As the theoretical result of Garbade and Silber (1983) and Figuerola-Ferretti and Gonzalo (2010), the measure selects the market with the higher number of participants as dominating the price discovery. An application on some metals of the London Metal Exchange shows that some markets are in Backwardation and others in Contago. And that, 3-month futures contract dominates the spot and the 15-month in price formation.

The third chapter tries to build a comprehensive framework for Price discovery analysis with High Frequency data. The literature exists only in a discrete time framework, we build a continuous-time framework that incorporates explicitly microstructure noises. We derive a measure of price discovery evaluating the permanent impact of a shock on a market's innovation. It has advantages on the literature in that: it is in continuous-time, deals with non-informative microstructure noises and accommodates a stochastic volatility. An application is done on the four Dow Jones stocks primary listed on NASDAQ and traded on NYSE: Apple, Intel, Microsoft and Cisco. The results show that for those stocks the NASDAQ dominates the continuous price discovery process

In the fourth chapter, as literature has focused on where information enters the price, I develop a framework to study how each markets' volatility contributes to the permanent volatility of interlinked assets. This allows answering questions such as: Where does new volatility enter the volatility of securities listed in many markets? Does volatility of futures markets dominate volatility of the Cash market in the formation of permanent volatility? I build a VECM with Autoregressive Stochastic Volatility estimated by MCMC method and Bayesian inference. I show empirically that not only prices of strongly related are cointegrated, but their conditional volatilities share a permanent factor at the daily and at the intraday level, and I propose measures of market's contribution to Volatility discovery. In the application, I study daily data of cash and 3Month futures markets of some metals traded on the London Metals Exchange, and intraday data of the EuroStoxx50 index and its futures. I find that for most of the securities, while price discovery happens on the cash market, the volatility discovery happens in the Futures market. Overall, the results suggest that Information discovery and volatility discovery do not necessarily have the same determinants. In the last part of the study, I build a framework that exploits High frequency data and avoid computational burden of MCMC. I show that Realized Volatilities are driven by a common component and I compute contribution of NYSE and NASDAQ to permanent volatility of Dow Jones stocks. I obtain a slight domination of NYSE. And among liquidity, Volume market Share by trade size, and volatility of volume. I obtain that volatility of the volume is the best determinant of volatility discovery, But low figures suggest others important factors.

Resumé

Au cours des années récentes, les marchés financiers ont connu de nombreux changements institutionnels et de nouvelles réglementations. Les developpemens dans les TIC, la multiplication et la fragmentation des marchés, ont accru l'activité de trading à Haute fréquence , et aussi la cotation simultanée des actifs dans plus en plus de villes ou pays. Les prix d'un titre donné sur ces differentes places sont liés par des activités d'arbitrage. Cette situation se présente aussi pour un titre et ses dérivés: Les prix spot sont liés aux prix futures, les CDS sont liés au Spread de crédit, le prix spot est lié au prix des options. Dans ces cadres de marchés "informationnellement reliés", il est intéressant pour le regulateur et les investisseurs de comprendre comment chaque marché contribue à la dynamique de la valeur fondamentale. Cette thèse développe de nouveaux outils pour mesurer la contribution, relativement à la fréquence, de chaque marché à la formation du prix et à la formation de la volatilité.

Dans le premier chapitre, Je montre que, en raison de bruits de microstructure, les mesures existantes de la découverte des prix conduisent à des conclusions trompeuses lorsque l'on utilise des données à haute fréquence. Je propose ensuite des mesures robustes au bruit, capables de détecter "quel marché intègre rapidement de nouvelles informations". En utilisant les titres du Dow Jones vendues sur le NYSE et le NASDAQ sur la période du 1er mars au 30 mai 2011, je montre que les données corroborent mes conclusions théoriques. De plus, lorsque les bornes de l' "Information Share" sont larges et inutilisables, mon robuste IS proposé a des bornes très serrées. J'obtiens ensuite que la découverte de prix se produit principalement sur le NYSE et est positivement corrélée avec sa liquidité et sa part de marché dans les transactions de petite taille. Le NASDAQ est dominant sur les stocks listés initialement au NASDAQ. La contribution du NASDAQ à la découverte des prix augmente légèrement les jours avec annonces macroéconomiques.

Dans le deusième chapitre, je propose une mesure de découverte prix sur les marchés reliés en construisant une fonction de réponse qui évalue l'impact permanent de l'innovation d'un marché, et je donne sa distribution asymptotique. Ce cadre semble être le premier à fournir des résultats testables pour les mesures de découverte des prix basées sur la variance d'innovation de Hasbrouck (1995), et il donne une justification à la borne supérieure de " l'Information Share". Je

présente ensuite un modèle d'équilibre des marchés à terme à différentes maturités avec rendement d'opportunité, et on montre qu'il soutient notre cadre: Conformement aux conclusions théoriques de Garbade and Silber (1983) et Figuerola-Ferretti and Gonzalo (2010), la mesure sélectionne le marché avec le plus de participants comme dominant le processus de découverte des prix. Une application sur certains métaux de la London Metal Exchange montre que le contrat à terme de 3 mois domine coinjointement le marché cash et le contrat à 15mois dans la formation des prix.

Le troisième chapitre introduit un cadre complet pour l'analyse de la découverte des prix sur données à haute fréquence. La littérature n'existe que dans un cadre de temps discret, nous construisons un cadre en temps continu qui incorpore explicitement des bruits de microstructure. Nous obtenons une mesure de la découverte des prix qui évalue l'impact permanent, d'un choc sur l'innovation d'un marché. Il présente des avantages sur la littérature en ce sens qu'il est en temps continu, traite des bruits de microstructure non informatifs et permet d'intégrer une volatilité stochastique. Une application est faite sur les quatre principales actions Dow Jones cotées au NASDAQ et négociées sur NYSE: Apple, Intel, Microsoft et Cisco. Les resultats montrent que pour ces actions le Nasdaq dominent le processus continu de découverte des prix.

Le quatrième chapitre s'intéresse à la volatilité de la volatilité. Alors que la littérature se concentre sur la quête du marché où l'information rentre dans les prix, je développe un cadre pour étudier comment la volatilité de chaque marché contribue à la volatilité permanente de l'actif. Ce qui permet de répondre à des questions telles que: La volatilité du marché futures contribue-t-elle plus que la volatilité du marché spot dans la formation de la volatilité du fondamental? Premierement, je construis un VECM avec Volatilité Stochastique estimé avec les MCMC et inférence bayésienne. Je montre empiriquement que les volatilités conditionnelles ont une composante communes et propose des mesures de découverte de la volatilité. Je l'applique aux données journalières de certains metaux de la London Metals Exchange, et aux données intrajournalières de l'EuroStoxx50 et son contrat futures. Je trouve qu'alors que la formation des prix a lieu sur le marché au comptant, la découverte de la volatilité a lieu sur le marché futures. Globalement, les résultats suggèrent que la découverte de l'information et la découverte de la volatilité n'ont pas nécessairement les mêmes déterminants. Dans une seconde partie, je construis un cadre d'analyse qui exploite les données à Haute fréquence et évite la charge de calcul des MCMC. Je montre que les Volatilités Réalisées sont formées par une composante commune et calcule la contribution du NYSE et NASDAQ à la volatilité permanente des titres du Dow Jones. J'obtiens que pour la majorité des titres, le NYSE domine la formation de la volatilité. Et, entre la liquidité, le poids du marché dans les transactions par taille, le volume, et la volatilité des volumes, la volatilité des volumes est le meilleur déterminant de la découverte de la volatilité. Mais les chiffres faibles obtenues suggèrent l'existence d'autres facteurs.

Contents

Chapter 1

Price discovery measures and High Frequency data

Abstract: For an asset traded in multiple venues, an outstanding problem is how those places individually contribute to the price discovery mechanism (the incorporation of information into prices). I show that existing measures of price discovery lead to misleading conclusions when using High-frequency data, due to uninformative microstructure noises. I then propose robust-to-noise measures, good at detecting "which market incorporates quickly new information". Using the Dow Jones stocks traded on NYSE and NASDAQ on the period March 1st to May 30th 2011, I show that the data are in line with my theoretical conclusions. In addition, when the Information Share measure gives wide bounds making it unusable, my proposed robust IS has very close bounds. I later obtain that price discovery mostly happens on NYSE and is positively correlated with its liquidity and its market share in small and big size transactions. For NASDAQ-listed stocks, large quantities trades do not convey information and NASDAQ contribution to price discovery increases slightly the days with macroeconomic announcements.

Keywords: *Price discovery, Information Share, Permanent-Transitory component, Microstructure noise, Realized Variance*

JEL: *C32, C58, G14*

1.1 Introduction

The institutional evolutions of financial markets and the development of High-frequency Trading generated a growing literature on the resulting consequences on market's outcomes. The multiplication of trading platforms coupled with the internationalization of financial markets resulted in some assets being listed simultaneously in many town or even many countries. Similarly traders can send orders in remotely located market places. The trading prices for a given security on those interrelated markets are strongly linked by arbitrage activities. A similar situation arises for one security and its derivatives: The spot prices are related to futures prices, the CDS prices are related to the credit spread.

The price discovery mechanism is generally understood as the process by which information is computed into prices, it is interesting in the multiple markets setup to understand how each market does so. An international investor for instance, choosing how to split the orders in different markets, might find it worthy to know where the price is close to the fundamental. The regulator also, in its quest to the best market organization, is interested in which market contributes to the price movement of an asset and for which reasons[1]. This quest of the market with the "best" information processing mechanism goes back to Garbade and Silber (1983) problem: which market is dominant and which market is satellite?

To determine in which market price discovery happens some tools (known as price discovery measures) are developed in the literature. Hasbrouck (1995) pioneer paper, using the Beveridge-Nelson permanent component, presented a measure called the Information Share (IS) and provided comparison of NYSE and regional exchanges in the quotes formation of Dow stocks. The main competing measure to Hasbrouck (1995) is the PT measure in Harris et al. (2002b), consisting of the common factor weight in the permanent-transitory (PT) decomposition of Gonzalo and Granger (1995). Those measures are intensively debated by De Jong (2002), Lehmann (2002), Hasbrouck (2002),Baillie et al. (2002), and Yan and Zivot (2010). One conclusion of the debate is that the IS accounts more for the variability in the price discovery process and the permanent (efficient) price identified by Hasbrouck (1995) has an economic relevance[2].

The other part of the debate lies in their view of price discovery. Hasbrouck (1995) sees it as *"who moves first"* in the process of price adjustment and Harris et al. (2002b) as the process by which security markets attempt to identify permanent changes in equilibrium transaction prices. Meanwhile, what their proposed measures actually capture is unclear. And as stated by Lehmann (2002), a market should dominate the price discovery if it is the best in incorporating information

[1] Eun and Sabherwal (2003) report that the Canadian authority was really worried about US-markets becoming the place where the Canadian's stock prices were computed

[2] The PT relies on a permanent price that is not a random walk

in a "timely and efficient" manner. This widely accepted characterization of market dominance presents two dimensions. The first dimension is the timing: a market reflecting quickly new information is close to efficiency. The second dimension is the avoidance of noises. A market with less noises is also efficiently incorporating information. The noises can come from uninformative sources as bid-ask bounce, price discreteness, and measurement errors. It is then not very clear which dimension is actually captured by the existing measures. For example, using Monte Carlo exercises, Putniņš (2013) obtains that IS and PT are actually assessing how markets avoid noises. Whereas Yan and Zivot (2010) obtains in a specific structural VECM that the PT assess how markets avoid noises while the IS captures both dimensions.

This study innovates in exposing new facts on price discovery measures, particularly linked to the utilization of High-frequency data. Using those data bring issues that are studied in the literature for volatility estimation in the presence of microstructure noises (see Andersen et al., 2000; Zhang, 2010; Jacod et al., 2009). I show that IS and PT are not related to the fundamental value but rather to information-uncorrelated noises. This could lead to misleading interpretations in applications. I also contribute to the literature by proposing new measures that are robust-to-noise and restore a clear interpretation of what is being measured: My robust IS (ISR) and robust PT (PTR) measures are good at detecting which market incorporates *quickly* new information. My framework incidentally provides values to compare the pure noise in the markets.

If both "speed" and "noise-avoidance" dimensions of price discovery are relevant and meaningful, confusions might come in utilization of price discovery measures as their nature can change given the frequency of data at hand. The analysis of price discovery should disentangle the previous two dimensions for the following reasons:

First, the way most papers consider a market to be informationally dominant is that, once new information is available, the price of the asset on this market is the first to reflect it. But this market might be more affected by information uncorrelated-noise, if it has a different tick size for example. It is thus unclear which effect will dominate in the measure or which market reveals more about the fundamental value. Let's take the extreme case where one market's price equals the efficient price plus a noise with infinite variance, and another market's price is the one-period lagged efficient price. The latter market is clearly more informative about the efficient price even if the first market is the fastest. It thus appears that another source of confusion about what the measures will do is the size of the noise in the data. On this matter, I provide analytical insights on how price discovery measures are related to microstructure noises and the sampling frequency.

Secondly, Hasbrouck (1995) defines its price discovery measure as the contribution of a market's innovation to the variance of the innovation in the efficient price. He then suggests that his Information Share is good at detecting which market moves first. This statement is somewhat giving more importance to the fact that a market is the first to incorporate information. In addition, the IS has an identification problem and is only able to produce bounds[3]. Sometimes, bounds can be wide making the IS useless. Hasbrouck (1995) recommended to sample at High-frequency to reduce the correlation and tighten the IS bounds, but this practice ignores that at High-frequency non informative part of the noise dominates the variances estimation[4]. Meanwhile in application, Chakravarty et al. (2004) use IS and are interested in the timing sequence when they justify their contribution to the literature by stating: "there is surprisingly little evidence that new information is reflected in option prices *before* stock prices". My paper emphasizes that at high frequency the IS is not related to the efficient price and rather measures which market avoids noise.

Lastly, an endogeneity problem could arise in a number of applications. The values provided by price discovery are used as dependent variables in regression to investigate the determinants of a market's dominance. Chakravarty et al. (2004) use IS to show that 17% of informed trading happens in the options markets, and that price discovery across strike price is determined by relative spread, leverage, and volume. Huang (2002) uses the IS to compare who has the most timely and informative quote, between Electronics communication Networks (ECNs) and Nasdaq; they find that measures of market liquidity do not necessarily explain the market maker's contribution to price innovation. Barclay et al. (2003) study the impact of trading costs variables on the Information Share of ECNs. Eun and Sabherwal (2003) regress the PT coefficients on the relative spread, volume, listing age, and market Cap, to explain the contribution of Toronto Stock Exchange (TSE) and U.S. exchanges to price discovery of cross-listed Canadian stocks. As an example, in Chakravarty et al. (2004), the price discovery of the option market, measured by the IS, tends to be greater when the effective bid-ask spread is narrow relative to the stock market. If by definition the IS were to fully capture the bid-ask spread noise, then there is full endogeneity in their regression of the IS on the bid-ask spread. By disentangling the two aspects of price discovery, my proposed robust measures can be used to avoid the endogeneity issue.

In the application, using data of NYSE TAQ database, I examine if my conclusions are in line with the data. I observe that indeed the data seem to present the patterns I highlighted, but the frequency of the transactions might not be high enough to show certain features. As quotes data are more frequent, I do the same analysis with mid-quotes of some assets and it confirms my theoretical conclusions. I then investigate the relative contribution of NYSE and NASDAQ to the price formation of Dow Jones assets. The robust IS measure performs well as it has very close bounds, when the standard IS bounds are wide and thus unusable. Descriptively, NYSE captures the big part of volume traded but NASDAQ is the most liquid with a high level of activity. This implies

[3] it is based on the Cholesky decomposition of variance matrix and is thus dependent of variables ordering.

[4] This is related to the signature plot of Andersen et al. (2000)

that NASDAQ mostly runs the orders of small quantities while NYSE runs big quantities orders. In terms of contribution to price discovery for the assets under investigation, NYSE is generally dominant. The contribution of a market appears to be positively correlated with its liquidity. I also analyze the correlation between market's contribution and markets share in each category of trade size. It reveals that the contribution of a market is correlated with its share in small and medium size transactions. For NASDAQ listed stocks, there is no correlation with market's share in big size transactions, so large quantities trades do not convey information.

The remainder of the paper is organized as follow. The second section reviews the main existing measures of price discovery. The third section presents some structural microstructure models and the price discovery measures are analytically computed at High-frequency. In the fourth section I propose the robust-to-noise measures and present their performances in some simulation exercises. In the fifth section, an application is done on assets of the Dow Jones that are listed and traded on NYSE and NASDAQ on the period March 1st to May 30th 2011.

1.2 Measuring price discovery

Constructing a price discovery measure would normally require that the object of interest be clearly identified. There is a current and permanent discussion in this respect with existing measures. This originates from the fact that they are defined on a reduced form model and not in a structural model. The approach to build prices discovery measures is to extract a common unobserved permanent price from the observed prices, and to attribute its characteristics to each market.

Let's consider an asset traded on markets 1 and 2 at the respective prices p_{1t} and p_{2t}[5]. This is done via the VECM representation of the cointegrated price vector $p_t = (p_{1t}, p_{2t})'$. The gap between the two prices $(p_{1t} - p_{2t})$ is stationary such that there exists only one common trend for the prices. In fact, because the prices in the two markets are from the same asset, a gap between them can not remain infinitely as there will be room for profits by arbitrage (for example buying continuously in the first market and selling in the second). Under the previous notations and restrictions implied by arbitrage, Johansen (1991) results imply that the price vector admits the following Vector Error Correction Model (VECM):

$$\Delta p_t = -\alpha\beta' p_{t-1} + \Gamma_1 \Delta p_{t-1} + \ldots + \Gamma_K \Delta p_{t-K} + e_t, \tag{1.1}$$

where the cointegrating matrix is $\beta' = \begin{pmatrix} 1 & -1 \end{pmatrix}$ as $\beta' p_t = p_{1t} - p_{2t}$ is stationary. e_t is an independent white noise with $var(e_t) = \Omega$.

[5]The results are easily obtained for more than 2 markets

The Granger representation theorem gives the following transformation of 1.1 where $\Psi(L)$ is a lag polynomial:

$$p_t = p_0 + \Psi(1)\sum_{s=1}^{t} e_s + \Psi^*(L)e_t, \tag{1.2}$$

where the matrix $\Psi(1)$ is given by

$$\Psi(1) = \beta_\perp \left(\alpha_\perp^{'} \left(I - \sum_{i=1}^{p} \Gamma_i\right) \beta_\perp\right)^{-1} \alpha_\perp^{'}. \tag{1.3}$$

The representation 1.2 entails a decomposition of the prices in a stationary component $p_0 + \Psi^*(L)e_t$ and a permanent component $\Psi(1)\sum_{s=1}^{t} e_s$. The matrix $\Psi(1)$ summarizes the long run impact of the innovation e_t on prices p_t .

1.2.1 The Information Share measure

Hasbrouck (1995) looks for a measure that will determine on which market the price discovery does happen. He proposes to use the contribution of each market to the variance of the innovation of the "efficient price" price.

As $\beta' = \begin{pmatrix} 1 & -1 \end{pmatrix}$, its orthogonal $\beta_\perp^{'} = \begin{pmatrix} 1 & 1 \end{pmatrix}$ and the formula 1.3can be written as

$$\Psi(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \psi = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix}.$$

The 2×1 row ψ replaced in equation 1.2 yields

$$p_t = p_0 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \psi \sum_{s=1}^{t} e_s + \Psi^*(L)e_t. \tag{1.4}$$

This representation displays a scalar random walk component of the prices $\psi\sum_{s=1}^{t} e_s$, and a stationary part $\Psi^*(L)e_t$ that might be attributed to transitory effects. The common permanent component is identified as the implicit fundamental price of the asset. Something to notice here is that e_t drives both the permanent and the transitory component. So, the construction does not distinguish the non-informative noise (due for example to tick size or measurement errors) from the information-correlated frictions that would be due to information asymmetry, market under/over reaction (Menkveld et al., 2007).

The new information entering the fundamental price is the innovation ψe_t, and its variance

$(\psi\Omega\psi')$ is the total Information Share. Hasbrouck (1995) defines the market contribution to price discovery in the following way.

If Ω is diagonal, then the total Information Share is

$$\psi\Omega\psi' = \psi_{11}^2\Omega_{11} + \psi_{22}^2\Omega_{22}$$

and the Information Share (IS) for the market j, defined as the relative contribution of this market in the variance of the new information, is obtained as:

$$IS_1 = \frac{\psi_1^2\Omega_{11}}{\psi_{11}^2\Omega_{11} + \psi_{22}^2\Omega_{22}} \quad \text{and} \quad IS_2 = \frac{\psi_2^2\Omega_{22}}{\psi_{11}^2\Omega_{11} + \psi_{22}^2\Omega_{22}}. \tag{1.5}$$

As Ω is not diagonal in general, Hasbrouck (1995) suggests using its Cholesky root to obtain a lower triangular matrix F, such that $\Omega = FF'$. An identification problem arises as the ranking of the variables matters for the Cholesky decomposition. That is the matrix F changes with the ordering of the variables in the prices vector. Thus, the Information Share measure can only provides an upper and a lower bounds.

When the market 1 is placed in the first position in p_t, then $\Omega = FF'$, and I have the bounds

$$IS_{u,1} = \frac{([\psi F]_1)^2}{\psi\Omega\psi'} \quad \text{and} \quad IS_{l,2} = \frac{([\psi F]_2)^2}{\psi\Omega\psi'}, \tag{1.6}$$

where $[\psi F]_j$ represents the jth element of the vector ψF.

Now if the market 1 is switched to the 2nd position in p_t, the new Cholesky root $\tilde{F}$ is obtained such that $\Omega = \tilde{F}\tilde{F}'$. The others bounds are

$$IS_{u,2} = \frac{\left([\psi\tilde{F}]_1\right)^2}{\psi\Omega\psi'} \quad \text{and} \quad IS_{l,1} = \frac{\left([\psi\tilde{F}]_2\right)^2}{\psi\Omega\psi'} \tag{1.7}$$

The non-uniqueness of the Information share is a problem for applications as the measure are used as dependent variable in regression. Many studies thus, simply consider the lower bound or take the mid-bounds (see Chakravarty et al., 2004; Putniņš, 2013).

The IS identification issue is related to the Macroeconomics VAR literature problem of identifying the structural shocks from the reduced form model. Relying on Hasbrouck (1995)'s efficient price, some authors tried to solve this by doing some transformations of the innovation variance matrix. The limit of those techniques is that they completely lose an economic meaning behind the mathematical operations. For example, Lien and Shrestha (2014) use an orthogonalization of the correlation matrix to propose a measure that is independent of the variables ordering. Meanwhile there is no economic intuition behind the orthogonalization of the correlation matrix. Grammig and

Peter (2013) exploit "tail dependence" for identification which is done through heteroskedasticity on two regimes as in Rigobon (2003), Lanne and Lütkepohl (2010). The drawback is that identification relies on the data and it is not always the case that they provide enough tail dependence to identify unique information share. Another limit of all the existing method based on Hasbrouck (1995) efficient price is that they lack a testing theory. This is not the case of the PT measure, which in turn, has the severe drawback that its efficient price is not a random walk.

1.2.2 The Permanent-Transitory measure

The main competitor to IS is the Gonzalo and Granger (1995) common factor weight in the Permanent-Transitory (PT) decomposition. This consists of decomposing a difference stationary time series as the sum of a permanent component Q_t and a transitory stationary component T_t. The identification of the two components of $p_t = Q_t + T_t$ relies on two assumptions:

- T_t does not Granger-cause Q_t in the long run,
- T_t is a linear combination of the observed variables.

In the context of one asset and many markets, the permanent component is driven by a difference stationary[6] factor (f_t) that is common to both markets, such that the observed prices vector can be written as

$$p_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} f_t + T_t.$$

The common factor is a linear combination of current prices $f_t = \gamma_1 p_{1t} + \gamma_2 p_{2t}$. It is easily shown that given the ECM equation 1.1, the weight $(\gamma_1 \gamma_2)$ are proportional to $\alpha_\perp$ such that:

$$f_t = c\alpha_{1\perp} \times p_{1t} + c\alpha_{2\perp} \times p_{2t}$$

with c constant.

Harris et al. (2002a) evaluate the relative contribution to price discovery of market 1 and market 2 by taking the weight of each market in the permanent component as

$$PT_1 = \frac{\alpha_{1\perp}}{\alpha_{1\perp} + \alpha_{2\perp}}, PT_2 = \frac{\alpha_{2\perp}}{\alpha_{1\perp} + \alpha_{2\perp}}.$$

The link between the permanent price extracted by Hasbrouck (1995) and the permanent price of Harris et al. (2002b) is studied by De Jong (2002). A difference between the PT measure with the

[6] Or integrated of order 1 denoted $I(1)$

IS measure is that f_t is a linear combination of only the current prices. Thus the permanent component of the Gonzalo and Granger (1995) decomposition is generally not a random walk. This is a serious limitation as this permanent component could not represent an efficient price and only gets an economic meaning in a structural model (see Lehmann, 2002). Baillie et al. (2002) show that IS and PT can be computed easily after the estimation of the VECM and they present the relationship linking PT to IS. In the case of a diagonal Ω, the PT squared coefficients are weighted by the innovations variances to obtain IS. This is seen by deducing $\begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix} = c \begin{pmatrix} \alpha_{1\perp} & \alpha_{2\perp} \end{pmatrix}$ from formula 1.3 and replacing for example in formulas 1.5 to obtain $IS_1 = \left(\alpha_{1\perp}^2 \Omega_{11}\right) / \left(\alpha_{1\perp}^2 \Omega_{11} + \alpha_{2\perp}^2 \Omega_{22}\right)$.

Instead of focusing on the innovation variance, the permanent component share relies on the error correction weighting matrix $\alpha_\perp$. In this respect Eun and Sabherwal (2003) also think of price discovery as the adjustment to the equilibrium and access it by the coefficient α summarizing how a market corrects a departure from the other market price. Building the measures with only a coefficient of the VECM allows those methods to have testable implications and thus test of statistical significance can be performed.

1.3 Microstructure models and sampling frequency

The price discovery measures presented in Section 2 are used in the literature to detect which model is likely to have generated the observed log prices $p_t \equiv (p_{1t}, p_{2t})'$. Are the two markets structurally identical? Is one market leading the information while the other is lagged? To compare the performances of the measures in answering those questions, literature relies on some structural microstructure models (see Hasbrouck, 2002; Harris et al., 2002a) representing the different situations that might arise on market . I rewrite versions of those models to make them dependent of the sampling interval h and a delay parameters δ. For those models, Δp_t generally admits a Vector Moving-Average of order 1 ($VMA(1)$) representation, allowing to compute analytically the values of the prices discovery metrics. The VMA(1) equation is

$$\Delta p_{th} = e_{th} + \Theta e_{th-h} \text{ with } \Theta = \begin{pmatrix} -1+c & 1+d \\ c & d \end{pmatrix}, \tag{1.8}$$

where e_{th} is the white noise innovation with variance $\Omega = var\left(e_{th}\right) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}$.

The long run impact matrix is thus

$$\Psi(1) = I + \Theta = \begin{pmatrix} c & 1+d \\ c & 1+d \end{pmatrix} = \psi \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ with } \psi = \begin{pmatrix} c & 1+d \end{pmatrix} \tag{1.9}$$

To compute the measures, one needs the values of the parameters Ω, c, d in terms of the structural parameters in p_t. For this, the values of the structural variance and autocovariance are matched with the ones of the VMA(1) equation 1.8. That is

$$\begin{aligned} C_0 = var(\Delta p_{th}) &= \Omega + \Theta\Omega\Theta' \\ C_1 = cov(\Delta p_{th}, \Delta p_{th-h}) &= \Theta\Omega \end{aligned} \tag{1.10}$$

Computing ΘC_0 and replacing $\Theta\Omega$ by C_1 gives the equation 1.11

$$C_1 - \Theta C_0 + \Theta C_1 \Theta' = 0 \tag{1.11}$$

For each of the model I will present, I computed Θ by solving this matrix equation via long and tedious calculations given in appendix, and then Ω is obtained as $\Omega = \Theta^{-1} C_1$. Next, I present the structural models of interest and study the behavior of the price discovery measures.

1.3.1 Model I: A two-market "Roll" model.

In model I, both markets incorporate the efficient price m_t. This situation could arise from markets with no private information, and an efficient price driven by public non-traded information. At the sampling interval h, the latent fundamental log price of the asset is

$$m_{th} = m_{th-h} + \eta_{th}$$

The innovation is $\eta_{th} = \sigma_h \mathcal{N}(0,1)$ and its variance $\sigma^2(h)$ converges to zero when h goes to zero. This is not a limitation as empirically the returns and their variance become very small at high frequency. It can also be viewed in the discretization of the often-used continuous time model $dm_t = \sigma dB_t$, implying $\sigma_h = \sigma\sqrt{h}$. The observed prices are contaminated by i.i.d non correlated microstructure noises

$$\begin{aligned} p_{1th} &= m_{th} + c_1 \varepsilon_{1th} \\ p_{2th} &= m_{th} + c_2 \varepsilon_{2th} \end{aligned} \tag{1.12}$$

$\varepsilon_{1t}, \varepsilon_{2t}, \sim \mathcal{N}(0,1)$ with $E(\varepsilon_{1t}\varepsilon_{2t}) = E(\eta_{th}\varepsilon_{1t}) = E(\eta_{th}\varepsilon_{2t}) = 0$. The constants c_1, c_2 represent the variances of the noise components. They could be made dependent of h and going to zero but less faster than σ_h. This will not change the main conclusions as all the facts I describe remain qualitatively the same.

In this setup, there is no market dominating the price discovery process considered as the predominance in incorporating the new information η_{th}.

I compute the variance and covariance 1.10 and obtain

$$C_0 = \begin{pmatrix} \sigma_h^2 + 2c_1^2 & \sigma_h^2 \\ \sigma_h^2 & \sigma_h^2 + 2c_2^2 \end{pmatrix} \text{ and } C_1 = \begin{pmatrix} -c_1^2 & 0 \\ 0 & -c_2^2 \end{pmatrix}.$$

Using C_0 and C_1 to solve equation 1.11, the values of ψ and Ω that are necessary to compute the different measures, are obtained in terms of the structural parameters σ_h^2, c_1^2, c_2^2.

Lemma 1.1. *In model I, solving equations 1.44 gives*

$$\psi = \kappa \begin{pmatrix} c_1^{-2} & c_2^{-2} \end{pmatrix} \quad \textit{and} \quad \Omega = K \begin{pmatrix} c_1^2\left(1 - c_2^{-2}\kappa\right) & \kappa \\ \kappa & c_2^2\left(1 - c_1^{-2}\kappa\right) \end{pmatrix} \tag{1.13}$$

$$\text{With} \quad \kappa = -\frac{1}{2}h\sigma^2 + \sqrt{h}\sigma\frac{1}{2}\sqrt{h\sigma^2 + \frac{4c_1^2c_2^2}{(c_1^2+c_2^2)}}, \tag{1.14}$$

$$K = \left[1 - \kappa\left(c_1^{-2} + c_2^{-2}\right)\right]^{-1}. \tag{1.15}$$

Proposition 1.2. *The PT measure*

Using the results in Lemma 1.1, the PT gives

$$PT_1 = \frac{c_1^{-2}}{c_1^{-2} + c_2^{-2}} \textit{ and } PT_2 = \frac{c_2^{-2}}{c_1^{-2} + c_2^{-2}}. \tag{1.16}$$

Proof. See Appendix □

The PT does not assess the priority to incorporate m_t but is completely dependent of noises. The contribution of a market is inversely proportional to its own noise. That is, the market with the lowest noise has the biggest contribution, and the PT is measuring the avoidance of noises at any frequency. It is only when the level of noise is the same in the two markets, that the measure can be coherently interpreted in term of fundamental information with an equal value for each market.

Proposition 1.3. *The IS measure,*

Using the results in Lemma 1.1, The IS bounds for market 1 and for market 2 are computed as

$$IS_{u,1} = \frac{c_1^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)} \text{ and } IS_{l,1} = \frac{c_1^{-2}K^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)}, \tag{1.17}$$

$$IS_{u,2} = \frac{c_2^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)} \text{ and } IS_{l,2} = \frac{c_2^{-2}K^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}. \tag{1.18}$$

At High-frequency (when $h \simeq 0$), the parameter $\kappa \simeq 0$ and

$$IS_{u,1} \simeq IS_{l,1} \to \frac{c_1^{-2}}{c_1^{-2}+c_2^{-2}} = PT_1, \tag{1.19}$$

$$IS_{u,2} \simeq IS_{l,2} \to \frac{c_2^{-2}}{c_1^{-2}+c_2^{-2}} = PT_2. \tag{1.20}$$

Proof. *See Appendix* □

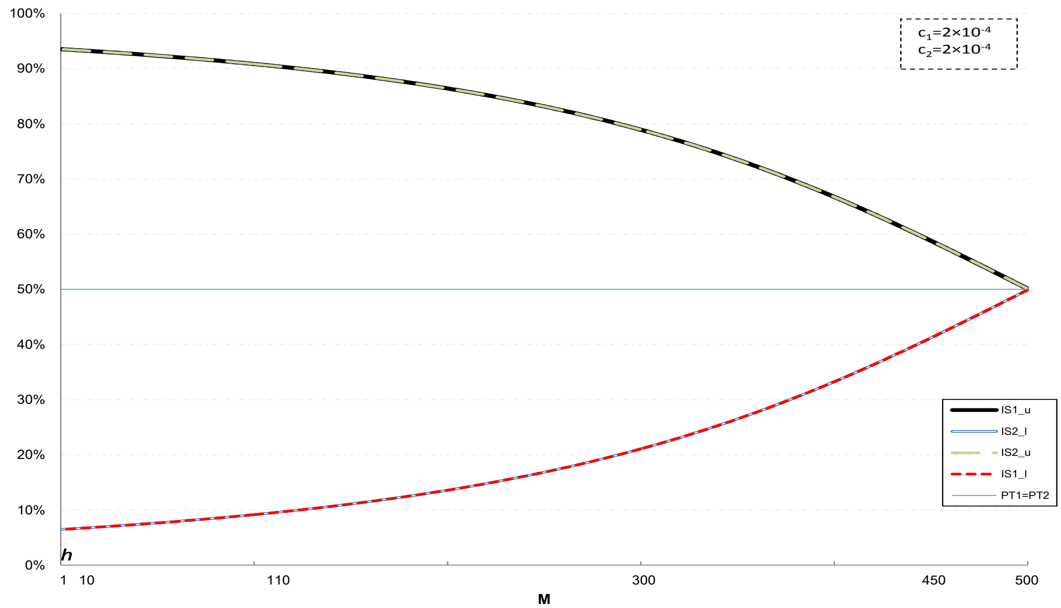
In this model at high frequency, when h is small, the bounds on the Information Share becomes tighter and close to the value of the PT measure. But the limiting values are dominated by information-uncorrelated microstructure noises and are not related to the fundamental value. This result challenges the interpretation of price discovery measure in term of the fundamental price. At high frequency, the parameter $\sigma^2$of the fundamental price disappears from the formulas and I are let with a comparison of the level of noises. So if c_1^2 is smaller than c_2^2, then $IS_1 = PT_1 > PT_2$, and ones might conclude that the Market 1 is fast to compound new information, while the market are actually equally fast. The formulas 1.19 and 1.20 could meanwhile be taken as positive result, in the sense that they provide items to compare the costs of trading in different markets for a cross listed asset.

To explore how the measures depend on the noise and the frequency I plot the IS and PT as a function of $M = 1/h$. In Figure 1.1 with equal level of noise the bounds are wide at lower frequency but go to 50% when the frequency ($M = 1/h$) increases. When the level of noise is different (Figure 1.2) the bounds are reduced but then the market with the smallest noise becomes dominant.

1.3.1.1 Time Varying noises

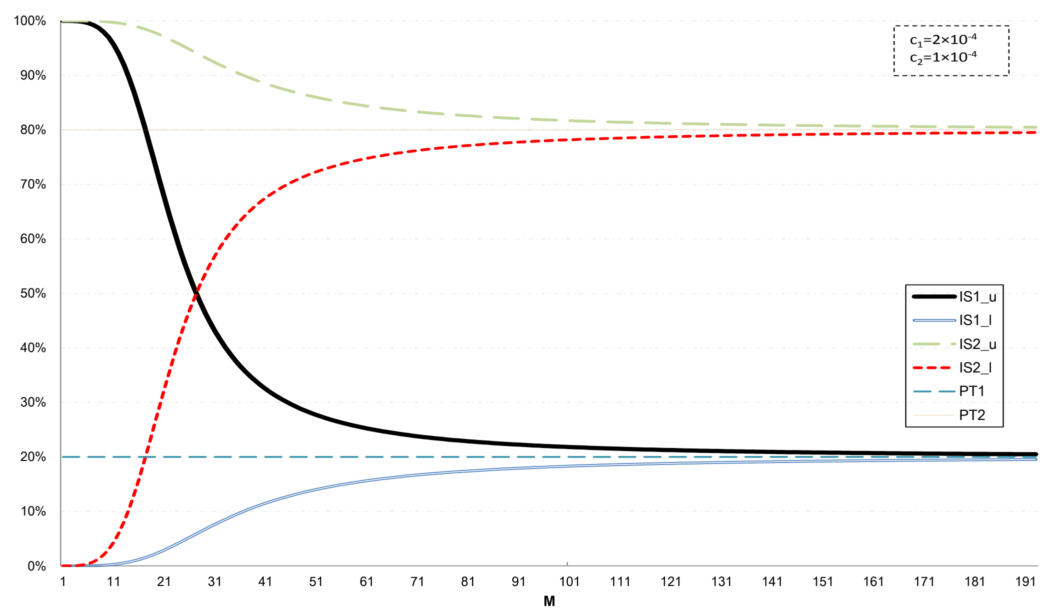
The previous results are derived under constant noises variances. The following theorem show that all the conclusions remain when the noises variance vary with the sampling frequency as long as

Figure 1.1: Model 1: Equal noises $c_1^2 = c_2^2$



Note: The figures plot the IS and the PT measures computed analytical on model I. The horizontal axis represents the sampling frequency $M = 1/h$.

Figure 1.2: Model I: Different noises $c_1^2 = 2c_2^2$



Note: The figures plot the IS and the PT measures computed analytical on model I. The horizontal axis represents the sampling frequency $M = 1/h$.

the fundamental return decreases less faster than h.

Proposition 1.4. *Time Varying noises: Let* $c_1^2 \equiv c_1'^2 h^{\alpha_1}$, $c_2^2 \equiv c_2'^2 h^{\alpha_2}$ *with* $\alpha_1, \alpha_2 > 0$

If $Max\left(\alpha_1, \alpha_2\right) < 1$, *then*

$$\begin{cases} \kappa = & -\frac{1}{2}h\sigma^2 + \sqrt{h}\frac{\sigma}{2}\sqrt{\sigma^2 h + \frac{4c_1'^2 c_2'^2 h^{\alpha_1} h^{\alpha_2}}{\left(c_1'^2 h^{\alpha_1} + c_2'^2 h^{\alpha_2}\right)}} \longrightarrow 0 \\ K = & \left[1 - \kappa\left(c_1'^{-2} h^{-\alpha_1} + c_2'^{-2} h^{-\alpha_2}\right)\right]^{-1} \longrightarrow 1 \end{cases}$$

And

$$IS_{u,i} \simeq IS_{l,i} \xrightarrow{h\to 0} PT_i = \frac{c_i'^{-2}}{c_1'^{-2} + c_2'^{-2}}, \; i = 1,2$$

Proof. See appendix ☐

1.3.2 Model II: The Roll model with a delayed market

The fundamental log price of the asset is still driven by the innovation $\eta_{th} = \sigma_h \mathcal{N}(0,1)$ with $\sigma(h) = \sigma\sqrt{h}$, and

$$m_{th} = m_{th-h} + \eta_{th}.$$

The first market incorporates m_t, but the second market is delayed of δ. The observe prices are

$$\begin{aligned} p_{1th} &= m_{th} + c_1 \varepsilon_{1th} \\ p_{2th} &= m_{th-\delta} + c_2 \varepsilon_{2th}. \end{aligned} \tag{1.21}$$

I compute the variance and covariance 1.10 as

$$C_0 = \begin{pmatrix} h\sigma^2 + 2c_1^2 & (h-\delta)\sigma^2 \\ (h-\delta)\sigma^2 & h\sigma^2 + 2c_2^2 \end{pmatrix} \text{ and } C_1 = \begin{pmatrix} -c_1^2 & 0 \\ \delta\sigma^2 & -c_2^2 \end{pmatrix}.$$

When $h > \delta$ the price admits a VMA(1) representation and I calculate the analytical solutions by solving the matrix equation 1.11

Proposition 1.5. *The PT measure*

In Model II, solving equations 1.44 gives

$$\begin{aligned} PT_1 &= \frac{\left(-\frac{1}{2}\frac{\sigma_l^2\left[\left(\delta\sigma_l^2+c_2^2\right)\left(\sigma_l^2\delta(h-\delta)+h\left(c_1^2+c_2^2\right)\right)+2c_1^2c_2^2\delta\right]}{c_2^4c_1^2+\left(\delta\sigma_l^2+c_1^2\right)^2c_2^2} \pm \frac{1}{2}\sqrt{\Delta}\right)}{\left(-\frac{1}{2}\frac{\sigma_l^2\left[\left(\delta\sigma_l^2+c_2^2\right)\left(\sigma_l^2\delta(h-\delta)+h\left(c_1^2+c_2^2\right)\right)+2c_1^2c_2^2\delta\right]}{c_2^4c_1^2+\left(\delta\sigma_l^2+c_1^2\right)^2c_2^2} \pm \frac{1}{2}\sqrt{\Delta}\right)\left(1+\frac{\delta\sigma_l^2+c_1^2}{c_2^2}\right) - \frac{\delta\sigma_l^2+c_2^2}{c_2^2}} \\ PT_2 &= 1 - PT_1 \end{aligned}$$

where

$$\Delta = \left[\frac{\sigma_l^2\left[\left(\delta\sigma_l^2+c_2^2\right)\left(\sigma_l^2\delta\left(h-\delta\right)+h\left(c_1^2+c_2^2\right)\right)+2c_1^2c_2^2\delta\right]}{c_2^4c_1^2+\left(\delta\sigma_l^2+c_1^2\right)^2c_2^2}\right]^2+4\frac{\sigma_l^2\left[\sigma_l^2\delta\left(h-\delta\right)+hc_2^2\right]}{c_2^2c_1^2+\left(\delta\sigma_l^2+c_1^2\right)^2}$$

Proof. See appendix □

The formula for the IS, which is also very cumbersome, is computed after obtaining Ω by the equation 1.36. The formulas are not really intuitive, but it displays the fact that PT depend on the information parameter σ, on the frequency parameter h, and on the delay δ. Here, the limit when h is small can not be easily obtained analytically. In fact, by computing the autocavariance for the process for $h < \delta$, the order of the VMA becomes bigger than 1 and increases when h decreases. I will rely on graphical analysis for more insights.

The behavior of the measures in Model II is summarized in Figures 1.3 and 1.4. The Panel A of the graph corresponds to $h > \delta$ is plotted using the analytical formulas. The Panel B is plotted for all h using simulations. In this setup by construction, the first market dominates structurally the price discovery mechanism as it is the first to compute new information. When the level of noise are equals (Figure 1.3), the measures succeed in designing market 1 as dominant when $h \geq \delta$. But at high frequency with $h < \delta$, the measures converges to 0.5, stating that the two markets are equally contributing to the price discovery mechanism. When the market 1 is noisier than market 2 (Figure 1.4), the measures in both panels seem to converge to values such that market 2 is dominant. Theses results simply reflect the relative size of noise in market 1, compared to noise in market 2.

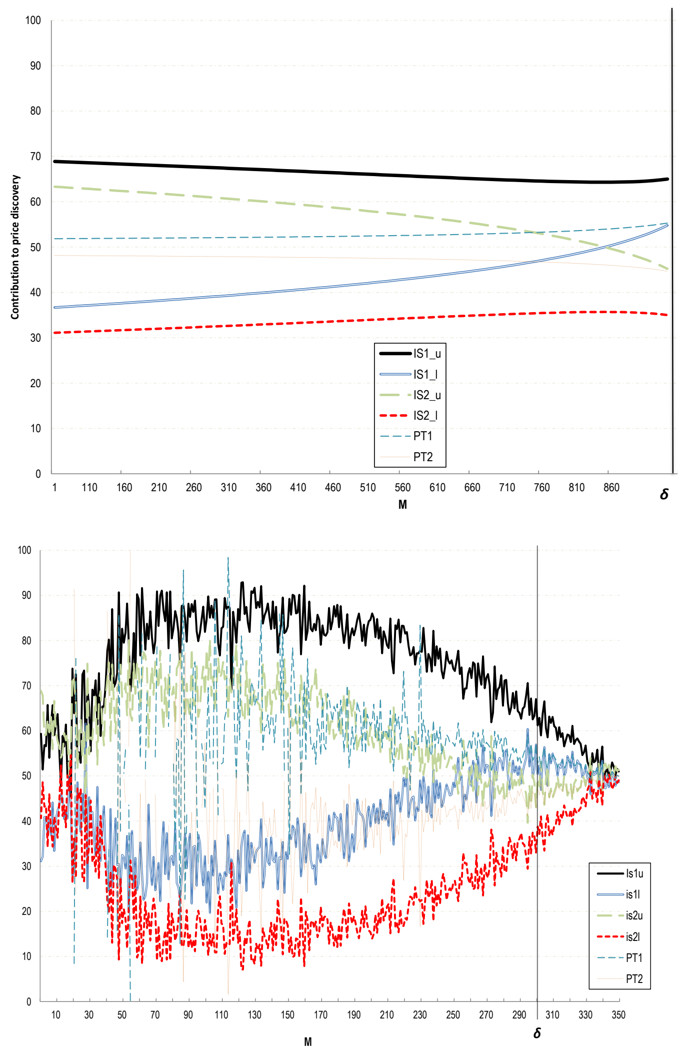
Note: The figure plots the IS and PT model II. The horizontal axis represents the sampling frequency $M = 1/h$. Panel A and Panel B are separated at the point where $h < \delta$. $c_2^2 = 0.002/2$.

1.3.3 Model III: A Two-market model with public and private information

In this model presented by Hasbrouck (2002), the efficient price is driven by informative trading on the market 1 (η_{1th}) and a non-traded public information $\eta_{th} = \sigma_h\mathcal{N}(0,1)$. The dynamic of price is described by the following system

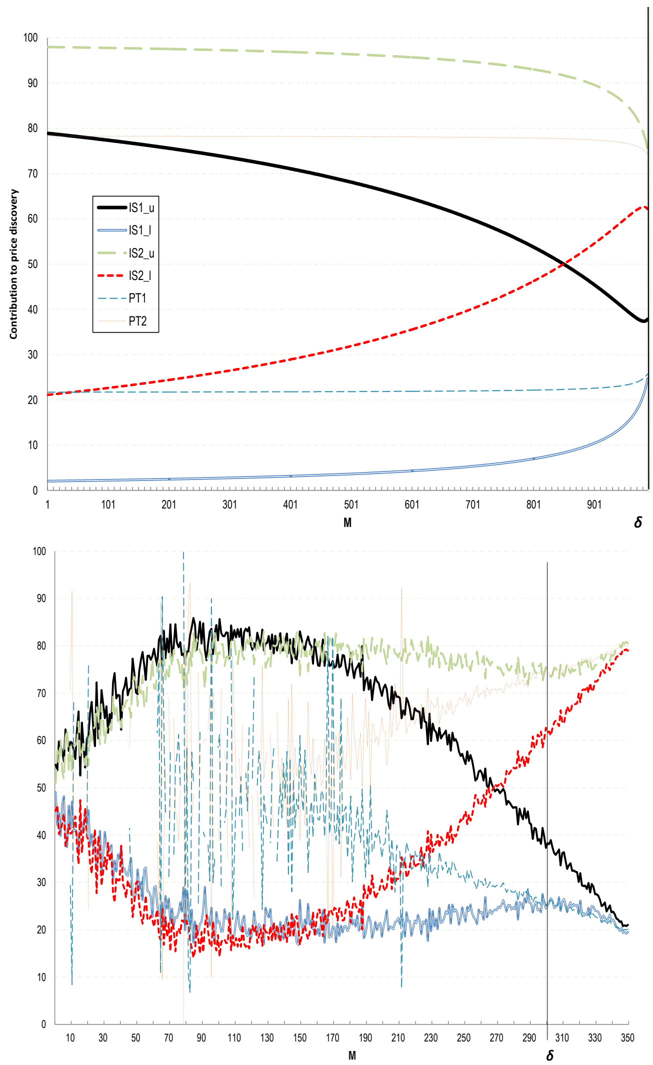
$$\begin{aligned} m_t &= m_{t-h} + \lambda_h\eta_{1th} + \eta_{th} \qquad (1.22)\\ p_{1th} &= m_{th} + \eta_{1th} + c_1\varepsilon_{1th} \\ p_{2th} &= m_{th-h} + c_2\varepsilon_{2th} \end{aligned}$$

Figure 1.3: Model II: Equal noises



Panel A Panel B

Figure 1.4: Model II: Different noises



Panel A Panel B

where λ_h, the liquidity parameter, goes to zero with h for the same reasons as σ_h in the previous sections. The Market 2 relies on a delayed value (with lag h) of m_t. The parts of microstructure noises that are information-uncorrelated are ε_{1th} and ε_{2th}.

As before, Market 1 is dominant from the structural point of view.

I solve for the equation 1.11 and the results at the order of $\sqrt{\lambda_h}$ are :

Proposition 1.6. *The PT Share*

In Model III, the solutions of equations 1.44 at the order of $\sqrt{\lambda_h}$ gives

$$PT_1 = \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} + o\left(\sqrt{\lambda_h}\right) \quad \text{and} \quad PT_2 = \frac{c_2^{-2}\left(1+c_1^2\right)}{1+c_2^{-2}\left(1+c_1^2\right)} + o\left(\sqrt{\lambda_h}\right)$$

Proof. See appendix ☐

Proposition 1.7. *The IS bounds*

In Model III, the solutions of equations 1.44 at the order of $\sqrt{\lambda_h}$ gives

$$\begin{aligned}
IS_{u,1} &= -\frac{1}{c_2^{-2}\left(1+c_1^2\right)D-1} \times \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} && +o\left(\sqrt{\lambda_h}\right)\\
IS_{l,1} &= \frac{D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1}{D-1} \times \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} && +o\left(\sqrt{\lambda_h}\right)\\
IS_{u,2} &= \frac{-\left(1+c_1^2\right)c_2^{-2}}{(-1+D)} \times \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} && +o\left(\sqrt{\lambda_h}\right)\\
IS_{l,2} &= \frac{D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1}{Dc_2^{-2}\left(1+c_1^2\right)-1} \times \frac{c_2^{-2}\left(1+c_1^2\right)}{1+c_2^{-2}\left(1+c_1^2\right)} && +o\left(\sqrt{\lambda_h}\right)
\end{aligned}$$

with $D = \sqrt{\lambda_h}\left(\sqrt{\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}\right)^{-1} \xrightarrow{h\to 0}.$

When $h \simeq 0$*,*

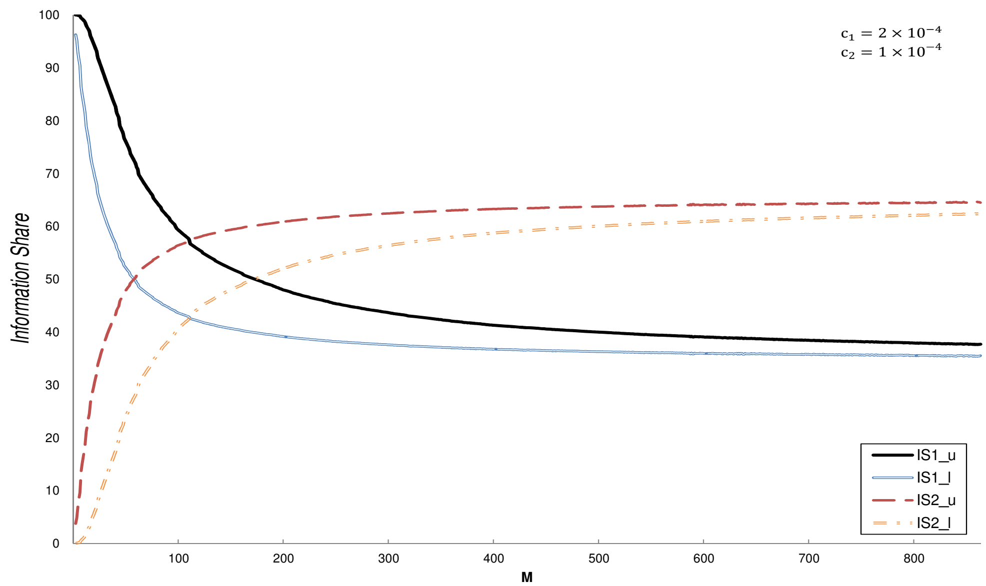
$$\begin{aligned}
IS_{u,1} \simeq IS_{l,1} \simeq PT_1 &\longrightarrow \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} = \frac{c_1^{-2}}{c_1^2c_2^{-2}+c_1^{-2}+c_2^{-2}}\\
IS_{u,2} \simeq IS_{l,2} \simeq PT_2 &\longrightarrow \frac{\left(1+c_1^2\right)c_2^{-2}}{1+c_2^{-2}\left(1+c_1^2\right)} = \frac{c_2^{-2}\left(c_1^2+1\right)}{c_1^2c_2^{-2}+c_1^{-2}+c_2^{-2}}
\end{aligned}$$

Proof. See appendix ☐

When h is small in this setup, IS and PT give the same value. The contribution of market 2 decreases with the noise variance in market 2, and increases with the noise variance in market 1. When the level of noise is equal in the two markets, $PT_2 > PT_1$ and market 2 is chosen as the dominant market, which is in opposition with the structural model. In Figure 1.5, I compare the measures for the model III computed numerically for h decreasing. Even if the model is changing by reducing the delay parameter $\delta = h$, market 1 remains dominant as it drives the efficient price.

At lower frequency, the IS of market 1 is almost 100% and the IS of market 2 is close to 0, even if market 1 is the noisiest. When the values of h is small, the contribution of market 2 is bigger than that of market 1, suggesting falsely that market 2 is dominant. The issues highlighted here are less important with small noises variances or with small noises difference between the two markets. The frequency at which the dominance commutes increases (see Figures 1.7a and 1.7b in appendix).

Figure 1.5: Model III: IS with sampling frequency and delay



Note: The figure plots the IS model III. The horizontal axis represents $M = 1/h$. The PT (not plotted here) has the same pattern

Remark: By fixing $\delta = 0$ and $\eta_{th} = 0.$, I obtain $m_t = m_{t-h} + \lambda_h \eta_{1th}$ and

$$\begin{aligned} p_{1th} &= m_{th} + \eta_{1th} + c_1 \varepsilon_{1th} \\ p_{2th} &= m_{th} + c_2 \varepsilon_{2th} \end{aligned} \tag{1.23}$$

corresponding to a Two-market model with overreaction. It is not very clear which market dominates the price discovery in this setup. The market 1 incorporates m_t but there is an overreaction to information in the observed prices. The market 2 also incorporates timely m_t. The computation using the formula 1.32 in appendix gives directly PT as

$$PT_1 = \frac{c_1^{-2}}{c_1^2 c_2^{-2} (\lambda_h \sigma^2 + 1) + c_1^{-2} + c_2^{-2}} \quad \text{and} \quad PT_2 = \frac{c_2^{-2} \left(c_1^2 \left(\lambda_h \sigma^2 + 1\right) + 1\right)}{c_1^2 c_2^{-2} (\lambda_h \sigma^2 + 1) + c_1^{-2} + c_2^{-2}}$$

I still see that the measures vary inversely proportional to noises. But the contribution of market 2 increases with the variance of the efficient price. For equal level of information uncorrelated-noises, market 2 has a greater contribution than market 1. So the PT captures the "efficient" facet of prices.

All the new facts just explained here are warnings about the interpretations when using existing price discovery measures on High-Frequency data. It is thus of interest to develop a price discovery measure that is adapted in high frequency data and clarify what aspect of the market is actually captured.

1.4 Robust-to-Noise price discovery measures

The different analytical computations showed that at High-frequency the price discovery measures are dominated by noises. In this sense, the measures seem to be better interpreted in terms of noises avoidance. This point is also made by Yan and Zivot (2010) who suggest combining IS and PT in one measure to reduce the noises effects. The issue here is related to the debate in the literature about the property that those price discovery measures are actually capturing. My results suggest that at lower frequency they might be capturing the speed at which markets incorporate information while at a high frequency they are capturing which market is less noisy. This creates a misleading interpretation caused only by the frequency of observations. To restore a consistency in the definition of the measures at all frequency, I propose a correction of the measures to reduce the effect of noises. For this, a look at the different formulas suggests that the measures are dominated by a factor equal to the inverse variance of the market microstructure noises. I thus propose to robustify the IS and the PT by multiplying them by the noise variance.

So the bounds on IS and the PT for the market 1 are multiplied by c_1^2, and the bounds on IS and PT for the market 2 are multiplied by c_2^2 . I re-normalize the robust to-noise versions of the measures to keep the sum to one:

$$ISR_{u,1} = \frac{c_1^2 IS_{u,1}}{c_1^2 IS_{u,1} + c_2^2 IS_{l,2}} \quad \text{and} \quad ISR_{l,1} = \frac{c_1^2 IS_{l,1}}{c_1^2 IS_{l,1} + c_2^2 IS_{u,2}} \tag{1.24}$$

$$ISR_{u,2} = \frac{c_2^2 IS_{u,2}}{c_1^2 IS_{l,1} + c_2^2 IS_{u,2}} \quad \text{and} \quad ISR_{l,2} = \frac{c_2^2 IS_{l,2}}{c_1^2 IS_{u,1} + c_2^2 IS_{l,2}} \tag{1.25}$$

$$PTR_1 = \frac{c_1^2 PT_1}{c_1^2 PT_1 + c_2^2 PT_2} \quad \text{and} \quad PTR_2 = \frac{c_2^2 PT_2}{c_1^2 PT_1 + c_2^2 PT_2} \tag{1.26}$$

Obviously, if Ω is diagonal, here too for each market, I have equality of its lower and upper bounds ($ISR_{u,1} = ISR_{l,1}$ and $ISR_{u,2} = ISR_{l,2}$).

To compute the previous quantities, estimations of the microstructure noise variances c_1^2 and c_2^2 are required. Fortunately, the literature on integrated volatility estimation in the presence of microstructure noises provides good ones. At High-frequency, the realized volatility (the sum of squared log return) divided by $(2n)$ is a good approximation of the noise variance (see Andersen et al., 2000; Zhang, 2010; Jacod et al., 2009). I thus consider the estimators

$$\widehat{c_1^2} = (2n)^{-1} \sum_{t=1}^{n} \Delta p_{1t}^2 \text{ and } \widehat{c_2^2} = (2n)^{-1} \sum_{t=1}^{n} \Delta p_{2t}^2$$

The properties of this estimator of the noise variance are proven in Zhang (2010). The intuition behind the results is the following. Let's Consider an observed price written as $p_{th} = m_{th} + c_0 \varepsilon_{th}$ with $\varepsilon_{th} \sim i.i.d\,\mathcal{N}(0,1)$ and $\Delta m_{th} = \eta_{th} = \sigma_h \mathcal{N}(0,1)$, $E(\eta_{th}\varepsilon_{th}) = 0$, the variance of the intraday return $\left(\sigma_h = O(\sqrt{h})\right)$ decreases with the sampling interval $h = 1/n$. The expectation of the realized volatility is

$$\begin{aligned}
E\left(\sum_{t=1}^{n} \Delta p_t^2\right) &= \sum_{t=1}^{n} E\left(\eta_{th}^2 + c_0^2 \Delta\varepsilon_{th}^2 + 2c_0 \eta_{th} \Delta\varepsilon_{th}\right) \\
&= n\sigma_h^2 + 2n \times c_0^2 \\
&= O(nh) + 2n \times c_0^2 \\
&\simeq 2n \times c_0^2
\end{aligned}$$

This development incidentally provides a way to evaluate the noise in the data. In fact, if one is to consider only how markets avoid noises, the values of c_1^2 and c_2^2 estimated previously could measure price discovery in the sense of "which market is not noisy".

1.5 Simulation

I analyze through Monte Carlo simulations the performances of the robust measures (ISR and PTR) relatively to IS and PT. For this, I simulate the structural models I, II and III (1.12,1.21, 1.22). For each model, I simulate a sample of 23 400 observations (to imitate a trading day in second), then the data are sampled at a given frequency (1, 2, 5, 10, 60) and the measures are computed in a VECM. The order of the VECM is chosen using the Akaike Information Criterion (AIC) which is typically what people do in practice. Each design is replicated 1000 times and the numbers in Tables 1.1-1.3 are the averages results and standard deviations (in parenthesis) over the 1000 replications. The gray-shaded columns of the Tables are the robust-to-noise estimates. The results are presented only for market 1.

In model I, both markets incorporate m_t, so a good estimate of price discovery in term of "where information enter the price first" should be 0.5. When the two markets have the same level of information-uncorrelated noises (Table 1.1, Panel A), all the measures perform well with values close to 1/2. When the market 1 is noisier than market 2 (see Table 1.1, Panel B), the estimated mid-bounds on IS and the PT (0.35 and 0.33) are far below 0.5 when data are sampled at High-frequency ($h = 1$). Meanwhile, my proposed robust mid-bounds ISR is 0.48 and PTR is 0.47, suggesting rightly that both markets are similar in incorporating m_t. All the metrics perform quite well at lower frequency.

In model II, the market 2 is slow and incorporates new information with a lag $\delta = 3s$, and market 1 noise's variance is set to $c_1^2 \equiv 0.0002$, bigger than $c_2^2 \equiv 0.0001$ of market 2. Price discovery happens in market 1, but the small values in Table 1.2, obtained for IS and PT falsely suggests that it happens in market 2. By using the robust measures the good interpretation is restored with values for ISR and PTR close to 0.89 and 0.76. The effect is more pronounced in Table 1.3 where the first market drives the fundamental price. While the other measures suggest an equal role for both markets in the price discovery process, the robust-to-noise measures are almost 0.99. The estimated values presented here depend on the size of the noise and on the sampling frequency. The performances of IS and PT are improved when the noise is reduced or when the difference in noises between the two markets diminishes. But the qualitative result remains unchanged: the robust measures are better than IS and PT to detect which market incorporates timely new information.

1.6 Empirical application

I study the daily relative part in the price discovery of assets of the Dow Jones Industrial Index that are listed and traded on NYSE and NASDAQ. I focus on the trade prices coming from the TAQ Database and covering the period from the 01 March to the 30 May 2011. Before using the data a cleaning job is done on the raw data: First, I suppress the data stamped before the opening (9h30) and after the closing (16h00) of the market. I also remove the data between 9h35 because the activity at the opening session creates a lot large values with respect to the daily continuous activity I aim to study. Second, to handle the synchronicity problem, I fill the data with the last trade price.

1.6.1 Descriptive analysis

The Dow Jones stocks data, on NYSE and NASDAQ, amount to 30 assets on a 3 month period for a total of 22,444,752 observations. NYSE and NASDAQ are the two biggest exchanges in the world

Table 1.1: Simulation Results: Model I

Panel A: $c_1^2 = c_2^2$								
h	$IS_{u,1}$	$IS_{l,1}$	$ISR_{u,1}$	$ISR_{l,1}$	IS_1	ISR_1	PT_1	PTR_1
1s	0.65	0.35	0.50	0.50	0.50	0.50	0.50	0.50
	(0.06)	(0.06)	(0.04)	(0.08)	(0.06)	(0.06)	(0.04)	(0.04)
5s	0.77	0.22	0.50	0.49	0.50	0.50	0.49	0.49
	(0.12)	(0.11)	(0.07)	(0.22)	(0.11)	(0.15)	(0.14)	(0.14)
10s	0.83	0.17	0.50	0.50	0.50	0.50	0.50	0.50
	(0.14)	(0.14)	(0.08)	(0.31)	(0.14)	(0.20)	(0.28)	(0.28)
Panel B: $c_1^2 = 2c_2^2$								
h	$IS_{u,1}$	$IS_{l,1}$	$ISR_{u,1}$	$ISR_{l,1}$	IS_1	ISR_1	PT_1	PTR_1
1s	0.51	0.19	0.54	0.42	0.35	0.48	0.33	0.47
	(0.06)	(0.05)	(0.05)	(0.09)	(0.06)	(0.06)	(0.04)	(0.05)
5s	0.7	0.11	0.54	0.36	0.41	0.45	0.34	0.42
	(0.13)	(0.08)	(0.07)	(0.23)	(0.1)	(0.15)	(0.15)	(0.17)
10s	0.78	0.09	0.53	0.36	0.44	0.44	0.34	0.37
	(0.15)	(0.11)	(0.08)	(0.32)	(0.13)	(0.2)	(0.3)	(0.34)
30s	0.82	0.14	0.51	0.45	0.48	0.48	0.33	0.26
	(0.2)	(0.17)	(0.1)	(0.38)	(0.18)	(0.19)	(2.76)	(2.02)

Estimates for market 1 of the Information Share (IS) bounds, the PT share, the robust ISR and PTR. It is computed on simulated prices of Model I:

$$\begin{array}{lll} m_{th} = m_{th-h} + \eta_{th}, & \eta_{th} \backsim \sigma_h \mathcal{N}(0,1), \sigma_h = T^{-0.5}, & panel\,A: c_1^2 = c_2^2 = 2.10^{-4} \\ p_{ith} = m_{th} + c_i \varepsilon_{ith}, & \varepsilon_{ith} \sim \mathcal{N}(0,1), i = 1,2, & panel\,B: c_1^2 = 2c_2^2 = 2.10^{-4} \end{array}$$

A path of T=23400 observations is generated, prices are sampled at each interval h and a VECM is estimated with lag chosen by AIC. The values presented are the averages and the standard deviation (in parenthesis) over 1000 simulated paths.The gray shaded columns are robust measures. The reference value is 0.50

Table 1.2: Simulation Results: Model II

h	$IS_{u,1}$	$IS_{l,1}$	$ISR_{u,1}$	$ISR_{l,1}$	IS_1	ISR_1	PT_1	PTR_1
1s	0.39	0.29	0.90	0.88	0.34	0.89	0.16	0.76
	(0.06)	(0.06)	(0.02)	(0.03)	(0.06)	(0.03)	(0.02)	(0.03)
2s	0.33	0.19	0.84	0.77	0.26	0.81	0.14	0.69
	(0.09)	(0.07)	(0.05)	(0.10)	(0.08)	(0.08)	(0.03)	(0.06)
3s	0.30	0.14	0.80	0.66	0.22	0.73	0.13	0.63
	(0.10)	(0.08)	(0.08)	(0.17)	(0.09)	(0.13)	(0.04)	(0.13)
5s	0.30	(0.10	0.74	0.51	0.20	0.63	0.12	0.53
	(0.12	(0.08)	(0.11)	(0.23)	(0.10)	(0.17)	(0.06)	(0.24)
10s	0.34	0.08	0.67	0.36	0.21	0.52	0.11	0.40
	(0.18)	(0.10)	(0.16)	(0.28)	(0.14)	(0.22)	(0.11)	(1.77)

The Table reports estimates for market 1 of the Information Share (IS) bounds, the PT share, the robust ISR and PTR). It is computed on simulated prices of Model II: $m_{th} = m_{th-h} + \eta_{th}$, $p_{1th} = m_{th} + c_1\varepsilon_{1th}, p_{2th} = m_{th-\delta} + c_2\varepsilon_{2th}$, $\varepsilon_{ith}, (\eta_{th}/\sigma_h) \sim \mathcal{N}(0,1), i = 1,2$, $\sigma_h = T^{-0.5}, c_1^2 = 0.002$, $c_2^2 = 0.0001$, $\delta = 3$. *A path of T=23400 observations is generated, prices are sampled at each interval h and a VECM is estimated with lag chosen by AIC. The values presented averages and standard deviations (in parenthesis) over 1000 simulated paths.The gray shaded columns are robust measures. The reference value is 1.*

Table 1.3: Simulation Results: Model III

h	$IS_{u,1}$	$IS_{l,1}$	$ISR_{u,1}$	$ISR_{l,1}$	IS_1	ISR_1	PT_1	PTR_1
1s	0.47	0.49	1.00	1.00	0.48	1	0.01	1
	(0.19)	(0.19)	(0.00)	(0.00)	(0.19)	(0.00)	(0.01)	(0.00)
2s	0.25	0.26	1.00	1.00	0.26	1	0.01	0.99
	(0.20)	(0.20)	(0.03)	(0.04)	(0.20)	(0.03)	(0.06)	(0.02)
3s	0.22	0.22	0.99	0.99	0.22	0.99	0.01	1.00
	(0.21)	(0.21)	(0.09)	(0.09)	(0.21)	(0.09)	(0.04)	(0.51)
5s	0.18	0.18	0.98	0.98	0.18	0.98	0.01	0.61
	(0.22)	(0.22)	(0.10)	(0.10)	(0.22)	(0.10)	(0.03)	(11.57)
10s	0.24	0.23	0.98	0.97	0.24	0.98	0.00	0.99
	(0.27)	(0.27)	(0.13)	(0.09)	(0.27)	(0.11)	(0.26)	(0.21)

The Table reports estimates for market 1 of the Information Share bounds (ISu,ISl), the robust IS bounds (ISlr,ISur), the PT and the robust PT (PTr). It is computed on simulated prices of Model III: $m_{th} = m_{th-h} + \lambda_h \eta_{1th}$, $p_{1th} = m_{th} + \eta_{1th} + c_1 \varepsilon_{1th}, p_{2th} = m_{th-\delta} + c_2 \varepsilon_{2th}$, $\varepsilon_{ith}, (\eta_{th}/\sigma_h) \sim \mathcal{N}(0,1), i = 1,2$, $\sigma_h = \lambda_h = T^{-0.5}, c_1^2 = 0.002$, $c_2^2 = 0.0001$, $\delta = 3$. *A path of T=23400 observations is generated, prices are sampled at each interval h and a VECM is estimated with lag chosen by AIC. The values presented averages and standard deviations (in parenthesis) over 1000 simulated paths.The gray shaded columns are robust measures. The reference value is 1.*

by capitalization and trade value. NYSE remains by far the first with a capitalization of around 14 USD trillion in 2011 (around 16 USD trillion in 2014). During this year, the trade value was about 20 USD trillions, which represents an average daily amount of 55 USD billions. NASDAQ has a market capitalization of 4.6 USD trillions, and a trade value of 13.5 USD trillions, corresponding to an average daily amount of 37 USD billions[7].

Concerning where the assets are traded, the domination is not that pronounced as shown by the average daily statistics in Table 1.8 in appendix. For JP Morgan (JMP) for example, around 5 millions of share are traded each day on NYSE, while 4.8 millions are traded on NASDAQ. This pattern is the same for most of the stocks, that is to say that NYSE concentrates the biggest part of share exchanged in a day. For few assets like PFE and GE, NASDAQ dominates the exchanges in term of volume. If I look at the liquidity (I think of liquidity as the frequency of transactions), NASDAQ dominates for almost all assets. For PFE I have around 16,153 trading times in one day on NASDAQ, while I have only 7 080 trading times on NYSE. This is not in contradiction with the analysis of volumes, it simply states that most of the trades of bigger size happens on NYSE, while NASDAQ is characterized by a lot of trades of small quantities (details in table 1.9 in appendix). For example, NASDAQ cumulates 43.3% of small size trades for American express (AXP) and only 23.4% of big size trades, while NYSE cumulates 57.3% of big size trades.

Those descriptive statistics also show that, if prior-belief is that price discovery is completely driven by the liquidity or by the volume of share traded, the answer is not straightforward as I have for each market depending on the asset: high-volume and high-liquidity, high-volume and low-liquidity, low-volume and high-liquidity.

1.6.2 Results on markets contribution

Before looking at market dominance, I compute the IS and the PT measures for the assets at different sampling frequency with the VECM-lag chosen by AIC. I obtain the same type of patterns described in Section 3 with the structural models. Figures 1.7 and 1.8 plot the results for American Express (AXP) and Exxon Mobil Corporation (XOM). It shows that the evolution of the measures with sampling frequency looks like the theoretical path up to a given frequency. It doesn't show the crossing of the lines, but this might be just that raw data are not frequent high enough to display all the interesting features. I can only have convincing guess by looking at the limit of the lines. In fact, for most of the stocks, the number of transactions per day is such that the interval h is between $4s$ and $7s$.

Now, let's consider the mid-quotes data at the microsecond frequency for Microsoft and Pfizer on the December 12th 2013. This choices are imposed only by data accessibility, and I consider

[7] Source: (http://www.i3investor.com/jsp/hti/usmarket.jsp).

NYSE ARca (market 1) and NASDAQ (market 2). This trading day corresponds to an amount of 424,876 observations for Microsoft, and 149090 for Pfizer. Figure 1.9 shows the results of the IS and the PT measures with respect to the sampling frequency. It confirms that the interpretation of the results can change with the sampling frequency, and that the IS bounds tighten to the same values at high frequency.

The results on markets contributions in Table 1.4 show that, for most of the stocks, NYSE appears to be the dominant market. The dominance of NYSE on NASDAQ is strong for MMM, NKE and TRV. NASDAQ dominates the price discovery mechanism for BA, CAT, GS, IBM and all Nasdaq primary listed stocks. The table 1.4 also reports the lower and the upper bounds on the IS. It shows that bounds are quite wide for all assets (for example NYSE has a 28% to 85% contribution for American Express-AXP) which clearly complicates the interpretation. Meanwhile the robust IS indicates that the contribution is between 55% and 56% coherent with the numbers for PT and PTR. The results are robust to the latency problem while recording the data. To check that, I redo the estimations by delaying the prices of 1 second and the results remain qualitatively the same. These results also indicate that the markets structure have really changed during the recent years. For comparison, in Hasbrouck (1995), NYSE concentrated most of the trades resulting in more than 90 % of the contribution to price discovery.

I now compute the correlation of each market's contribution to price discovery with its share in different categories of transaction size. I see that (table 1.5) for all the exchanges the correlation between their contribution and their market share in small-size transactions is 1/3. The correlation with big-size trades is 0.27 for NYSE-listed share, while it is only -0.04 for the set of NASDAQ-listed shares. The correlation of the ISR with the liquidity does not show a specific pattern. In summary, price discovery happens generally on NYSE for the stocks under investigation, the contribution of a market is correlated with its market share for small and medium size transactions. In the results, The robust IS measures present another advantage over the IS. When the bounds on IS are wide, the robust IS provides very close bounds that facilitate the interpretation.

1.6.3 Macroeconomics announcements days

The releases of Macroeconomic indicators constitute some of the times where fundamental information arrive in the markets. Interesting insights could thus be investigated by looking at how the different markets behave the days of major macroeconomic news compared to normal days. For this, I identify a set of events from the literature (Andersen et al., 2003; Frijns et al., 2015) and the corresponding dates at which they are released in the sample. I mention that almost all the announcements here happen at 8:30 AM, which is before the markets open. An important comparison of the markets could be done for the news that are released during the trading session, to see for

Table 1.4: Contribution to price discovery of NYSE

	$IS_{u,1}$	$IS_{l,1}$	$ISR_{u,1}$	$ISR_{u,1}$	IS_1	ISR_1	PT_1	PTR_1
NYSE-listed stocks								
AXP	0.85	0.28	0.55	0.56	0.59	0.67	0.60	0.63
BA	0.64	0.14	0.36	0.39	0.47	0.33	0.39	0.40
CAT	0.79	0.21	0.48	0.50	0.54	0.52	0.51	0.53
CVX	0.89	0.31	0.63	0.60	0.57	0.72	0.63	0.65
DD	0.82	0.28	0.54	0.55	0.57	0.63	0.58	0.60
DIS	0.89	0.36	0.63	0.63	0.63	0.79	0.67	0.71
GE	0.79	0.36	0.53	0.58	0.64	0.71	0.62	0.68
GS	0.73	0.23	0.47	0.48	0.50	0.47	0.49	0.49
HD	0.88	0.33	0.59	0.60	0.62	0.75	0.64	0.69
IBM	0.74	0.25	0.48	0.49	0.53	0.51	0.51	0.52
JNJ	0.91	0.41	0.66	0.66	0.67	0.84	0.71	0.75
JPM	0.89	0.29	0.59	0.59	0.61	0.75	0.64	0.68
KO	0.87	0.29	0.58	0.58	0.60	0.72	0.63	0.66
MCD	0.87	0.40	0.63	0.63	0.62	0.75	0.66	0.69
MMM	0.94	0.54	0.75	0.74	0.67	0.88	0.75	0.78
MRK	0.84	0.36	0.58	0.60	0.63	0.74	0.64	0.69
NKE	0.84	0.45	0.62	0.65	0.64	0.74	0.65	0.69
PFE	0.72	0.32	0.47	0.52	0.62	0.64	0.58	0.63
PG	0.85	0.30	0.56	0.57	0.60	0.69	0.61	0.65
TRV	0.87	0.44	0.64	0.65	0.64	0.77	0.67	0.71
UNH	0.87	0.34	0.60	0.61	0.61	0.74	0.63	0.67
UTX	0.80	0.31	0.54	0.56	0.57	0.62	0.57	0.60
VZ	0.88	0.38	0.62	0.63	0.64	0.79	0.67	0.71
WMT	0.87	0.33	0.58	0.60	0.63	0.75	0.64	0.69
XOM	0.94	0.31	0.68	0.62	0.61	0.83	0.71	0.72
Total	**0.84**	**0.33**	**0.69**	**0.60**	**0.57**	**065**	**0.58**	**0.62**
Nasdaq-listed stocks								
AAPL	0.64	0.08	0.42	0.19	0.36	0.30	0.32	0.33
CSCO	0.60	0.22	0.34	0.27	0.41	0.31	0.48	0.38
INTC	0.60	0.15	0.34	0.22	0.38	0.28	0.42	0.34
MSFT	0.60	0.16	0.33	0.22	0.38	0.28	0.44	0.35
Total	**0.61**	**0.15**	**0.36**	**0.23**	**0.38**	**0.29**	**0.41**	**0.35**

Table 1.5: Correlation of ISR (NYSE) with transactions size and liquidity

	NYSE-Listed	NASDAQ-Listed
small trade	0.33	0.30
medium trade	0.10	0.19
big trade	0.27	-0.04
Liquidity	0.33	0.27

liquidity= number of trades per day.

Table 1.6: Macroeconomic News days on the period of study

Macroeconomics Announcement	Source	Release dates
GDP (Advance, preliminary, final) estimate	BEA	March 25, April 28, May 26
Personal Income, Personal Consumption Expenditures	BEA	March 28, April 29, May 27
International Trade Balance in Goods and Services	BEA	March 10, April 12, May 11
Nonfarm Payroll Employment	BLS	March 4, April 21, May 6
Producer Price Index PPI	BLS	March 16, April 14, May 12
Consumer Price Index CPI	BLS	March 17, April 15, May 13
Industrial Production, Capacity Utilization	FRB	May 17, April 15, March 17
Consumer Credit	FRB	March 7, April 7, May 6
Federal Funds Rate	FRB	March 15, April 27

example which market reacts quickly. However, this would required a very long sample as they are typically published only one day per month. Table 1.6 presents the macroeconomics indicators that I consider and the announcement dates in the sample.

I compute the measures for the announcement days and for the non-announcements days, I obtain that on average NASDAQ's contribution to information share is slightly bigger the days where there is a news (from 0.41 to 0.42). The contribution of NYSE is still greater than for the NASDAQ but slightly decreases compared to non-announcement days. Since the changes in the numbers are to small it is difficult to convince of a particularity for these news days. If I believe the contribution of NASDAQ has significantly increased, it is difficult to explain why but a reason might be found in the liquidity of NASDAQ. Traders wanting to exploit quickly those public pre-scheduled news, could prefer to do so on the most liquid market. More details, per asset, on which market increases its contribution to price discovery can be found in table 1.10.

1.7 Conclusion

Among the assets traded on markets places, some are strongly related by arbitrage relationships. This is the case of securities and their derivatives, and assets listed simultaneously in many coun-

Table 1.7: Markets' contribution to Price discovery on Macro News days

	NYSE				NASDAQ			
	IS_u	IS_l	IS	PT	IS_u	IS_l	IS	PT
Announcement	0.84	0.32	0.58	0.57	0.68	0.16	0.42	0.43
Non Announcements	0.84	0.33	0.59	0.58	0.67	0.16	0.41	0.42

tries. To determine in which market the efficient price is determined, some measures of price discovery were proven useful in the literature. In this paper, I started by studying the behavior of the popular prices discovery metrics in their relationship with sampling frequency and market microstructure noises. I showed analytically, in some standard microstructure models, that the Information Share measure (IS) of Hasbrouck (1995) and the Permanent-Transitory component measure (PT) of Harris et al. (2002b) are driven by non-informative noises when the sampling interval is small. The IS is identified only between bounds and, when the frequency is low, the bounds are too wide to provide straight conclusions. When the frequency is particularly high, the IS bounds tighten and converge to a unique value which is the same as the PT. But this value is dominated by noises and is not affected by the informative innovation. Using data of NYSE TAQ database, I examined if my conclusions are in line with the data. I observed that indeed the data seem to present the patterns I highlighted. The frequency of the transaction prices might not be high enough to show certain features, but the analysis with mid-quotes of Microsoft confirms my theoretical conclusions.

The price discovery measures are typically used to decide which price is close to the fundamental price. The "closeness" involves two interesting dimensions, the "speed" at which a market incorporates news and the "noise-avoidance" in the mechanism. The two dimensions are economically relevant but confusions come from that a market is not necessarily the best in both dimensions. A market can be the fastest and the noisiest. At lower frequency the measures capture a mix of the two aspects, while at high frequency, my results showed that they rather capture the avoidance of noise. This is a serious problem because first, many papers use and think of price discovery as the rapidity to process new information. Second, the measures are used in regression to investigate the determinant of a market's efficiency. Those papers conclude for example that market with relative small bid-ask spread are dominating price discovery process. If price discovery were to measure only how markets avoid noises, this conclusion amounts to stating that "Noise is small in this market because noise is small". I then presented new measures of price discovery, that I named Robut IS (ISR) and Robust PT (PTR), that disentangle the two dimensions and clarify the interpretation. They are good at detecting "which market incorporates quickly new information". My overall contribution constitutes many steps forward into the debate on what the price discovery measures are actually capturing.

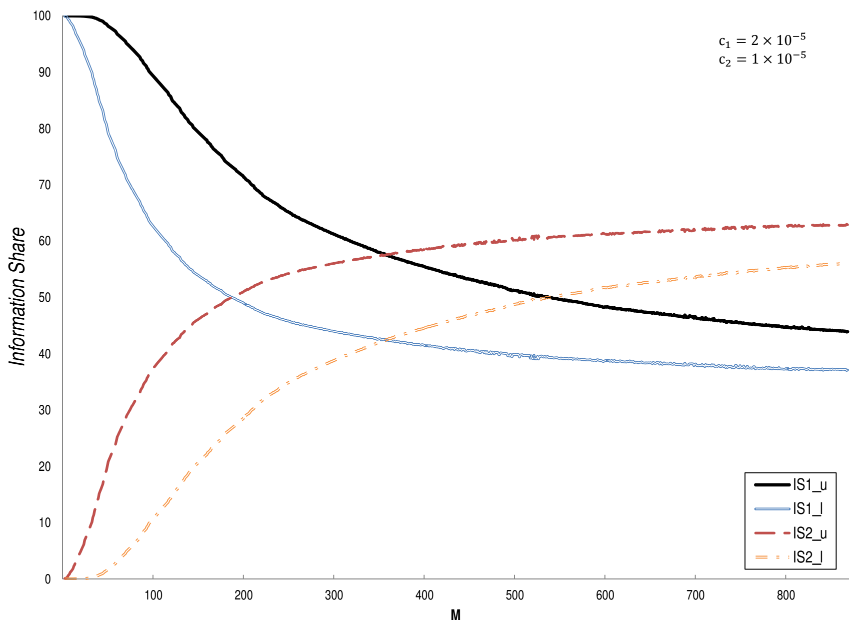
In the application, I investigated the relative contribution of NYSE and NASDAQ to the price formation of Dow Jones assets. The robust measures seem to improve a little bit on the IS and PT. I found that NYSE captures the big part of volume traded, but NASDAQ is the most liquid with a high level of activity. This implies that NASDAQ mostly runs the orders of small quantities while NYSE runs big quantities orders. In terms of contribution to price discovery for the assets under investigation, NYSE is generally dominant and Nasdaq dominates for teh four Nasdaq-listed stocks. The contribution of a market appears to be positively correlated with its liquidity. I also computed the correlation between market's contribution and markets share in each category of trade size. It reveals that the contribution of a market is correlated with its share in small size transactions. For NASDAQ listed stocks, there is no correlation with market's share in big size transactions, so large quantities trades do not convey information. I analyze the performance of the measures the days with major macroeconomic announcements and the contribution of NASDAQ to price discovery increases only slightly the days with news. As the announcements considered here are typically done when the market is closed, it is not possible to conclude about the behavior of price discovery measures around the news. There are more insights to investigate with a good database of news released during trading session.

1.8 APPENDIX A

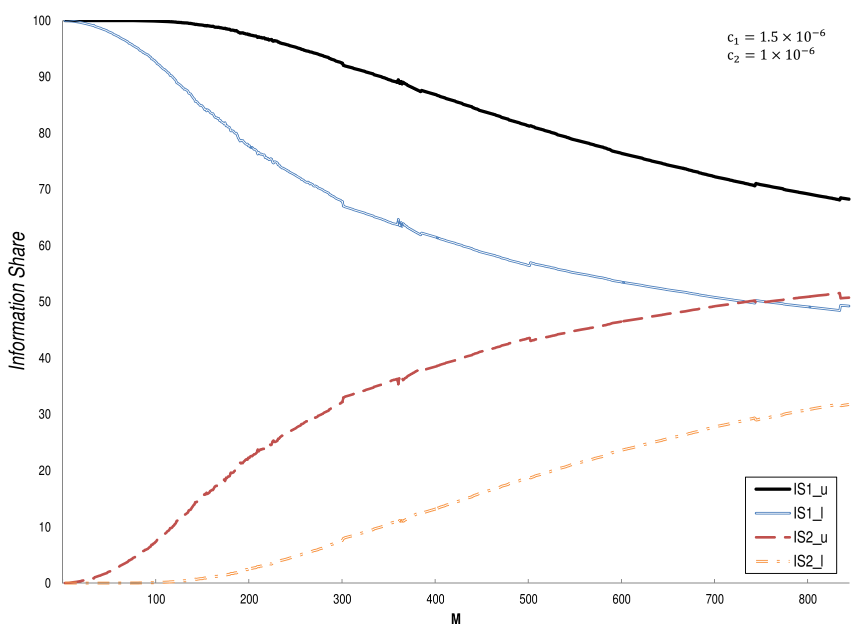
1.8.1 Tables and Figures

1.8.1.1 Figures

Figure 1.6: Model III: IS performance with noise and frequency



(a) Small noises variance

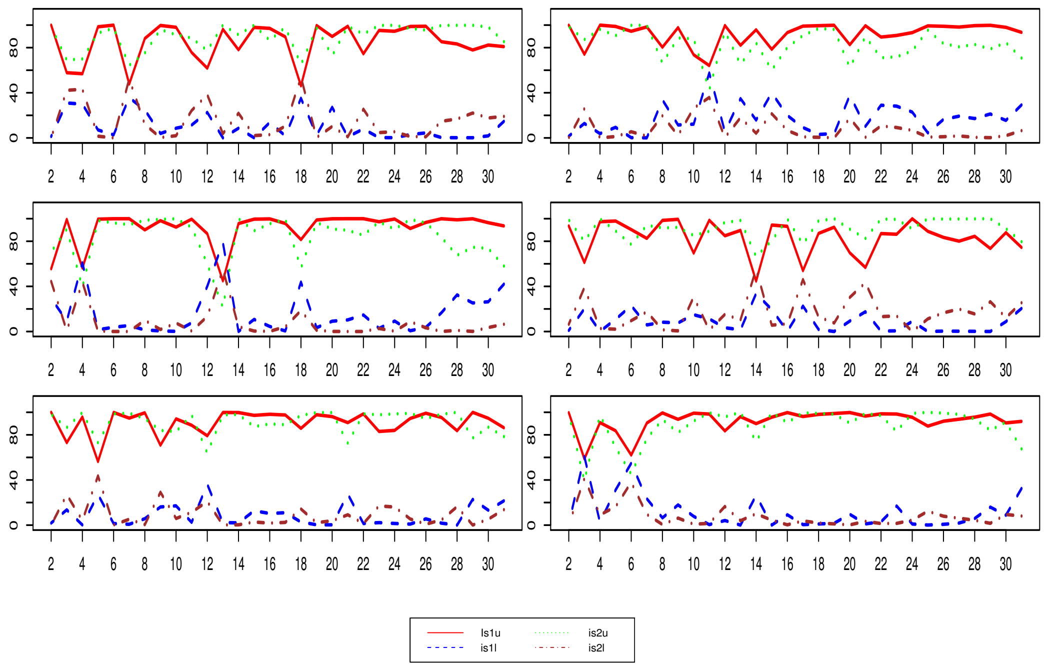


(b) mall difference between the noises

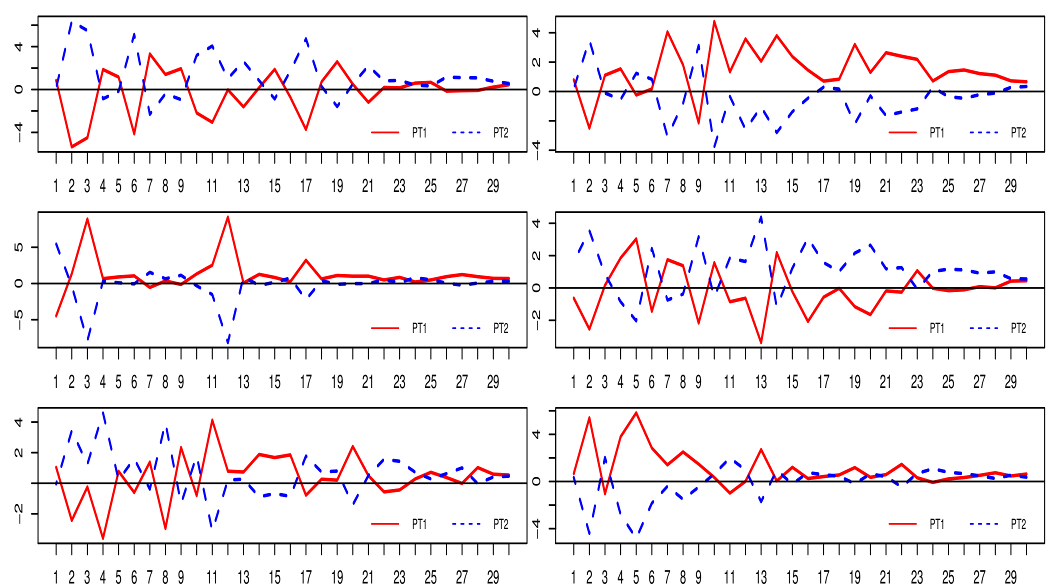
Note: The figure plots the IS model III. The horizontal axis represents $M = 1/h$. The PT (not plotted here) has the same pattern

Figure 1.7: IS and PT by sampling frequency for American Express

(a) IS bounds AXP



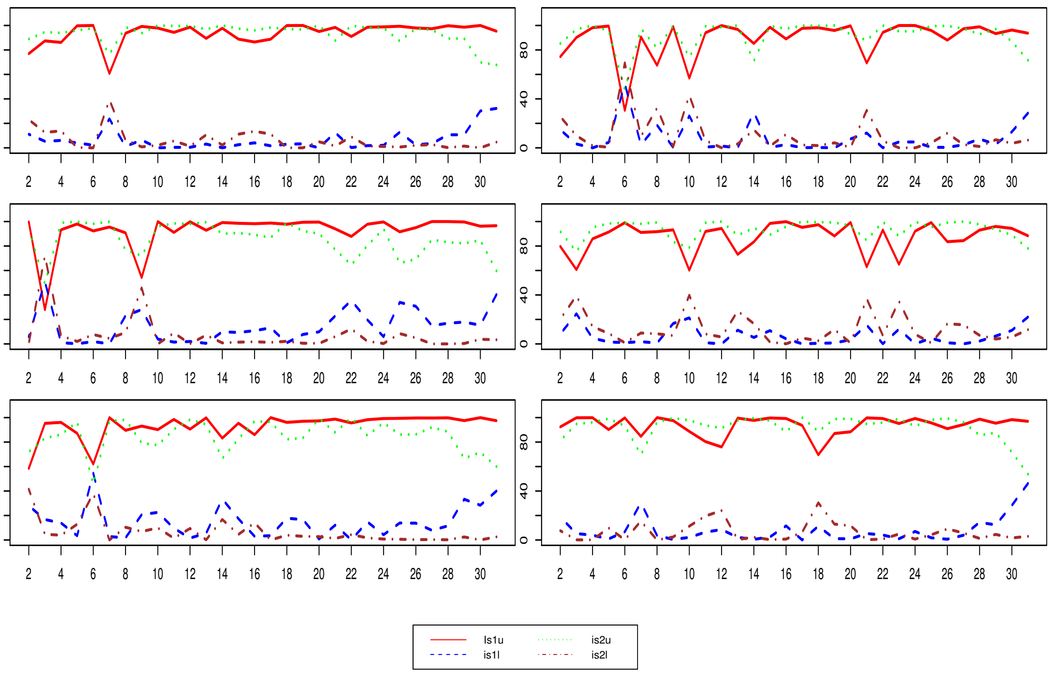
(b) PT for AXP



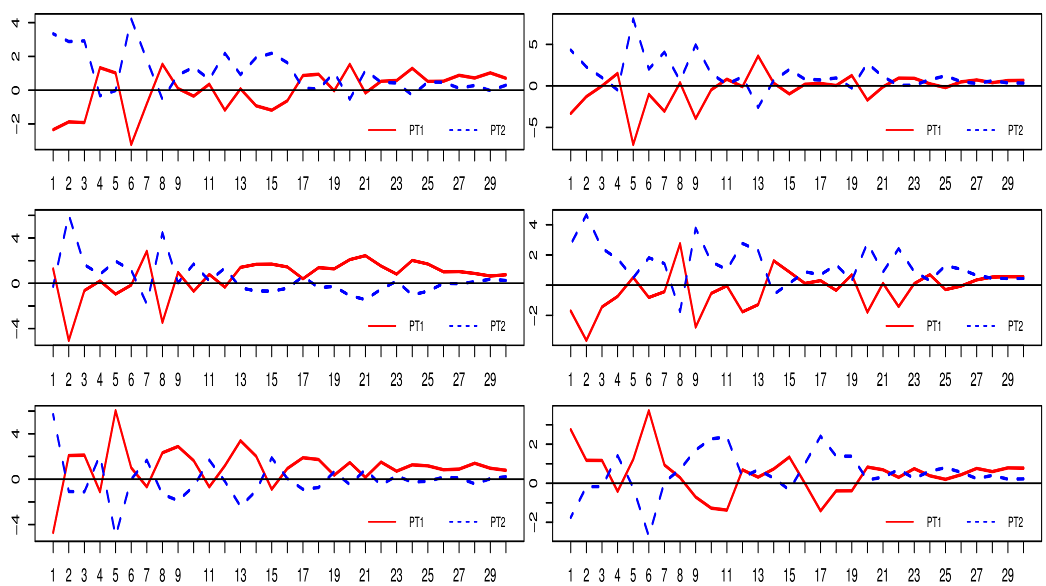
Note: The figures plot 6 chosen day in the database. For each day the data are sampled at different frequency and the measures at computed in a VECM with lag selected by AIC. The horizontal axis represents the sampling frequency M.

Figure 1.8: IS and PT by sampling frequency for Exxon Mobil (XOM)

(a) IS bounds for XOM

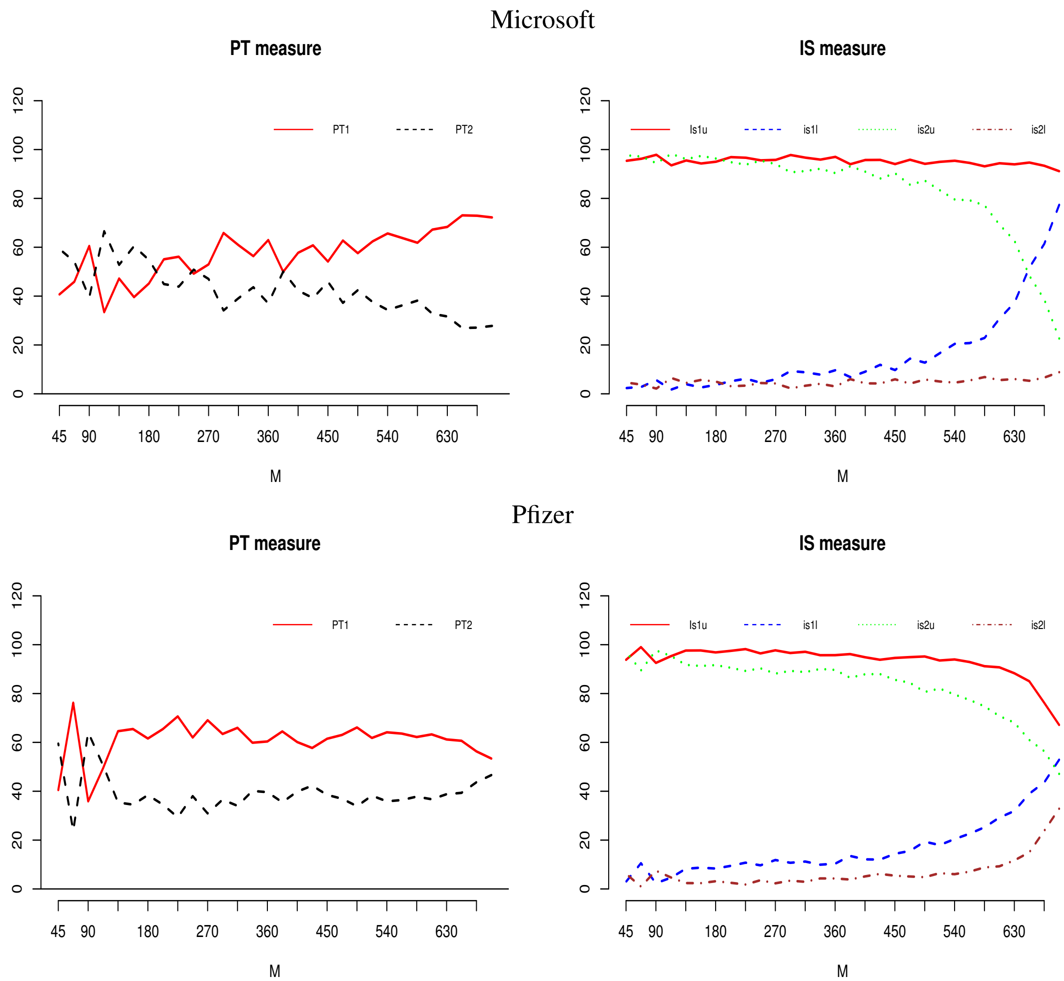


(b) PT for XOM



Note: The figures plot 6 chosen day in the database. For each day the data are sampled at different frequency and the measures at computed in a VECM with lag selected by AIC. The horizontal axis represents the sampling sampling frequency M. e.g: $M = 1$ means 1 observation per 200 s , $M = 30$ means 3 obs. per 20s

Figure 1.9: IS and PT by sampling frequency for Microsoft and Pfizer



Note: IS and PT for Microsoft and Pfizer Inc on 02/12/2013. The mid-quotes are sampled at different frequency and the measures at computed in a VECM with lag selected by AIC. The horizontal axis represents the sampling frequency M. E.g $M = 1$ *means 1 observation per 200 s ,* $M = 600$ *means 3 obs. per second*

1.8.1.2 Tables

Table 1.8: Average daily number and volume of transactions by markets and assets

	Volume		Liquidity	
Stocks	NYSE	NASDAQ	NYSE	NASDAQ
AXP	1,390,154	1,333,287	5,675	9,175
BA	761,827	939,104	4,265	7,518
CAT	1,509,825	1,258,157	8,587	9,909
CVX	1,638,540	1,387,986	9,221	11,646
DD	1,197,841	932,970	5,618	7,544
DIS	2,205,886	1,279,417	7,123	8,919
GE	7,303,315	7,452,367	8,162	18,419
GS	852,376	957,759	5,407	7,373
HD	1,885,715	1,489,919	5,804	9,231
IBM	1,065,407	979,830	6,294	7,523
JNJ	2,606,299	1,614,322	8,618	10,354
JPM	5,036,588	4,866,551	**11,438**	**22,572**
KO	1,833,110	1,282,143	6,942	9,534
MCD	1,217,757	835,071	5,398	6,390
MMM	877,629	506,236	5,389	4,154
MRK	2,028,357	2,106,988	4,626	9,580
NKE	808,253	479,999	4,164	3,713
PFE	6,601,470	7,418,801	**7,080**	**16,153**
PG	2,010,251	1,732,199	5,725	10,131
TRV	803,587	478,962	3,893	3,877
UNH	1,331,712	974,829	5,753	7,084
UTX	852,967	766,713	4,778	6,367
VZ	2,261,393	2,160,213	5,614	10,639
WMT	2,136,040	1,521,625	6,743	9,916
XOM	4,649,096	2,546,214	15,937	17,663
	NYSE Arca	NASDAQ	NYSE Arca	NASDAQ
AAPL	2,318,123	3,944,326	17,123	27,099
CSCO	7,573,610	13,954,385	14,945	22,991
INTC	6,455,527	16,635,294	15,125	30,653
MSFT	5,513,734	14,580,294	14,825	27,540

Note: The period is from the 01/03 to 30/05/2011. liquidity=number of transactions per day; volume=volume of trades per day.

Table 1.9: Share of markets each category of transactions size

	Small size		**Medium size**		**Big size**	
Stock	**NYSE**	**NASDAQ**	**NYSE**	**NASDAQ**	**NYSE**	**NASDAQ**
AXP	0.32	0.68	0.51	0.49	0.71	0.29
BA	0.32	0.68	0.55	0.45	0.65	0.35
CAT	0.41	0.59	0.66	0.34	0.72	0.28
CVX	0.37	0.63	0.69	0.31	0.79	0.21
DD	0.35	0.65	0.66	0.34	0.81	0.19
DIS	0.39	0.61	0.67	0.33	0.86	0.14
GE	0.29	0.71	0.28	0.72	0.57	0.43
GS	0.40	0.60	0.60	0.40	0.58	0.42
HD	0.33	0.67	0.55	0.45	0.71	0.29
IBM	0.41	0.59	0.62	0.38	0.69	0.31
JNJ	0.41	0.59	0.59	0.41	0.78	0.22
JPM	0.30	0.70	0.38	0.62	0.66	0.34
KO	0.33	0.67	0.68	0.32	0.79	0.21
MCD	0.38	0.62	0.67	0.33	0.79	0.21
MMM	0.52	0.48	0.75	0.25	0.81	0.19
MRK	0.28	0.72	0.36	0.64	0.56	0.44
NKE	0.47	0.53	0.72	0.28	0.80	0.20
PFE	0.30	0.70	0.27	0.73	0.48	0.52
PG	0.32	0.68	0.61	0.39	0.75	0.25
TRV	0.44	0.56	0.77	0.23	0.85	0.15
UNH	0.38	0.62	0.63	0.37	0.75	0.25
UTX	0.37	0.63	0.68	0.32	0.76	0.24
VZ	0.31	0.69	0.33	0.67	0.67	0.33
WMT	0.35	0.65	0.57	0.43	0.78	0.22
XOM	0.41	0.59	0.69	0.31	0.83	0.17
Total	0.37	0.63	0.58	0.42	0.73	0.27

Note: *Let DM the average transactions size in a day: Small size* $\equiv$ *quantity* $< DM$*, medium size:* $\equiv DM \leq quantity \leq 2DM$*, big size:* $\equiv quantity > 2DM$*. The numbers are Daily* Averages.

Table 1.10: NYSE contribution by asset on the days of Macroeconomic announcements

	ISR_u		ISR_l		ISR		PTR	
	A	**N**	**A**	**N**	**A**	**N**	**A**	**N**
AXP	0.85	0.85	0.26	0.28	0.55	0.57	0.54	0.56
BA	0.65	0.63	0.15	0.14	0.40	0.39	0.37	0.35
CAT	0.80	0.78	0.22	0.20	0.51	0.49	0.49	0.47
CVX	0.89	0.89	0.28	0.33	0.59	0.61	0.61	0.63
DD	0.82	0.82	0.27	0.28	0.55	0.55	0.55	0.54
DIS	0.89	0.90	0.32	0.38	0.61	0.64	0.61	0.65
GE	0.83	0.78	0.38	0.36	0.60	0.57	0.57	0.52
GS	0.75	0.71	0.25	0.23	0.50	0.47	0.49	0.46
HD	0.86	0.88	0.29	0.35	0.57	0.62	0.56	0.61
IBM	0.76	0.73	0.26	0.24	0.51	0.49	0.50	0.47
JNJ	0.89	0.91	0.39	0.42	0.64	0.67	0.64	0.67
JPM	0.90	0.88	0.28	0.29	0.59	0.59	0.59	0.59
KO	0.87	0.87	0.27	0.30	0.57	0.58	0.57	0.58
MCD	0.87	0.87	0.40	0.40	0.63	0.63	0.64	0.62
MMM	0.93	0.94	0.51	0.55	0.72	0.75	0.72	0.77
MRK	0.85	0.84	0.33	0.38	0.59	0.61	0.57	0.58
NKE	0.85	0.84	0.48	0.44	0.66	0.64	0.64	0.62
PFE	0.75	0.71	0.31	0.33	0.53	0.52	0.49	0.46
PG	0.84	0.85	0.27	0.31	0.55	0.58	0.54	0.57
TRV	0.85	0.88	0.40	0.46	0.63	0.67	0.62	0.66
UNH	0.87	0.87	0.34	0.34	0.60	0.61	0.60	0.60
UTX	0.80	0.81	0.30	0.31	0.55	0.56	0.53	0.55
VZ	0.86	0.89	0.34	0.41	0.60	0.65	0.59	0.64
WMT	0.87	0.86	0.33	0.33	0.60	0.60	0.59	0.58
XOM	0.94	0.94	0.31	0.31	0.62	0.62	0.69	0.68
Total	0.84	0.84	0.32	0.33	0.58	0.59	0.57	0.58

$IS_l = (IS_u + IS_l)/2$,,A=announcement, N=Non announcement

1.8.2 Proofs: Analytical formulas of the measures

I present calculations by skipping some details. The detailed calculations can be found (Online here)

1.8.2.1 General results

Consider $p_t = \begin{pmatrix} p_{1t} & p_{2t} \end{pmatrix}'$ admitting the VMA(1) : $\Delta p_t = e_t + \Theta e_{t-1}$

$$\text{With } \Omega = var\left(\varepsilon_t\right) = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \text{ and } \Psi\left(1\right) = I + \Theta = \begin{pmatrix} c & 1+d \\ c & 1+d \end{pmatrix}.$$

The goal is to solve for Θ and Ω given the structural parameters. Let

$$C_0 \equiv var(\Delta p_t) = \begin{pmatrix} v_1^2 & v_{12} \\ v_{12} & v_2^2 \end{pmatrix} \text{ and } C_1 \equiv cov(\Delta p_t, \Delta p'_{t-h}) = \begin{pmatrix} m_1 & m_{12} \\ m_{21} & m_2 \end{pmatrix} \tag{1.27}$$

Using the VMA(1) 1.8 gives

$$C_0 = \Omega + \Theta\Omega\Theta' \tag{1.28}$$

$$C_1 = \Theta\Omega \tag{1.29}$$

By multiplying 1.28 by Θ and using 1.29 then

$$C_1 - \Theta C_0 + \Theta C_1 \Theta' = 0 \tag{1.30}$$

$$\begin{aligned}
\Theta C_1 \Theta' &= \begin{pmatrix} -1+c & 1+d \\ c & d \end{pmatrix} \begin{pmatrix} m_1 & m_{12} \\ m_{21} & m_2 \end{pmatrix} \begin{pmatrix} -1+c & c \\ 1+d & d \end{pmatrix} \\
&= \begin{pmatrix} c^2 m_1 + d^2 m_2 + cd\left(m_{12}+m_{21}\right) + c\left(-2m_1 + m_{12} + m_{21}\right) + d\left(2m_2 - m_{12} - m_{21}\right) \\ \qquad + m_1 + m_2 - m_{12} - m_{21} \\ c^2 m_1 + d^2 m_2 + cd\left(m_{12}+m_{21}\right) + c\left(-m_1 + m_{21}\right) + d\left(m_2 - m_{12}\right) \\ c^2 m_1 + d^2 m_2 + cd\left(m_{12}+m_{21}\right) + c\left(-m_1 + m_{12}\right) + d\left(-m_{21} + m_2\right) \\ c^2 m_1 + d^2 m_2 + cd\left(m_{12}+m_{21}\right) \end{pmatrix}
\end{aligned}$$

Set $Q = c^2 m_1 + d^2 m_2 + cd\,(m_{12} + m_{21})$ then

$$\Theta C_1 \Theta' = \begin{pmatrix} Q + c\,(-2m_1 + m_{12} + m_{21}) + d\,(2m_2 - m_{12} - m_{21}) + m_1 + m_2 - m_{12} - m_{21} \\ Q + c\,(-m_1 + m_{21}) + d\,(m_2 - m_{12}) \\ Q + c\,(-m_1 + m_{12}) + d\,(m_2 - m_{21}) \\ Q \end{pmatrix}$$

I stack the lines to ease the presentation ΘC_0.

Using equation 1.30

$$\begin{aligned} 0 &= \Theta C_1 \Theta' + C_1 - \Theta C_0 \\ &= \begin{pmatrix} Q + c\left(-2m_1 + m_{12} + m_{21} - v_1^2\right) + d\,(2m_2 - m_{12} - m_{21} - v_{12}) \\ \quad +2m_1 + m_2 - m_{12} - m_{21} + v_1^2 - v_{12} \\ Q + c\,(-m_1 + m_{21} - v_{12}) + d\left(m_2 - m_{12} - v_2^2\right) + m_{12} + v_{12} - v_2^2 \\ Q + c\left(-m_1 + m_{12} - v_1^2\right) + d\,(m_2 - m_{21} - v_{12}) + m_{21} \\ Q - cv_{12} - dv_2^2 + m_2 \end{pmatrix} \end{aligned}$$

Subtracting the 2nd from the 3rd line

$$d\left(m_{12} - m_{21} + v_2^2 - v_{12}\right) = c\left(-m_{12} + m_{21} + v_1^2 - v_{12}\right) + m_{12} - m_{21} - v_2^2 + v_{12} \tag{1.31}$$

$$\begin{aligned} d &= c\frac{\left(-m_{12} + m_{21} + v_1^2 - v_{12}\right)}{m_{12} - m_{21} + v_2^2 - v_{12}} + \frac{m_{12} - m_{21} - v_2^2 + v_{12}}{m_{12} - m_{21} + v_2^2 - v_{12}} \\ &= cF + G \end{aligned} \tag{1.32}$$

with

$$F = \frac{-m_{12} + m_{21} + v_1^2 - v_{12}}{m_{12} - m_{21} + v_2^2 - v_{12}} \text{ and } G = \frac{m_{12} - m_{21} - v_2^2 + v_{12}}{m_{12} - m_{21} + v_2^2 - v_{12}} \tag{1.33}$$

Which is plugged into the quadratic equation (4th line):

$$c^2 m_1 + d^2 m_2 + cd\,(m_{12} + m_{21}) - cv_{12} - dv_2^2 + m_2 = 0 \tag{1.34}$$

$$
\begin{aligned}
0 &= c^2 m_1 + d^2 m_2 + cd\left(m_{12}+m_{21}\right) - cv_{12} - dv_2^2 + m_2 \\
&= c^2 m_1 + (cF+G)^2 m_2 + c\,(cF+G)\left(m_{12}+m_{21}\right) - cv_{12} - (cF+G)\,v_2^2 + m_2 \\
&= c^2\left(m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)\right) + c\left[2FGm_2 + G\left(m_{12}+m_{21}\right) - v_{12} - Fv_2^2\right] \\
&\quad + G^2 m_2 - Gv_2^2 + m_2
\end{aligned}
$$

$$
c^2 + c\frac{\left[2FGm_2 + G\left(m_{12}+m_{21}\right) - v_{12} - Fv_2^2\right]}{m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)} + \frac{G^2 m_2 - Gv_2^2 + m_2}{\left(m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)\right)}
$$

$$
\Delta = \left[\frac{2FGm_2 + G\left(m_{12}+m_{21}\right) - v_{12} - Fv_2^2}{m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)}\right]^2 - 4\frac{G^2 m_2 - Gv_2^2 + m_2}{\left(m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)\right)}
$$

$$
\boxed{
\begin{aligned}
c &= -\frac{1}{2}\frac{2FGm_2 + G\left(m_{12}+m_{21}\right) - v_{12} - Fv_2^2}{m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right)} \pm \frac{1}{2}\sqrt{\Delta} \\
d &= cF + g
\end{aligned}
} \tag{1.35}
$$

then

$$
\Omega = \Theta^{-1} C_1 = -\begin{pmatrix} -1+c & d \\ c & -1+d \end{pmatrix}^{-1} \begin{pmatrix} m_1 & m_{12} \\ m_{21} & m_2 \end{pmatrix} \tag{1.36}
$$

The PT measure I have computed $\psi \equiv \begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix} = \begin{pmatrix} c & 1+d \end{pmatrix}$

$$
PT_1 = \frac{c}{1+c+d} \text{ and } PT_2 = \frac{1+d}{1+c+d} \tag{1.37}
$$

The information share bounds The total IS is

$$
\begin{aligned}
\psi\Omega\psi' &= \begin{pmatrix} c & 1+d \end{pmatrix} \Omega \begin{pmatrix} c \\ 1+d \end{pmatrix} \\
&= \left[c^2 \sigma_{11} + 2\sigma_{12} c\,(1+d) + (1+d)^2 \sigma_{22}\right]
\end{aligned} \tag{1.38}
$$

The IS bounds for market 1 are

$$\begin{aligned} IS1u &= \left(\psi_{11}\sqrt{\sigma_{11}}+\psi_{12}\rho\sqrt{\sigma_{22}}\right)^2/\psi\Omega\psi \\ &= \left(c\sqrt{\sigma_{11}}+(1+d)\,\sigma_{12}\left(\sqrt{\sigma_{11}}\right)^{-1}\right)^2/\psi\Omega\psi \\ IS1l &= \left(\psi_{11}\sqrt{\sigma_{11}}\sqrt{(1-\rho^2)}\right)^2/\psi\Omega\psi \\ &= c^2\sigma_{11}\left(1-\sigma_{12}^2/\sigma_{11}\sigma_{22}\right)/\psi\Omega\psi \end{aligned} \tag{1.39}$$

and for market 2

$$\begin{aligned} IS2u &= \left(\psi_{12}\sqrt{\sigma_{22}}+\psi_{11}\rho\sqrt{\sigma_{11}}\right)^2/\psi\Omega\psi \\ &= \left((1+d)\sqrt{\sigma_{22}}+c\sigma_{12}\left(\sqrt{\sigma_{22}}\right)^{-1}\right)^2/\psi\Omega\psi \\ IS2l &= \left(\psi_{12}\sqrt{\sigma_{22}}\sqrt{(1-\rho^2)}\right)^2/\psi\Omega\psi \\ &= (1+d)^2\sigma_{22}\left(1-\sigma_{12}^2/\sigma_{11}\sigma_{22}\right)/\psi\Omega\psi \end{aligned} \tag{1.40}$$

1.8.2.2 Model I: A two-market "Roll" model.

Here $m_{th}=m_{th-h}+\eta_{th}$, the is innovation $\eta_{th}=\sigma_h\mathcal{N}(0,1)$ and $\sigma(h)$converges to zero with$h$

$$\begin{aligned} p_{1th} =& m_{th}+c_1\varepsilon_{1th} \\ p_{2th} =& m_{th}+c_2\varepsilon_{2th} \end{aligned} \tag{1.41}$$

With $\varepsilon_{it}\sim\mathcal{N}(0,1)$, $E(\eta_{th}\varepsilon_{it})=0$, i=1,2. $c_1,c_2>0$

Equation 1.33 gives $G=-1$, $F=c_1^2c_2^{-2}$, thus $1+d=cc_1^2c_2^{-2}$

In the 2nd degree equation 1.34

$$
\begin{aligned}
\Delta &= \left[\frac{2c_1^2c_2^{-2}c_2^2-\sigma_h^2-c_1^2c_2^{-2}\left(\sigma_h^2+2c_2^2\right)}{-c_1^2-c_1^4c_2^{-4}c_2^2}\right]^2-4\frac{-c_2^2+\sigma_h^2+2c_2^2+-c_2^2}{\left(-c_1^2-c_1^4c_2^{-4}c_2^2\right)}\\
&= \left[\frac{-\sigma_h^2-c_1^2c_2^{-2}\sigma_h^2}{-c_1^2-c_1^4c_2^{-2}}\right]^2+4\frac{\sigma_h^2}{\left(-c_1^2-c_1^4c_2^{-2}\right)}\\
&= c_1^{-4}\sigma_h^4+4c_1^{-4}\frac{\sigma_h^2}{\left(c_1^{-2}+c_2^{-2}\right)}\\
&= c_1^{-4}\sigma_h^4\left[\sigma_h^2+\frac{4}{\left(c_1^{-2}+c_2^{-2}\right)}\right]\\
&= c_1^{-4}\sigma_h^4\left[\sigma_h^2+4\left(c_1^{-2}+c_2^{-2}\right)^{-1}\right]
\end{aligned}
$$

$$
c = -\frac{1}{2}c_1^{-2}\sigma_h^2+\frac{1}{2}c_1^{-2}\sigma_h\sqrt{\sigma_h^2+4\left(c_1^{-2}+c_2^{-2}\right)^{-1}}
$$

set $\kappa=-\frac{1}{2}\sigma_h^2+\frac{\sigma_h}{2}\sqrt{\sigma_h^2+4\left(c_1^{-2}+c_2^{-2}\right)^{-1}}$ then

$$
\begin{aligned}
c &= c_1^{-2}\kappa\\
1+d &= c_2^{-2}\kappa
\end{aligned}
$$

Computation of Ω

$$
\begin{aligned}
\Omega=\Theta^{-1}C_1 &= -\begin{pmatrix} -1+c_1^{-2}\kappa & c_2^{-2}\kappa \\ c_1^{-2}\kappa & -1+c_2^{-2}\kappa \end{pmatrix}^{-1}\begin{pmatrix} c_1^2 & 0 \\ 0 & c_2^2 \end{pmatrix}\\
&= -\left[\left(-1+c_1^{-2}\kappa\right)\left(-1+c_2^{-2}\kappa\right)-c_1^{-2}\kappa c_2^{-2}\kappa\right]^{-1}\begin{pmatrix} -1+c_2^{-2}\kappa & -c_2^{-2}\kappa \\ -c_1^{-2}\kappa & -1+c_1^{-2}\kappa \end{pmatrix}\begin{pmatrix} c_1^2 & 0 \\ 0 & c_2^2 \end{pmatrix}\\
&= -\left[1-c_1^{-2}\kappa-c_2^{-2}\kappa\right]^{-1}\begin{pmatrix} c_1^2\left(-1+c_2^{-2}\kappa\right) & -\kappa \\ -\kappa & c_2^2\left(-1+c_1^{-2}\kappa\right) \end{pmatrix}\\
&= K\begin{pmatrix} c_1^2\left(1-c_2^{-2}\kappa\right) & \kappa \\ \kappa & c_2^2\left(1-c_1^{-2}\kappa\right) \end{pmatrix}
\end{aligned}
$$

With $K=\left[1-c_1^{-2}\kappa-c_2^{-2}\kappa\right]^{-1}$

The PT measure I have $\begin{pmatrix} c & 1+d \end{pmatrix} = \begin{pmatrix} c_1^{-2}\kappa & c_2^{-2}\kappa \end{pmatrix}$ so

$$PT_1 = \frac{c_1^{-2}}{c_1^{-2}+c_2^{-2}} \text{ and } PT_2 = \frac{c_2^{-2}}{c_1^{-2}+c_2^{-2}}$$

The IS bounds The total IS 1.38 using 1.39 and 1.40 is:

$$\begin{aligned}
\psi\Omega\psi' &= K_{12}\kappa^2 \begin{pmatrix} c_1^{-2} & c_2^{-2} \end{pmatrix} \begin{pmatrix} c_1^2\left(1-c_2^{-2}\kappa\right) & \kappa \\ \kappa & c_2^2\left(1-c_1^{-2}\kappa\right) \end{pmatrix} \begin{pmatrix} c_1^{-2} \\ c_2^{-2} \end{pmatrix} \\
&= K_{12}\kappa^2\left(c_1^{-2}+c_2^{-2}\right)
\end{aligned}$$

Then the Cholesky decomposition matrix F:

$$F = \sqrt{K_{12}} \begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \rho\sqrt{\sigma_{22}} & \sqrt{\sigma_{22}}\sqrt{(1-\rho^2)} \end{pmatrix}$$

I have

- The Upper bound for market 1 is

$$\begin{aligned}
IS1u &= \frac{([\psi F]_1)^2}{\psi\Omega\psi'} = \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{11}\sqrt{\sigma_{11}} + \psi_{12}\rho\sqrt{\sigma_{22}}\right)^2 \\
&= \frac{K_{12}}{\psi\Omega\psi}\kappa^2\left(c_1^{-2}\sqrt{\sigma_{11}} + c_2^{-2}\sigma_{12}\left(\sqrt{\sigma_{11}}\right)^{-1}\right)^2 \\
&= \frac{1}{c_1^{-2}+c_2^{-2}}\left(c_1^{-4}\sigma_{11} + c_2^{-4}\sigma_{12}^2\sigma_{11}^{-1} + 2c_1^{-2}c_2^{-2}\sigma_{12}\right) \\
&= \frac{1}{c_1^{-2}+c_2^{-2}}\left(c_1^{-4}c_1^2\left(1-c_2^{-2}\kappa\right) + c_2^{-4}(\kappa)^2\left(c_1^2\left(1-c_2^{-2}\kappa\right)\right)^{-1} + 2c_1^{-2}c_2^{-2}(\kappa)\right) \\
&= \frac{1}{c_1^{-2}+c_2^{-2}}\left(c_1^{-2}\left(1-c_2^{-2}\kappa\right) + c_2^{-4}\kappa^2c_1^{-2}\left(1-c_2^{-2}\kappa\right)^{-1} + 2c_1^{-2}c_2^{-2}\kappa\right) \\
&= \frac{c_1^{-2}}{c_1^{-2}+c_2^{-2}}\left(1 + c_2^{-2}\kappa + c_2^{-4}\kappa^2\left(1-c_2^{-2}\kappa\right)^{-1}\right) \\
&= \frac{c_1^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}\left(\left(1+c_2^{-2}\kappa\right)\left(1-c_2^{-2}\kappa\right) + c_2^{-4}\kappa^2\right) \\
&= \frac{c_1^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}
\end{aligned}$$

- The lower bound for market 2 is *IS2l*

$$
\begin{aligned}
IS2l &= \frac{\left([\psi F]_2\right)^2}{\psi\Omega\psi'} = \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{12}\sqrt{\sigma_{22}}\sqrt{(1-\rho^2)}\right)^2 \\
&= \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{12}^2\sigma_{22}\left(1-\rho^2\right)\right) \\
&= \frac{K_{12}}{\psi\Omega\psi}\kappa^2 c_2^{-4}c_2^2\left(1-c_1^{-2}\kappa\right)\frac{c_1^2c_2^2-c_1^2\kappa-c_2^2\kappa+\kappa^2-\kappa^2}{c_1^2c_2^2-c_1^2\kappa-c_2^2\kappa+\kappa^2} \\
&\quad \frac{K_{12}}{\psi\Omega\psi}\kappa^2 c_2^{-2}\left(1-c_1^{-2}\kappa\right)\frac{c_1^2c_2^2-c_1^2\kappa-c_2^2\kappa}{c_1^2c_2^2\left(1-c_2^{-2}\kappa\right)\left(1-c_1^{-2}\kappa\right)} \\
&= \frac{c_2^{-2}}{c_1^{-2}+c_2^{-2}}\frac{1-c_1^{-2}\kappa-c_2^{-2}\kappa}{\left(1-c_2^{-2}\kappa\right)} \\
&= \frac{c_2^{-2}}{c_1^{-2}+c_2^{-2}}\frac{\left(1-\kappa\left(c_1^{-2}+c_2^{-2}\right)\right)}{\left(1-c_2^{-2}\kappa\right)} \\
&= \frac{c_2^{-2}K_{12}^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}
\end{aligned}
$$

By switching the variable this gives $\begin{pmatrix}\psi_{11} & \psi_{22}\end{pmatrix} = \kappa\begin{pmatrix}c_2^{-2} & c_1^{-2}\end{pmatrix}$ and $\Omega = K_{12}\begin{pmatrix} c_2^2\left(1-c_1^{-2}\kappa\right) & \kappa \\ \kappa & c_1^2\left(1-c_2^{-2}\kappa\right)\end{pmatrix} = \tilde{F}\tilde{F}'$ with

$$
\tilde{F} = \sqrt{K_{12}}\begin{pmatrix}\sqrt{\sigma_{22}} & 0 \\ \rho\sqrt{\sigma_{11}} & \sqrt{\sigma_{11}}\sqrt{(1-\rho^2)}\end{pmatrix}
$$

- The Lower bound for market 1 is

$$
\begin{aligned}
IS1l &= \frac{\left([\psi\tilde{F}]_2\right)^2}{\psi\Omega\psi} = \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{11}\sqrt{\sigma_{11}}\sqrt{(1-\rho^2)}\right)^2 \\
&= \frac{c_1^{-2}K_{12}^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)}
\end{aligned}
$$

- The Upper bound for market 2 is

$$
\begin{aligned}
IS2u &= \frac{\left([\psi\tilde{F}]_1\right)^2}{\psi\Omega\psi'} = \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{11}\sqrt{\sigma_{22}}+\psi_{12}\rho\sqrt{\sigma_{11}}\right)^2 \\
&= \frac{c_2^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)}
\end{aligned}
$$

To summarize I have

The bounds for market 1 are

- $IS1u = \frac{c_1^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}$ and $IS1l = \frac{c_1^{-2}K^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)}$

And for market 2

- $IS2u = \frac{c_2^{-2}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_1^{-2}\kappa\right)}$ and $IS2l = \frac{c_2^{-2}K^{-1}}{\left(c_1^{-2}+c_2^{-2}\right)\left(1-c_2^{-2}\kappa\right)}$

The Lemma 1 and the Propositions 1,2,3 are proven.

1.8.2.3 Model II: The Roll model with a delayed market

The prices system is

$$\begin{aligned} m_{th} &= m_{th-h} + \eta_{th} \\ p_{1th} &= m_{th} + c_1 \varepsilon_{1th} \\ p_{2th} &= m_{th-\delta} + c_2 \varepsilon_{2th} \end{aligned}$$

The second market is delayed of δ.

To specify the how h moves with respect to δ, I set $\delta = b \times l$, $h = k \times l$. l is a short time pace and there is a white noise u_l, with $var\left(\mu_l\right) = \sigma^2$ and

$$\begin{aligned} m_{tkl} &= m_{tkl-l} + u_{tkl} \\ &= m_{tkl-2l} + u_{tkl} + u_{tkl-l} \\ &\vdots \\ m_{tkl} &= m_{tkl-kl} + \sum_{j=0}^{k-1} u_{tkl-jl} \end{aligned}$$

Thus $m_{th} = m_{th-h} + \eta_{th}$ with $\eta_{th} = \sum_{j=0}^{k-1} u_{tkl-jl}$ and $var\left(\eta_{th}\right) = h\sigma^2$.

When $h \geq \delta$, Δp_t is a VMA (1) and I have

$$\begin{aligned} \Delta p_{1th} = \ & \Delta m_{th} + c_1 \Delta \varepsilon_{1th} &= \ & \textstyle\sum_{j=0}^{k-1} u_{tkl-jl} + c_1 \Delta \varepsilon_{1th} \\ \Delta p_{2th} = \ & \Delta m_{th-\delta} + c_2 \Delta \varepsilon_{2th} &= \ & \textstyle\sum_{j=b}^{k+b-1} u_{tkl-jl} + c_2 \Delta \varepsilon_{2th} \end{aligned}$$

Then I easily compute

$$C_0 = \begin{pmatrix} h\sigma^2 + 2c_1^2 & (h-\delta)\sigma^2 \\ (h-\delta)\sigma^2 & h\sigma^2 + 2c_2^2 \end{pmatrix} \text{ and } C_1 = \begin{pmatrix} -c_1^2 & 0 \\ \delta\sigma^2 & -c_2^2 \end{pmatrix}$$

I compute F,G from 1.33

$$F=\frac{-m_{12}+m_{21}+v_1^2-v_{12}}{m_{12}-m_{21}+v_2^2-v_{12}} \text{ and } G=\frac{m_{12}-m_{21}-v_2^2+v_{12}}{m_{12}-m_{21}+v_2^2-v_{12}}$$

$$\begin{aligned}
F &= \frac{\delta\sigma_l^2+h\sigma_l^2+2c_1^2-(h-\delta)\,\sigma_l^2}{-\delta\sigma_l^2+h\sigma_l^2+2c_2^2-(h-\delta)\,\sigma_l^2}\\
&= \frac{2\delta\sigma_l^2+2c_1^2}{2c_2^2}\\
&= \frac{\delta\sigma_l^2+c_1^2}{c_2^2}
\end{aligned}$$

$$\begin{aligned}
G &= \frac{-\delta\sigma_l^2-h\sigma_l^2-2c_2^2+(h-\delta)\,\sigma_l^2}{-\delta\sigma_l^2+h\sigma_l^2+2c_2^2-(h-\delta)\,\sigma_l^2}\\
&= \frac{-\delta\sigma_l^2-2c_2^2-\delta\sigma_l^2}{2c_2^2}\\
&= -\frac{\delta\sigma_l^2+c_2^2}{c_2^2}
\end{aligned}$$

set $A=c_2^2$

$$C=\begin{pmatrix} h\sigma_l^2+2c_1^2 & (h-\delta)\,\sigma_l^2\\ (h-\delta)\,\sigma_l^2 & h\sigma_l^2+2c_2^2\end{pmatrix} \text{ and } M=\begin{pmatrix} -c_1^2 & 0\\ \delta\sigma_l^2 & -c_2^2\end{pmatrix}$$

Which is plugged into the elements of 1.35:

$$\begin{aligned}
&= 2A^2FGm_2+A^2Gm_{12}-A^2v_{12}-A^2Fv_2^2\\
&= 2\left(\delta\sigma_l^2+c_1^2\right)\left(\delta\sigma_l^2+c_2^2\right)c_2^2-\left(\delta\sigma_l^2+c_2^2\right)\delta\sigma_l^2-c_2^4\,(h-\delta)\,\sigma_l^2-c_2^2\left(\delta\sigma_l^2+c_1^2\right)\left(h\sigma_l^2+2c_2^2\right)\\
&= c_2^2\left[2\left(\delta\sigma_l^2+c_1^2\right)\left(\delta\sigma_l^2+c_2^2\right)-c_2^2\,(h-\delta)\,\sigma_l^2-\left(\delta\sigma_l^2+c_1^2\right)\left(h\sigma_l^2+2c_2^2\right)\right]-\left(\delta\sigma_l^2+c_2^2\right)\delta\sigma_l^2\\
&= c_2^2\left[2\delta\sigma_l^2\left(c_1^2+c_2^2\right)+2\left(\delta\sigma_l^2\right)^2+2c_1^2c_2^2-c_2^2\,(h-\delta)\,\sigma_l^2-\delta\sigma_l^2h\sigma_l^2-\delta\sigma_l^2 2c_2^2-c_1^2h\sigma_l^2-2c_1^2c_2^2\right]\\
&\quad -\left(\delta\sigma_l^2+c_2^2\right)\delta\sigma_l^2\\
&= c_2^2\sigma_l^2\left[2\delta c_1^2+2\delta^2\sigma_l^2-c_2^2\,(h-\delta)-\delta h\sigma_l^2-c_1^2h\right]-\left(\delta\sigma_l^2+c_2^2\right)\delta\sigma_l^2
\end{aligned}$$

$$
\begin{aligned}
&= A^2G^2m_2 - A^2Gv_2^2 + A^2m_2 \\
&= \left(-c_2^2\right)\left(\delta\sigma_l^2 + c_2^2\right)^2 + c_2^2\left(\delta\sigma_l^2 + c_2^2\right)\left(h\sigma_l^2 + 2c_2^2\right) - c_2^4c_2^2 \\
&= \left(-c_2^2\right)\left[\left(\delta\sigma_l^2 + c_2^2\right)^2 - \left(\delta\sigma_l^2 + c_2^2\right)\left(h\sigma_l^2 + 2c_2^2\right) + c_2^4\right] \\
&= \left(-c_2^2\right)\left[2\delta\sigma_l^2c_2^2 + \left(\delta\sigma_l^2\right)^2 + \left(c_2^2\right)^2 - \left(\delta\sigma_l^2 + c_2^2\right)h\sigma_l^2 - \left(\delta\sigma_l^2 + c_2^2\right)2c_2^2 + c_2^4\right] \\
&= \left(-c_2^2\right)\sigma_l^2\left[\delta^2\sigma_l^2 - \left(\delta\sigma_l^2 + c_2^2\right)h\right] \\
&= \left(-c_2^2\right)\sigma_l^2\left[\delta\sigma_l^2\left(\delta - h\right) - c_2^2h\right] \\
&= c_2^2\sigma_l^2\left[\delta\sigma_l^2\left(h - \delta\right) + c_2^2h\right]
\end{aligned}
$$

$$
\begin{aligned}
&= A^2m_1 + A^2F^2m_2 + A^2Fm_{12} \\
&= -c_2^4c_1^2 - \left(\delta\sigma_l^2 + c_1^2\right)^2c_2^2
\end{aligned}
$$

$$
\begin{aligned}
\Delta &= \left[\frac{2\left(F\right)\left(G\right)m_2 + \left(G\right)\left(m_{12}\right) - v_{12} - \left(F\right)v_2^2}{m_1 + \left(F\right)^2m_2 + \left(F\right)\left(m_{12}\right)}\right]^2 - 4\frac{\left(G\right)^2m_2 - \left(G\right)v_2^2 + m_2}{m_1 + \left(F\right)^2m_2 + \left(F\right)\left(m_{12}\right)} \\
&= \left[\frac{-\sigma_l^2\left[\left(\delta\sigma_l^2 + c_2^2\right)\left(\sigma_l^2\delta\left(h - \delta\right) + h\left(c_1^2 + c_2^2\right)\right) + 2c_1^2c_2^2\delta\right]}{-c_2^4c_1^2 - \left(\delta\sigma_l^2 + c_1^2\right)^2c_2^2}\right]^2 + 4\frac{c_2^2\sigma_l^2\left[\sigma_l^2\delta\left(h - \delta\right) + hc_2^2\right]}{c_2^4c_1^2 + \left(\delta\sigma_l^2 + c_1^2\right)^2c_2^2} \\
&= \left[\frac{\sigma_l^2\left[\left(\delta\sigma_l^2 + c_2^2\right)\left(\sigma_l^2\delta\left(h - \delta\right) + h\left(c_1^2 + c_2^2\right)\right) + 2c_1^2c_2^2\delta\right]}{c_2^4c_1^2 + \left(\delta\sigma_l^2 + c_1^2\right)^2c_2^2}\right]^2 + 4\frac{\sigma_l^2\left[\sigma_l^2\delta\left(h - \delta\right) + hc_2^2\right]}{c_2^2c_1^2 + \left(\delta\sigma_l^2 + c_1^2\right)^2}
\end{aligned}
$$

$$
\begin{aligned}
c &= -\frac{1}{2}\frac{\sigma_l^2\left[\left(\delta\sigma_l^2 + c_2^2\right)\left(\sigma_l^2\delta\left(h - \delta\right) + h\left(c_1^2 + c_2^2\right)\right) + 2c_1^2c_2^2\delta\right]}{c_2^4c_1^2 + \left(\delta\sigma_l^2 + c_1^2\right)^2c_2^2} \pm \frac{1}{2}\sqrt{\Delta} \\
d &= c\frac{\delta\sigma_l^2 + c_1^2}{c_2^2} - \frac{\delta\sigma_l^2 + c_2^2}{c_2^2}
\end{aligned}
$$

then

$$\begin{aligned}
\Omega = \Theta^{-1} C_1 &= -\begin{pmatrix} -1+c & d \\ c & -1+d \end{pmatrix}^{-1} \begin{pmatrix} -c_1^2 & 0 \\ \delta\sigma_l^2 & -c_2^2 \end{pmatrix} \\
&= (1-c-d)^{-1} \begin{pmatrix} -1+d & -d \\ -c & -1+c \end{pmatrix} \begin{pmatrix} -c_1^2 & 0 \\ \delta\sigma_l^2 & -c_2^2 \end{pmatrix} \\
&= (1-c-d)^{-1} \begin{pmatrix} -c_1^2(-1+d) - d\delta\sigma_l^2 & dc_2^2 \\ cc_1^2 + (-1+c)\,\delta\sigma_l^2 & -(-1+c)\,c_2^2 \end{pmatrix} \\
&= \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}
\end{aligned}$$

And then the PT and the IS are computed by replacing in 1.39, 1.40, 1.37.

The PT measure

$$PT_1 = \frac{c}{1+c+d} \text{ and } PT_2 = \frac{1+d}{1+c+d} \tag{1.42}$$

The information share bounds The total IS is

$$\begin{aligned}
\psi\Omega\psi' &= \begin{pmatrix} c & 1+d \end{pmatrix} \Omega \begin{pmatrix} c \\ 1+d \end{pmatrix} \\
&= \left[c^2\sigma_{11} + 2\sigma_{12}c\,(1+d) + (1+d)^2\,\sigma_{22}\right] \\
&= c^2 \times (-1+c+d)^{-1}\left(c_1^2(-1+d) + d\delta\sigma^2\right) + 2dc_2^2\,(1-c-d)^{-1}\,c\,(1+d) - (1-c-d)^{-1}\,(-1+c)\,c_2^2
\end{aligned}$$

The IS bounds for market 1 are

$$\begin{aligned}
IS1u &= \frac{\left(c^2(-1+c+d)^{-1}\left(c_1^2(-1+d)+d\delta\sigma^2\right) + \frac{(1+d)^2\left(dc_2^2(1-c-d)^{-1}\right)^2}{\left((-1+c+d)^{-1}\left(c_1^2(-1+d)+d\delta\sigma^2\right)\right)} / + 4c(1+d)dc_2^2\,(1-c-d)^{-1}\right)}{c^2 \times (-1+c+d)^{-1}\left(c_1^2(-1+d)+d\delta\sigma^2\right) + 2dc_2^2\,(1-c-d)^{-1}\,c\,(1+d) - (1-c-d)^{-1}\,(-1+c)\,c_2^2} \\
IS1l &= \frac{c^2(-1+c+d)^{-1}\left(c_1^2(-1+d)+d\delta\sigma^2\right)\left(1 - \frac{(1+d)^2\left(dc_2^2(1-c-d)^{-1}\right)^2}{\left(c_1^2(-1+d)+d\delta\sigma^2\right)\left((-1+c)c_2^2\right)}\right)}{c^2 \times (-1+c+d)^{-1}\left(c_1^2(-1+d)+d\delta\sigma^2\right) + 2dc_2^2\,(1-c-d)^{-1}\,c\,(1+d) - (1-c-d)^{-1}\,(-1+c)\,c_2^2}
\end{aligned}$$

and for market 2

$$
\begin{aligned}
IS2u &= 1-\frac{c^{2}\left(-1+c+d\right)^{-1}\left(c_{1}^{2}\left(-1+d\right)+d\delta\sigma^{2}\right)\left(1-\frac{(1+d)^{2}\left(dc_{2}^{2}(1-c-d)^{-1}\right)^{2}}{\left(c_{1}^{2}(-1+d)+d\delta\sigma^{2}\right)\left((-1+c)c_{2}^{2}\right)}\right)}{c^{2}\times\left(-1+c+d\right)^{-1}\left(c_{1}^{2}\left(-1+d\right)+d\delta\sigma^{2}\right)+2dc_{2}^{2}\left(1-c-d\right)^{-1}c\left(1+d\right)-\left(1-c-d\right)^{-1}\left(-1+c\right)c_{2}^{2}} \\
IS2l &= 1-\frac{\left(c^{2}\left(-1+c+d\right)^{-1}\left(c_{1}^{2}\left(-1+d\right)+d\delta\sigma^{2}\right)+\frac{(1+d)^{2}\left(dc_{2}^{2}(1-c-d)^{-1}\right)^{2}}{\left((-1+c+d)^{-1}\left(c_{1}^{2}(-1+d)+d\delta\sigma^{2}\right)\right)}/+4c(1+d)dc_{2}^{2}\left(1-c-d\right)^{-1}\right)}{c^{2}\times\left(-1+c+d\right)^{-1}\left(c_{1}^{2}\left(-1+d\right)+d\delta\sigma^{2}\right)+2dc_{2}^{2}\left(1-c-d\right)^{-1}c\left(1+d\right)-\left(1-c-d\right)^{-1}\left(-1+c\right)c_{2}^{2}}
\end{aligned}
$$

1.8.2.4 Model III: two markets with public and private information

I have $\lambda_h, \sigma\left(h\right) \xrightarrow{h\to 0} 0$ and

$$
\begin{aligned}
m_t &= m_{t-h}+\lambda_h\eta_{1th}+\eta_{th} \\
p_{1th} &= m_{th}+\eta_{1th}+c_1\varepsilon_{1th} \\
P_{2th} &= m_{th-h}+c_2\varepsilon_{2th}
\end{aligned}
\tag{1.43}
$$

thus

$$
C_0=\begin{pmatrix}\left(\lambda_h+1\right)^2+1+\sigma_h^2+2c_1^2 & -\lambda_h \\ -\lambda_h & \lambda_h^2+\sigma_h^2+2c_2^2\end{pmatrix} \text{ and } C_1=\begin{pmatrix}-\left(\lambda_h+1\right)-c_1^2 & 0 \\ \lambda_h\left(\lambda_h+1\right)+\sigma_h^2 & -c_2^2\end{pmatrix}
\tag{1.44}
$$

Using the equation 1.33

$$
\begin{aligned}
d\left(m_{12}-m_{21}+v_2^2-v_{12}\right) &= c\left(-m_{12}+m_{21}+v_1^2-v_{12}\right)+m_{12}-m_{21}-v_2^2+v_{12} \\
d\left(-\lambda_h\left(\lambda_h+1\right)-\sigma_h^2+\lambda_h^2+\sigma_h^2+2c_2^2+\lambda_h\right) &= c\left(\lambda_h\left(\lambda_h+1\right)+\sigma_h^2+\left(\lambda_h+1\right)^2+1+\sigma_h^2+2c_1^2+\lambda_h\right) \\
&\quad -\lambda_h\left(\lambda_h+1\right)-\sigma_h^2-\lambda_h^2-\sigma_h^2-2c_2^2-\lambda_h \\
d\left(2c_2^2\right) &= c\left(2\left(\lambda_h+1\right)^2+2\sigma_h^2+2c_1^2\right)-2\sigma_h^2-2\lambda_h^2-2c_2^2-2\lambda_h \\
dc_2^2 &= c\left[\left(\lambda_h+1\right)^2+\sigma_h^2+c_1^2\right]-\left(\sigma_h^2+\lambda_h^2+c_2^2+\lambda_h\right) \\
d &= cc_2^{-2}\left[\left(\lambda_h+1\right)^2+\sigma_h^2+c_1^2\right]-c_2^{-2}\left(\sigma_h^2+\lambda_h^2+c_2^2+\lambda_h\right) \\
d &= cF+G
\end{aligned}
$$

$$
F=c_2^{-2}\left[\left(\lambda_h+1\right)^2+\sigma_h^2+c_1^2\right] \quad \text{and} \quad G=-c_2^{-2}\left(\sigma_h^2+\lambda_h^2+c_2^2+\lambda_h\right)
$$

For $h \simeq 0$ I consider the development at the order of λ_h and σ_h. That is $F = c_2^{-2}\left(1+c_1^2\right)$, $G = -\left(1+c_2^{-2}\lambda_h\right)$ thus $d = c_2^{-2}\left(1+c_1^2\right)c - \left(1+c_2^{-2}\lambda_h\right)$ and I have

$$C_0 = \begin{pmatrix} 2\lambda_h + 2 + 2c_1^2 & -\lambda_h \\ -\lambda_h & 2c_2^2 \end{pmatrix} \quad \text{and} \quad C_1 = \begin{pmatrix} -1-c_1^2 & 0 \\ \lambda_h & -c_2^2 \end{pmatrix}$$

In equation 1.34

$$m_1 + F^2 m_2 + F\left(m_{12}+m_{21}\right) = -\left(1+c_1^2\right)\left[1+c_2^{-2}\left(1+c_1^2\right) - c_2^{-2}\lambda_h\right]$$

$$2FGm_2 + G\left(m_{12}+m_{21}\right) - v_{12} - Fv_2^2 = \left(c_2^{-2}\lambda_h\right)\left(2\left(1+c_1^2\right)-\lambda_h\right)$$

$$G^2 m_2 - Gv_2^2 + m_2 = -c_2^2\left(2+c_2^{-2}\lambda_h\right)^2$$

$$\begin{aligned}
\Delta &= \left(c_2^{-2}\lambda_h\right)^2\left(2\left(1+c_1^2\right)-\lambda_h\right)^2 - 4c_2^2\left(2+c_2^{-2}\lambda_h\right)^2\left(1+c_1^2\right)\left[1+c_2^{-2}\left(1+c_1^2\right)-c_2^{-2}\lambda_h\right] \\
&= c_2^{-4}\lambda_h^2\left(4\left(1+c_1^2\right)^2+\lambda_h^2-4\left(1+c_1^2\right)\lambda_h\right) - 4\left(4+4c_2^{-2}\lambda_h+c_2^{-4}\lambda_h^2\right)c_2^2\left(1+c_1^2\right)\left[1+c_2^{-2}\left(1+c_1^2\right)\right] \\
&\quad +4\left(4+4c_2^{-2}\lambda_h+c_2^{-4}\lambda_h^2\right)c_2^2\left(1+c_1^2\right)c_2^{-2}\lambda_h
\end{aligned}$$

$$\begin{aligned}
c &= \frac{c_2^{-2}\lambda_h\left(1+c_1^2\right)+\sqrt{\lambda_h\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}}{\left(1+c_1^2\right)\left[1+c_2^{-2}\left(1+c_1^2\right)-c_2^{-2}\lambda_h\right]} \\
d &= c_2^{-2}\left(1+c_1^2\right)c - \left(1+c_2^{-2}\lambda_h\right) \\
&= c_2^{-2}\left(1+c_1^2\right)\frac{c_2^{-2}\lambda_h\left(1+c_1^2\right)+\sqrt{\lambda_h\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}}{\left(1+c_1^2\right)\left[1+c_2^{-2}\left(1+c_1^2\right)-c_2^{-2}\lambda_h\right]} - c_2^{-2}\lambda_h - 1 \\
1+d &= c_2^{-2}\frac{c_2^{-2}\lambda_h\left(1+c_1^2\right)+\sqrt{\lambda_h\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}}{\left[1+c_2^{-2}\left(1+c_1^2\right)-c_2^{-2}\lambda_h\right]} - c_2^{-2}\lambda_h
\end{aligned}$$

I go at the order $\sqrt{\lambda_h}$

$$c = \frac{\sqrt{\lambda_h}}{\sqrt{\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}} \quad \text{and} \quad 1+d = \frac{c_2^{-2}\left(1+c_1^2\right)\sqrt{\lambda_h}}{\sqrt{\left(1+c_1^2\right)}\sqrt{1+c_2^{-2}\left(1+c_1^2\right)}}$$

The variance 1.36 $\Omega = [c+d]^{-1}\begin{pmatrix} d\left(1+c_1^2\right) & -(1+d)c_2^2 \\ -c\left(1+c_1^2\right) & (-1+c)\,c_2^2 \end{pmatrix}$

The PT measure

$$PT_1 = \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} \quad \text{and} \quad PT_2 = \frac{c_2^{-2}\left(1+c_1^2\right)}{1+c_2^{-2}\left(1+c_1^2\right)}$$

The information share bounds The total IS is

$$\begin{aligned}
\psi\Omega\psi' &= [c+d]^{-1}\begin{pmatrix} c & 1+d \end{pmatrix}\begin{pmatrix} d\left(1+c_1^2\right) & -c\left(1+c_1^2\right) \\ -c\left(1+c_1^2\right) & (-1+c)\,c_2^2 \end{pmatrix}\begin{pmatrix} c \\ 1+d \end{pmatrix} \\
&= [c+d]^{-1}\begin{pmatrix} c & 1+d \end{pmatrix}\begin{pmatrix} d\left(1+c_1^2\right)c - c(1+d)\left(1+c_1^2\right) \\ -c^2\left(1+c_1^2\right)+(1+d)(-1+c)\,c_2^2 \end{pmatrix} \\
&= [c+d]^{-1}\begin{pmatrix} c & c_2^{-2}\left(1+c_1^2\right)c \end{pmatrix}\begin{pmatrix} -c\left(1+c_1^2\right) \\ -c\left(1+c_1^2\right) \end{pmatrix} \\
&= -[c+d]^{-1}c^2\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right) \\
&= -[c+d]^{-1}\lambda_h
\end{aligned}$$

Then the Cholesky matrix F:

$$F = \sqrt{K}\begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \rho\sqrt{\sigma_{22}} & \sqrt{\sigma_{22}}\sqrt{(1-\rho^2)} \end{pmatrix}$$

I have

- The Upper bound for market 1 is

$$
\begin{aligned}
IS1u &= \frac{([\psi F]_1)^2}{\psi\Omega\psi'} = \frac{K}{\psi\Omega\psi} D^2 \left(\sqrt{\sigma_{11}} + c_2^{-2}\left(1+c_1^2\right)\rho\sqrt{\sigma_{22}}\right)^2 \\
&= \frac{KD^2}{\psi\Omega\psi}\left(\sqrt{\sigma_{11}} + c_2^{-2}\left(1+c_1^2\right)\sigma_{12}\left(\sqrt{\sigma_{11}}\right)^{-1}\right)^2 \\
&= \frac{KD^2}{\psi\Omega\psi}\left(\sigma_{11} + \left(c_2^{-2}\left(1+c_1^2\right)\right)^2 \sigma_{12}^2\sigma_{11}^{-1} + 2\left(c_2^{-2}\left(1+c_1^2\right)\right)\sigma_{12}\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\left(d\left(1+c_1^2\right) + \left(c_2^{-2}\left(1+c_1^2\right)\right)^2\left(1+c_1^2\right)^2 d^{-1}\left(1+c_1^2\right)^{-1}\right) \\
&\quad -2\left(c_2^{-2}\left(1+c_1^2\right)\right)\left(1+c_1^2\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\left(1+c_1^2\right)\left(d + \left(c_2^{-2}\left(1+c_1^2\right)\right)^2 d^{-1} - 2c_2^{-2}\left(1+c_1^2\right)\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\left(1+c_1^2\right)\left(d + \left(1+2d+d^2\right)d^{-1} - 2\left(1+d\right)\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\left(1+c_1^2\right)\left(d^{-1}\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\frac{\left(1+c_1^2\right)}{d}
\end{aligned}
$$

- The lower bound for market 2 is $IS2l$

$$
\begin{aligned}
IS2l &= \frac{([\psi F]_2)^2}{\psi\Omega\psi'} = \frac{K}{\psi\Omega\psi}\left(\psi_{12}\sqrt{\sigma_{22}}\sqrt{(1-\rho^2)}\right)^2 \\
&= \frac{K}{\psi\Omega\psi}\psi_{12}^2\sigma_{22}\left(1-\rho^2\right) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2\sigma_{22}\left(1-\sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1}\right) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2(-1+c)c_2^2\left(1-c^2\left(1+c_1^2\right)^2 d^{-1}\left(1+c_1^2\right)^{-1}(-1+c)^{-1}c_2^{-2}\right) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2(-1+c)c_2^2(-1+c)^{-1}d^{-1}\left(d(-1+c) - c^2\left(1+c_1^2\right)c_2^{-2}\right) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2(-1+c)c_2^2(-1+c)^{-1}d^{-1}\left(d(-1+c) - c(1+d)\right) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2(-1+c)c_2^2(-1+c)^{-1}d^{-1}(-c-d) \\
&= \frac{K}{\psi\Omega\psi}(1+d)^2 c_2^2 d^{-1}(-c-d) \\
&= -\frac{K}{\psi\Omega\psi}\frac{(c+d)}{d}(1+d)^2 c_2^2
\end{aligned}
$$

By switching the variable this gives $\begin{pmatrix} \psi_{11} & \psi_{22} \end{pmatrix} = \begin{pmatrix} 1+d & c \end{pmatrix}$
and $\Omega = K \begin{pmatrix} (-1+c)\,c_2^2 & -c\left(1+c_1^2\right) \\ -c\left(1+c_1^2\right) & d\left(1+c_1^2\right) \end{pmatrix} = \tilde{F}\tilde{F}'$ with $\tilde{F} = \sqrt{K} \begin{pmatrix} \sqrt{\sigma_{22}} & 0 \\ \rho\sqrt{\sigma_{11}} & \sqrt{\sigma_{11}}\sqrt{(1-\rho^2)} \end{pmatrix}$

- The Lower bound for market 1 is

$$
\begin{aligned}
IS1l &= \frac{\left([\psi\tilde{F}]_2\right)^2}{\psi\Omega\psi} = \frac{K}{\psi\Omega\psi}\left(\psi_{11}\sqrt{\sigma_{11}}\sqrt{(1-\rho^2)}\right)^2 \\
&= \frac{K}{\psi\Omega\psi}\psi_{11}^2\sigma_{11}\left(1-\rho^2\right) \\
&= \frac{KD^2}{\psi\Omega\psi}\sigma_{11}\left(1-\sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1}\right) \\
&= \frac{KD^2}{\psi\Omega\psi}d\left(1+c_1^2\right)(-1+c)^{-1}d^{-1}(-c-d) \\
&= -\frac{KD^2}{\psi\Omega\psi}\left(1+c_1^2\right)\left(\frac{c+d}{-1+c}\right)
\end{aligned}
$$

- The Upper bound for market 2 is

$$
\begin{aligned}
IS2u &= \frac{\left([\psi\tilde{F}]_1\right)^2}{\psi\Omega\psi'} = \frac{K_{12}}{\psi\Omega\psi}\left(\psi_{22}\sqrt{\sigma_{22}}+\psi_{11}\rho\sqrt{\sigma_{11}}\right)^2 \\
&= \frac{K}{\psi\Omega\psi}\left((1+d)\sqrt{\sigma_{22}}+c\sigma_{12}\left(\sqrt{\sigma_{22}}\right)^{-1}\right)^2 \\
&= \frac{K}{\psi\Omega\psi}\left((1+d)^2\sigma_{22}+c^2\sigma_{12}^2(\sigma_{22})^{-1}+2c(1+d)\sigma_{12}\right) \\
&= \frac{K}{\psi\Omega\psi}\left((1+d)^2(-1+c)\,c_2^2+c^2c^2\left(1+c_1^2\right)^2(-1+c)^{-1}c_2^{-2}-2c^2(1+d)\left(1+c_1^2\right)\right) \\
&= \frac{K}{\psi\Omega\psi}\left(c_2^{-4}\left(1+c_1^2\right)^2c^2(-1+c)\,c_2^2+c^2c^2\left(1+c_1^2\right)^2(-1+c)^{-1}c_2^{-2}-2c^2c_2^{-2}\left(1+c_1^2\right)c\left(1+c_1^2\right)\right) \\
&= \frac{K}{\psi\Omega\psi}\left(1+c_1^2\right)^2c^2c_2^{-2}\left((-1+c)+c^2(-1+c)^{-1}-2c\right) \\
&= \frac{K}{\psi\Omega\psi}\left(1+c_1^2\right)^2c^2c_2^{-2}(-1+c)^{-1}\left((-1+c)^2+c^2-2c(-1+c)\right) \\
&= \frac{K}{\psi\Omega\psi}\left(1+c_1^2\right)^2c^2c_2^{-2}(-1+c)^{-1}\left(c^2-2c+1+c^2+2c-2c^2\right) \\
&= \frac{K}{\psi\Omega\psi}\left(1+c_1^2\right)^2c^2c_2^{-2}(-1+c)^{-1} \\
&= \frac{K}{\psi\Omega\psi}\frac{c^2}{(-1+c)}\left(1+c_1^2\right)^2c_2^{-2}
\end{aligned}
$$

To summarize I have $K = [c+d]^{-1} = \left[D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1\right]^{-1}$, $c = D$, $1+d = c_2^{-2}\left(1+c_1^2\right)D$
, $\psi\Omega\psi = -K\lambda_h$

$$D=\sqrt{\lambda_h}\left(\sqrt{\left(1+c_1^2\right)\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}\right)^{-1}$$

- $IS1u=-\frac{1}{c_2^{-2}\left(1+c_1^2\right)D-1}\times\frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}$
- $IS1l=\left(\frac{D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1}{D-1}\right)\frac{1}{1+c_2^{-2}\left(1+c_1^2\right)}$
- $IS2u=\frac{-\left(1+c_1^2\right)c_2^{-2}}{(-1+D)}\times\frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}$
- $IS2l=\frac{\left[D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1\right]}{Dc_2^{-2}\left(1+c_1^2\right)-1}c_2^{-2}\left(1+c_1^2\right)\frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}$

When $h\longrightarrow 0$ then $D\longrightarrow 0$ and

$$\begin{array}{rclcll}
IS1u & = & -\frac{1}{c_2^{-2}\left(1+c_1^2\right)D-1}\times\frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)} & = & \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} & = PT_1\\
IS1l & = & \left(\frac{D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1}{D-1}\right)\frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} & = & \frac{1}{1+c_2^{-2}\left(1+c_1^2\right)} & = PT_1\\
IS2u & = & \frac{-\left(1+c_1^2\right)c_2^{-2}}{(-1+D)}\times\frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)} & = & \frac{\left(1+c_1^2\right)c_2^{-2}}{1+c_2^{-2}\left(1+c_1^2\right)} & = PT_2\\
IS2l & = & \frac{1}{\left(1+c_2^{-2}\left(1+c_1^2\right)\right)}\frac{\left[D\left(1+c_2^{-2}\left(1+c_1^2\right)\right)-1\right]}{Dc_2^{-2}\left(1+c_1^2\right)-1}c_2^{-2}\left(1+c_1^2\right) & & \frac{\left(1+c_1^2\right)c_2^{-2}}{1+c_2^{-2}\left(1+c_1^2\right)} & = PT_2
\end{array}$$

I obtain also here that the IS bound and PT are similar at high frequency.

1.8.2.5 Proof of Proposition 1.4

Proof. If $1\geq\alpha_1$ then

$$\kappa=-\frac{\sigma^2h}{2}+\frac{\sigma h}{2}\sqrt{4c_1'^2c_2'^2h^{\alpha_1}\left(c_1'^2h^{\alpha_1-\alpha_2}+c_2'^2\right)^{-1}}\simeq-\frac{\sigma^2h}{2}+\sigma c_1'c_2'h^{0.5+\alpha_1}\left(c_1'^2h^{\alpha_1-\alpha_2}+c_2'^2\right)^{-1/2}$$
$$\longrightarrow 0$$

$$\begin{aligned}
K&=\left[1-\kappa h^{-\alpha_1}\left(c_1'^{-2}+c_2'^{-2}h^{\alpha_1-\alpha_2}\right)\right]^{-1}\\
&=\left[1-\left(-\frac{1}{2}\sigma^2h+\sigma c_1'c_2'h^{0.5+\alpha_1}\left(c_1'^2h^{\alpha_1-\alpha_2}+c_2'^2\right)^{-1/2}\right)h^{-\alpha_1}\left(c_1'^{-2}+c_2'^{-2}h^{\alpha_1-\alpha_2}\right)\right]^{-1}\\
&\backsimeq\left[1-\left(-\frac{1}{2}\sigma^2hh^{-\alpha_1}+\sigma c_1'c_2'h^{0.5+\alpha_1}h^{-\alpha_1}\left(c_1'^2h^{\alpha_1-\alpha_2}+c_2'^2\right)^{-1/2}\right)\left(c_1'^{-2}+c_2'^{-2}h^{\alpha_1-\alpha_2}\right)\right]^{-1}\\
&\backsimeq\left[1+\frac{1}{2}\sigma^2c_1'^{-2}h^{1-\alpha_1}-\sigma c_1c_2h^{0.5}c_2'^{-1}c_1'^{-2}\right]\\
&\longrightarrow 1
\end{aligned}$$

□

Chapter 2

Adjustment of the permanent price on interlinked markets

Abstract: I provide a new way to evaluate price adjustment across linked markets by building an Impulse Response measuring the permanent impact of market's innovation and I give its asymptotic distribution. The framework seems to be the first to provide testable results for price discovery measures based on Hasbrouck (1995) innovation variance and gives a rationale to the Information Share Upper bound. I later present an equilibrium model of different maturities futures markets with convenience yield and I show that it supports my measure: As Garbade and Silber (1983), adjustment is driven by the number of participants in each market. An application on some metals of the London Metal Exchange shows that some markets are in Backwardation and others in Contago. And that, 3-month futures contract dominates the spot and the 15-month in price formation.

Keywords: *Backwardation, Contago, Generalized Impulse Responses, Information Share, Price discovery,*

JEL: C32,G13,G14

2.1 Introduction

The multiplication of financial derivatives and markets places resurges the interest in understanding the path of information shocks in financial markets. Information becomes critically important when it is not shared by all the market participants and they would like prices to reflect what is known by others liquidity suppliers at a given point in time. Investors trading at that time would benefit from a fair price, without a shortfall resulting from adverse selection when other side of the market has a private information. The markets authorities activity involves engaging reforms to make information available to market participants and ameliorating the informational quality of prices. This put them in the need of a permanent assessment of information diffusion trough the market and strengthen the importance of having measures of the information carried by prices.

For assets listed in many markets or securities linked by strong arbitrage relationships, the concern is to evaluate how each market contributes to the adjustment of the permanent price: the so-called price discovery mechanism. For this purpose, some measures were developed in the literature triggered by Hasbrouck (1995). He presented a measure of price discovery called Information Share (IS) and provides comparison of New York Stock Exchange and the Regional exchanges in the quotes formation of thirty Dow stocks. The main competing measure is the common factor weight of Gonzalo and Granger (1995) permanent-transitory (PT) decomposition (see Harris et al., 2002b). Those methods are intensively discussed by De Jong (2002), Lehmann (2002), Hasbrouck (2002),Baillie et al. (2002),Yan and Zivot (2010). The main lesson is that the IS is more concerned with the variability in the process with an economic sensitive identification of its efficient price. The IS measure suggests to evaluate the market contribution to price discovery by the relative part of this market in the variance of the innovation in the efficient price. Meanwhile IS has some drawbacks; it is not identified and is only able to produce upper and lower bound, and sometimes those bounds can be very wide.

Many authors tried to solve this identification issue by doing some transformations of the innovation variance matrix. But the limit of those techniques is that they completely lost an economic meaning behind the mathematical operations. For example, Lien and Shrestha (2014) use an orthogonalization of the correlation matrix to propose the Generalized Information Share (GIS), this measure has the advantage of being independent of the variable ordering and is applicable to CDS and Bond as in their application. Meanwhile, the orthogonalization procedure of the correlation matrix lacks some economic intuition. This is the same limitation with methods based on heteroskedasticity as in Grammig and Peter (2013). They exploit "tail dependence" for identification through heteroskedasticity on two regimes (Rigobon, 2003; Lanne and Lütkepohl, 2010).

In this study, I build the Impulse response Information Share (IRIS) based on the definition of Generalized Impulse Response Function (GIRF) of Pesaran and Shin (1998). I propose to evaluate the informational content of the price of a given market by measuring the response of the permanent component of prices to a shock in this market. My framework has the advantage that it provides a unique value and has a straightforward economic interpretation. In contrary to other methods based on Information Share, I provide a limiting distribution to test the significance of the responses. While also making use of the cointegration properties, the setup doesn't impose the cointegrating coefficient to be one, and it provides a rationale for the practical choice of Hasbrouck (1995) IS upper bound when facing the identification issue. Then, I present an equilibrium cost-of- carry model of futures markets with convenience yield. I show that the model supports my measure: As the theoretical result of Garbade and Silber (1983) andFiguerola-Ferretti and Gonzalo (2010), my measure selects the market with the higher number of participants as dominating the price discovery.

The remainder of the paper is organized as follow. The section 2 presents the modeling framework; the IRIS measure, the estimation and asymptotic theory are constructed. It finishes by some Monte-Carlo exercises showing how it performs. In section 3 a theoretical cost-of-carry model is presented and shows that IRIS has theoretical relevance. In section 4 I present an application to study the spot, 3 months and 15 months futures contracts on some metals traded at the London Metal Exchange (LME). One of the primary functions of futures contract design is to improve price discovery. I obtain that only the 3-month contract serves that purposes, the 15-month contract being less relevant.

2.2 Modeling and estimation

2.2.1 Setup

There is d strongly related securities that are traded at the respective prices $p_{1t}\dots$ and p_{dt}. For example, for one asset listed on two markets, p_{1t} is the price of the asset on the first market and p_{2t} the price on the second. I denote the vector of prices by $P_t = (p_{1t}, p_{2t}\dots p_{dt})'$. In this situation it is classical that P_t is assumed to be cointegrated and the gap between every pair of prices is stationary such that there exist only one common trend for all prices. Using Johansen (1991) results, P_t can be shown to admit the following Vector error correction model (VECM) representation:

$$\Delta P_t = -\alpha\beta' P_{t-1} + \Gamma_1 \Delta P_{t-1} + \dots + \Gamma_K \Delta P_{t-K} + \varepsilon_t \tag{2.1}$$

where the cointegrating matrix is normalized here to be

$$\beta' = \begin{bmatrix} 1 & -\beta_1 & \dots & 0 & 0 \\ 0 & 1 & -\beta_2 & & 0 \\ & & \dots & & \\ 0 & 0 & & 1 & -\beta_{d-1} \end{bmatrix} : (d-1)\times d \tag{2.2}$$

The infinite moving average representation of the price vector difference and the Granger representation theorem give the following relationships where $\Psi(L)$ is a lag polynomial and ε_t independent white noise with $var(\varepsilon_t) = \Omega$:

$$\Delta P_t = \Psi(L)\varepsilon_t = (\Psi(1) + \Psi^*(L)(1-L))\,\varepsilon_t \tag{2.3}$$

$$P_t = P_0 + \Psi(1)\sum_{s=1}^{t}\varepsilon_s + \Psi^*(L)\varepsilon_t \tag{2.4}$$

The matrix of the long run impact is given by

$$\Psi(1) = \beta_\perp \left(\alpha_\perp^{'}\left(I - \sum_{i=1}^{p}\Gamma_i\right)\beta_\perp\right)^{-1}\alpha_\perp^{'} \tag{2.5}$$

From now on, to simplify the presentation and keep the focus on intuitions and arguments, I restrict to $d = 2$ markets. Meanwhile the proof are done for $d > 2$.

2.2.2 The Information Share measure

For an asset that is traded on two or more venues, Hasbrouck (1995) is looking for a measure that will determine on which market the price discovery does happen. He proposed to use the contribution of each market in the variance of the innovation of the fundamental value (or the "efficient price" which is common to all the markets). His method relies on the assumption that the cointegrating equation is $\beta = (1\ -1)$.

So the matrix $\Psi(1)$ can be expressed with a the unique ψ as

$$\Psi(1) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\psi = \begin{pmatrix} 1 \\ 1 \end{pmatrix}\begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix}$$

Replacing in equation 2.4 yields

$$P_t = P_0 + \begin{pmatrix} 1 \\ 1 \end{pmatrix} \psi \sum_{s=1}^{t} \varepsilon_s + \Psi^*(L)\varepsilon_t \tag{2.6}$$

The random walk component of the price is $\psi \sum_{s=1}^{t} \varepsilon_s$, it is a scalar random variable and it is common to market 1 and market 2. It is identified as the implicit fundamental price of the asset. The new information entering the fundamental price is $\psi\varepsilon_t$ and its variance $(\psi\Omega\psi')$ is the total information share. he defines the market contribution in the following way:

If Ω were to be diagonal, then the total information share will be

$$\psi\Omega\psi' = \psi_{11}^2\Omega_{11} + \psi_{22}^2\Omega_{22}$$

and the information share of the market j will be

$$IS_j = \frac{\psi_{jj}^2\Omega_{jj}}{\psi\Omega\psi'}$$

As Ω is not diagonal in general, its Cholesky decomposition root is computed as $\Omega = FF'$ with F lower triangular. And I have

$$IS_j = \left([\psi F]_j\right)^2 / \psi\Omega\psi' \tag{2.7}$$

$[\psi F]_j$ is the jth element of the matrix ψF.

An identification problem arises because the ranking of the variables matters for the result. Then by switching variables position in the price vector, he provides lower and upper bounds on the Information Share of each market.

2.2.3 Invariant information Share measures

The IS identification problem can be summarized in the structural shock identification problem in the SVAR literature. In fact having the reduced form shock ε_t the goal is to look for structural shock η_t and B such that $\eta_t = B^{-1}\varepsilon_t$. Briefly, the problem is to solve for the matrix B in equation $\Omega = BB'$. Unfortunately an infinity of solutions exist and the Hasbrouck (1995) choice is to consider the lower triangular matrix F obtained from the Cholesky root of Ω. But as seen previously the IS doesn't give the same value for a market if it is placed in first and in second position in the price vector.

A solution to this is presented in Lien and Shrestha (2014) where instead of focusing on the covariance matrix Ω, they consider the eigenvalues decomposition of its correlation matrix Φ. Let

G the matrix of eigenvectors, Λ is the diagonal matrix of eigenvalues and $V = diag\left(\Omega_{11}, \Omega_{22}\right)$ the diagonal matrix of standard deviations, and

$$F^{*} = \left[G\Lambda^{-1/2}G^{T}V^{-1}\right]^{-1} \tag{2.8}$$

It happens that $\Omega = F^{*}\left(F^{*}\right)^{T}$. They thus define their Generalized Information share for market j using the matrix $B = F^{*}$

$$GIS_{j} = \left(\left[\psi F^{*}\right]_{j}\right)^{2} / \psi\Omega^{2}\psi$$

Where ψ is a line of the matrix $\Psi(1)$[1].

This method has the advantage of being independent of the variables ordering, but it strongly lacks an economic relevance behind the decomposition of the correlation matrix.

Others attempts to compute a unique Information Share may be to use non-gaussianity or heteroskedasticity to identify structural shocks as it is done in the Macroeconomics literature. Those procedures have the advantage of allowing identification of the two structural shocks. Meanwhile it is not possible to say which shock comes from which market and the parameters are identified only up to a permutation matrix. In addition to the fact that they are purely statistical identification schemes with no economics motivation, this is a severe problem for the purpose of assigning market's contribution to price discovery. To overcome this problem, Grammig and Peter (2013) after considering heteroskedasticity on two regimes of structural innovations to identify uniquely the matrix B, assign shocks in a way that the coefficient of a shock on its market should be bigger than its coefficient on the other market.

Another non less important disadvantage of method based on information share is that they lack asymptotic theory and testing. The current practice is to use some bootstrap procedures to provide standard errors on Information Share.

2.2.4 α-based measures

The main competitor to IS measure in the literature is the Gonzalo and Granger (1995) common factor weight in the Permanent-Transitory (PT) decomposition. This consist of decomposing a difference stationary time series as the sum of a permanent I(1) component Q_t and a transitory stationary component T_t. The identification of the two components of $P_t = Q_t + T_t$ relies on two assumptions:

[1]Here they don't assume the Cointegrating value to be -1 like for the Modified Information Share (MIS) in Lien and Shrestha (2009) . But GIS and MIS are analytically equal, the only difference remains in the estimation of the GIS where there is no constraints on the coefficients β_1.

- Q_t and T_t form a PT decomposition,
- T_t is a linear combination of the observed variables,

In the contest of one asset and many markets the permanent component is driven by an $I(1)$ factor (f_t) common to both markets such that the observed price vector can be written as

$$P_t = \begin{bmatrix} 1 \\ 1 \end{bmatrix} f_t + T_t$$

And it is shown that given the ECM equation 2.1, the weight in the $I(1)$ component are proportional to $\alpha_\perp$ such that:

$$f_t = c \times \alpha_\perp P_t = c \begin{pmatrix} \alpha_{1\perp} & \alpha_{2\perp} \end{pmatrix} P_t \quad \text{,with } c \text{ constant}$$

The relative contribution to price discovery of market 1 and market 2 is thus computed by taking the weight of each market in the permanent component (Harris et al., 2002a) as

$$PT_1 = \frac{\alpha_{1\perp}}{\alpha_{1\perp} + \alpha_{2\perp}}, PT_2 = \frac{\alpha_{1\perp}}{\alpha_{1\perp} + \alpha_{2\perp}}$$

A difference between the PT and the IS is that f_t is a linear combination of only the current prices. Thus the permanent component of the Gonzalo and Granger (1995) decomposition is generally not a random walk. This is a serious drawback as this permanent component could not represent an efficient price. Baillie et al. (2002) show that both can be easily computed after the estimation of the ECM and they present the relationship linking CS to IS.

Instead of focusing on the innovation variation, the permanent component Share relies on the error correction weighting matrix $\alpha_\perp$. In this respect Eun and Sabherwal (2003) also think of price discovery as the adjustment to the equilibrium, but assess price discovery of a market directly by its coefficient in α, summarizing its speed of adjustment toward the long run equilibrium. Building the measures with only a coefficient of the VECM allows those methods to have testable implications and thus statistical significance checking of the contribution to price discovery.

2.2.5 A new measure: The Impulse response Information share

The question of measuring price discovery for cross-listed assets comes from the need to know on which market information enter the prices. As information is supposed to affect permanently the prices, an appealing intuition is to say that: if information comes trough market 1 (and not through market 2), the efficient price should react to innovation in market 1 (and not to innovation in market

2). Price discovery can thus be well evaluated by the response of the efficient price to each market's innovation.[2]

2.2.5.1 The Generalized impulse response

In the case of linear VAR and Cointegrated systems, Pesaran and Shin (1998) analyzed the generalized impulse response function (GIRF) by relying on Koop et al. (1996). For a vector Z_t the Generalized Impulse Response defines the reaction of Z_t to a shock δ_j on ε_{jt}, conditional on the information set (I_{t-1}) at time $t-1$ as

$$GI_z\left(n, \delta_j, I_{t-1}\right) = E\left(Z_{t+n} | \varepsilon_{jt} = \delta_j, I_{t-1}\right) - E\left(Z_{t+n} | I_{t-1}\right)$$

This formula doesn't rely on an orthogonalization procedure (e.g Cholesky), and the interpretation is straight-forward. In fact instead of shocking all the system, only the jth variable is shocked and the effect of the other variables are integrated out.

If Z_t is d-dimensional vector having the following Moving Average representation

$$Z_t = \sum_{i=1}^{\infty} A_i \varepsilon_{t-i}$$

with ε_t has a normal distribution, the integration is easily done using the formula

$$E\left(\varepsilon_t | \varepsilon_{jt} = \delta_j\right) = \left(\sigma_{1j}, \sigma_{2j}, \ldots, \sigma_{dj}\right)' \sigma_{jj}^{-1} \delta_j = \Omega e_j \sigma_{jj}^{-1} \delta_j$$

where e_j is the vector having 1 at the jth position and 0 elsewhere.

The vector of the unscaled impulse response of the effect of a shock in the jth equation at time t on Z_{t+n} is given by

$$GI_z(n) = A_n \delta_j = A_n \Omega e_j \sigma_{jj}^{-1} \delta_j \ , n = 0,\ 1,\ 2,\ \ldots$$

Then normalizing the size of the shock to a one standard deviation $\delta_j = \sqrt{\sigma_{jj}}$ gives

$$GI_z(n) = \sigma_{jj}^{-\frac{1}{2}} A_n \Omega e_j$$

[2] The impulse response function here differs totally from the analysis of Yan and Zivot (2010). They assume two structural shocks (informational and noisy) and they look at the reaction of existing measures (IS, PT) to the informational shock. Here after identifying the same permanent price as Hasbrouck (1995) I are interested in the response of this efficient price to an innovation shock in each market.

2.2.5.2 Definition of the measure

Let's write the long run impact matrix as $\Psi(1) = \begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix}$ and design its two rows by $\psi_1 = \begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix}$ and $\psi_2 = \begin{pmatrix} \psi_{21} & \psi_{22} \end{pmatrix}$.

With the cointegrating vector $\beta = (1\ -\beta_1)$ and the properties $\beta'\Psi(1) = 0$ I have

$$\begin{pmatrix} \psi_{11} & \psi_{12} \\ \psi_{21} & \psi_{22} \end{pmatrix} = \begin{pmatrix} \psi_{11} - \beta_1\psi_{12} \\ \psi_{12} - \beta_1\psi_{22} \end{pmatrix} = 0$$

Thus $\psi_{11} = \beta_1\psi_{12}$ and $\psi_{12} = \beta_1\psi_{22}$, which implies that the second row is multiple of the first row $\psi_2 = \beta_1^{-1}\psi_1$. There is 1 cointegrating relation so the space of permanent component is of dimension 1 and $\Psi(1) = \begin{pmatrix} 1 \\ \beta_1^{-1} \end{pmatrix}\psi_1$.

The permanent component entering the first price p_{1t} is given by $Q_{1t} = \psi_1 \sum_{s=1}^{t} \varepsilon_s$ it is a random walk, the same identified by the information share when $\beta_1 = 1$. To define the measure, I compute the generalized impulse response of Q_{1t}, to a shock in the first and the second market. The response of Q_{1t} after n periods to a shock to the jth market is given by

$$\begin{aligned} GI_{Q_1}\left(n, \varepsilon_j\right) &= E\left(Q_{1t+n} | \varepsilon_{jt} = \delta_j, \Omega_{t-1}\right) - E\left(Q_{1t+n}, \Omega_{t-1}\right) \\ &= \psi_1 E\left(\varepsilon_t | \varepsilon_{jt} = \delta_j\right) \\ &= \psi_1 \Omega e_j \sigma_{jj}^{-\frac{1}{2}} \\ &= \psi_1 \sigma_{jj}^{-\frac{1}{2}} \left(\sigma_{1j}, \sigma_{2j}\right)' \\ &= \sigma_{jj}^{-\frac{1}{2}} \left(\psi_{11}\sigma_{1j} + \psi_{12}\sigma_{2j}\right) \end{aligned}$$

The horizon n disappears from the formula at the second equality thanks to the random walk nature of Q_{1t}.

The square of the impulse response gives the variance of the permanent component forecast error resulting from the shock in the jth market. As the IS using the variance, it is a good summary of the permanent information entering the prices by market j:

$$GI_{q_1}(\varepsilon_j)^2 = \sigma_{jj}^{-1}\left(\psi_{11}\sigma_{1j} + \psi_{12}\sigma_{2j}\right)^2$$

This value can be compare for different market to see where information enter the price. The market with the biggest value of $GI_{q_1}^2$ is designed as the dominant market. To express the result in term of percentage (such that the results sum-up to one) and to compare with other measures, the contribution of the jth market to price discovery that I called Impulse Response Information Share (*IRIS*) is defined by

$$IRIS_j = \frac{\sigma_{jj}^{-1}\left(\psi_{11}\sigma_{1j} + \psi_{12}\sigma_{2j}\right)^2}{\sum_{l=1}^{2}\sigma_{ll}^{-1}\left(\sum_{i=1}^{d}\psi_{1i}\sigma_{il}\right)^2} \tag{2.9}$$

Remember that the *IRIS* was computed using Q_{1t} the permanent component entering the first price. If I consider the permanent component entering the second market: $Q_{2t} = \psi_2 \sum_{s=1}^{t} \varepsilon_s = \beta_1^{-1} Q_{1t}$. It is a multiple of Q_{1t} so the impulse response of Q_{2t} to a shock to the jth price is:

$$GI_{Q_2}(\varepsilon_j) \quad = E\left(Q_{2t+n}|\varepsilon_{jt} = \delta_j, \Omega_{t-1}\right) - E\left(Q_{2t+n}, \Omega_{t-1}\right) = \beta_1^{-1} \quad GI_{Q_1}(\varepsilon_j)$$

and

$$IRIS_j = \frac{\beta_1^{-1} GI_{Q_2}(\varepsilon_j)^2}{\sum_{j=1}^{2}\beta_1^{-1} GI_{Q_2}(\varepsilon_j)^2} = \frac{GI_{Q_1}(\varepsilon_j)^2}{\sum_{j=1}^{2} GI_{Q_1}(\varepsilon_j)^2}$$

So *IRIS* doesn't depend on which permanent component you choose[3] and the estimation of the VECM can be done without imposing the unit restriction on the cointegrating equation.

2.2.5.3 Relationship between *IRIS* and the Information Share measures.

To present the link between *IRIS* and the *IS* of Hasbrouck (1995), I write explicitly the formulas for market 1.

$$IRIS_1 = \frac{\sigma_{11}^{-1}\left(\psi_{11}\sigma_{11} + \psi_{12}\sigma_{12}\right)^2}{\sigma_{11}^{-1}\left(\psi_{11}\sigma_{11} + \psi_{12}\sigma_{12}\right)^2 + \sigma_{22}^{-1}\left(\psi_{11}\sigma_{12} + \psi_{12}\sigma_{22}\right)^2}$$

If Ω is diagonal then the *IRIS* measure gives the *IS* measure:

$$IRIS_1 = \frac{\sigma_{11}^{-1}\psi_{11}^2\sigma_{11}^2}{\sigma_{11}^{-1}\psi_{11}^2\sigma_{11}^2 + \sigma_{22}^{-1}\psi_{12}^2\sigma_{22}^2} = \frac{\psi_{11}^2\sigma_{11}}{\psi_{11}^2\sigma_{11} + \psi_{12}^2\sigma_{22}}$$

If Ω is not diagonal, let $\sigma_{12} = \rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}$ and let's consider the expression of the Cholesky roots of Ω when the market 1 is placed first in vector of prices (Baillie et al., 2002):

$$F = \begin{pmatrix} \sqrt{\sigma_{11}} & 0 \\ \rho\sqrt{\sigma_{22}} & \sqrt{\sigma_{22}}\sqrt{(1-\rho^2)} \end{pmatrix}$$

Then the numerator of the Information Share of the market 1 which correspond to Hasbrouck

[3] This properties is easily seen for more than two markets, there is $d-1$ cointegration relations so only 1 permanent component entering each market multiplied by the corresponding β_i^{-1}

(1995) upper bound is

$$\begin{aligned} ([\psi F]_1)^2 &= (\psi_{11}\sqrt{\sigma_{11}} + \psi_{12}\rho\sqrt{\sigma_{22}})^2 \\ &= \psi_{11}^2\sigma_{11} + \psi_{12}^2\rho^2\sigma_{22} + 2\psi_{11}\psi_{12}\rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \end{aligned}$$

Let's now focus on the numerator of $IRIS_1$, the square of the Impulse Response of the permanent component to a shock in market 1.

$$\begin{aligned} \sigma_{11}^{-1}(\psi_{11}\sigma_{11} + \psi_{12}\sigma_{12})^2 &= \sigma_{11}^{-1}(\psi_{11}\sigma_{11} + \psi_{12}\rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}})^2 \\ &= \psi_{11}^2\sigma_{11} + \psi_{12}^2\rho^2\sigma_{22} + 2\psi_{11}\psi_{12}\rho\sqrt{\sigma_{11}}\sqrt{\sigma_{22}} \end{aligned}$$

So the numerator of the *IRIS* for the first market corresponds to the numerator of the IS for the first market when it is in the first position in the orthogonalization procedure. I can thus write

$$IRIS_1 = \frac{Upper.IS1 \times \psi\Omega\psi'}{Upper.IS1 \times \psi\Omega\psi' + Upper.IS2 \times \psi\Omega\psi'} = \frac{Upper.IS1}{Upper.IS1 + Upper.IS2}$$

In applications studies, where one value of the IS is needed for regression purposes, one bound or the mid-bounds is chosen without justification. The framework provide thus an economic rationale for the use of Hasbrouck (1995) upper-bound. In appendix 2.6.2.1 I show this result for $d > 2$ markets.

This result is not really surprising given the relationship between Cholesky factorization and the Generalized Impulse response functions (Kim, 2013). It might lead to conclusion that there is no difference in the definitions of the information Share and IRIS. Actually there is a difference: The IRIS doesn't not try to identify the origin of the shock and its contribution to the permanent price variance. But it is looking at where this information enter the fundamental price. A concrete example is to think of a Canadian company that is traded in Toronto and NYSE. If a relevant information is produced in Canada but is reflected in the price by fast NYSE traders, then I consider NYSE as dominating the price discovery process.

2.2.6 Estimation and Testing

The computation of the impulse response is easy once the parameters of the VECM representation of the prices are identified. On real data after selecting the order K with the help of information criteria, the VECM (2.1) is estimated for $\hat{\Omega}$ and for the parameters $\Gamma_1, \ldots, \Gamma_K$. Then $\hat{\Psi}(1)$ is computed and the elements of the impulse response are identified.

To obtain standard errors and the limiting distribution of the response to jth market, I use of the limiting distribution of the coefficients. I have

$$\hat{GI}_j \quad = \quad \hat{\psi}_1 \hat{\Omega} e_j \hat{\sigma}_{jj}^{-\frac{1}{2}}$$

For the deduction of asymptotic the VECM (2.1) is represented as

$$\Delta Y = -\alpha\beta' Y_{-1} + \Gamma \Delta X + U$$

with T= sample size and
$\Delta Y = [\Delta P_1, \ldots \Delta P_T]$
$Y_{-1} = [P_0, \ldots P_{T-1}]$
$\Gamma = [\Gamma_1, \ldots, \Gamma_K]$
$U = [\varepsilon_0, \ldots \varepsilon_T]$
$\Delta X = [\Delta X_0, \ldots \Delta X_T]$ with $\Delta X_{t-1} = \begin{bmatrix} \Delta P_{t-1} \\ \vdots \\ \Delta P_{t-K} \end{bmatrix}$

The theorem 2.1 gives the asymptotic distribution for the response of the permanent component to a shock in the jth market.

Theorem 2.1. *Let P_t the vector of prices satisfying VECM (2.1). then*

$$\sqrt{T}\left(\hat{GI}_j - GI_j\right) \xrightarrow{d} \mathcal{N}\left(0, \Sigma_{\hat{iris}}\right)$$

With

$$\Sigma_{iris} \quad = \quad \sigma_{jj}^{-1}\left(e_j^{'}\Omega \otimes e_j^{'}\right) F \Sigma_\gamma F^{'}\left(\Omega e_j \otimes e_j\right) + \sigma_{jj}^{-1}\left(e_j^{'}\Psi(1) \otimes e_j^{'}\right) \Sigma_{\hat{\sigma}}\left(\Psi^{'}(1) e_j \otimes e_j\right)$$

Where the notations are defined in the following:
$\Sigma_{\hat{\sigma}} = 2D_K\left(D_K^{'}D_K\right)^{-1} D_K^{'}\left(\Omega \otimes \Omega\right)$
$\Sigma_\gamma = plim T \begin{bmatrix} \beta' Y_{-1} Y_{-1}^{'} \beta & \beta' Y_{-1} \Delta X^{'} \\ \Delta X Y_{-1}^{'} \beta & \Delta X \Delta X^{'} \end{bmatrix}^{-1} \otimes \Omega$
$F = \left(\left(\Psi^{'}(1)\left(I - \sum_{i=1}^{K} \Gamma_i\right)^{'} - I_d\right) H_{\alpha_1^{'}}\left(\alpha^{'} H_{\alpha_1^{'}}\right)^{-1}, \left(\iota_K^{'} \otimes \Psi^{'}(1)\right)\right) \otimes \Psi^{'}(1)$
D_K (duplication matrix) and $H_{\alpha_1^{'}}$ are matrix of 0-1 (defined in appendix). ι_K represents a column vector with ones of length K

The asymptotic variance can be used to compute standard errors and test the significance of the permanent response to a shock in one market. For this purpose, the different expressions in the variance are replaced by their feasible estimator. The matrix $H_{\alpha_1'}$ might be bit a tricky to build especially if the cointegration rank is not known. Fortunately in the setup the rank of cointegration is known to be $d-1$ and provided that α' is put in reduced echelon form I have the $[d \times (d-1)]$ matrix

$$H_{\alpha_1'} = \begin{bmatrix} 0_{1\times d-1} \\ I_{d-1} \end{bmatrix}$$

For example in the bivariate case I have $H_{\alpha_1'} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

To test the significance of GI_j the following testing statistic is computed

$$\hat{ST}_j = \sqrt{T}\frac{|\hat{GI}_j|}{\hat{\Sigma}_{iris}}$$

and can be compared with the critical values of the classical t-test.

2.2.7 Simulation

In this simulation exercise, I evaluate the performance of *IRIS* measure and the test statistic. For this, I rely on some structural models. The parameters of the microstructure noise (driven there by trade direction effect) are small as in the literature with respect to the fundamental innovation variance[4] .

Model-1: The two-markets "Roll" model

I have one asset that is traded on two markets at the respective prices p_{1t} and p_{2t}. The unobserved efficient price (fundamental value) of the asset is m_t is driven by non-trade public information. As Hasbrouck (2002), $\Delta m_t = u_t$ is the is the public and non traded information. The observed transaction prices are the fundamental price plus a microstructure component driven for example by signed trade direction. Formally it can be represented as:

- $m_t = m_{t-1} + u_t$ with $u_t \sim \mathcal{N}\left(0, \sigma^2\right)$
- Transaction prices: $p_{it} = m_t + cq_{it}$ and trade directions for q_{it}; $i = 1, 2$,

With $q_{it} \sim \mathcal{N}(0,1)$, $E(q_{1t}q_{2t}) = 0$, $E(q_{it}u_t) = 0$, for $i = 1, 2$

In this model, in term of fundamental information, one market can not be said to dominate the other. A good price discovery measure should thus give the same weight to both markets.

[4] Harris et al. (2002a) also criticizes Hasbrouck (2002) simulation by arguing that it considers only high and unrealistic values for the parameters of the trade direction (which give a small signal-to-noise ratio)

Model-2: Two markets with private information

This model presented in Hasbrouck (2002) considers that traders on market 1 have private information and their trade drive the efficient price. There is also a public information that is not traded and enter the efficient price each period.

The second market relies on lagged value of the fundamental price m_t. This might signify that traders on this market are slower, or that they learn about the fundamental from the traders actions on the first market. The observed price is equal to the fundamental price plus and impact of the trade direction.

- $m_t = m_{t-1} + \lambda c_1 q_{1t} + u_t$ with $u_t \sim \mathcal{N}\left(0, \sigma^2\right)$
- $p_{1t} = m_t + c_1 q_{1t}$
- $p_{2t} = m_{t-1} + c_2 q_{2t}$

With $q_{it} \sim \mathcal{N}\left(0,1\right)$, $E\left(q_{1t} q_{2t}\right) = 0$, $E\left(q_{it} u_t\right) = 0$, for $i = 1,2$

It is clear in this situation that all price discovery happens in market 1. A good price discovery measure should design market 1 as dominant.

Results

I simulate 1000 replications of Model-1 with 100,000 observations. A good price discovery measure is expected to give the same contribution to both market, Table 2.1 shows the measures perform pretty well for this and *IRIS* is more close to 50% than the others with the smallest standard deviation. The values of the test statistics show that both markets are highly significant.

Table 2.1: *Model 1. The values represent the measures for market 1 averaged over 1000 replications. The value of the test statistics $\hat{ST}$ are given for market 1 and market 2. Standard errors are in parenthesis.*

c	$IRIS_1$	IS_1U	IS_1L	PT_1	$\hat{ST}_1$	$\hat{ST}_2$
0.005	50.00	99.998	0.001	52.58	83.5**	83.9**
	(7e-4)	(1.5e-3)	(1.5e-3)	(17.8)	(59.01)	(58.24)
0.0005	50.00	99.99	0.00	50.1	7.30**	7.50**
	(1e-06)	(3e-6)	(3e-6)	(6.8)	(10.25)	(11.05)

** :significance at 1%

In the case of Model-2, the whole price discovery happens on market 1. I simulate 1000 replications of the model with 100,000 observations. I vary the noise variance as in the previous case. The measures perform quite well by designing market 1 as dominant (Table 2.2), *IRIS* having a small standard deviation . The value of the test statistic for market 1 is highly significant using student critical values, and for the second market it is not significant.

Table 2.2: *Model 2. The values represent the measures averaged over 1000 replications. For $\lambda = 1$, $c_2 = 0$, and different values of c_1. The value of the test statistics $\hat{ST}$ are given for market 1 and market 2. Standard errors are in parenthesis.*

c_1	$IRIS_1$	IS_1U	IS_1L	PT_1	$\hat{ST}_1$	$\hat{ST}_2$
0.005	99.96	99.97	99.97	42.35	154**	1.65
	(0.04)	(0.03)	(0.04)	(0.03)	(9.47)	(0.7)
0.0005	97.65	97.66	97.54	0.84	134**	0.016
	(3.947)	(3.949)	(3.945)	(0.63)	(29.8)	(0.008)

** :significance at 1%

2.3 A cost-of-carry model of futures market

Here I present an extension of the model of futures and spot market of Garbade and Silber (1983). Figuerola-Ferretti and Gonzalo (2010) extended this model and provide a justification for the PT measure of price discovery. I will show that the *IRIS* measure leads to the same conclusion stating that the market with the highest number of participants is informationally dominant.

Consider that I have a storable commodity that is traded on the spot market and let s_t be the log price at time t. There is a futures contract on this commodity that mature at date T (e.g. $T = 15$ months), and is sold at time t at the log price f_t. There is another futures contract with maturity date $T' < T$ (e.g. $T' = 3$ months) and is sold at log price q_t.

The model relies on the storage theory with convenience yield. Under some classical assumptions: no transaction costs; no limitations on borrowing; no limitation on short sales, The equilibrium relation between spot market and the futures market is given by the following cost-and-carry no-arbitrage relationship:

$$F_t = S_t e^{(r_t + c_t - y_t)(T-t)}$$

The same reasoning applies for two futures contracts with different maturities. The partial equilibrium relationship is

$$F_t = Q_t e^{(r_t + c_t - y_t)\left(T - T'\right)}$$

Where y_t is the convenience yield. It represents the fact that beyond carrying costs, people might have a benefit of having a physical commodity rather than a futures contract on it. I can fix $T - T^{'} = 1$, and taking log the relationship gives

$$f_t = q_t + r_t + c_t - y_t \tag{2.10}$$

Following Figuerola-Ferretti and Gonzalo (2010), the prices are assumed to be random walk $\Delta q_t = q_t - q_{t-1} = I(0)$ and $\Delta f_t = f_t - f_{t-1} = I(0)$. The total interest rate (interest rate+storage cost) is stationary and written $r_t + c_t = -w + I(0)$, then

$$q_t = f_t + y_t + w + I(0)$$

The convenience yield is assumed endogenous and written as $y_t = \gamma_1 q_t - \gamma_2 f_t + I(0)$ with $\gamma_1, \gamma_2 > 0$. Then I have

$$q_t = \beta_1 f_t + \beta_0 + I(0)$$

This is a cointegration relation with a cointegrating vector $\begin{pmatrix} 1 & -\beta_1 & -\beta_0 \end{pmatrix}$, where $\beta_1 = \frac{1-\gamma_2}{1-\gamma_1}$ and $\beta_0 = \frac{w}{1-\gamma_1}$. The market is said to be under long run Backwardation when the cash price is smaller than the futures price coefficient $(\beta_1 > 1)$, and under long run Contago when the cash price is bigger than the futures price $(\beta_1 > 1)$.

The description of interaction between spot and futures market is done following Garbade and Silber (1983). There are N_f participants in the long maturity futures market and N_s participants in the spot or the short maturity futures market. Let E_{it} the endowment of the i_{th} participant immediately prior to period t and r_{it} the reservation price at which that participant is willing to hold E_{it}. The demand schedule of the i_{th} participant in the short maturity futures market in period t is

$$E_{it} - A\left(q_t - r_{it}\right), \qquad A > 0,\ t = 1, \ldots, n$$

Where A is the elasticity of demand assumed to be the same for all participants. The aggregate short-maturity futures market demand schedule of arbitrageurs in period t

$$H\left(\beta_1 f_t + \beta_0 - q_t\right),\ H > 0$$

H is the elasticity of short-maturity futures market demand by arbitrageurs. It is infinite when arbitrage activities between markets are riskless, this corresponds to the previous model with a

$y_t = 0$. The market on the short maturity contract will clear at the value of q_t that solves

$$\sum_{i=1}^{N_s} E_{it} = \sum_{i=1}^{N_s} \left[E_{it} - A\left(q_t - r_{it}\right)\right] + H\left(\beta_1 f_t + \beta_0 - q_t\right) \tag{2.11}$$

The long maturity futures market will clear the value of f_t such that

$$\sum_{i=1}^{N_s} E_{it} = \sum_{i=1}^{N_s} \left[E_{it} - A\left(q_t - r_{it}\right)\right] + H\left(\beta_1 f_t + \beta_0 - q_t\right) \tag{2.12}$$

The two previous equations give q_t and f_t as a function of the mean reversion price of the two markets ($R_t^s = N_s^{-1}\sum_{i=1}^{N_s} r_{it}$ and $R_t^f = N_f^{-1}\sum_{i=1}^{N_f} r_{it}$)

$$\begin{aligned} q_t &= \frac{\left(AN_f + H\right) N_s R_t^s + HN_f R_t^f + HN_f\beta_0}{\left(AN_s + H\right) N_f + HN_s} \\ f_t &= \frac{HN_s R_t^s \left(AN_s + H\right) N_f R_t^f - HN_f\beta_0}{\left(AN_s + H\right) N_f + HN_s} \end{aligned} \tag{2.13}$$

The next step is to describe the evolution of reservation prices. A reasonable dynamic is to say that immediately after the market clearing at period $t-1$ the i^{th} short-maturity futures market participant was willing to hold amount E_{it} at the price q_{t-1}. This implies that q_{t-1} was his reservation price after that clearing. This reservation price change to R_{it} according to equation

$$\begin{aligned} r_{it} &= q_{t-1} + v_t + w_{it}, \;\; i = 1, \ldots, N_s \\ r_{jt} &= f_{t-1} + v_t + w_{jt}, \;\; j = 1, \ldots, N_f \end{aligned} \tag{2.14}$$

Where $cov\left(v_t, w_{it}\right) = 0$, $cov\left(w_{it}, w_{kt}\right) = 0$ for $\forall k \neq i$, and $\left(v_t, w_{it}, w_{kt}\right)$ is a vector white noise.

The price change $R_{i,t} - q_{t-1}$ reflects the arrival of new information between period $t-1$ and period t which changes the price at which the i^{th} participant is willing to hold the quantity E_{it} of the commodity. The price change has a common component (v_t) and an idiosyncratic component (w_{it}). Summing 2.14 by market, the mean reservation price in period t will be

$$\begin{aligned} R_t^s &= q_{t-1} + v_t + w_t^s \\ R_t^f &= f_{t-1} + v_t + w_t^f \end{aligned}$$

Where $w_t^s = N_s^{-1}\sum_{i=1}^{N_s} w_{it}$ and $w_t^f = N_f^{-1}\sum_{j=1}^{N_f} w_{jt}$. Substituting into 2.13 yields

$$\begin{pmatrix} q_t \\ f_t \end{pmatrix} = \Pi \begin{pmatrix} q_{t-1} \\ f_{t-1} \end{pmatrix} + \frac{H\beta_0}{d} \begin{pmatrix} N_f \\ -N_s \end{pmatrix} + \begin{pmatrix} u_t^s \\ u_t^f \end{pmatrix} \tag{2.15}$$

Where $\begin{pmatrix} u_t^s \\ u_t^f \end{pmatrix} = M \begin{pmatrix} v_t + w_t^s \\ v_t + w_t^f \end{pmatrix}$, $\Pi = \frac{1}{d} \begin{pmatrix} N_s\left(\beta_1 H + AN_f\right) & \beta_1 HN_f \\ HN_s & (H + AN_s)N_f \end{pmatrix}$ and $d = (H + AN_s)N_f + \beta_1 HN_f$

Rewriting the previous equation in a VECM form gives

$$\begin{pmatrix} \Delta q_t \\ \Delta f_t \end{pmatrix} = \frac{H\beta_0}{d} \begin{pmatrix} N_f \\ -N_s \end{pmatrix} + (\Pi - I) \begin{pmatrix} q_{t-1} \\ f_{t-1} \end{pmatrix} + \begin{pmatrix} u_t^s \\ u_t^f \end{pmatrix} \tag{2.16}$$

with the reduce rank matrix $\Pi - I = \frac{1}{d} \begin{pmatrix} -HN_f & \beta_1 HN_f \\ HN_s & -\beta_1 HN_s \end{pmatrix}$ and rearranging the terms

$$\begin{pmatrix} \Delta q_t \\ \Delta f_t \end{pmatrix} = \frac{H}{d} \begin{pmatrix} N_f \\ -N_s \end{pmatrix} \begin{pmatrix} 1 & -\beta_1 & -\beta_0 \end{pmatrix} \begin{pmatrix} q_{t-1} \\ f_{t-1} \\ 1 \end{pmatrix} + \begin{pmatrix} u_t^s \\ u_t^f \end{pmatrix} \tag{2.17}$$

The weighting matrix $\alpha^T = \begin{pmatrix} N_f & -N_s \end{pmatrix}$ shows immediately the link with the component share measure of price discovery. This leads to the same conclusion of GS that the market with the highest number of participant is dominant. The proposition 2.2 shows that *IRIS* has the same feature.

Proposition 2.2. *Given the previous structural economy and equation 2.16:*

If $N_s > N_f$ then the IRIS measure of the short-maturity futures market is bigger than the IRIS of the long-maturity futures market: $IRIS_s > IRIS_f$.

The previous proposition, proven in appendix, also provides some theoretical support to the IRIS measure. The main difference with the PT measure is that for the PT, the relative contribution of a market is proportional to its number of participants. This is not necessarily a best representation of the reality in the sense that marginal effect or marginal risk might behave differently for big and small market size.

2.4 Empirical application

In this application I investigate the relative contribution to price discovery of futures contracts on some metals and alloys: Cobalt (Co), Molybdenum (Mo), Steel billet (St), Tin (Ti), Aluminium (Al), Copper (cu). They are traded on the London Metals Exchange (LME) which is the biggest exchange for Metals in the world. The advantage of LME is that in addition to have futures market transactions, the cash market also operate at the same place. I will focus on 3 timed securities, the cash market, the 3-month futures contract and the 15-month futures contract. The goal is to study the relative relevance of theses contracts in term of price discovery.

The data for each of the metals under study is comprised of the cash official prices, the 3-month futures official prices and the 15-month futures official prices, extracted from the database Eikon of Thomson Reuters. The Official prices of some base metals stopped being produced by LME for 15-month contract in April 2012, so the study period is 7th July 1993 to 01 February 2012 for Ti, Al and Cu. The data for Co, Mo and St are available from 19th May 2010 to 31th March 2016.

2.4.1 Results

The first step of the empirical strategy to compute the elements of interest accordingly to my framework, is to perform some statistical analysis on the data. For all the series presented in graphs 2.1 to 2.6 the 3 prices under investigation are closely moving together, in fact arbitrage operations will reduce any tendency of the spread to diverge. The stationarity examination on the time series reveals that for all prices the unit root hypothesis cannot be rejected (see table 2.5). When considering the first difference they all appear stationary. The cointegration test is applied for each metal two-by-two: except for copper, for all the commodities I can not reject cointegration between Cash and Futures markets, and between 3-month and 15 month futures prices. The Lag length of the Vector autoregressive is selected using the Akaike Criterion (AIC), then estimation of the VECM and the necessary checks on the residuals are done.

The trace test failed to find a cointegrating relation for Copper between the cash price and 15-month futures and between the 3-month futures and 15-months. I nevertheless estimate the long term relationships in the VECM and the plots (graph 2.7) shows that the problem is not really severe. In fact by applying the ADF test to the long term relationships (table 2.7), I obtain that I can accept cointegration at a level of 7% and 9%.

2.4.1.1 Analysis of long run Contago and Backwardation

The estimation of the long term relation between two related securities shows mixing conclusion for different type of markets. For the Co, Mo and St I obtain cointegrating coefficients that are greater than 1 between cash and 3-month futures, between cash and 15-month, between 3-month futures and 15 months. I also compute the statistics and a comparison with critical values of the unilateral test ($H_0 : \beta = 1 \, vs \, H_1 : \beta > 1$) shows that the alternative hypothesis that they are strictly greater than 1 can be accepted. There is a consistent long-run backwardation pattern in those

Table 2.3: *Impulse response Information Share computed using the Vector Autoregressive equation on the following four vectors of variables:*

$I : (pcash, p03m)$ $III : (p03m, p15m)$
$II : (pcash, p15m)$ $IV : (p03m, p15m, pcash)$

	I	II	III	IV	I	II	III	IV
	Cobalt				Molybdenum			
Cash	46.75	57		40.74	47.8	92.3		46.93
3M	53.25		55.7	46.58	52.2		93.3	51.61
15M		43	44.3	12.66		7.7	6.7	1.45
	Steel billet				Tin			
Cash	49.9	38.1		43.17	50.14	54.93		35.57
3M	50.1		38.8	42.49	49.86		54.75	34.93
15M		61.9	61.2	14.33		45.07	45.25	29.48
	Aluminium				Copper			
Cash	49.97	54.76		33.53	47.89	51.13		33.21
3M	50.03		54.56	34.61	52.11		55.15	36.35
15M		45.24		31.84		48.87	44.85	30.42

markets. That is in the long run, having taken into account all the storage costs, forward prices are decreasing with the maturity: the cash price is bigger than the 3-month forward price, which is in turn bigger than the 15-month forward prices. This result is not obtained by comparing the observed prices series. In fact, while the observed prices are mostly decreasing with maturity for Cobalt, it is mostly increasing for Molybdenum. A look at the graphs of the series also show that the relative position of the 3 curves depends on the metal.

For Aluminium and Copper and Tin, coefficients are smaller than 1. As previously in the test $H_0 : \beta = 1 VSH_1 : \beta < 1$; the null hypothesis is rejected. The markets are in long run Contago that is in the long run the cash price is smaller than the 3-month forward price, and the 3-month forward price is smaller than the 15-month forward prices. These results are the opposite of Figuerola-Ferretti and Gonzalo (2010) for Aluminium and copper. A justification of this difference is that their analysis of non-ferrous metal is on the period January 1989 to October 2006. Since then the market has drastically changed as can be seen in the graph 2.6, the spread between the different prices is huge in early 2000 compare to 2010. A deep understanding of what happens is important but not in the scope of this paper.

2.4.1.2 Analysis of price discovery

The results in table 2.3 on the permanent impact of market differ according to each metal and maturity. I estimate the IRIS measure in bivariate VECM (I,II,III) and in a trivariate VECM (IV) and obtain the following: For the Cobalt, Copper and Aluminium, in the two-by-two comparison, the

3-month dominates the spot market and 15-Month, the spot market also dominates the 15-month contract. When the three markets are compared it appears that permanent price reacts more to a shock in cash and 3-Month than a shock to the 15-Month future. All markets here appears to be informationally relevant with highly significant responses (see testing at Table 2.4). The same pattern is observed for the molybdenum where the 15-Month market is strongly dominated by the two others, the 3-month market being the leading ones. For the Tin there is only a minor difference with the previous results, for model I the cash market dominates very slightly the 3-month future. The Steel billet presents a different configuration, I have that 15-month future dominates the cash market and dominates the 3-months contract in the bivariate analysis even if the permanent component of the 3 markets reacts less to a shock in 15-month prices

Globally the empirical results for the sample of LME metals under investigation seems to suggest that even if in theory the function of the future market is to provide price discovery, only the 3-month contract seems to fully play this role. This suggest that in those markets the 15-Month futures contract are less used than the 3-month to exploit new information.

Table 2.4: *The table presents the value of the Test statistic $t_{\hat{g}i}$ for Model I,II,III.*

$I : (pcash, p03m)$ $\quad III : (p03m, p15m)$
$II : (pcash, p15m)$ $\quad IV : (p03m, p15m, pcash)$

	I	II	III	I	II	III
	Cobalt			Molybdenum		
Cash	19.5**	10.0**		28.9**		4.0**
3M	18.2**		7.4**	6.12**	6.97**	
15M		43.4**	19.8**		14.4**	15**
	Steel billet			Tin		
Cash	17.8**	7.3**		17.8**	7.3**	
3M	20.6**		6.9**	20.6**		6.9**
15M		33.6**	13.6**		33.6**	13.6**
	Aluminium			Copper		
Cash	19.5**	10.0**		28.9**	6.9**	
3M	18.8**		7.4**	6.1**		4.0**
15M		43.5**	19.8**		14.4**	15.6**

** :significance at 1%

2.5 Conclusion

The study of assets that are strongly related requires tools to evaluate the information content of prices. In this paper I proposed to study the impact of a shock in one market on the permanent component of prices by using the Generalized impulse response function as defined by Pesaran and

Shin (1998). I provide a measure that I named Impulse Response Information Share, to determine which market is the leading one. For me a market is leading if the effect on the permanent component of prices, of a shock in that market is greater than the permanent effect of the shock on the other markets. IRIS has some advantages over the existing measure as it provides a unique value of the information without loosing its economic sensitive definition. Some Monte-Carlo exercises also show that it can produce lower standard deviation compared to other measures.

In the application I study price discovery phenomenon on spot and futures markets of some base metals traded on the London Metals Exchange. I found that 3-month futures market dominates the spot market in term of price discovery suggesting that the futures contracts is a good expectation of the future cash prices, and that the futures market can be well used for hedging purposes. When comparing the spot and the 15-month futures with the 3M market or the spot market it appears that the 15-month futures contract is always dominated by the others. The join comparison of the threes securities in the same VECM confirms that when the 3-month contract importance in term of price discovery the 15-month contract is not relevant. Of course there is a big jump to go far and say that 15-month contract are useless, since there are many used of this contract by economics agents. A deep study of its importance should include a broad view of the actors and their strategies. Some events studies might also help with this respect, like understanding why for the copper 15 month LME production of (Official, Settlement and Unofficial) prices ceased in April 2012.

2.6 APPENDIX B

2.6.1 Tables and figures

Table 2.5: ADF unit root test results.

	level		$\Delta(Series)$			level		$\Delta(Series)$	
Series	stat	Prob	stat	Prob	**Series**	stat	Prob	stat	Prob
CoCash	-1.40	0.58	-44.73	0.00	TiCash	-1.30	0.63	-37.61	0.00
Co3m	-1.40	0.59	-45.59	0.00	Ti3m	-1.26	0.65	-37.71	0.00
Co15m	-1.63	0.47	-45.23	0.00	Ti15m	-1.22	0.67	-37.92	0.00
MoCash	0.11	0.97	-44.03	0.00	AlCash	-1.26	ă0.65	-40.31	0.00
Mo3m	0.10	0.97	-43.92	0.00	Al3m	-1.13	ă0.70	-40.63	0.00
Mo15m	0.06	0.96	-43.80	0.00	Al15m	-0.97	ă0.77	-40.98	0.00
StCash	-1.04	0.74	-43.64	0.00	CuCash	-0.52	0.88	-40.14	0.00
St3m	-1.00	0.76	-43.39	0.00	Cu3M	-0.48	0.89	-40.0	0.00
St15m	-0.96	0.77	-42.53	0.00	Cu3M	-0.39	0.90	-40.21	0.00

The table shows the statistic and prob for the level and first difference of each series

Figure 2.1: Cobalt: Daily cash , 3-month and 15-month Futures prices

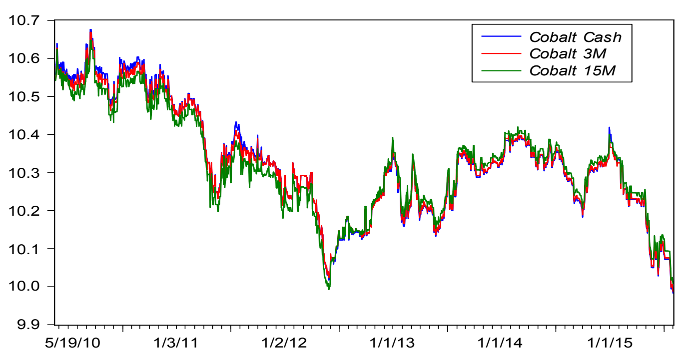


Figure 2.2: Molybdenum: Daily cash , 3-month and 15-month Futures prices

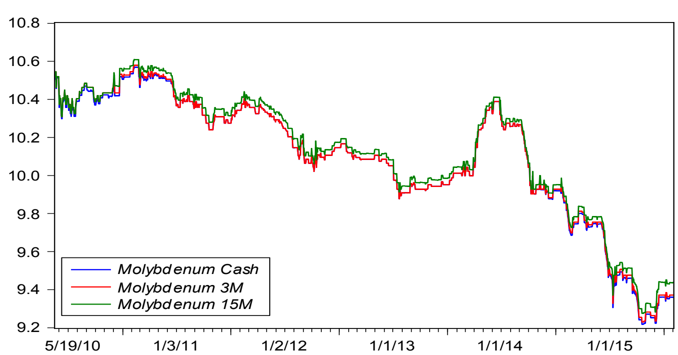


Table 2.6: Trace Test for the rank(r) of Cointegration.

Model	H_0	**I**	**II**	**III**
Cobalt	$r = 0$	0.0000	0.0000	0.0000
	$r \geq 1$	0.5162	0.4550	0.4390
Molybdenum	$r = 0$	0.0002	0.0000	0.0000
	$r \geq 1$	0.4358	0.4513	0.4367
Steel	$r = 0$	0.0013	0.0126	0.0053
	$r \geq 1$	0.9040	0.6871	0.6600
Tin	$r = 0$	0.0006	0.0379	0.0160
	$r \geq 1$	0.7744	0.7801	0.7905
Aluminium	$r = 0$	0.0000	0.0062	0.0074
	$r \geq 1$	0.3904	0.3302	0.2922
Copper	$r = 0$	0.0227	0.498*	0.439*
	$r \geq 1$	0.8267	0.8491	0.8451

Table 2.7: ADF Test results on the cointegrating equations for Copper:

	ADF test	
	Statistic	Prob
Moded II	-2.7835	0.0607
Model III	-2.6656	0.0802

Table 2.8: Long term relationships in the VECM.

$I: pcash = c + \beta \times p03m + \varepsilon$ $\quad III: p03m = c + \beta \times p15m + \varepsilon$
$II: pcash = c + \beta \times p15m + \varepsilon$

	I	II	III	I	II	III
		Cobalt			Molybdenum	
β	1.03	1.24	1.31	1.007	1.03	1.02
	(0.005)	(0.08)	(0.08)	(0.008)	(0.01)	(0.01)
c	0.16	-1.11	1.39	0.034	0.03	0.13
	(0.02)	(0.37)	(0.39)	(0.035)	(0.03)	(0.06)
		Steel billet			Tin	
β	1.09	1.31	1.20	0.99	0.99	0.98
	(0.009)	(0.02)	(0.01)	(0.001)	(0.009)	(0.01)
c	-0.25	-0.92	-0.61	0.02	0.02	0.05
	(0.02)	(0.06)	(0.03)	(0.001)	(0.04)	(0.04)
		Aluminium			Copper	
β	0.995	0.94	0.98	0.995	0.97	0.98
	(0.013)	(0.058)	(0.03)	(0.006)	(0.02)	(0.01)
c	0.01	0.15	0.02	0.02	0.1	0.07
	(0.04)	(0.19)	(0.12)	(0.02)	(0.08)	(0.05)

Standard errors are in parenthesis:

Figure 2.3: Steel Billet: Daily cash , 3-month and 15-month Futures prices

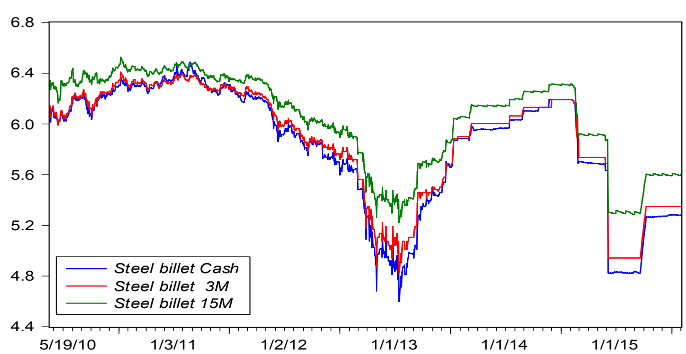


Figure 2.4: Tin: Daily cash , 3-month and 15-month Futures prices

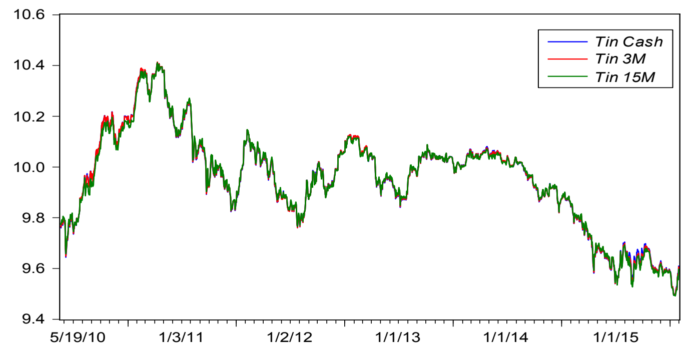


Figure 2.5: Aluminium: Daily cash , 3-month and 15-month Futures prices

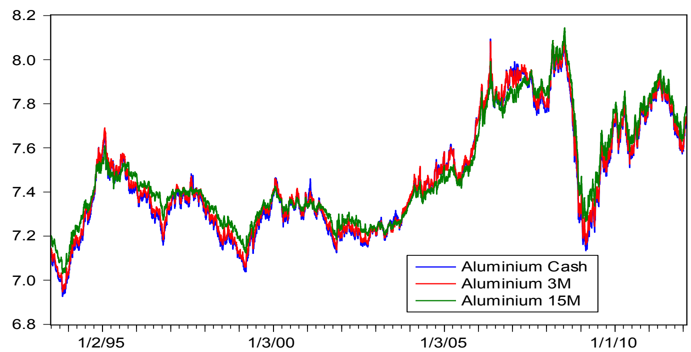


Figure 2.6: Copper: Daily cash , 3-month and 15-month Futures price

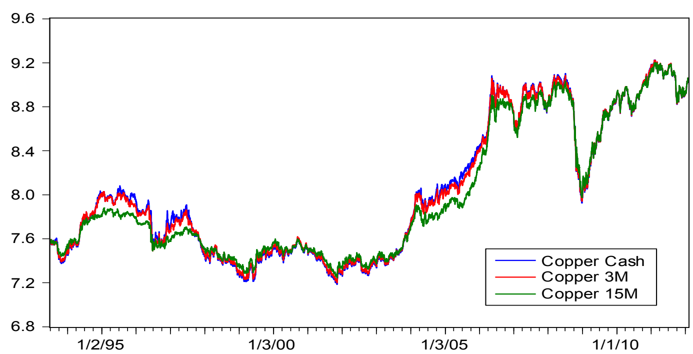
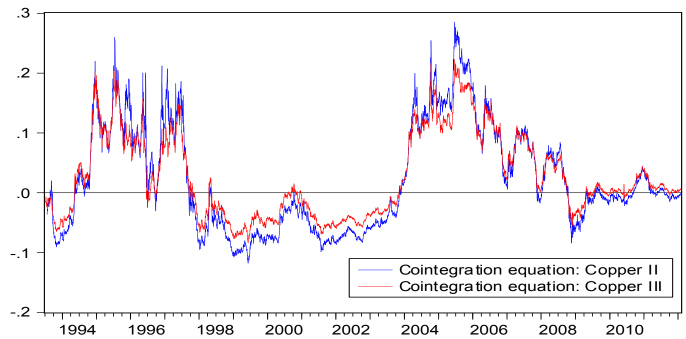


Figure 2.7: Long term relationship for the Copper: Model I and model III



2.6.2 Proofs

2.6.2.1 Link between IRIS and IS

Proof. For d markets, the IRIS measure for the market j is given by.

$$IRIS_j = \frac{\sigma_{jj}^{-1}\left(\sum_{i=1}^{d}\psi_{1i}\sigma_{il}\right)^2}{\sum_{l=1}^{d}\sigma_{ll}^{-1}\left(\sum_{i=1}^{d}\psi_{1i}\sigma_{il}\right)^2} = \frac{\left(\sigma_{jj}^{-\frac{1}{2}}\psi\Omega e_j\right)^2}{\sum_{l=1}^{d}\left(\sigma_{ll}^{-\frac{1}{2}}\psi\Omega e_l\right)^2} \tag{2.18}$$

I will focus on the numerator of the IS measure for the first market when it is placed first in the Cholesky decomposition. For this let F be the Cholesky root of $\Omega = FF^T$, I have the following relationships

$$\Omega = \begin{bmatrix} f_{11} & & & \\ f_{21} & f_{22} & & \\ \vdots & \vdots & \ddots & \\ f_{d1} & f_{d2} & \cdots & f_{dd} \end{bmatrix} \begin{bmatrix} f_{11} & f_{21} & \cdots & f_{d1} \\ & f_{22} & \cdots & f_{d2} \\ & & \ddots & \vdots \\ & & & f_{dd} \end{bmatrix}$$
$$= \begin{bmatrix} f_{11}^2 & f_{11}f_{21} & \cdots & \\ f_{21}f_{11} & f_{21}^2 + f_{22}^2 & \cdots & \\ \vdots & \vdots & \cdots & \vdots \\ f_{d1}f_{11} & f_{d1}f_{21} + f_{d2}f_{22} & \cdots & \end{bmatrix}$$

By equalizing the first columns I have

$$\begin{cases} \sigma_{11} & = f_{11}^2 \\ \sigma_{21} & = f_{11} f_{21} \\ & \vdots \\ \sigma_{d1} & = f_{11} f_{d1} \end{cases} \Longrightarrow \begin{cases} f_{11} & = \sqrt{\sigma_{11}} \\ f_{12} & = \sigma_{21}/\sqrt{\sigma_{11}} \\ & \vdots \\ f_{d1} & = \sigma_{d1}/\sqrt{\sigma_{11}} \end{cases}$$

So the numerator of IS for the market 1 is

$$\begin{aligned} ([\psi F]_1)^2 &= \left(\psi \begin{bmatrix} F_{11} & F_{21} & \dots & F_{d1} \end{bmatrix}^T \right)^2 \\ &= \left(\psi \begin{bmatrix} \sqrt{\sigma_{11}} & \sigma_{21}/\sqrt{\sigma_{11}} & \dots & \sigma_{d1}/\sqrt{\sigma_{11}} \end{bmatrix}^T \right)^2 \\ &= \sigma_{11}^{-1} \left(\sigma_{11}^{-\frac{1}{2}} \psi \begin{bmatrix} \sigma_{11} & \sigma_{21} & \dots & \sigma_{d1} \end{bmatrix}^T \right)^2 \\ &= \left(\sigma_{11}^{-\frac{1}{2}} \psi \Omega e_1 \right)^2 \end{aligned}$$

Which is exactly the numerator of the IRIS measure for the first market. □

2.6.2.2 Proof of theorem 2.1

Proof. As it is shown in Lütkepohl (2007) the asymptotic is the same considering that β is known the reason being that $\hat{\beta}$ is estimated at the rate T better than the $\sqrt{T}$ of $\hat{\alpha}$. To simplify the formulas I use Ψ to denote $\Psi(1)$.

Let $\gamma \equiv vec[\alpha : \Gamma]$ where the *vec* operator stacks the columns of matrix into one column. The *vech* operator stacks the elements on and below the diagonal of a square matrix.

The following asymptotics and the expression for Σ_γ and $\Sigma_{\hat{\sigma}}$ are derived from Lütkepohl (2007) and Pesaran and Shin (1998) :

- $\sqrt{T}(\hat{\gamma} - \gamma) \xrightarrow{d} \mathcal{N}(0, \Sigma_\gamma)$
- $\sqrt{T} vec\left(\hat{\Omega} - \Omega\right) \xrightarrow{d} \mathcal{N}(0, \Sigma_{\hat{\sigma}})$
- The duplication matrix D_K is the $\left(K^2 \times \frac{1}{2} K (K+1)\right)$ matrix of 0-1 such that for any $(K \times K)$ matrix A, $vec(A) = D_K vech(A)$.
- The estimators of $[\alpha : \Gamma]$ and Ω are asymptotically independent.

As Ψ depends only on the $[\alpha : \Gamma]$ I have

$$\sqrt{T}\begin{bmatrix} vec\left(\hat{\Psi}-\Psi\right) \\ vec\left(\hat{\Omega}-\Omega\right) \end{bmatrix} \xrightarrow{d} \mathcal{N}\left(0, \begin{bmatrix} \Sigma_{\psi} & 0 \\ 0 & \Sigma_{\hat{\sigma}} \end{bmatrix}\right) \tag{2.19}$$

To obtain the asymptotic variance of Ψ as a function of γ the Delta-method gives

$$\Sigma_{\psi} = \frac{\partial vec\Psi(1)}{\partial \gamma'}\Sigma_{\gamma}\frac{\partial vec\Psi(1)'}{\partial \gamma} \tag{2.20}$$

and the expression for $\frac{\partial vec\Psi(1)}{\partial \gamma'}$ is derived from Vlaar (2004):

$$F = \frac{\partial vec\Psi(1)}{\partial \gamma'} = \left(\left(\Psi'\left(I-\sum_{i=1}^{K}\Gamma_i\right)' - I_d\right)H_{\alpha_1'}\left(\alpha' H_{\alpha_1'}\right)^{-1}, \left(\iota_K' \otimes \Psi'\right)\right)\otimes \Psi' \tag{2.21}$$

The matrix $H_{\alpha_1'}$ is a matrix of 0-1 selecting the column of α' such that $\left(\alpha' H_{\alpha_1'}\right)$ is non singular and $H_{\alpha_2'}$ selects the remaining columns. Those matrix allows the representation $\alpha_{\perp} = H_{\alpha_2'} - H_{\alpha_1'}\left(\alpha' H_{\alpha_1'}\right)^{-1}\alpha' H_{\alpha_2'}$.

The response of the market j is given by

$$GI_j = \psi_1 \Omega e_j \sigma_{jj}^{-\frac{1}{2}} = \left(e_j'\Omega e_j\right)^{-\frac{1}{2}} \times e_j'\Psi\Omega e_j$$

its estimator is thus

$$\begin{aligned} \hat{GI}_j &= \hat{\psi}_1 \hat{\Omega} e_j \hat{\sigma}_{jj}^{-\frac{1}{2}} \\ &= e_j'\hat{\Psi}\hat{\Omega}e_j \times \left(e_j'\hat{\Omega}e_j\right)^{-\frac{1}{2}} \\ &= \left(e_j'\left(\hat{\Psi}-\Psi\right)\left(\hat{\Omega}-\Omega\right)e_j + e_j'\left(\hat{\Psi}-\Psi\right)\Omega e_j + e_j'\Psi\left(\hat{\Omega}-\Omega\right)e_j + e_j'\Psi\Omega e_j\right) \times \left(e_j'\hat{\Omega}e_j\right)^{-\frac{1}{2}} \\ &= Ng/Dg \end{aligned}$$

For the denominator I have from Pesaran and Shin (1998) that given the consistency of the ML estimators there is a scalar R which is $o_p(1)$ such that

$$Dg = \left(e_j^{'}\Omega e_j\right)^{\frac{1}{2}} + \frac{1}{2}\left(e_j^{'}\Omega e_j\right)^{\frac{1}{2}}\left(e_j^{'}\otimes e_j^{'}\right) vec\left(\hat{\Omega}-\Omega\right) + R = \sigma_{jj}^{\frac{1}{2}} + o_p(1) \tag{2.22}$$

I now compute the distribution of the numerator

$$\begin{aligned}
\sqrt{T}\left(Ng - e_j^{'}\Psi\Omega e_j\right) &= e_j^{'}\left(\hat{\Psi}-\Psi\right)\left(\hat{\Omega}-\Omega\right)e_j + e_j^{'}\left(\hat{\Psi}-\Psi\right)\Omega e_j + e_j^{'}\Psi\left(\hat{\Omega}-\Omega\right)e_j \\
&= \left(e_j^{'}\Omega\otimes e_j^{'}\right)\sqrt{T}vec\left(\hat{\Psi}-\Psi\right) + \left(e_j^{'}\Psi\otimes e_j^{'}\right)\sqrt{T}vec\left(\hat{\Omega}-\Omega\right) + o_p(1) \qquad (2.23) \\
&= \left[\left(e_j^{'}\Omega\otimes e_j^{'}\right), \left(e_j^{'}\Psi\otimes e_j^{'}\right)\right]\left[\begin{array}{c}\sqrt{T}vec\left(\hat{\Psi}-\Psi\right)\\ \sqrt{T}vec\left(\hat{\Omega}-\Omega\right)\end{array}\right] + o_p(1) \qquad (2.24)
\end{aligned}$$

The second equality uses the following straightforward relations

$$\begin{cases} e_j^{'}\left(\hat{\Psi}-\Psi\right)\Omega e_j & = \left(e_j^{'}\Omega\otimes e_j^{'}\right) vec\left(\hat{\Psi}-\Psi\right) \\ e_j^{'}\Psi\left(\hat{\Omega}-\Omega\right)e_j & = \left(e_j^{'}\Psi\otimes e_j^{'}\right) vec\left(\hat{\Omega}-\Omega\right) \\ \left(\hat{\Psi}-\Psi\right)\left(\hat{\Omega}-\Omega\right) & = o_p\left(1/\sqrt{T}\right) \end{cases}$$

From the results in formulas 2.22, 2.24and equation 2.19, $\sqrt{T}\left(\hat{GI}_j - GI_j\right)$ is asymptotically normal with variance

$$\begin{aligned}
\Sigma_{iris} &= \sigma_{jj}^{-1}\left[\left(e_j^{'}\Omega\otimes e_j^{'}\right), \left(e_j^{'}\Psi\otimes e_j^{'}\right)\right]\left[\begin{array}{cc}\Sigma_\psi & 0\\ 0 & \Sigma_{\hat{\sigma}}\end{array}\right]\left[\left(e_j^{'}\Omega\otimes e_j^{'}\right), \left(e_j^{'}\Psi\otimes e_j^{'}\right)\right]^T \\
&= \sigma_{jj}^{-1}\left(e_j^{'}\Omega\otimes e_j^{'}\right)\Sigma_\psi\left(\Omega e_j\otimes e_j\right) + \sigma_{jj}^{-1}\left(e_j^{'}\Psi\otimes e_j^{'}\right)\Sigma_{\hat{\sigma}}\left(\Psi^{'}e_j\otimes e_j\right)
\end{aligned}$$

□

2.6.2.3 Proof of proposition 2.2

Proof. I want to show here that when $N_s > N_f$ then $IRIS_s > IRIS_f$. For this I focus only on the numerators of the IRIS, the denominators being the same. According to formula 2.9, the numerator of the IRIS are given by

$$\begin{aligned}
Num_s &= \sigma_{11}^{-1}\left(\psi_{11}\sigma_{11} + \psi_{12}\sigma_{12}\right)^2 \\
Num_f &= \sigma_{22}^{-1}\left(\psi_{11}\sigma_{12} + \psi_{12}\sigma_{22}\right)^2
\end{aligned}$$

I will show that if $N_s > N_f$ then $Num_s - Num_f > 0$. I have

$$\begin{aligned}
Num_s - Num_f &= \sigma_{11}^{-1}\left(\psi_{11}\sigma_{11}+\psi_{12}\sigma_{12}\right)^2 - \sigma_{22}^{-1}\left(\psi_{11}\sigma_{12}+\psi_{12}\sigma_{22}\right)^2 \\
&= \psi_{11}^2\sigma_{11} + \sigma_{11}^{-1}\psi_{12}^2\sigma_{12}^2 + 2\psi_{11}\psi_{12}\sigma_{12} - \sigma_{22}^{-1}\psi_{11}^2\sigma_{12}^2 - \psi_{12}^2\sigma_{22} - 2\psi_{11}\psi_{12}\sigma_{12} \\
&= \left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) + \sigma_{12}^2\left(\sigma_{11}^{-1}\psi_{12}^2 - \sigma_{22}^{-1}\psi_{11}^2\right) \\
&= \left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) + \sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1}\left(\sigma_{22}^{-1}\psi_{12}^2 - \sigma_{11}^{-1}\psi_{11}^2\right) \\
&= \left(1-\sigma_{12}^2\sigma_{11}^{-1}\sigma_{22}^{-1}\right)\left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) \\
&= \left(1-\rho^2\right)\left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right)
\end{aligned}$$

Where $\rho = \sigma_{12}\sqrt{\sigma_{11}^{-1}\sigma_{22}^{-1}}$ is a correlation coefficient. $\rho^2 \leq 1$ so it is sufficient to show that

$$\left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) \geq 0$$

The VECM representation 2.16 gives

$$\begin{pmatrix} \psi_{11} & \psi_{12} \end{pmatrix} = \begin{pmatrix} N_s & N_f \end{pmatrix} \tag{2.25}$$

To Compute the variance (Ω) of the VECM errors I use the following notations to simplify the presentation:

$$\begin{aligned}
a &= HN_s \\
b &= HN_f \\
c &= AN_sN_f \\
e_t^s &= v_t + w_t^s \\
e_t^f &= v_t + w_t^f \\
N &= N_s + N_f \\
K &= (a+b+c)^2
\end{aligned}$$

Thus $\Pi = \frac{1}{d}\begin{pmatrix} a+c & b \\ a & b+c \end{pmatrix}$ and the VECM errors are

$$\begin{pmatrix} u_t^s \\ u_t^f \end{pmatrix} = \Pi \begin{pmatrix} e_s \\ e_f \end{pmatrix} = \frac{1}{d}\begin{pmatrix} (a+c)\,e_s + be_f \\ ae_s + (b+c)\,e_f \end{pmatrix}$$

I have

$$\begin{array}{lll} \sigma_s^2 & = var(e_s) & = v^2 + w^2/N_s \\ \sigma_f^2 & = var(e_f) & = v^2 + w^2/N_f \\ \sigma_{sf} & = cov(e_s, e_f) & = v^2 \end{array} \tag{2.26}$$

So I compute the variance of the VECM error Ω

$$\begin{aligned} var\left(u_t^s\right) &= (a+c)^2\sigma_s^2 + b^2\sigma_f^2 + 2b(a+c)\sigma_{sf} \\ var\left(u_t^f\right) &= a^2\sigma_s^2 + (b+c)^2\sigma_f^2 + 2a(b+c)\sigma_{sf} \end{aligned} \tag{2.27}$$

Replacing 2.26 in 2.27 gives

$$\begin{aligned} var\left(u_t^s\right) &= v^2\left[(a+c)^2 + b^2 + 2b(a+c)\right] + w^2\left[\frac{1}{N_s}(a+c)^2 + \frac{1}{N_f}b^2\right] \\ &= v^2(a+b+c)^2 + w^2\left[\frac{1}{N_s}(a+c)^2 + \frac{1}{N_f}b^2\right] \\ &= v^2K^2 + w^2\left[\frac{1}{N_s}\left(N_s\left(AN_f + H\right)\right)^2 + \frac{1}{N_f}\left(HN_f\right)^2\right] \\ &= v^2K^2 + w^2\left[N_s\left(AN_f + H\right)^2 + H^2N_f\right] \\ &= v^2K^2 + w^2\left(N_sA^2N_f^2 + 2AHN_sN_f + H^2N_s + H^2N_f\right) \\ &= v^2K^2 + w^2\left(N_sA^2N_f^2 + 2AHN_sN_f + H^2N\right) \end{aligned} \tag{2.28}$$

Similarly

$$var\left(u_t^f\right) = v^2K^2 + w^2\left(N_fA^2N_s^2 + 2AHN_sN_f + H^2N\right) \tag{2.29}$$

Using 2.27 and 2.25 gives

$$\begin{aligned} \left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) &= N_s^2 var(u_t^s) - N_f^2 var(u_t^f) \\ &= N_s^2\left[v^2K^2 + w^2\left(N_sA^2N_f^2 + 2AHN_sN_f + H^2N\right)\right] \\ &\quad - N_f^2\left[v^2K^2 + w^2\left(N_fA^2N_s^2 + 2AHN_sN_f + H^2N\right)\right] \\ &= v^2K^2\left(N_s^2 - N_f^2\right) + w^2A^2N_s^2N_f^2\left(N_s - N_f\right) + w^2N\left(N_s^2 - N_f^2\right) \end{aligned}$$

When $N_s > N_f$ all the the terms on right-hand side are strictly positive so $\left(\psi_{11}^2\sigma_{11} - \psi_{12}^2\sigma_{22}\right) \geq 0$

and

$$IRIS_S > IRIS_f$$

$\square$

Chapter 3

Continuous time Analysis of Price Discovery

Abstract: For assets that are traded simultaneously on many markets, the problem of measuring their contribution to price discovery has mainly been studied in discrete time setup. We provide a new way to study this problem in the continuous time setup. We propose a measure evaluating the permanent impact of a shock on a market's innovation. It has advantages on the literature in that: it is in continuous-time and deals with non-informative microstructure noises; it provides a unique meaningful measure of information processing.

Keywords: *continuous-time cointegration, Generalised Impulse Response, preaverraging, Price discovery, Modulated Realized Covariance*

JEL: *C32, C58, G14*

3.1 Introduction

The study of asset prices dynamic has been accelerated by the availability of data and the development of adapted statistical techniques. The development of High Frequency trading have made available data sampled at a frequency close to the continuous time. In addition, the listing of assets on many markets and the competition between them, maintain the focus on the understanding of price discovery mechanism in this cross market setting. The main interest remaining to study the relative contribution of a market to the formation of the fundamental price. The pioneer study of Hasbrouck (1995) presented a measure of price discovery called Information Share (IS) and provides comparison of New York Stock Exchange and the Regional exchanges in the quotes formation of thirty Dow stocks. The IS has some drawbacks and an intensive literature discusses improvements, properties and alternatives to this measure: Harris et al. (see 2002b), De Jong (2002), Lehmann (2002), Hasbrouck (2002),Baillie et al. (2002),Yan and Zivot (2010), Lien and Shrestha (2014). Some papers use the model-free price discovery measures provided by the literature to study the determinants of market performance. Eun and Sabherwal (2003), Harris et al. (2002b),Blume and Goldstein (1997); Chakravarty et al. (2004); Huang (2002), Barclay et al. (2003).

With the availability of high frequency data, the tendency is to use all of them to exploit all available information. In fact, on one hand, Hasbrouck (1995) recommended to sample data at high frequency in order to tighten the IS bounds. This is done by many papers, but this practice ignores that at high frequency non informative part of the noise dominates the variances estimation. On the other hand, it is statistically "absurd" to not use the huge amount of data once we are lucky to have them. Meanwhile, as shown in the previous chapter, price discovery measures using high frequency data can be seriously misleading if there are too much non informative noises in those data. The intuition, coming from the literature on integrated volatility estimation in the presence of microstructure noise, is that: Using high frequency data, the realized variance of the observe prices is completely driven by the microstructure noises and not by the fundamental price (See Andersen et al., 2000; Jacod et al., 2009; Zhang, 2006).

Even if they differ in the way they define price discovery mechanism, all the previous studies identify their price discovery measure by relying on a discrete framework. They rely on a Vector Error Correction Model of the non stationary price processes. In addition, their construction is based on a model that considers only microstructure noises related to information sources: Information asymmetry, market under/over reaction (Menkveld et al., 2007). It is not concerned with non-informative noises due for example to tick size or measurement errors. There is up to our knowledge no literature dealing with with price discovery measure of cross listed assets in continuous time.

In this paper, we model the price vector by using a Continuous time VECM, it is simple but considered as general enough in finance literature. We add some microstructure noises at the observation times in the model. By relying on Hasbrouck (1995) identification of the unobserved efficient price, we measure market contribution by writing a continuous time version of the IRIS measure presented in the first chapter. It consists of accessing how this permanent unobserved price reacts to a shock in one market. It is related to the generalized impulse response of Pesaran and Shin (1998) in the VAR literature, but instead of looking the response function of each market, we look at the response function of the permanent common component to the markets. We showed in the first chapter in a discrete time framework that it is a sensible way of defining price discovery. We propose the **High-frequency Impulse Response (HIR)** that is best defined as ***the variance of the change in the fundamental price over a period of time, resulting from a shock to the innovation in one market***. Our framework has the following advantages over the literature:

- It uses a continuous time-setup to deal with high frequency data.
- It accommodates a stochastic volatility, important for example to capture clustering effects.
- It explicitly deals with non-informative part of microstructure noises.

The remainder of the paper is organized as follow: The second section presents the high frequency measure of price discovery in a continuous time framework. In the third section, the estimation strategy is discussed. In the fourth section and Monte-Carlo simulations are performed an an application is done on Apple stocks prices on NYSE and NASDAQ. And the conclusion is presented in the fifth section.

3.2 Price discovery measure in Continuous time

The aim of this paper is to provide a continuous time framework for price discovery measures in a continuous time allowing a coherent use of High-frequency data. Using those data involves features that are extensively studied in the literature for volatility estimation in the presence of microstructure noises[1]. By relying on a continuous time VECM, we present the framework and derive a measure that we name High frequency Impulse Response (HIR).

[1]The importance of designing a measure for high frequency data comes principally from the distortion in the variance estimation when noises are present.

3.2.1 Setup

There is one asset that is traded on two markets 1 and 2. Let $X_t = (X_{1t}, X_{2t})$ be the price we would observed on market 1 and 2 in the absence of non-informative noise.

We make the following assumption which is simplifying, but is general enough in continuous time finance literature.

Assumption 3.1. *X_t admits the following CVAR(1) representation*

$$DX_t = -\Pi X_t + \Omega_t DB_t \tag{3.1}$$

where B_t is a standard Brownian motion. Ω_t is a zero-mean stationary stochastic volatility with $E(\Omega_t^2) = \Omega^2$.

Ω_t is the volatility matrix and $\Omega_t^2 \equiv \Omega_t \Omega_t'$ is the Covariance matrix.

This assumption is relaxed later in a dedicated section, after the construction of the measure and the estimation strategy. The following assumption simply clarifies that X_{1t} and X_{2t} are cointegrated.

Assumption 3.2. *X_{1t}, X_{2t} and X_t are non-stationary but their increments are covariance stationary. The spread between the two prices $X_{1t} - X_{2t}$ is covariance stationary*

The trade prices are recorded at discrete points $(t_i)_{i=1,..,n} \in [0, T]$ and are dirtied by microstructure noises not related to information (tick size, measurement errors...) $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t})$. We have the following observation equations:

$$\begin{aligned} P_{1th} =& X_{1th} + \varepsilon_{1th} \\ P_{2th} =& X_{2th} + \varepsilon_{2th} \end{aligned} \tag{3.2}$$

Formally this wight be justified by

writing $X_{1t} = Y_t + \mu_{1t}$ and $X_{2t} = Y_t + \mu_{2t}$

The fundamental price of the asset is Y_t, this is the price which in finance literature will result in perfect world , with no arbitrage.

$\mu_t = (\mu_{1t}, \mu_{2t})$ is the part of microstructure noise that is completely related to information, this is for example due to asymmetric information, market's over-reaction, or under-reaction to information[2]

[2]This is consistent with Menkveld et al. (2007) where they estimate a price equation comprising the sum of three elements. The first term is the efficient price, the second term is the market over/under-reaction to information proportional to the efficient price innovation, and the third part is the microstructure noise arising from bid ask spread and price discreteness.

Briefly said, the observed prices are equal to the efficient price plus information-correlated noise and plus information-uncorrelated noise. We make the following assumption on the noise:

Assumption 3.3. *The noise process (ε_t) is i.i.d and $X \perp\!\!\!\perp \varepsilon_t$.*

The symbol $\perp\!\!\!\perp$ means stochastic independence. The i.i.d noise is tricky to define in continuous time, but Christensen et al. (2010) gives probability space and filtration in which ε_t is well defined.

Under assumptions 1 to 3, X_t is cointegrated and admits the following Granger representation in continuous time

$$\begin{array}{lcrcl} MA(\infty) & : & DX_t & = & \mathscr{L}_c(D)\Omega_t DB_t \quad = \int_{\tau=0^-}^{\infty} c(\tau)e^{-D\tau}\Omega_t dB_t \\ VECM & : & \mathscr{L}_d(D)\Omega_t DB_t & = & -\alpha\beta' X_t + \Omega_t DB_t \end{array} \tag{3.3}$$

dB_t is the innovation and ($\mathscr{L}$) is the Laplace transform operator function

$$\mathscr{L}_c(D) = \int_{\tau=0^-}^{\infty} c(\tau)e^{-D\tau}d\tau$$

The Difference operator ($DX_t = \frac{1}{dt}dX_t$) is useful to write ARMA class processes in continuous time. As presented in Cochrane (2012) , working with D in continuous time rather than L renders simplify the manipulations. [3]

3.2.2 Construction of the measure

We construct the measure under a generality that X_t is cointegrated with cointegrating vectors $\beta = \begin{pmatrix} 1 & -\lambda_1 \end{pmatrix}$.

3.2.2.1 Identification of the permanent component

The Beveridge-Nelson decomposition of the Laplace operator function is

$$\mathscr{L}_c(D) = \mathscr{L}_c(0) + D\mathscr{L}_b(D) \tag{3.4}$$

replacing in 3.3 we get

$$DX_t = \mathscr{L}_c(0)\Omega_t DB_t + D\mathscr{L}_b(D)\Omega_t DB_t$$

[3]For example the $MA(\infty)$ representation of a Gaussian stationary process y_t is: $y_t = \int_{\tau=0}^{\infty} c(\tau)\sigma dB_{t-\tau} = \mathscr{L}_c(D)\Omega DB_t$

This parallels the discrete time version where $D\mathscr{L}_b(D)$ is the operator function of a stationary process. It permit a decomposition in term of common trend and stationary components:

$$DX_t = DZ_t + Dw_t \tag{3.5}$$

where $w_t = \mathscr{L}_b(D)DB_t$ is a stationary process, and $Z_t = X_t - w_t$ is a martingale (even a pure random walk) satisfying

$$DZ_t = \mathscr{L}_c(0)\Omega_t DB_t \tag{3.6}$$

We restate the difference with IS where X_t is the observed process and then the transitory part of price w_t has its innovations correlated with the innovation in the martingale component. Here X_t is unobserved and will be dirtied by the other sources of microstructure noise.

If $\mathscr{L}_c(0) \neq 0$ and is not full rank $r < n$ then X_t is cointegrated, and there exist a matrix β' such that $\beta' X_t$ is stationary. There is also another matrix α summarizing the impact of the long common trend on the variable X_t.

Cointegration required $\beta'\mathscr{L}_c(0) = 0$, and thus the row of $\mathscr{L}_c(0)$ are collinear. They are the same when the cointegrating coefficient is equal to 1. Otherwise, the second row $\psi_{2r} = \lambda_1^{-1}\psi_{1r}$ and using one row or the other will not change the value of our proposed measure. Let us consider the first row and define $\psi = \psi_{1r}$, we have

$$\mathscr{L}_c(0) = J\psi\,, with J = (1\,\lambda_1^{-1})^\top$$

From the formula 3.6 we have

$$dZ_t = J\psi\Omega_t dB_t \tag{3.7}$$

The fundamental price (common component to all the markets) of the asset is then

$$z_t = z_0 + \psi \int_0^t \Omega_s dB_s \tag{3.8}$$

3.2.2.2 The Impulse response function

Consider the permanent component of prices given by equation 3.8. In the spirit of IRIS we define the impulse response of this permanent price to a shock in market 1. For this we rewrite the permanent prices in term of original shocks $dW_t = \Omega_t dB_t$ and consider a shock of value $\delta_{1t} = dW_{1t}$, so the response of dz_t to a shock in market 1 is

$$GI_1(t) = E\left(dz_t | dW_{1t} = \delta_{1t}\right) \tag{3.9}$$

Similarly the response of dz_t to a shock $\delta_{2t} = dW_{2t}$ in market 2 is

$$GI_2(t) = E\left(dz_t | dW_{2t} = \delta_{2t}\right) \tag{3.10}$$

This measures how the permanent price reacts to news in the first and in the second market. This is a good intuitive property to define a measure of price discovery, in the sense that fundamental information impacts permanently the prices.

3.2.2.3 Constant Volatility case

We start our development by assuming that the volatility matrix Ω is constant, which allows a comprehensive presentation of our construction. Using the normality properties of the Brownian motion and the conditional expectation formula we derive

$$\begin{aligned} GI_1(t) &= \psi E\left(dW_t | dW_{1t} = \delta_{1t}\right) \\ &= \psi \Omega e_1 \sigma_{11}^{-1} \times \delta_{1t} \\ &= \psi \begin{pmatrix} \sigma_{11}, & \sigma_{12} \end{pmatrix}' \sigma_{11}^{-1} \times \delta_{1t} \end{aligned}$$

Thus the cumulative response on a period $[0,T]$ is given by

$$\begin{aligned} CGI_1(t) &= \psi \int_0^T \begin{pmatrix} \sigma_{11}, & \sigma_{12} \end{pmatrix}' \sigma_{11}^{-1} \times \delta_{1t} \\ &= \psi \int_0^T \begin{pmatrix} \sigma_{11}, & \sigma_{12} \end{pmatrix}' \sigma_{11}^{-1} dW_{1t} \end{aligned} \tag{3.11}$$

As Hasbrouck (1995) using the variance, the quadratic variation of this cumulative response is good summary of the amount of information due to market 1:

$$\begin{aligned} < CGI_1 > &= \left[\psi \begin{pmatrix} \sigma_{11}, & \sigma_{12} \end{pmatrix}' \sigma_{11}^{-1}\right]^2 \times < dW_{1t} > \\ &= \left[\psi \begin{pmatrix} \sigma_{11}, & \sigma_{12} \end{pmatrix}' \sigma_{11}^{-1}\right]^2 \times \sigma_{11} T \end{aligned} \tag{3.12}$$

Similarly for a shock on the second market we obtain that the cumulative response variation of the permanent component is

$$< CGI_2 > \;=\; \left[\psi\begin{pmatrix}\sigma_{12}, & \sigma_{22}\end{pmatrix}'\sigma_{22}^{-1}\right]^2 \times \sigma_{22}T \tag{3.13}$$

Thus we define the ***"High-frequency Impulse Response"*** **(HIR)** measure of price discovery as

$$HIR_1 \;=\; \frac{< CGI_1 >}{< CGI_1 > + < CGI_2 >}$$

If we replace the value it gives

$$HIR_j \;\equiv\; \frac{(T\sigma_{11})^{-1}\left[\psi_{11}(T\sigma_{11})+\psi_{12}(T\sigma_{12})\right]^2}{(T\sigma_{11})^{-1}\left[\psi_{11}(T\sigma_{11})+\psi_{12}(T\sigma_{12})\right]^2+(T\sigma_{22})^{-1}\left[\psi_{11}(T\sigma_{12})+\psi_{12}(T\sigma_{22})\right]^2} \tag{3.14}$$

We add T in front of each of the volatility coefficient to keep visible the continuous-time feature of the framework, this will be important for the next section and to understand the estimation strategy.

3.2.2.4 The stochastic volatility case

When we relax the assumption of a constant volatility, the spot variance matrix is now dependent on t and in previous framework the generalized response of the permanent component is

$$GI_1(t) \;=\; \psi\begin{pmatrix}\sigma_{11t}, & \sigma_{12t}\end{pmatrix}'\sigma_{11t}^{-1} \times \delta_{1t} \tag{3.15}$$

Let's denote by $\bar{\sigma}_{ijt} = \int_0^T \sigma_{ijt}dt$ and $\bar{\Omega}^2 = \int_0^T \Omega_t^2 dt$. By mimicking the previous construction we propose the following formula for our HIR.

$$HIR_1 = \frac{\bar{\sigma}_{11t}^{-1}\left[\psi_{11}\bar{\sigma}_{11t}+\psi_{12}\bar{\sigma}_{12t}\right]^2}{\bar{\sigma}_{11t}^{-1}\left[\psi_{11}\bar{\sigma}_{11t}+\psi_{12}\bar{\sigma}_{12t}\right]^2+\bar{\sigma}_{22t}^{-1}\left[\psi_{11}\bar{\sigma}_{12t}+\psi_{12}\bar{\sigma}_{22t}\right]^2} \tag{3.16}$$

For d markets with $d \geq 2$, the contribution of market j is

$$HIR_1 = \frac{\left[\psi\bar{\Omega}^2 e_j\right]^2 \bar{\sigma}_{jjt}^{-1}}{\sum_{j=1}^{d}\left[\psi\bar{\Omega}^2 e_j\right]^2 \bar{\sigma}_{jjt}^{-1}} \tag{3.17}$$

Where e_j is the vector having 1 at the j^{th} position and 0 elsewhere.

Some comments need to be done on this formula. In fact as there is the inverse of the variance in the formula of CGI_1, the formula 3.16 is not what appears exactly when computing the quadratic variation. Applying strictly the quadratic variation of the cumulative process of GI_1 in equation 3.15 will give a formula where the coefficient $\int_0^T \sigma_{12t} / \int_0^T \sigma_{11t}$ is replaced by $\int_0^T (\sigma_{12t}/\sigma_{11t})\,dt$.

This is not really an issue as first it is a definition and it doesn't have an impact on our relative measure. Second it will complicate the presentation and makes the estimation unfeasible without additional assumptions. Assuming for example that the volatility parameters are constant piece-wise, an estimation of this ratio can be done block-by-block using the methods presented in section 3.

Another important feature of our framework is that it provides a generalization to continuous time of the IS of Hasbrouck (1995). In fact the total information entering z_t on $[0,T]$ can be represented by its quadratic variation:

$$S_T = \frac{1}{T}\psi\left(\int_0^T \Omega_t^2 dt\right)\psi'$$

and by taking F as the Cholesky root of $\left(\int_0^T \Omega_t^2 dt\right)$, the **High-frequency Information Share (HIS)** of market j defined by formula 2.7 is:

$$HIS_j = \left([\psi F]_j\right)^2 / S_T \tag{3.18}$$

and by switching ordering one gets lower and upper bound on information share.

Remark: The intensive set of simulations that we perform shows that the price discovery is well estimated when the formula is adjusted by the variance of the microstructure noise. The corrected formula we suggest is then

$$HIR_1 = \frac{\bar{\sigma}_{11t}^{-1}\left[\psi_{11}\bar{\sigma}_{11t} + \psi_{12}\bar{\sigma}_{12t}\right]^2 \times \sigma_{\varepsilon_1}^2}{\bar{\sigma}_{11t}^{-1}\left[\psi_{11}\bar{\sigma}_{11t} + \psi_{12}\bar{\sigma}_{12t}\right]^2 \times \sigma_{\varepsilon_1}^2 + \bar{\sigma}_{22t}^{-1}\left[\psi_{11}\bar{\sigma}_{12t} + \psi_{12}\bar{\sigma}_{22t}\right]^2 \times \sigma_{\varepsilon_2}^2}$$

The variance of the non-informative noise will be estimated using the realized volatility by

$$\widehat{\sigma_{\varepsilon_1}^2} = (2n)^{-1}\sum_{t=1}^{n}\Delta p_{1t}^2 \text{ and } \widehat{\sigma_{\varepsilon_1}^2} = (2n)^{-1}\sum_{t=1}^{n}\Delta p_{2t}^2$$

3.3 Estimation

The estimation of the measure is done by computing each element of the formula 3.16 which includes components of the integrated covariance matrix $(\int_0^T \Omega_t^2 dt)$ and elements of the vector ψ. The volatility parameters will be estimated using existing covolatility estimators in recent econometric of high frequency data and ψ is estimated through the weighting matrix of the classical VECM equation.

3.3.1 Estimation of $\int_0^T \Omega_t^2 dt$

From Assumptions 1 to 5, using the observed noisy values, the system to estimate is:

$$\begin{cases} dX_t & = \mu_t dt + \Omega_t dB_t \\ P_t & = X_t + \varepsilon_t \end{cases} \tag{3.19}$$

With $\mu_t = \Pi X_t$

The integrated volatility $\int_0^T \Omega_t^2 dt$ of the process 3.19 is estimated in the recent high frequency econometric framework. We propose with proposition 3.4 to use the Modulated Realized Covariance (MRC) estimator of Christensen et al. (2010) which is robust to microstructure noise. We use the sub-optimal estimator which is always definite positive. It is minor issue here since the second part of the estimator is to remove the asymptotic bias coming entirely from noise.

Proposition 3.4. *Let* $\frac{k_n}{n^{1/2+\delta}} = \theta + o\left(n^{-1/4+\delta/2}\right)$, *with* $0 < \delta < 0.5$

In equation 3.26,

$$MRC[P]_n^{\delta} = \frac{n}{n-k_n+2} \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{P}_n^i (\bar{P}_n^i)' \tag{3.20}$$

$$MRC[P]_n^{\delta} \xrightarrow{h \to 0} \int_0^T \Omega_t^2 dt \; with\, h: \; discretization\; pace$$

where the different notations are presented in appendix.

In practice in this framework, the drift is known after the estimation of α. Removing the drift before computing the preaveraged return can improve the results.

3.3.2 Estimation of ψ

With cointegration properties the assumption 3.1 corresponds to the ECM

$$dX_t = -\alpha\beta' X_t dt + \Omega_t dB_t \tag{3.21}$$

For the estimation, a discretisation should be made. Advantages of the " exact discretisation " scheme form of process are highlighted by Phillips (1991), Comte (1999), Chambers (1999, 2011). And we have the following discrete time VECM representation

Proposition 3.5. *Using the exact discrete form of equation 3.21, the following VECM representation applies to* P_t

$$\Delta P_{th} = g(h,d)\, \alpha\beta' P_{th-h} + \xi_{th} \tag{3.22}$$

where $\xi_{th} = e_{th} + \Delta u_{th} + g(h,d)\,\alpha\beta' u_{th-h}$,
$e_{th} = \int_{th-h}^{th} e^{-(th-s)\alpha\beta'} \Omega_s dB_{th-s}$
and $g(h,d) = 1 - e^{-hd}$

Using the representation 3.25, $\hat{\alpha}$ is estimated consistently (see proposition 3.6), $\mathscr{L}_c(0)$ is computed using the formula

$$\mathscr{L}_c(0) = I - \alpha(\beta'\alpha)^{-1}\beta'$$

and then ψ is identified as a row of $\mathscr{L}_c(0)$.

Proposition 3.6. *Let*

- $Z_{th} = \beta' P_{th} = P_{1th} - P_{2th}$
- $\hat{\alpha} = h^{-1}\left(\sum_{t=1}^{T} Z_{th-1h} Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^{T} Z_{th-2h}\Delta P_{th}\right)$ [4]

when $Th \to \infty$ *and* $Th \times h^2 \to 0$

$$\sqrt{Th}\,(\hat{\alpha} - \alpha) \Longrightarrow O(1)\,\sqrt{2\Omega_u}W(1)$$

Where W *is the standard wiener process*

Proof. See Appendix □

3.3.3 Generalization

We will present how the previous framework is adapted to more general setting.

Estimation of ψ

According to the Granger representation theorem of cointegrated time series, there exists a representation in vector error correction form:

$$\mathscr{L}_d(D)DX_t = -\alpha\beta' X_t + \Omega_t DB_t \tag{3.23}$$

$$\int_{\tau=0}^{\infty} d(\tau)dX_{t-\tau} = -\alpha\beta' X_t + \Omega_t dB_t$$

The assumption 3.1 corresponds to the case where $\mathscr{L}_d(D) = I$.

[4] This is like an IV estimator of α in equation 3.25 using Z_{th-2h} as Instrument to solve for endogeneity due to measurements errors

Discretizing this equation in the simple scheme to estimate is α,

$$\Delta X_t = -\alpha\beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \ldots + \Gamma_K \Delta X_{t-K} + e_t \tag{3.24}$$

The last lag K can be chosen by information criteria.

With the observed value P_t we obtain the following linear regression model with autocorrelated error

$$\Delta P_t = -\alpha\beta' P_{t-1} + \Gamma_1 \Delta P_{t-1} + \ldots + \Gamma_K \Delta P_{t-K} + \xi_t \tag{3.25}$$

α (thus ψ) can be again estimated by linear regression, IV can be used with past value of $\beta' P_t$ as instruments. Consistency is difficult to established formally in this case when $T \to \infty$ and $h \to 0$. And especially because K should be moving, but as in the discrete time literature this is the same problem as choosing the K.

Estimation of $\int_0^T \Omega_t^2 dt$

From equation 3.3, to impose a restriction on the first term, it is assume a Dirac Delta in $c(\tau)$ at $\tau = 0$ such that its Laplace transform is $c_0 = 1$. that is the contemporaneous impact of the noise of the price is one,

$$DX_t = c(0)\Omega_t DB_t + \int_{\tau=0}^{\infty} c(\tau) e^{-D\tau} \Omega_t DB_t$$

$$dX_t = \left(\int_{\tau=0}^{\infty} c(\tau)\Omega_t dB_{t-\tau} \right) dt + c(0)\Omega_t dB_t = \mu_t dt + \Omega_t dB_t \tag{3.26}$$

So replacing in 3.2 the system to estimate is

$$\begin{cases} dX_t &= \mu_t dt + \Omega_t dB_t \\ P_t &= X_t + \varepsilon_t \end{cases} \tag{3.27}$$

Which is estimated using the MRC.

3.4 Simulations an application

3.4.1 Simulation setting

To evaluate the bias in the estimated price discovery measures, it is necessary to initialize the value of market's contribution and compare it with the estimations. This required specifying the model in the reduced error correction form.

We will consider designs of equation 3.21 written as

$$d\begin{pmatrix} X_{1t} \\ X_{2t} \end{pmatrix} = \begin{pmatrix} -\alpha_1 \\ \alpha_2 \end{pmatrix} Z_t dt + \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}^{0.5} dB_t \tag{3.28}$$

$Z_t = X_{1t} - X_{2t}$ is the stationary component

The observed prices are contaminated by i.i.d noises: $P_{ith} = X_{ith} + \varepsilon_{ith}$, $\varepsilon_{it} \sim D(0, \sigma^2_{\varepsilon_i})$, $i = 1, 2$.

We simulate a path of observation with equation 3.28 corresponding to $N = 100,000$, and $T = \Delta_{ti}^{-1} = 1000$. We also consider 2 designs, and for each we compute the average and standard deviation of the measures over 1000 replications.

Design 1: $\alpha_1 = \alpha_2 = 0.9$, $\sigma_1^2 = \sigma_2^2 = 1,\ \rho = 0$

In this design both markets have the same parameters and are structurally identical. The contribution of market 1 should be equal 50%.

Design 2: $\alpha_1 = 0.25$, $\alpha_2 = 0.75$ $\sigma_1^2 = \sigma_2^2 = 1, \rho = 0$

With those parameters, the first market dominates the price discovery process with a contribution of 90%.

Results:

The simulation results in Table 3.1 show that the HIR captures pretty well the dominance in price discovery. For design 1, the markets are structurally similar, and the true contribution of market 1 is equal to 50%. When the level of noises is equal in the two markets, the simulation gives 49.99% for market 1 with a standard deviation of 5.36. This is clearly better than the Hasbrouck Information share bounds 57.11 and 39.13 with a standard deviation of 20.56. The high standard deviations that appear in the simulations come from the instability in simulating cointegrated variables. When the level of noises is different, the previous conclusions remain. In design 2, the values of the parameters for simulations are chosen such that the contribution of market 1 is 90%. When the level of noise is the same with a variance of 0 or 5e-4, our measure gives a close value of 85%. When the level of noise is different the measure seems to worsen, but very slightly than what happens to ISu and ISl.

Table 3.1: Simulation results:

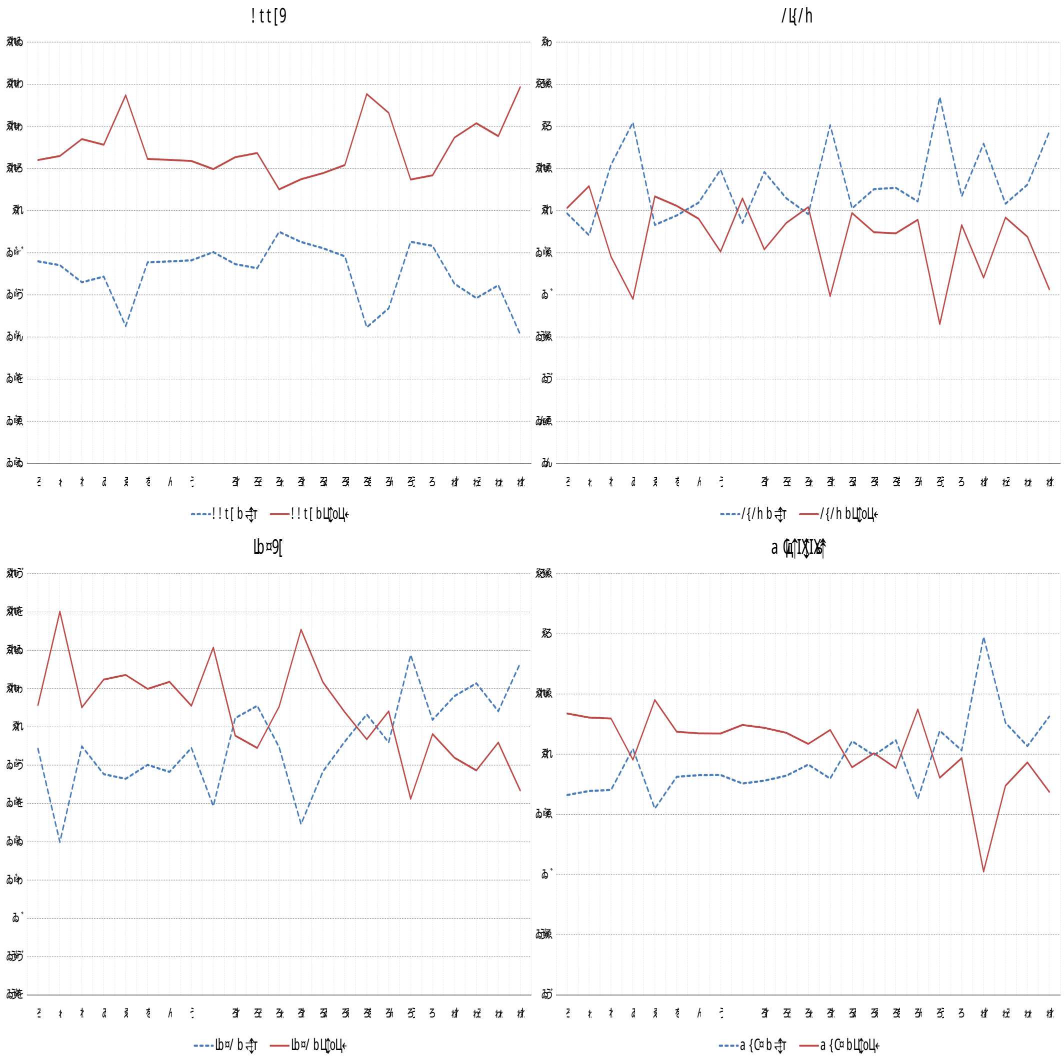
Noise c_1	c_2	***HIR***	HIS	$HISu_1$	$HISl_1$	$HISu_2$	$HISl_2$		ISu_1	ISl_1
		A. Value: ***50%***								
0	0	**49.99**	**49.91**	56.6	43.1	56.82	43.3		57.11	39.13
		(5.36)	(5.58)	(8.42)	(8.45)	(8.45)	(8.42)		(20.56)	(19.8)
0.5	0.5	**50.09**	**50.12**	57.0	43.29	56.7	42.9		52.3	52.9
		(6.7)	(6.6)	(8.8)	(8.7)	(8.7)	(8.8)		(32.6)	(32.7)
0.5	5	**49.89**	**49.88**	64.4	35.33	64.6	35.59		55.55	38.51
		(6.35)	(1.51)	(2.44)	(2.48)	(2.48)	(2.44)		(22.58)	(22.9)
		B. Value: ***90%***								
0	0	**84.9**	**85.1**	87.84	83.7	17.8	12.2		79.6	79.6
		(8.8)	(8.7)	(9.51)	(8.13)	(8.13)	(9.51)		(18.7)	(18.7)
0.5	0.5	**85.41**	**85.66**	84.81	89.73	10.28	15.2		82.1	81.92
		(7.8)	(7.8)	(8.5)	(7.05)	(7.05)	(8.5)		(17.52)	(17.57)
0.5	5	**84.28**	**84.30**	87.89	83.48	16.51	12.10		76.26	76.41
		(8.36)	(8.39)	(7.74)	(9.01)	(9.01)	(7.74)		(22.85)	(22.82)

c_1, c_2 in 10^{-4}. Note: The table reports the High frequency measures averaged over 1000 replications. Standard deviation in parenthesis. HISu and HISl are the upper and the lower bound on Information share computed in the continuous time framework. ISu_1, ISl_1 are the Hasbrouck discrete time IS.

3.4.2 Application

In this short application, we estimate the price discovery contribution of NYSE an NASDAQ for the Dow Jones stocks primary listed on NASDAQ. There are four stocks: Apple (AAPL), Microsoft (MSFT), Cisqo (CSCO) and Intel (INTC) during the 23 trading days of March 2011. The results are plotted in Figure 3.1 and shows the evolution of the daily contribution of each market. We obtain that price discovery mostly happens in NASDAQ, and there are very few days where NYSE dominates NASDAQ. These results confirm the finding in the previous chapter application, where in addition we found that information for those assets was driven by orders of small quantities.

Figure 3.1: Evolution of the Daily Contribution to price discovery



Note: The figures show the daily High Frequency Impulse Response during the 23 trading days of March 2011.

3.5 Conclusion

This paper aims at providing the literature with a new framework for assessing price discovery in the context of High-frequency data. The literature on this topic have mainly used a discrete time representation of the price dynamic, with some drawbacks highlighted in the previous sections: non-unique value of price discovery measure, non-intuitive identification of the fundamental price, absence of non-informative microstructure noise. We tried to build a framework that is not concerned by all these issues. In a continuous time setup, we proposed to consider the Generalized impulse response of the permanent price, to a shock in one market. Thus, constructing an invariant measure of price discovery with economic relevance. We proposed the High-Frequency Impulse Response measure which is not affected by information-uncorrelated noise. Our strategy is to build the error correction model of latent prices with stochastic volatility and additive non-informative noise. Thanks to the recent literature on integrated volatility estimation in the noisy diffusion setup, we are able to estimate the objects of interest by getting rid of the noise.

Using some simulations, we obtained that our framework displays good properties in term of capturing price discovery. The application is done on the Dow Jones stock listed on NASDAQ: Apple, Cisco, Intel and Microsoft. It showed that NASDAQ dominates the NYSE in the prices formation of those stocks. The High frequency Impulse response HIR can ultimately be used on empirical analysis to investigate the determinants of a market's efficiency. For example by a regression of HIR on a set of explanatory variables: number of market makers, latency, liquidity, trade size, location, trading fees...etc. The framework also considered the information as being smoothly introduced into price, but some empirical evidences suggest that information surprises enter prices as jumps, further research will attempt to consider price discovery in a models comprising a jump process.

3.6 APPENDIX C

3.6.1 Proofs

3.6.1.1 Proof of proposition 3.5

Proof. By solving the differential equation represented by 3.21 for a given h

$$\begin{array}{rcl} X_{th} & = & e^{-h\alpha\beta}X_{th-h}+e_{th} \\ \Delta X_{th} & = & \left(e^{-h\alpha\beta}-I\right)X_{th-h}+e_{th} \end{array} \tag{3.29}$$

with $e_{th}=\int_{th-h}^{th}e^{-(th-s)\alpha\beta'}\Omega_s dB_{th-s}$.
$exp\left(-h\alpha\beta'\right)=\sum_{l=0}^{\infty}\left(-h\right)^l\left(\alpha\beta'\right)^l$
$\left(\alpha\beta'\right)^2=\alpha\beta'\alpha\beta'=d\times\alpha\beta'$ where $d=\beta'\alpha=\alpha_1-\alpha_2$
by recurrence $\left(\alpha\beta'\right)^l=d^{l-1}\times\alpha\beta'$ and we have

$$\begin{array}{rcl} exp\left(-h\alpha\beta'\right)-I & = & \left(-h\alpha\beta'+\left(-h\right)^2\left(\alpha\beta'\right)^2/2+\ldots\right) \\ & = & \alpha\beta'\left(-h+\left(-h\right)^2d/2+\ldots\right) \\ & = & \frac{\alpha\beta'}{d}\left(-hd+\left(-h\right)^2d^2/2+\ldots\right) \\ & = & \frac{\alpha\beta'}{d}\left(-1+1-hd+\left(-h\right)^2d^2/2+\ldots\right) \\ & = & \alpha\beta'\frac{\left(-1+e^{-hd}\right)}{d} \\ & = & -\alpha\beta'g(h,d) \end{array}$$

With $g(h,d)=\frac{\left(1-e^{-hd}\right)}{d}$,
replacing in the expression 3.29 gives

$$\Delta X_{th}=-g\left(h,d\right)\alpha\beta'X_{th-h}+e_{th}$$

With the observed value P_t :

$$\begin{array}{rcl} \Delta P_{th} & = & \Delta X_{th}+\Delta u_{th} \\ & = & -g(h,d)\alpha\beta'X_{th-h}+\Delta u_{th}+e_{th} \\ & = & -g(h,d)\alpha\beta'P_{th-h}-g(h,d)\alpha\beta'u_{th-h}+\Delta u_{th}+e_{th} \\ & = & -g(h,d)\alpha\beta'P_{th-h}+\xi_{th} \end{array}$$

with $\xi_{th}=e_{th}+\Delta u_{th}-g(h,d)\alpha\beta'u_{th-h}$ □

Lemma 3.7.

1. $g\left(h,d\right)=h+O\left(h^2\right)$

2. $Var(e_{th}) = h\Omega^2 + O(h^2)$

3. $Var(\xi_{th}) = 2\Omega_u^2 + O(h)$

4. $E\left(\xi_{th}\xi_{th-h}'\right) = \Omega_u^2 + g(h,d)\alpha\beta'\Omega_u^2 = \Omega_u^2 + O(h)$

Proof. .

1. $g(h,d) = \frac{\left(1-e^{-hd}\right)}{d} = h - (-h)^2 d/2 + \ldots = h + O(h^2)$

2. $Var(e_{th}) = Var\left(\int_{th-h}^{th} e^{-(th-s)\alpha\beta'}\Omega_s dB_s\right) = Var\left(\int_0^h e^{-u\alpha\beta'}\Omega_s dB_u\right)$
   $Var(e_{th}) = \int_0^h (I - g(h,u)\alpha\beta')\Omega^2 (I - g(h,u)\alpha\beta')' du$
   $Var(e_{th}) = h\Omega^2 - \left(\int_0^h g(h,u)du\right)\alpha\beta'\Omega^2 - \left(\int_0^h g(h,u)du\right)\Omega^2\beta\alpha' + \int_0^h g(h,u)^2 \alpha\beta'\Omega^2\beta\alpha' du$
   Using 1) we have the result: $Var(e_{th}) = h\Omega^2 + O(h^2)$

3. $\xi_{th} = (e_{th} + u_{th} - (1 - g(h,d)\alpha\beta')u_{th-h})$

$$\begin{aligned}
Var(\xi_{th}) &= Var(e_{th}) + \Omega_u^2 + \left(1 - g(h,d)\alpha\beta'\right)\Omega_u^2\left(1 - g(h,d)\beta\alpha'\right)\\
&= h\Omega^2 + O\left(h^2\right) + 2\Omega_u^2 - g(h,d)\left(\alpha\beta'\Omega_u^2\right) - g(h,d)\left(\Omega_u^2\beta\alpha'\right) + g(h,d)^2\left(\alpha\beta'\Omega_u^2\beta\alpha'\right)\\
&= 2\Omega_u^2 + \left(\Omega^2 - \alpha\beta'\Omega_u^2 - \Omega_u^2\beta\alpha'\right)O(h) + O\left(h^2\right)
\end{aligned}$$

4. .

$$\begin{aligned}
E\left(\xi_{th}\xi_{th-h}'\right) &= E\left(\left(e_{th} + u_{th} - \left(1 - g(h,d)\alpha\beta'\right)u_{th-h}\right)\left(e_{th-h} + u_{th-h} - \left(1 - g(h,d)\alpha\beta'\right)u_{th-2h}\right)'\right)\\
&= \left(1 - g(h,d)\alpha\beta'\right)\Omega_u^2\\
&= \Omega_u^2 - g(h,d)\alpha\beta'\Omega_u^2
\end{aligned}$$

□

Lemma 3.8. *Let $Z_{th} = \beta' P_{th} = P_{1th} - P_{2th}$ then*
$Var(Z_{th}) = O\left(h^{-1}\right)$
$E(Z_{th}Z_{th-h}) = O\left(h^{-1}\right)$

Proof. $Z_{th} = \beta' P_{th} = P_{1th} - P_{2th}$

$$\begin{aligned}
\Delta P_{th} &= g(h,d)\alpha Z_{th-h} + \xi_{th} \\
\Delta\beta' P_{th} &= g(h,d)\beta'\alpha Z_{th-h} + \beta'\xi_{th} \\
\Delta Z_{th} &= g(h,d)d \times Z_{th-h} + \beta'\xi_{th} \\
Z_{th} &= (1+g(h,d)d) \times Z_{th-h} + \beta'\xi_{th} \\
Z_{th} &= e^{-hd} \times Z_{th-h} + \beta'\xi_{th}
\end{aligned}$$

$$\begin{aligned}
Var(Z_{th})\left(1-e^{-2hd}\right) &= Var(\beta'\xi_{th}) + Cov\left(\beta'\xi_{th}Z'_{th-h}\right) \\
&= \beta' Var(\xi_{th})\beta + \beta' Cov\left(\xi_{th}\xi'_{th-h}\right)\beta \\
&= \beta'\left(2\Omega_u^2 + O(h)\right)\beta + \beta'\left(\Omega_u^2 + O(h)\right)\beta \\
Var(Z_{th})\left(-2hd + O\left(h^2\right)\right) &= 3\beta'\Omega_u^2\beta + O(h) \\
Var(Z_{th}) &= \frac{3\beta'\Omega_u^2\beta + O(h)}{h(-2d+O(h))} \\
&= O\left(h^{-1}\right)
\end{aligned}$$

$$\begin{aligned}
Cov\left(Z_{th}, Z'_{th-h}\right) &= e^{-hd} Var(Z_{th-h}) + Cov\left(\beta'\xi_{th}Z'_{th-h}\right)\beta' Var(\xi_{th})\beta + \beta' Cov\left(\xi_{th}\xi'_{th-h}\right)\beta \\
&= e^{-hd} O\left(h^{-1}\right) + \beta'\left(\Omega_u^2 + g(h,d)\alpha\beta'\Omega_u^2\right)\beta\beta'\left(2\Omega_u^2 + O(h)\right)\beta + \beta'\left(\Omega_u^2 + O(h)\right)\beta \\
&= e^{-hd} O\left(h^{-1}\right) + \beta'\Omega_u^2\beta + O(h) \\
&= O\left(h^{-1}\right)
\end{aligned}$$

$$\begin{aligned}
E\left(Z_{th-2h}\xi'_{th}\right)^2 &= Var(Z_{th-h}) \times Var(\xi_{th}) \\
&= O\left(h^{-1}\right) \times \left(2\Omega_u^2 + O(h)\right)
\end{aligned}$$

$\square$

3.6.1.2 Proof of proposition 3.6

Now we prove the proposition using the Lemma 3.7 to replace the variable by the bigs O results

$$\begin{aligned}
\hat{\alpha} &= h^{-1}\left(\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^T Z_{th-2h}\Delta P_{th}\right) \\
h\hat{\alpha} &= \left(\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^T Z_{th-2h}\Delta P_{th}\right) \\
&= -g(h,d)\alpha + \left(\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^T Z_{th-2h}\xi_{th}\right)
\end{aligned}$$

$$\begin{aligned}
h\hat{\alpha} - h\alpha &= \left(-g(h,d)+h\right)\alpha + \left(\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^T Z_{th-2h}\Delta P_{th}\right) \\
h(\hat{\alpha}-\alpha) &= O\left(h^2\right) + \left(\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\left(\sum_{t=1}^T Z_{th-2h}\xi_{th}\right) \\
\sqrt{Th}(\hat{\alpha}-\alpha) &= O\left(\sqrt{T}h^2\right) + \left(T^{-1}\sum_{t=1}^T Z_{th-1h}Z_{th-2h}\right)^{-1}\frac{1}{\sqrt{Th}}\left(\sum_{t=1}^T Z_{th-2h}\xi_{th}\right) \\
&\approx O\left(\sqrt{T}h^2\right) + \left(Cov\left(Z_{th}, Z'_{th-h}\right)\right)^{-1}\frac{1}{\sqrt{h}}\sqrt{Var(Z_{th-2h})Var(\xi_{th})} \times W(1) \\
&\approx O\left(\sqrt{T}h^2\right) + \left(O\left(h^{-1}\right)\right)^{-1}\sqrt{O(h)}\frac{1}{\sqrt{h}}\sqrt{2\Omega_u^2 + O(h)} \times W(1) \\
&\approx O\left(\sqrt{Th}h\right) + O(1) \times \sqrt{2\Omega_u^2 + O(h)} \times W(1) \\
&\Longrightarrow O(1)\sqrt{2\Omega_u^2} \times W(1)
\end{aligned}$$

Provided that $\sqrt{Th} \to \infty$ and $\sqrt{Th} \times h \to 0$.

3.6.1.3 Notations of proposition 3.4:

The notations in 3.4 originally come from the pre-averaging method of Jacod et al. (2009), which provides and estimator of the integrated volatility,

Assuming that the true log price is generated by an Îto process of the form

$$X_t = X_0 + \int_0^t u_s ds + \int_0^t \sigma_s dW s$$

where W is a standard wiener process and $\mu = (\mu_t)$ and $\sigma = (\sigma_t)$ are adapted processes.

and the noisy observed process is given by

$$Z_t = X_t + \varepsilon_t$$

let k_n be the size of each group in the length of each period at the first stage. it is chosen such that

$k_n = \theta * n^{0.5+\delta} + o(n^{\frac{1}{4}})$

We have a function g on $[0,1]$, continuous, piece-wise C^1 with a piece-wise Lipschitz derivative g'satisfying $g(0) = g(1) = 0$, and $\int_0^1 g(x)^2 dx > 0$. let's define some quantities and notation.

$g_i^n = g(i/k_n)$,

$\phi_1(s) = \int_s^1 g'(u) g'(u-s) du, \qquad \phi_2(s) = \int_s^1 g(u) g(u-s) du,$

$fors > 1, \phi_1(s) = 0, \quad \phi_2(s) = 0,$

$\phi_{ij} = \int_0^1 \phi_i(s)\phi_j(s) ds, \qquad \psi_i = \phi_i(0),\ i,j = 1,2$

their empirical equivalent:

$\widehat{\psi}_1 = k_n \sum_{j=1}^{k_n} (g_{j+1}^n - g_j^n)^2 \qquad \widehat{\psi_2}(j) = \frac{1}{k_n} \sum_{j=1}^{k_n-1} (g_j^n)^2 \widehat{\phi}_1(j) = \sum_{i=j+1}^{k_n} (g_{j-1}^n - g_i^n)(g_{i-j-1}^n - g_{i-j}^n) \qquad \widehat{\phi}_2(j) = \sum_{i=j+1}^{k_n} g_i^n g_{j-1}^n$

$\widehat{\phi}_{11}(j) = k_n \left(\sum_{j=0}^{k_n-1} (\widehat{\phi}_1(j))^2 - \frac{1}{2} (\widehat{\phi}_1(0))^2 \right) \quad \widehat{\phi}_{12}(j) = \frac{1}{k_n} \left(\sum_{j=0}^{k_n-1} \widehat{\phi}_1(j) \widehat{\phi}_2(j) - \frac{1}{2} \widehat{\phi}_1(0) \widehat{\phi}_2(0) \right)$

Then, pre-averaged return are defined as:

$$Z_i^n = Z_{i\Delta_n}, \quad \Delta_i^n Z = Z_i^n - Z_{i-1}^n, \quad \overline{Z}_i^n = \sum_{j=1}^{k_n-1} g_i^n \Delta_{i+j}^n Z$$

Chapter 4

Volatility discovery across interlinked securities

Abstract: Where does new volatility enter the volatility of securities listed in many countries? While literature has focused on where information enters the price. I develop a framework to study how each markets' volatility contributes to the permanent volatility of the Asset. I build a VECM with Autoregressive Stochastic Volatility framework estimated by MCMC method and Bayesian inference. This specification allows defining measures of a market's contribution to Volatility discovery. In the application, I study cash and 3Month futures markets of some metals traded on the London Metals Exchange. I also study the EuroStoxx50 index and its futures. I find that for most the securities, while price discovery happens on the cash market, the volatility discovery mostly happens in the Futures market. Overall, the results suggest that Information discovery and volatility discovery do not necessarily have the same determinants. In a second part of the study. I build a framework that exploits High frequency data and avoid computational burden of MCMC. I show that Realized variances are driven by a permanent component and I compute contribution NYSE and NASDAQ to permanent volatility of Dow Jones stocks. It appears for most of the stocks that, from March 2011 to May 2011. NYSE dominates the Volatility discovery process. I later check the correlation between Volatility Discovery measures, liquidity, Volume market Share by trade size, and volatility of volume. I obtain that volatility of the volume is the best determinant of volatility discovery, but low figures suggest others important factors.

Keywords: *Multivariate Stochastic Volatility. Monte Carlo Markov Chain. VECM.*

JEL: *C32. C58. G14*

4.1 Introduction

The recent decades have seen huge investigations and debates about how a market contributes to the fundamental price discovery of cross-listed assets and derivatives prices. The literature triggered by Hasbrouck (1995) and Harris et al. (2002b) proposed price discovery measures to evaluate the relative informativeness of one market's prices and to perform regressions aiming at accessing the determinants of a market's efficiency. Meanwhile, concerning Volatility formation, there is almost no literature nor such statistical measures.

The motivation for this literature is manifold: An investor who wants to avoid being adverse selected would prefer to trade in the markets where prices are close-to-efficiency. Regulators are also interested in having an efficient market for their stocks that are listed abroad (Eun and Sabherwal, 2003). Then, the derivatives products as futures contracts are generally implemented, among other reasons, to improve the price discovery mechanism. Papers are thus interested to which market, between cash and futures, conveys more fundamental information. For instance Chakravarty et al. (2004) show that informed trading happens in the spot market compared to option market. For the case of futures market, a bunch of papers study to which extend price of the asset on cash market reflects the information in the futures markets. Lien and Shrestha (2009) show that Price discovery takes place mostly in the futures market. Fricke and Menkhoff (2011) obtain that the 10-year Euro bond future contract on German sovereign debt is important but does not dominate two futures with shorter maturity.

This literature do not look at which market's volatility is dominant in forming the volatility of the fundamental price, while there are many reasons to analyze the volatility formation. First the theoretical literature shows that the information arrival and the volatility on information have different implications for price and price volatility formation (Ross, 1989). Second, given the arbitrage activities between spot and derivatives, missing the link between their volatilities can lead to incorrect inferences and misunderstanding of the relationship between their returns. Chan et al. (1991) advocates that for the case of cash and futures markets. Third, the volatility itself is an object of interest for investors and there are tradable volatility indexes like the VIX, and derivatives on the volatility. An actively traded derivative is the volatility of the volatility index (VVIX) which captures the uncertainty of investors about future volatility. In risk management terms also, it is obviously important to know which market's risk contributes the most to the fundamental risk of the asset.

The previous elements provide sufficient motivation to go far than information, and investigate on which market the trading with uncertainty on information is happening. This objective is by formulation an empirical problem This paper contributes to the literature by developing a simple framework in a cross-market setting, that allows measuring a relative contribution of a market's volatility to the volatility of the fundamental price.

The literature is not empty concerning the analysis of volatility transmission between derivatives and spot markets. Some papers investigate how the volatility in the futures market impacts the cash market's volatility and vice versa. This is done by using multivariate GARCH models see Bhar (2001), Antoniou and Holmes (1995), Antoniou and Holmes (1995), Zhong et al. (2004), Yang et al. (2012), Chan et al. (1991). Those papers are more concerned with a lead lag relationship, This paper also shows the volatility contagion, but in addition tackles a question that those papers do not look at: Does the volatility of futures market contributes more than the volatility of the cash market in the formation of the permanent price volatility? The paper innovates in this respect. A very small number of papers are approaching the topic in this way. Dias et al. (2016) assess the contribution to the permanent volatility of Dow Jones asset by assuming that realized volatilities are fractionally cointegrated. Baule et al. (2017) evaluate how Warrants market and classical option markets contribute to the formation of volatility. They use information share of each market computed in bivariate system of implied volatility. Our paper distinguishes from their in several aspects that are explained below.

In the first part, I rely on a classical Vector Error Correction Model (VECM) which is the simple powerful tool for interrelated assets with cointegrated prices. I also add an Autoregressive Stochastic Volatility dynamic to the innovations to form what I call next VECM-ASV. In the specification, the conditional volatilities of the both markets are also driven by common persistent component. I show that this is the case in the data. This allows then building the Volatility Share (VS), defined as ***the contribution of the volatility's innovation of a given market to the volatility of the permanent volatility,*** and the ***Permanent Volatility Share (PV)***, defined as ***the weight of a market in the permanent volatility***[1]**.** This paper has the advantage of defining simultaneously price discovery and volatility discovery in the same framework.

To estimate the model, I use MCMC simulation method in a Bayesian framework. This estimation method appears to be the adapted workhorse for time series models introducing non observable components like a stochastic variance (Tsay, 2005).

In the application on daily data, I estimate the VECM-ASV model to compute the contribution of 3-month futures and the cash markets to volatility discovery of some metals traded on the London Metal exchange. They are extracted from the database Eikon of Thomson Reuters. The data range is from January 2010 to December 2015. The literature on futures contracts has mixed results concerning price discovery. Most of the studies find that price discovery happens in the futures, but some find that there is informed trading in the cash market. Concerning volatility, I find that for most of the metals the contribution of the futures market to volatility discovery is big-

[1] See the definitions of IS and PT in Hasbrouck (1995) and Lehmann (2002)

ger than that of the cash market. For Aluminum and Tin, the Cash market dominates the futures in volatility discovery with contribution of 56% and 52.47%. To check if there is link between Volatility discovery and the price discovery, I compute the main price discovery measures IS and PT on those assets. The results I obtain vary by stocks. For Aluminum, Tin and Molybdenum, the IS and PT of the Cash market is the highest. Overall the results shows that for Aluminum and Tin, price discovery and Volatility discovery happen in the cash market; for copper and Molybdenum, price discovery happens in the cash market and volatility discovery in the futures market; For Lead, price discovery and Volatility discovery happen in the Futures market. I then consider the intraday EuroStoxx 50 Index and futures contracts on this index. The Futures market contributes for more than 90% to volatility discovery while price discovery happens for around 90% on the Cash market. Thus for Euro50Stoxx. informed trading happens on Cash market while Volatility trading happens in Futures market. Those results point out a that price discovery and volatility discovery do not necessarily have the same determinants.

In a second part. I develop a framework that exploits High frequency data[2] and avoids computational burden of MCMC. When High frequency data are available, Realized volatility is another way of capturing the total volatility in a day. Using Dow Jones stocks on NYSE and NASDAQ, I first show that volatilities appear also to be driven by a permanent component when measured by Realized Volatilities. This allows estimating contributions to Volatility discovery by simply applying Hasbrouck (1995) and Harris et al. (2002c) methodology to time series of Realized volatility. Dias et al. (2016) also propose a measure of volatility discovery based on IS and features of realized volatility. This paper proposes two measures of volatilities based on innovations variance and weighting matri and studies the determinant of volatility discovery.

In the application, I obtain for most of the stock that from March 2011 to May 2011, NYSE dominates the Volatility discovery process. I later check the correlation between Volatility Discovery measures, liquidity, Volume market Share by trade size, and volatility of volume. I obtain that volatility of the volume is the best determinant of volatility discovery. It is followed by the market share in big size trade. This result is coherent with the literature in a one market setting showing strong linkages between the volume activity and price volatility (Epps and Epps, 1976; Tauchen and Pitts, 1983). In absolute terms, those coefficients remain meanwhile low; suggesting the existence of other sources of volatility discovery.

The remainder of the paper is organized as follow: The second section presents the VECM-ASV model with its features and defines the Volatility discovery measures. Section 3 presents the MCMC method and the estimation strategy using Gibbs sampling. In the fourth section a simulation exercise is performed. In the fifth section an application is done on base metal traded on

[2]The first part of this chapter is concerned with VECM-ASV framework. mostly suitable for Low frequency data.

LME and the EuroStoxx 50 Index. The sixth section presents the analysis of volatility discovery with Realized volatilities on Dow jones stocks. Conclusion is presented in the seventh section.

4.2 Modeling

There are 2 strongly related securities that are traded at the respective prices p_{1t}. p_{2t}. For example for one asset listed on two markets, p_{1t} is the price of the asset on the first market and p_{2t} the price on the second. It is classical to consider these prices as cointegrated such that they have only one common trend representing the "latent efficient price". In fact arbitrageurs' activities prevent the prices between those markets from diverging.

4.2.1 A VECM with Stochastic-Volatility

By denoting the vector of prices by $P_t = (p_{1t}, p_{2t})'$, Johansen (1991) results show that the Vector Error Correction Model (VECM) 4.1 applies. I then add an Autoregressive dynamic to the error's volatility to form the VECM with Autoregressive Stochastic Volatility (VECM-ASV):

VECM-ASV(1):

$$\begin{aligned}
\Delta P_t &= -\alpha\beta' P_{t-1} + \Gamma_1 \Delta P_{t-1} + \ldots + \Gamma_K \Delta P_{t-K} + \varepsilon_t && (4.1)\\
\varepsilon_t &= H_t^{1/2}\eta_t && (4.2)\\
H_t &= \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} && (4.3)\\
log\left(\sigma_{1t}^2\right) &= a_{01} + a_{11} V_{t-1} + v_{1t} && (4.4)\\
log\left(\sigma_{2t}^2\right) &= a_{02} + a_{12} V_{t-1} + v_{2t} && (4.5)\\
V_t &= a + \lambda V_{t-1} + u_t && (4.6)
\end{aligned}$$

$$\textbf{with } u_t \sim i.i.d\left(0, \sigma_u^2\right),\ \eta_t \sim i.i.d\left(0, I\right),\ v_t = \begin{pmatrix} v_{1t} \\ v_{2t} \end{pmatrix} \sim i.i.d(0, \begin{pmatrix} \sigma_{v_1}^2 & 0 \\ 0 & \sigma_{v_2}^2 \end{pmatrix}).$$

The cointegrating vector is normalized to be $\beta' = \begin{bmatrix} 1 & -\beta_1 \end{bmatrix}$.

The model is classical in the formulation of the error term where the variance H_t is then modeled as a Factor with SV. First, I use a decomposition in a Constant Correlation framework. This will ensure that the Volatility matrix is positive semi-definite. For the dynamic of the individual volatility, since the innovations v_{1t} and v_{2t} can be negative, I use a log form that forces the individual conditional variances to be positive. The correlation coefficient ρ is constant so the model is restricted in some senses compared to free varying correlation or covariance (see Dynamic Conditional Correlation Garch). Meanwhile, the specification is general enough to include non diagonal matrix Ω in the price equation.

4.2.1.1 Identification of the common component

The specification of the common factor dynamic displays a parameter λ which measure persistence in the factor. As the conditional volatility of financial returns are known to be highly persistent. The common factor component is expected to be highly persistent and capturing all the persistence in both series. The coefficient λ is thus expected to be close to 1. If λ is set equal to 1, then the factor component is identify in a Berveridge Nelson decomposition of the vector of log volatiliy. Given the strongness of this assumption for volatility, I dont set a value for lambda. Meanwhile, assumptions are useful to extract the factor:

Assumptions:

- The Factor is a assumed to be a linear combination of the conditionnal variances

$$V_t = \psi_1 \times log\left(\sigma_{1t}^2\right) + \psi_2 \times log\left(\sigma_{2t}^2\right) \tag{4.7}$$

- The factor V_t is the axis that maximises the inertia of the projection of $log\left(\sigma_{1t-1}^2\right)$, $log\left(\sigma_{2t-1}^2\right)$; It can be obtained as their first principal component.

This imposes a long run relationship between the conditional variances.[3] This assumption is supported by the presence of arbitrageurs on volatility. For a cross-listed asset, traders on volatility can ultimately make profit on differences in the volatility for the asset on two markets. At the end of this section, I show empirically that the conditional variances are indeed driven by a common component. After removing the highly persistent component, the remaining parts show no persistence.

By replacing the expressions of $log\left(\sigma_{1t}^2\right)$ and $log\left(\sigma_{2t}^2\right)$ in 4.7 we have

$$V_t = (\psi_1 a_{01} + \psi_2 a_{02}) + (\psi_1 a_{11} + \psi_2 a_{12}) V_{t-1} + \psi_1 v_{1t} + \psi_2 v_{2t} \tag{4.8}$$

which allow extracting the persistence parameter

$$\lambda = \psi_1 a_{11} + \psi_2 a_{12} \tag{4.9}$$

[3] It also allows non-stationary volatility but this is not in contradiction with Cointegration. The non-stationary volatility can only slightly biased the Cointegration test (Cavaliere et al., 2010). This is not an issue here as the series I analyze are strongly cointegrated.

And the innovation in the common component

$$w_t = \psi_1 v_{1t} + \psi_2 v_{2t} \tag{4.10}$$

The specification can finally provides some nice interpretations in term of volatility spillovers. By replacing the factor formula 4.7 in the variances equations. They become

$$log\left(\sigma_{1t}^2\right) = a_{10} + a_{11}\psi_1 \times log\left(\sigma_{1t-1}^2\right) + a_{11}\psi_2 \times log\left(\sigma_{2t-1}^2\right) + v_{1t} \tag{4.11}$$

$$log\left(\sigma_{2t}^2\right) = a_{20} + a_{21}\psi_1 \times log\left(\sigma_{1t-1}^2\right) + a_{21}\psi_2 \times log\left(\sigma_{2t-1}^2\right) + v_{2t} \tag{4.12}$$

Which is a classical Autoregressive representation. The coefficients $(a_{11}\psi_2)$. $(a_{21}\psi_1)$.$(a_{11}\psi_1)$ and $(a_{12}\psi_2)$ have interesting interpretation in terms of individual volatility spillovers and persistence.

4.2.2 The volatility discovery measures

The common factor equation 4.7 is similar to the setup in which the price discovery measures are generally defined. By considering this factor as the permanent component in the volatility, a notion of contribution to volatility discovery can be build. The coefficient ψ_1(respectively ψ_2) measures the weight of market 1 (respectively market 2) in the formation of the common persistent volatility factor. Accordingly. as done with the Permanent Transitory measure, the contribution of market 1 to volatility discovery can be defined by the ***"Permanent Volatility Share"***:

$$PV_1 = \frac{\psi_1}{\psi_1 + \psi_2} \text{ and } PV_2 = \frac{\psi_2}{\psi_1 + \psi_2} \tag{4.13}$$

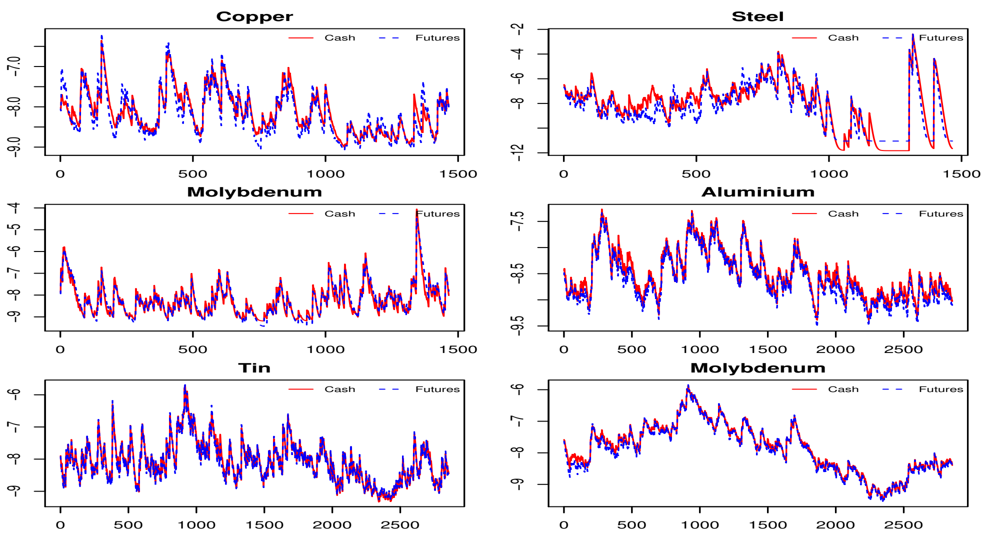
Considering now the innovation in the volatility component in equation 4.10. Its variance $\psi_1^2\sigma_{v1}^2 + \psi_2^2\sigma_{v2}^2$ can be considered as the total volatility of volatility entering the volatility at time. This allows another construction of a measure of volatility discovery. The ***"Volatility Share"*** is defined as the contribution of a market volatility to the volatility of the common volatility:

$$VS_1 = \frac{\psi_1^2\sigma_{v1}^2}{\psi_1^2\sigma_{v1}^2 + \psi_2^2\sigma_{v2}^2} \text{ and } VS_2 = \frac{\psi_2^2\sigma_{v2}^2}{\psi_1^2\sigma_{v1}^2 + \psi_2^2\sigma_{v2}^2} \tag{4.14}$$

4.2.3 Common factor in Volatilities

The previous specification supposes that the conditional variances are driven by a common factor. To check the validity of this assumption, I estimate a Multivariate GARCH model on the data, and

Figure 4.1: Conditional volatilities of Cash and Futures markets

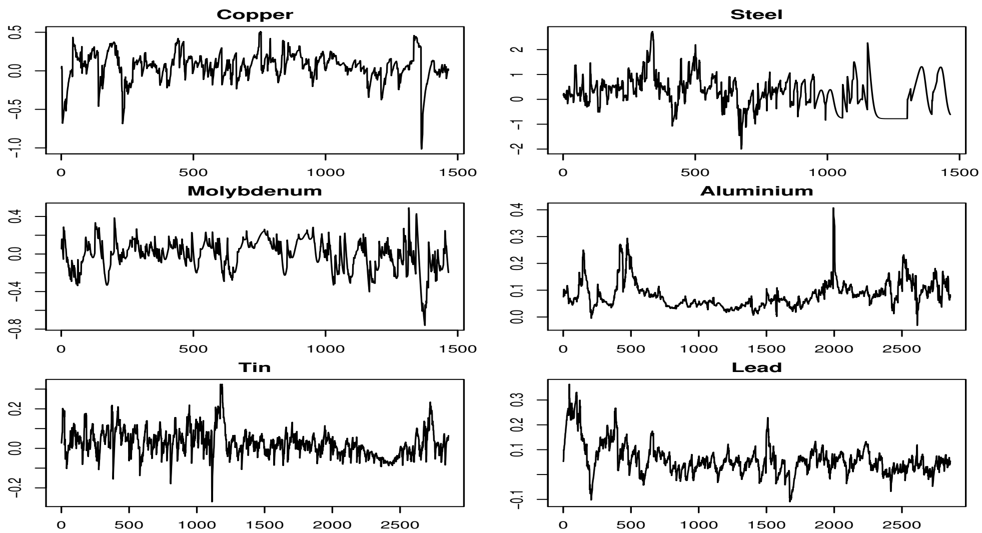


The graphs plot the conditional volatilities of Cash and Futures market estimated in a Bivariate VECM with CCC-Garch errors.

I extract the log conditional variances. With those conditional variances. I do a classical factor analysis.

The data are comprised of futures and cash markets prices of some metals traded on the London Metals Exchange (LME). It is the same commodities that were studied in chapter 1: Tin (Ti), Aluminium (Al), Nickel (Ni) and Lead (Le), Copper (cu). Here I focus on the comparison between the two main securities that are the cash market and the 3-month futures contract. The cash market is the market 1 and the futures market is the market 2. I estimate Bivariate VECM models with errors driven by a CCC-Garch model. This corresponds to the VECM-ASV without the volatility terms v_{1t} and v_{2t} in the variances equations 4.1. Figure 4.1 shows some instances of common evolution of conditional variances. It is clear from those graphs that the conditional volatilities of cash and futures move perfectly closely and exhibits common factor behavior. I compute the volatility gap ($Volatility_{Cash} - Volatility_{Futures} = residuals$). Their plot in Figure 4.2 and ADF test in Table **??** shows that the non stationarity of those residuals can strongly be rejected. Later after the estimation of the VECM-ASV, a more refined analysis is done. Figure 4.13 shows the common persistence in the conditionnal volatility. Then after the removal of the estimated common factor, figure 4.14 shows that there is no persistence anymore in the remaining part.

Figure 4.2: Volatility gap between Cash and Futures markets



4.3 Estimation: MCMC method and Gibbs sampler

The VECM-ASV model specified in equation 4.2 incorporates unobserved random variables. Markov Chain Monte Carlo (MCMC) procedure in Bayesian framework is adapted for estimation purpose (Tsay, 2005). For this, I have to specify details of the likelihood and the prior distributions of all the parameters.

4.3.1 MCMC method

I start by regrouping the parameters using Tsay (2005) notations. Let $G_i = (h_{i1}, \dots, h_{in})^{'}$ and $G = [G_1, G_2]$. The parameters of the mean equation are $B \equiv vec\,[\alpha : \Gamma]$ where the *vec* operator stacks the columns of matrix into one column.

The volatility equation parameters $\omega = (a_{01}, a_{11}, a_{02}, a_{12})$

The VECM (4.1) is represented as

$$R = \widetilde{B}X + U$$

with T= sample size . $U \equiv [\varepsilon_0, \dots \varepsilon_T]$

$R \equiv [\Delta P_1, \dots \Delta P_T]$ $\widetilde{B} = [\alpha : \Gamma]$. $\Gamma \equiv [\Gamma_1, \dots, \Gamma_K]$

$X = \left[\beta' Y_{-1},\ \ \Delta X\ \right]^{'}$. $Y_{-1} \equiv [P_0, \dots P_{T-1}]$.

$\Delta X \equiv [\Delta X_0, \dots \Delta X_T]$ with $\Delta X_{t-1} \equiv \left[\ \Delta P^{'}_{t-1} \dots \Delta P^{'}_{t-K}\ \right]$

The likelihood is defined as

$$f\left(R|X,B,w\right)=\int f\left(R|X,B,\mathbf{H}\right)f\left(H|w\right)dH$$

A difficulty is added to this Bayesian framework by the presence of the volatility vector H playing the role of augmented parameters. As suggested Tsay (2005), by Conditioning on $\mathbf{H}$, I have the distribution functions $f\left(R|X,H,B\right)$. the distribution $f\left(\mathbf{H}|,\omega\right)$. The prior distribution $p\left(B,w\right)=p(B)p\left(w\right)$ is separated such that the prior distributions for the mean and volatility are independent. To estimate the parameters using a Gibbs sampling, it is the straightforward to draw random samples and compute the empirical moments from the following conditional posterior distributions:

$$f\left(\beta|R,X,B,w\right),\quad f\left(\mathbf{H}|R,X,B,w\right),\quad f\left(w|R,X,B,H\right)$$

To perform the simulations. a multivariate GARCH model is first estimated on the residuals of an estimated VECM model. The estimated values are used to initialize the prior distributions of the parameters of interest. Then a Gibbs sampling algorithm is performed.

4.3.2 Estimation procedure

To estimate the VECM-ASV model, I start by estimating a VECM with GARCH errors. The mean equation parameters are used to initialize $\alpha=\mathbf{B}$ and the conditional variances to initialize $\mathbf{H}$. Using the extracted σ_{1t} and σ_{2t}, the common factor is extracted as the first principal component of their PCA and to obtain initial values of ψ_1 and ψ_2. I then do the regression 4.4 and 4.5 by OLS to initialize the parameters of the variance equation.

The following steps are repeated 1000 times, and the firsts 100 values are discarded, to generate empirical distribution of the parameters to estimate:

1. The prior distribution of the α is assumed to be a bivariate normal distribution with mean AL_0 and variance C_0. To obtain the posterior, I estimate $\hat{\alpha}$ from the Weighted Least-Squared regression:

$$\begin{bmatrix}\Delta P_{1t}/\sigma_{1t}\\ \Delta P_{2t}/\sigma_{2t}\end{bmatrix}=\alpha\begin{bmatrix}\beta' Y_{-1t}/\sigma_{1t}\\ \beta' Y_{-1t}/\sigma_{2t}\end{bmatrix}+U,\ U\sim\mathcal{N}\left(0,I\right)$$

Then the posterior distribution of α is a bivariate normal distribution with mean AL^* and variance C^* such that

$$AL^* = C^* \left(C_0^{-1} AL_0 + Var(\hat{\alpha})^{-1}\hat{\alpha}\right)^{-1} \text{ and variance } C^* = \left(Var(\hat{\alpha})^{-1} + C_0^{-1}\right)^{-1}.$$

The realizations of α are drawn from this posterior.

2. The volatility vector $\mathbf{H}$ is drawn element by element from the following conditional posterior distribution

$$\begin{aligned} f(h_t|R,X,H_t,\alpha,\omega) &\propto f(a_t|h_t,\alpha,\Delta P_t,\beta' P_{t-1}) \times f(h_t|\omega) \\ &\propto \phi_{\mathscr{N}(0,H_t)}(a_t) \times \phi_{\mathscr{N}(A_0 log(h_t),|\Sigma_v)}(log(h_{t+1})) \times \phi_{\mathscr{N}(A_0 log(h_{t-1}),|\Sigma_v)}(log(h_t)) \end{aligned}$$

with $\Sigma_v = \begin{pmatrix} \sigma_{v_1}^2 & 0 \\ 0 & \sigma_{v_2}^2 \end{pmatrix}$ and $\phi_{\mathscr{N}(m,V)}(x)$ the density function at x of a $\mathscr{N}(m,V)$.

To draw from this density, I do a Griddy-Gibbs sampling by fixing h_2 and drawing h_1 from the density, and then fixing the h_1, h_2 is drawn from the density. The process is repeated 500 times to obtain h_1 and h_2 that are independent. To draw the h_1 for instance, I compute the values of the previous density on a grid of the interval $[0, 1.5 \times var(a_{1t})]$ to obtain an empirical density of h_1. Then I draw randomly a value in $[0,1]$. and I obtain the draw of h_1 by inverting the cumulative empirical density at this value.

3. A PCA is performed to obtain new values of ψ_1, ψ_2 and the common factor.

4. The prior distribution of $\omega = A_0$ is a multivariate normal distribution with mean ω_0 and variance Cw_0. The posterior distribution is $f(A_0|R,X,H_t,\alpha) = f(A_0|R,H_t,\alpha)$ is a bivariate normal distribution with mean ω^* and variance Cw^* such that

$$\omega^* = Cw^* \left(Cw_0^{-1}\omega_0 + n\Sigma_v^{-1}\overline{h_t}\right)^{-1} \text{ and variance } C^* = \left(n\Sigma_v^{-1} + C_0^{-1}\right)^{-1}.$$

5. Using the drawn ω, $\mathbf{H}$ and V_t , the variance errors is computed as $v_{1t} = log\left(\sigma_{1t}^2\right) - a_{01} + a_{11} log(V_{t-1})$ and $v_{2t} = log\left(\sigma_{2t}^2\right) - a_{02} + a_{12} log(V_{t-1})$.

6. The conjugate prior distribution for each $\left(\sigma_{v_i}^2\right)_{i=1,2}$ is $(m\lambda)/\sigma_{v_i}^2 \sim \chi_m^2$. The posterior distribution is the inverted Chi-squared distribution with $m+n-1$ degrees of freedom $(m\lambda + \sum_{l=2}^n v_{il})/\sigma_{v_i}^2 \sim \chi_{m+n-1}^2$. The new variances $\left(\sigma_{v_i}^2\right)_{i=1,2}$ are drawn from this distribution.

Finally the value of each parameter is estimated as the mean of the empirical distribution. Thoses values are used in the formulas to obtain the different Volatility discovery measures.

Table 4.1: Simulation results

		Volatility Share (VS)		Permanent Volatility (PV)	
a_{11}	a_{12}	Mk1	Mk2	Mk1	Mk 2
0.98	0.50	59.66	40.34	63.52	36.48
		(0.34)	(0.34)	(0.01)	(0.01)

The Table reports the price and volatility discovery measure estimated in a bivariate VECM-ASV model.

4.4 Simulation

To analyse the pefomance of the estimation strategy in computing volatility discovery in the VECM-ASV framework, a simulation exercise is performed in the setting is the following

VECM-ASV(1):

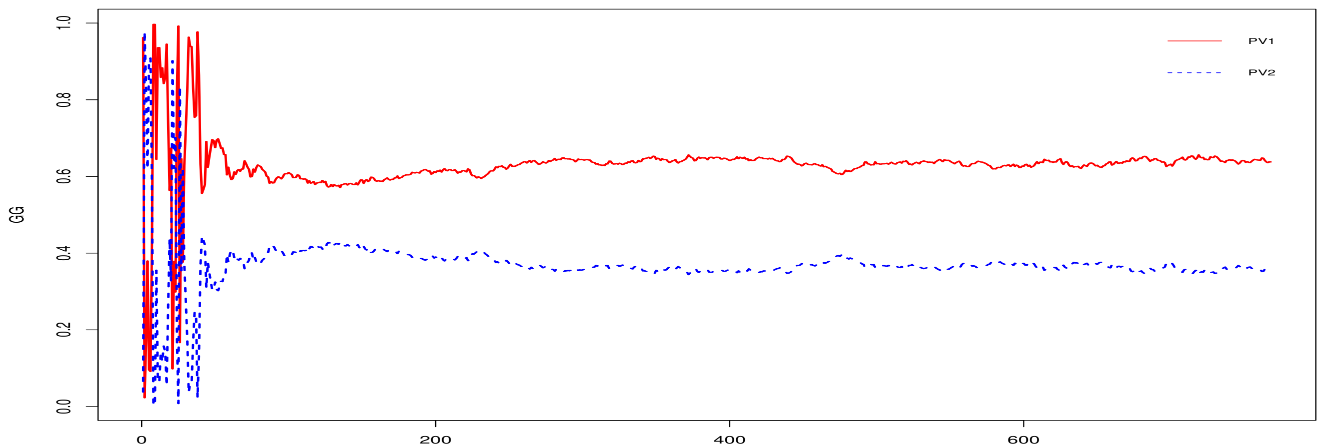
$$
\begin{aligned}
\Delta P_t &= -\alpha\beta' P_{t-1} + \varepsilon_t \\
\varepsilon_t &= H_t^{1/2}\eta_t \\
H_t &= \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\begin{pmatrix} \sigma_{1t} & 0 \\ 0 & \sigma_{2t} \end{pmatrix} \\
log\left(\sigma_{1t}^2\right) &= 0 + a_{11} \times log\left(\sigma_{1t-1}^2\right) + v_{1t} \\
log\left(\sigma_{2t}^2\right) &= 0 + a_{12} \times log\left(\sigma_{1t-1}^2\right) + v_{2t}
\end{aligned}
$$

with $\beta_1 = 1$**,** $\sigma_{v_1}^2 = \sigma_{v_2}^2 = 0.05$, $\begin{pmatrix} \alpha_1 & \alpha_2 \end{pmatrix} = \begin{pmatrix} 0.025 & -0.025 \end{pmatrix}$, $\begin{pmatrix} a_{11} & a_{12} \end{pmatrix} = \begin{pmatrix} 0.98 & 0.50 \end{pmatrix}$,

In this setting, the prices are cointegrated and the coefficient of cointegrated is 1. The common factor in the volatility equation is completly driven by the market 1. So volatility discovery happens in market 1. We simulate the setting 1000 times, in each setting the MCMC is simulated 100 times and average values are taken.

The table presents the simulation results. It shows that the estimation strategy select market 1 as dominating the volatility discovery process.

Figure 4.3: Convergence of MCMC algorithm: Permanent Volatility Share



4.5 Empirical application

4.5.1 Cash and Futures markets volatility

Using the VECM-ASV framework, I investigate the relative contribution to volatility discovery of futures and cash markets on some metals traded on the London Metals Exchange (LME). It is the same commodities that were studied in chapter 1: Tin. Aluminium, Lead, Copper, Molybdenum and Steel billet. Here I focus on the comparison between the two main securities that are the cash market and the 3-month futures contract. I study the relative relevance of these contracts in transmitting volatility. The overnight information

The data for each of the metals under study is comprised of the cash official prices. the 3-month futures official prices. the daily volume. the high and low daily official prices. They are extracted from the database Eikon of Thomson Reuters. The data range is from January 2010 to December 2015.

The data on the base metals have been analyzed descriptively and are shown to be cointegrated.

Analysis of volatility discovery The Table 4.2 presents the results of the VECM-ASV models estimated for each metal. For most of the metals. the contribution of the futures market to volatility discovery is bigger than that of the cash market. For Aluminum and Tin, the Cash market dominates the futures in volatility discovery with contribution of 56% and 52.47%. To check if there is link between Volatility discovery and the price discovery, I compute the main price discovery measures IS and PT on those assets. The results I obtain vary by stocks. Overall the results shows that for

Table 4.2: Estimation results

	Volatility Share (VS)		Permanent Volatility (PV)		Information Share (IS)				Permanent Transitory (PT)	
					Upper		Lower			
	Cash	Fut	Cash	Fut	Cash	Fut	Cash	Fut	Cash	Fut
Euro50	04.11	95.89	13.86	86.13	99.20	60.69	39.30	0.79	90.9	9.1
Copper	23.46	76.54	49.10	50.90	65.03	96.28	3.71	34.96	24.2	75.7
Molybdenum	49.00	51.00	49.00	51.00	99.7	95.98	4.01	0.26	79.4	20.5
Aluminum	56.2	43.8	50.40	49.6	89.9	81.04	18.9	10.05	56.9	43.1
Tin	52.47	47.53	50.5	49.5	94.13	90.44	9.55	5.86	55.8	44.1
Lead	17.53	82.46	32.55	67.44	98.30	99.85	0.14	1.69	22.2	77.8
Steel Billet	4.19	95.80	17.23	82.76	64.93	97.7	2.28	35.06	18.1	81.8

The Table reports the price and volatility discovery measure estimated in a bivariate VECM-ASV model.

Aluminum and Tin, price discovery and Volatility discovery happen in the cash market; for copper and Molybdenum, price discovery happens in the cash market and volatility discovery in the futures market; For Lead, price discovery and Volatility discovery happen in the Futures market.

4.5.2 Euro Stoxx 50 index Futures

The Euro Stoxx 50 index is comprised of 50 major European equity prices and futures contracts on this index are intensively traded. The data are the index values and the continuous nearby futures contract series extracted from the modules Times and Sales of the database Eikon of Thomson Reuters. I have tick-by-tick data for January 23th of 2017.

The results of the measures computed on the VECM-ASV models are presented in Table 4.2. The contribution of the volatility from the Futures market is 95.89% for Volatility Share and 86% for Permanent Volatility measure. At the same time on this data the contribution to price formation is dominated by cash market with Information Share of the Cash market in the interval 66% to 99%. and a PT measure of 90.9%. The literature discusses a lot about where between Index and its Futures, informed trading happens. The results show that for Euro50Stoxx informed trading happens on Cash market, and suggest that Volatility trading happens in Futures market.

4.6 Volatility Discovery and Realized Variance

The first part of this chapter presented a VECM-ASV framework that is suitable for Low frequency data. When High frequency data are available, an analysis of volatility discovery can be somewhat simplified by using non parametric estimation of volatility. In this section, I study volatility discovery on intraday data for assets listed and traded on multiples markets. I consider the Dow Jones stocks; they are traded simultaneously on NYSE and NASDAQ. The data come from the NYSE TAQ Database and cover the period from March 2011 to May 2011.

4.6.1 Realized Variance Cointegration

The observed prices of cross-listed assets on different markets are cointegrated as highlighted in the literature on price discovery measure (Hasbrouck 1995). In addition, as claimed by the framework developed in this study, their volatilities should also share a common persitent component cointegrated. I showed it using conditional variances extracted from GARCH models of daily data for futures contracts. It appears also to be true for High frequency data and even when using other volatility measures. Figures 4.7. 4.4 and 4.8 plot the daily realized volatility of the Dow Jones stocks on NYSE and NASDAQ. For all the stocks it appears clearly that indeed their volatilities are driven by a common component.

Realized volatility is another way of capturing the total volatility in a day. I use them as input to compute Volatility discovery for each stock. That is for each stock, I build the time series of daily Realized variances. And I estimate the Volatility Share by simply applying Hasbrouck (1995) and Harris et al. (2002a) methodology to a VAR of Realized volatility:

$$\begin{pmatrix} \Delta RV_{1t} \\ \Delta RV_{2t} \end{pmatrix} = -\alpha\beta' \begin{pmatrix} RV_{1t-1} \\ RV_{2t-1} \end{pmatrix} + \Gamma_1 \begin{pmatrix} \Delta RV_{1t-1} \\ \Delta RV_{2t-1} \end{pmatrix} + \ldots + \Gamma_K \begin{pmatrix} \Delta RV_{1t-K} \\ \Delta RV_{2t-K} \end{pmatrix} + \varepsilon_t \tag{4.15}$$

The daily RV is computed with intraday prices at the sampling interval of 2 min to limit microstructure noises effects.

The results for each stock are in the Table 4.3. The Table presents the Volatility Share estimated with the VECM-ASV model on the different assets.

I observe that using realized volatility, for most of the stock the NYSE dominates the Volatility discovery process.

Figure 4.4: Realized volatility of Nyse and Nasdaq

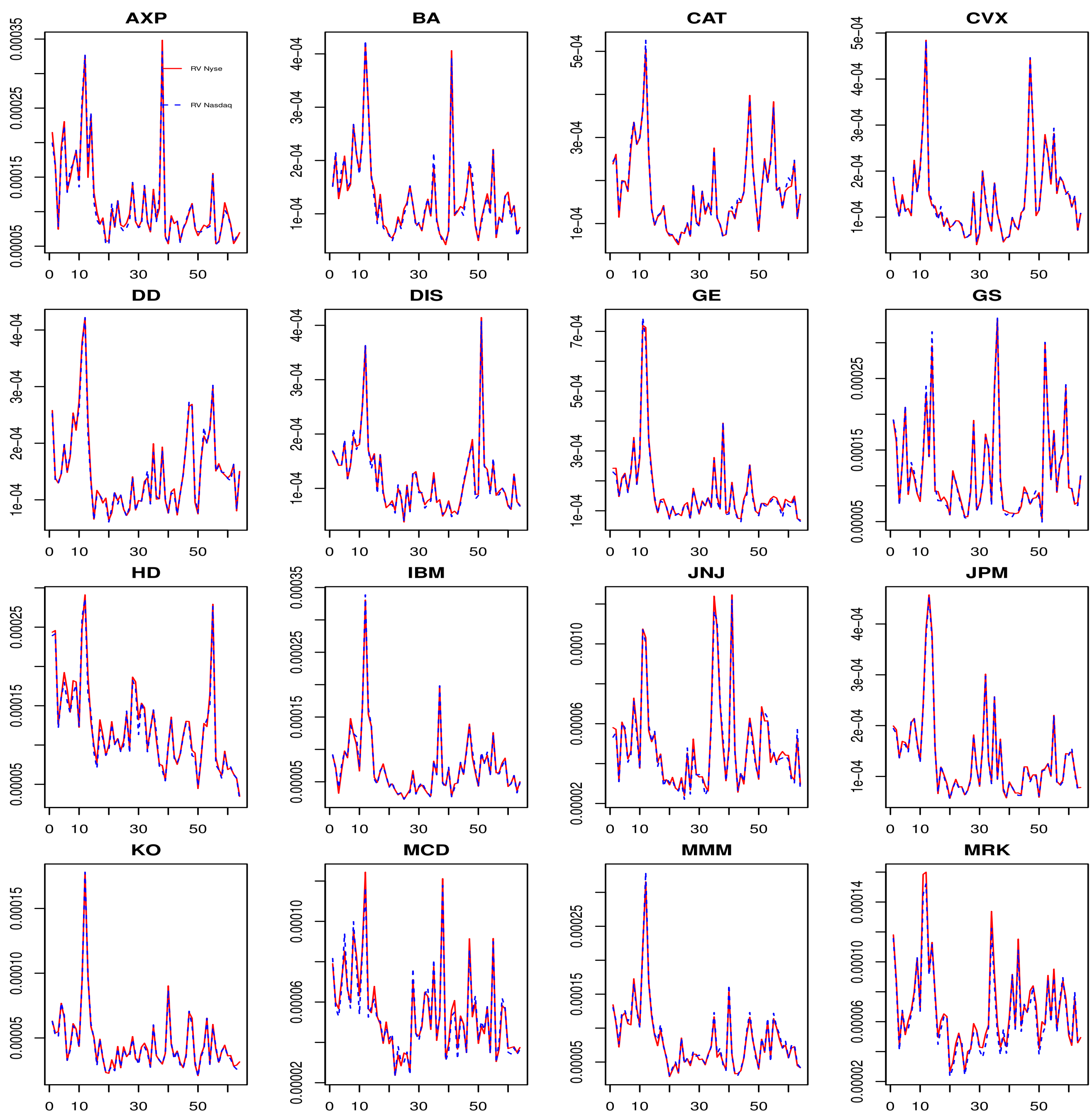


Table 4.3: Volatility Share and Permanent Volatility measures for NYSE on Realized Volatility

Stocks	IRIS	VSu	VSl	PV	VSdiag
AXP	49.866	99.340	0.127	28.857	16.116
BA	50.008	99.955	0.077	57.775	63.188
CAT	50.320	99.489	1.776	62.131	77.656
CVX	49.993	99.880	0.093	49.607	43.703
DD	50.018	99.791	0.279	55.519	57.232
DIS	50.002	99.893	0.117	51.862	52.191
GE	49.871	98.818	0.672	38.371	36.252
GS	50.017	99.855	0.211	52.466	59.291
HD	50.044	99.802	0.374	55.796	65.328
IBM	49.957	99.770	0.058	35.612	20.237
JNJ	49.986	99.755	0.191	47.889	43.712
JPM	50.058	99.957	0.273	73.819	86.429
KO	50.245	100.0	0.977	99.790	100.000
MCD	51.182	99.100	5.477	71.129	85.888
MMM	50.008	99.849	0.183	54.984	54.930
MRK	49.802	98.477	0.741	46.870	32.732
NKE	50.005	96.194	3.825	44.884	50.125
PFE	49.998	99.930	0.062	48.509	46.642
PG	49.994	99.767	0.207	45.790	47.093
TRV	52.237	99.727	8.816	85.349	96.994
UNH	49.438	96.646	1.157	31.511	25.653
UTX	50.229	99.976	0.935	86.317	97.446
VZ	50.058	99.609	0.623	63.429	61.470
WMT	49.990	99.864	0.094	46.576	40.923
XOM	51.989	99.521	8.093	78.466	94.417
AAPL	49.996	99.890	0.095	47.194	0.462
CSCO	50.119	99.286	1.187	58.957	0.624
INTC	49.442	88.345	9.661	43.493	0.453
MSFT	50.011	98.228	1.817	49.403	0.506

Note: The Volatility Share is the "Information Share" measure and the Permanent Volatility is The Permanent Transitory measure estimated on a VAR of Realized Volatility. VSdiag is the VS computed with zero correlation. IRIS is defined in the first chapter.

4.6.2 Determinants of volatility discovery

Table 4.4: Correlation between Volatility Discovery measures. liquidity and Volume market Share by trade size. and volatility of volume for NYSE

	Small size	Medium size	Big size	Liquidity	Volatility of Volume
IRIS	-0.059	0.192	0.273	0.097	**0.214**
PV	-0.011	0.191	0.168	0.099	**0.252**
VS	-0.06	0.47	**0.56**	0.27	**0.67**
VS lower	0.013	0.133	**0.273**	0.124	0.116
VS upper	-0.092	0.013	-0.124	-0.088	0.067

Correlation of the different Volatility Discovery measures (computed using Realized Volatility) with liquidity and Trade size variables

It appears clearly that for the variables tested. the Volatility of the volume is the best determinant of volatility discovery. It is followed by the market share in big size trade. This result is coherent with some studies in the literature that suggest strong linkages between the volatility of volume and price volatility, eventough they are not done in the context of cross listed assets (Epps and Epps, 1976; Tauchen and Pitts, 1983). In absolute terms, those coefficients remains meanwhile low, implying the existence of others important sources of volatility discovery. In the literature for example, Wang (2014). Ranaldo (2004) and Ni et al. (2008) show strong linkages between the aggressiveness of the limit order book and the future volatility.

4.6.3 Comparison of VECM-ASV and Realized Variance approaches

Here we compare the results on volatility discovery measures when using the two approaches previously presented. An intensive comparison is very difficult due to the timing and complexity of MCMC for large datasets. For this, we use the data of AAPL and Euro50 as they are highly liquid but we will rely only on one week of data. For the VECM-ASV approach, the series of 2 min Returns are used to produce for a given day, the contribution to volatility discovery. To compute the contribution for the same day using the Realized variance cointegration, Realized variances are computed for each 2 min. The RV obtained are then processed in the VECM of RV and volatility discovery measure are computed. For the asset under investigation, the results from the two approaches are similar. For Euro50 the futures markets dominates the volatility discovery with a Volatility Share of 95% and PV of 86%, when computed in the VECM-ASV setup. The contribution to volatility discovery computed in the second framework is also high for Futures market. The Cash market contributes to only 26.48% of volatility.

Table 4.5: Estimation results

	VECM-ASV				VECM of RV			
	Volatility Share (VS)		Permanent Volatility (PV)		Volatility Share (VS)			Permanent Volatility (PV)
	VS_1	VS_2	PV_1	PV_2	VS_{1u}	VS_{1l}	VS_1	PV_1
APPL	52.14	47.86	50.46	49.54	93.90	20.53	57.51	61.29
Euro50	04.11	95.89	13.86	86.13	25.38	27.59	26.48	40.56

4.7 About Seasonality

The literature almost agreed on the empirical fact that intraday variance have a U-shaped pattern. That is the volatility is high at the begining of the day and at the end of the day.

Andersen and Bollerslev (1997)showed how this intraday seasonality can be captured, and presented a Flexible Fourier Form (FFF) approach to obtain intraday periodic components of the volatility. Nuria et al. (2017) uses this FFF to study the impact of using of this seasonality on volatility transmission studies. They compare the behavior of VAR model of daily variances and Impulse response function when seasonality is ignored versus when seasonality is removed. Their main result is that the persistence in volatility and in Impulse Response Functions is for a part due to the seasonal component. Such that when the retruns are deseasonalized before analyses, the lag of the VAR is significantly reduced.

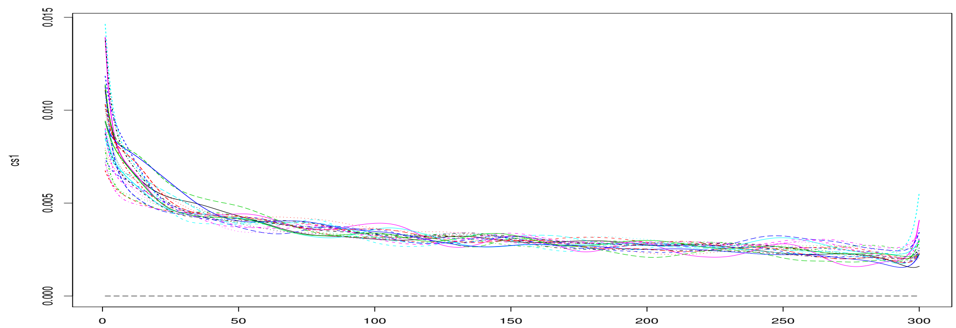
Here, I want to check if taking into account this seasonality can impact the results of contribution to volatility discovery. For this, I compute the periodic component for each asset using the FFF presented in Andersen and Bollerslev (1997).

The study of the intraday variance shows that it is not strictly U-shaped, but still we observed that it is high at the begining of the day and stable during the rest of the day. The figure shows the seasonal component for the 25 assets of the NYSE, we see that they have the same type of periodicity.

The study of the intraday variance shows that it is not strictly U-shaped, but still we observed that it is high at the begining of the day and stable during the rest of the day. The figure shows the seasonal component for the 25 assets of the NYSE, we see that they have the same type of periodicity.

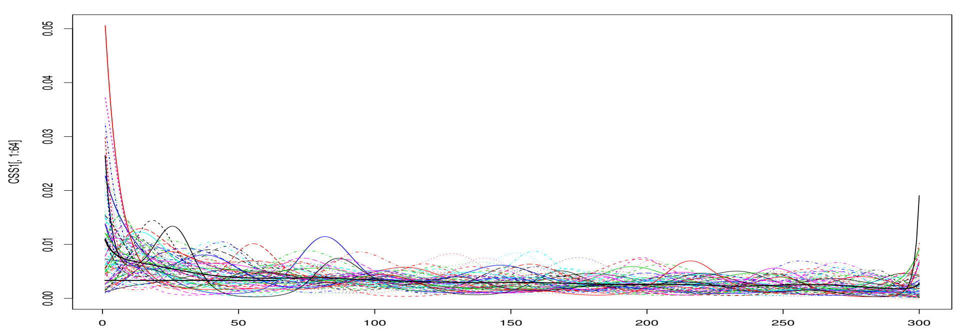
The results (Table 4.5) show globally that NYSE remains the dominant market in volatility discovery, for most of the assets there is a coherence between initial results and the deseasonalized

Figure 4.5: Intraday volatility Seasonal components



Each curve (color) is for one asset, it is the final volatility intraday seasonal (periodic) component .

Figure 4.6: Seasonal components of AXP for 64 days



Each curve (color) is for one day, it is the final volatility intraday seasonal (periodic) component .

study. Meanwhile, there are some assets for which NYSE is dominant with raw returns, while NASDAQ becomes dominant when seasonality is removed. Many things can be the cause of these shifts: First, if the intraday periodic component is not well represented in the data deseasonlising can change the structure of the data. For example the figure 4.6 shows each day estimated seasonal components for American Express (AXP). It can be seen that it is not very stable across the days. A second reason might be that if the seasonal component is not well identified and estimated then removing it might delete more than the periodicity in the observations.

4.8 Conclusion

The derivatives instruments of securities are generally considered as improving the price discovery mechanism. For example, futures markets are known to facilitate the price discovery and the liquidity of a security. Meanwhile, little is known on how the uncertainty on information is transmitted to the asset. The literature is rich in investigating on which market informed trading is happening and this paper innovates in proposing a framework to study where volatility on information enters the market. I build a VECM with stochastic volatility that specifies explicitly the common persistent behavior of volatilities among the two markets. I show that indeed the data supports the assumption of common permanent factor between the conditional volatilities. This allows to define measures of markets' contribution to volatility discovery. The estimation of the model relies on an intensive MCMC simulations and Gibbs sampling.

In a first application on daily data, I estimate the VECM-ASV model to compute the contribution of 3-month futures and the cash market to volatility discovery of some metals traded on the London Metal exchange. The data range is from January 2010 to December 2015. The results show that price discovery and volatility discovery do not necessarily have the same determinants. In fact, for most of the metals, the contribution of the futures market to volatility discovery is bigger than that of the cash market. Only for Molybdenum, the Cash market dominates the futures in volatility discovery. To check if there is link between Volatility discovery and the price discovery. I compute the main price discovery measures IS and PT on those assets. The results I obtain vary by stocks. For Aluminum. Tin and Molybdenum, the IS and PT of the Cash market is the highest. Overall the results shows that for Aluminum and Tin, price discovery and Volatility discovery happen in the cash market; for copper and Molybdenum, price discovery happens in the cash market and volatility discovery in the futures market; For Lead, price discovery and Volatility discovery happen in the Futures market. I then consider the intraday EuroStoxx 50 Index and futures contracts on this index. The Futures market contributes for more than 90% to volatility discovery while price discovery happens for around 90% on the Cash market.

There are many ways to evaluate the volatility of financial assets prices. When High frequency data are available. the Realized volatility offers a good summary of the volatility in a period. I show that Realized volatility are driven by a common component and how Contribution of a market to volatility discovery can be measured. By applying Hasbrouck (1995) and Harris et al. (2002c) methodologies on the time series of realized volatility, I compared NYSE and NASDAQ in the formation of permanent volatility of Dow Jones stocks. And NYSE appears to be dominant market in this respect. An analysis of the correlation between the Volatility discovery measure and some trading activity variables shows that the domination in volatility of volume and in Market share of big volume trade are determinants of prices volatility discovery dominance. But the low figures obtained also suggests the existence of others important determinants.

Overall the results shows that dominating the price discovery does not necessarily implies dominating the volatility discovery. This opens rooms for further research in order to understand the determinant of volatility discovery. There are certainly more to say about the volatility transmission process by a cross section analysis with high frequency data. A deep and complete analysis will require the availability of a rich datasets of trade and quotes activities that can provide enough variables to investigate the determinants of volatility discovery.

Table 4.6: Estimation results

		a_{01}	a_{11}	a_{02}	a_{12}	λ	σ_v	PV_1	PV_2	VS_1	VS_2
Euro50	Mean	-0.021	0.002	-14.016	0.124	1.000	0.005	0.411	0.9589	0.1386	0.86
	SD	0.479	0.031	0.504	0.035	0.000	0.001	0.018	0.018	0.433	0.433
Copper	Mean	-8.785	0.318	-8.172	0.372	0.981	6.490	0.491	0.509	0.460	0.540
	SD	0.520	0.031	0.411	0.027	0.004	1.565	0.003	0.003	0.035	0.035
Molybdenum	Mean	-7.950	0.450	-7.461	0.491	0.981	8.564	0.490	0.510	0.490	0.510
	SD	0.424	0.024	0.391	0.030	0.005	2.576	0.003	0.003	0.018	0.018
Alumni	Mean	-11.312	0.106	-11.561	0.093	0.979	6.941	0.504	0.496	0.562	0.438
	SD	0.440	0.028	0.408	0.027	0.004	1.402	0.003	0.003	0.169	0.169
Tin	Mean	-8.913	0.203	-9.062	0.190	0.973	6.974	0.505	0.495	0.525	0.475
	SD	0.441	0.029	0.464	0.031	0.003	1.122	0.000	0.000	0.074	0.074

Figure 4.7: Realized volatility of Nyse and Nasdaq

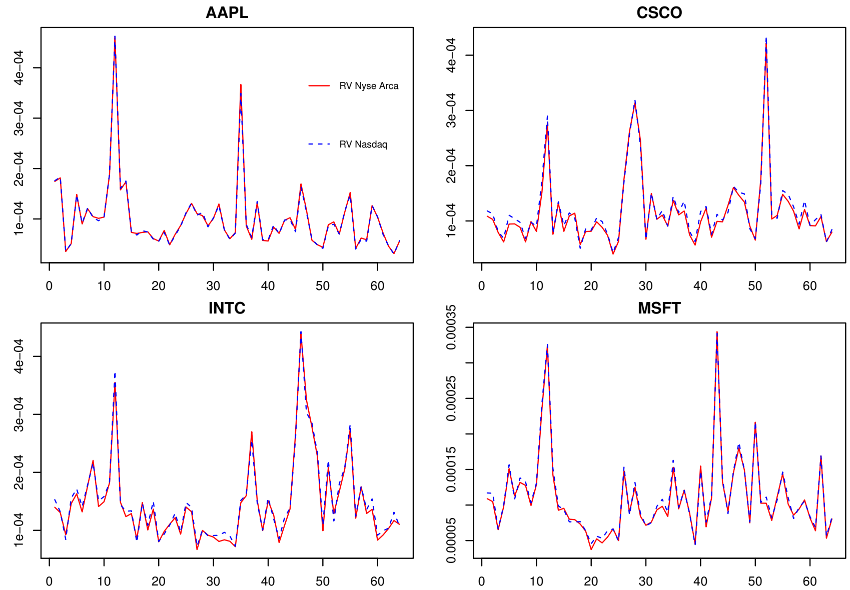
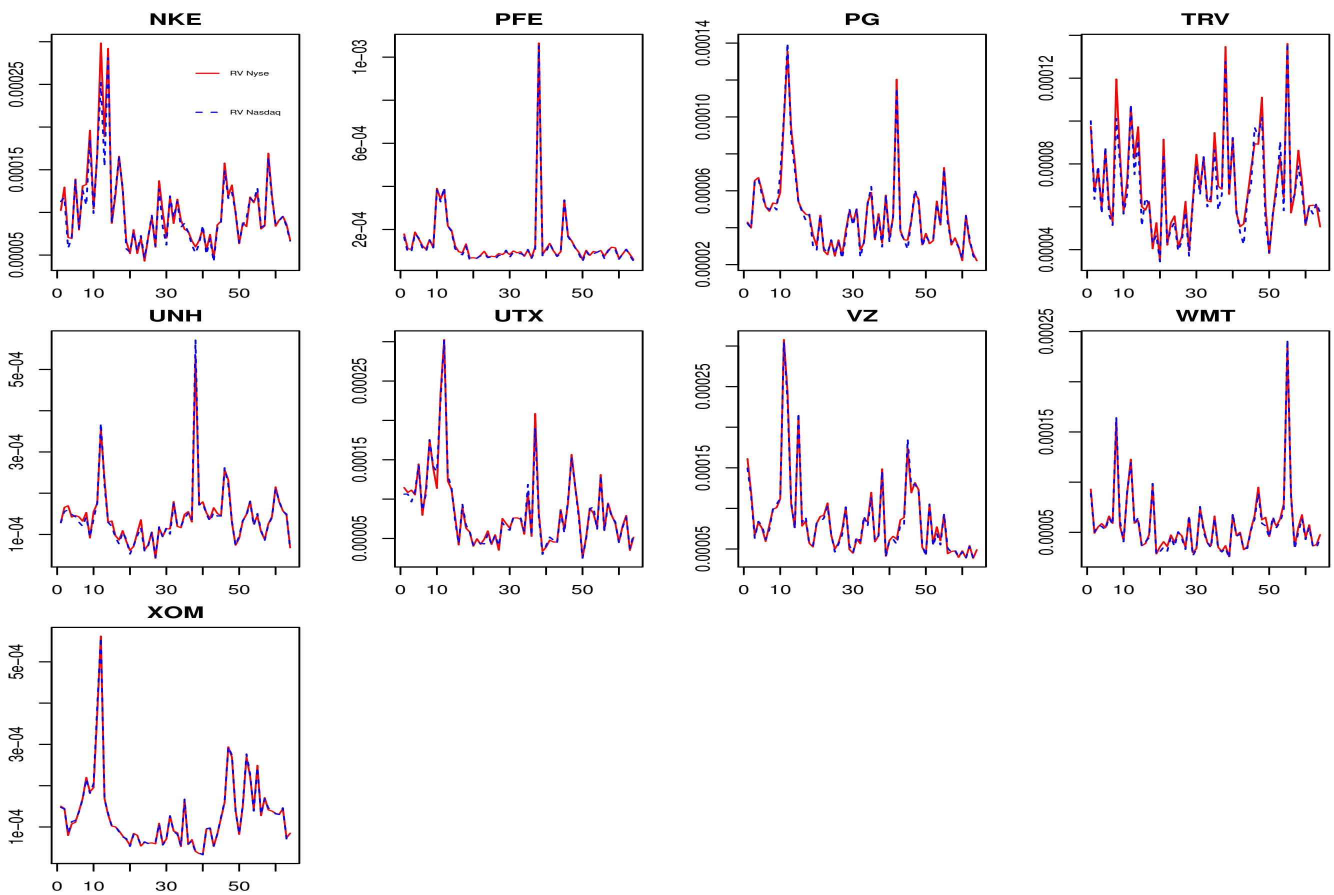


Figure 4.8: Realized volatility of Nyse and Nasdaq



4.9 APPENDIX D

4.9.1 MCMC results

4.9.2 Realized volatility Results

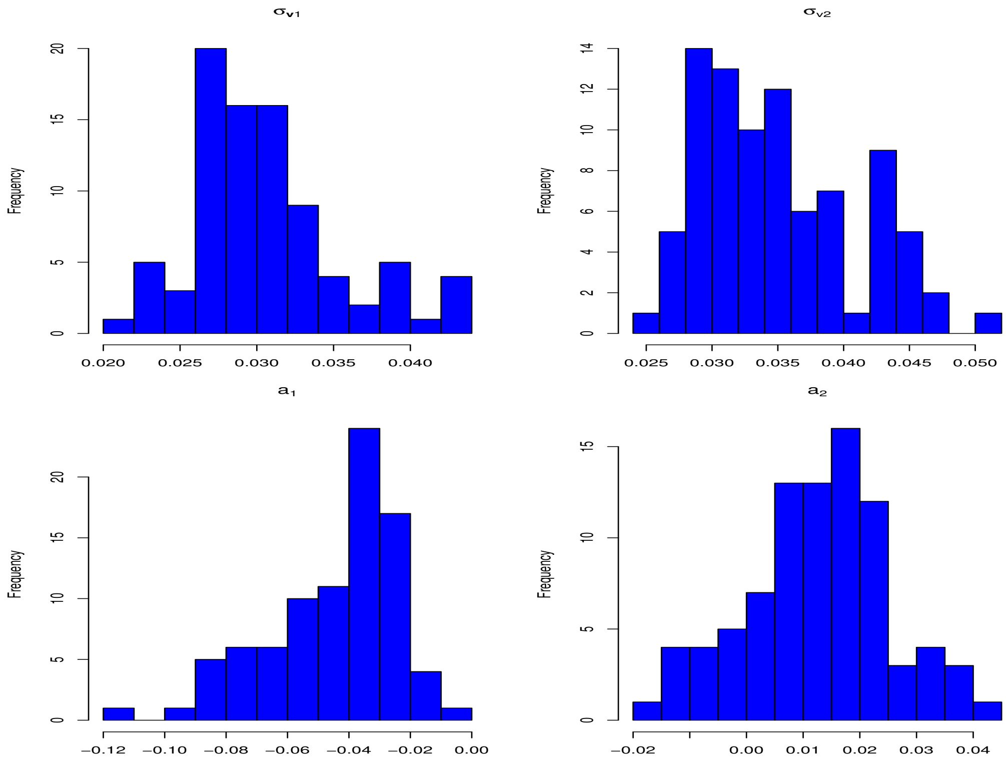
Table 4.7: Deseasonalised , Volatility Share and Permanent Volatility measures for NYSE

Stocks	IRIS	VSu	VSl	PV	VSdiag
AXP	50.009	99.646	0.392	52.239	52.521
BA	49.876	98.685	0.826	43.282	38.565
CAT	50.054	99.757	0.458	55.094	65.331
CVX	49.994	99.931	0.047	46.339	40.504
DD	50.007	99.875	0.152	52.147	54.813
DIS	50.038	99.456	0.696	50.586	56.142
GE	50.412	98.722	2.891	48.940	69.347
GS	50.016	99.962	0.103	60.893	73.139
HD	50.051	99.021	1.179	52.051	54.642
IBM	50.077	99.848	0.459	62.848	75.173
JNJ	49.781	98.669	0.462	36.640	25.750
JPM	49.968	99.806	0.067	38.122	25.663
KO	49.844	99.100	0.282	40.488	23.823
MCD	50.102	99.345	1.059	56.248	61.785
MMM	50.179	99.964	0.748	82.811	95.359
MRK	51.788	98.933	7.898	66.424	88.096
NKE	50.588	92.751	9.404	52.075	56.470
PFE	50.022	99.938	0.149	57.151	70.821
PG	49.785	98.691	0.457	34.390	25.901
TRV	49.904	99.209	0.411	40.342	34.172
UNH	50.003	99.887	0.126	47.943	52.709
UTX	49.927	99.504	0.205	41.415	29.208
VZ	51.364	99.908	5.400	88.404	98.318
WMT	49.950	99.690	0.111	36.454	26.413
XOM	49.933	99.339	0.393	43.359	37.322
AAPL	50.132	99.982	0.546	84.945	0.968
CSCO	50.364	99.974	1.472	88.537	0.982
INTC	49.988	99.731	0.222	46.905	0.452
MSFT	49.951	98.126	1.683	47.605	0.473

*Note: The results ard obtained using deseasonnalized return as in*Andersen et al. (2000). *Volatility Share is the "Information Share" measure and the Permanent Volatility is The Permanent Transitory measure estimated on a VAR of Realized Volatility. VSdiag is the VS computed with zero correlation. IRIS is defined in the first chapter.*

Figure 4.9: Histograms of the parameters (a)

Euro Stoxx 50



Aluminum

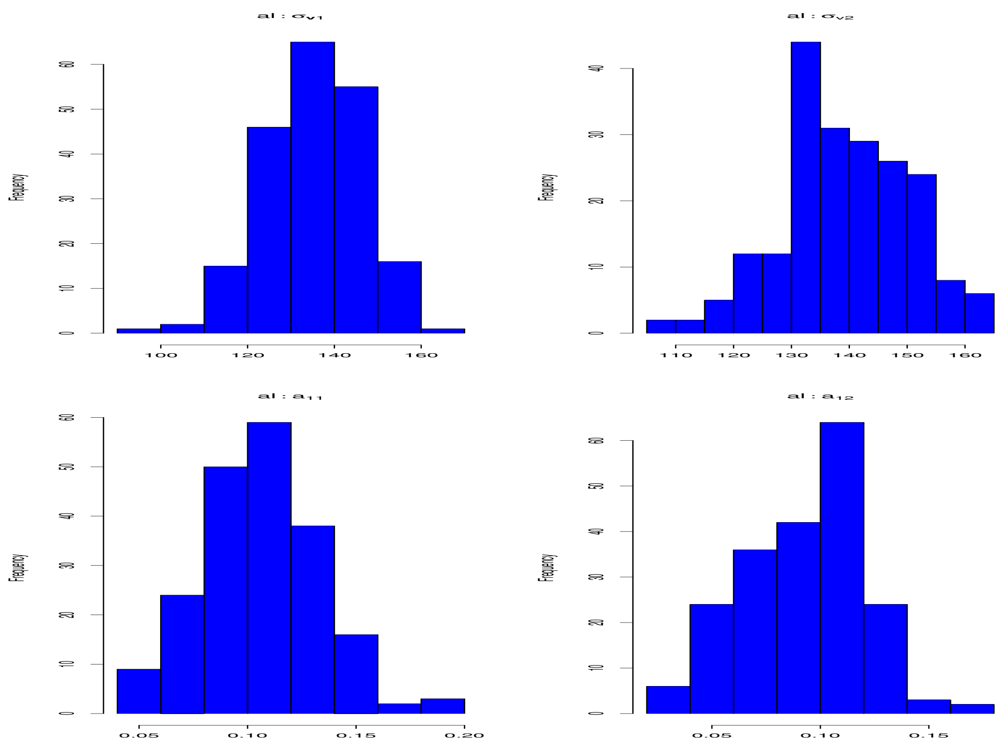
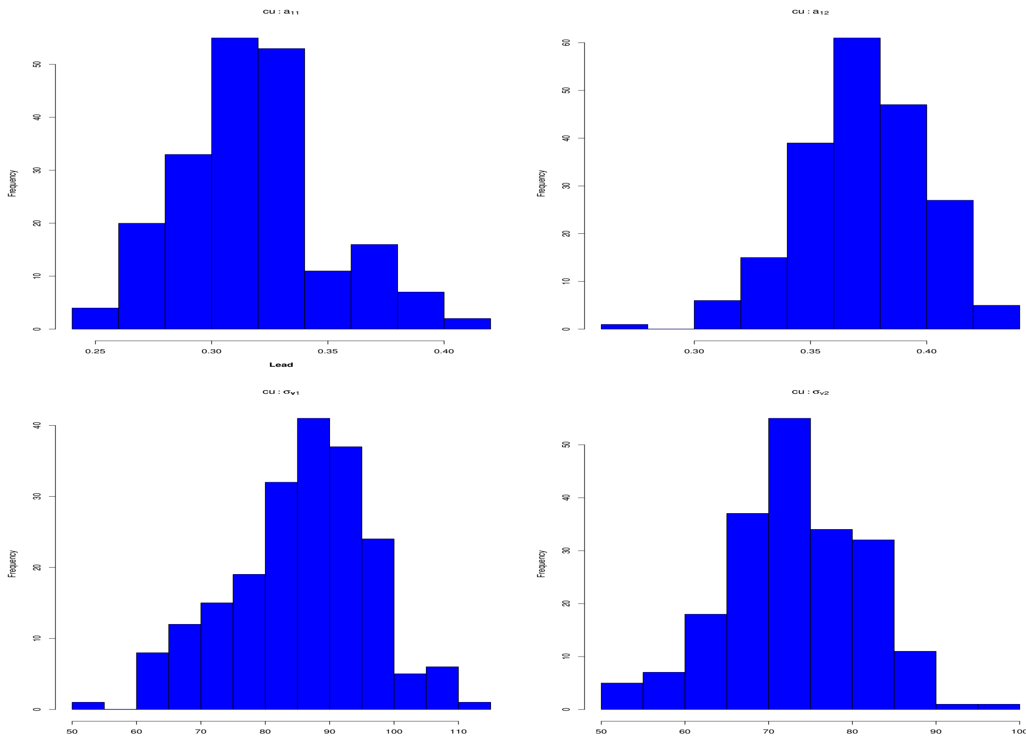


Figure 4.10: Histograms of the parameters (b)

Copper



Molybdenum

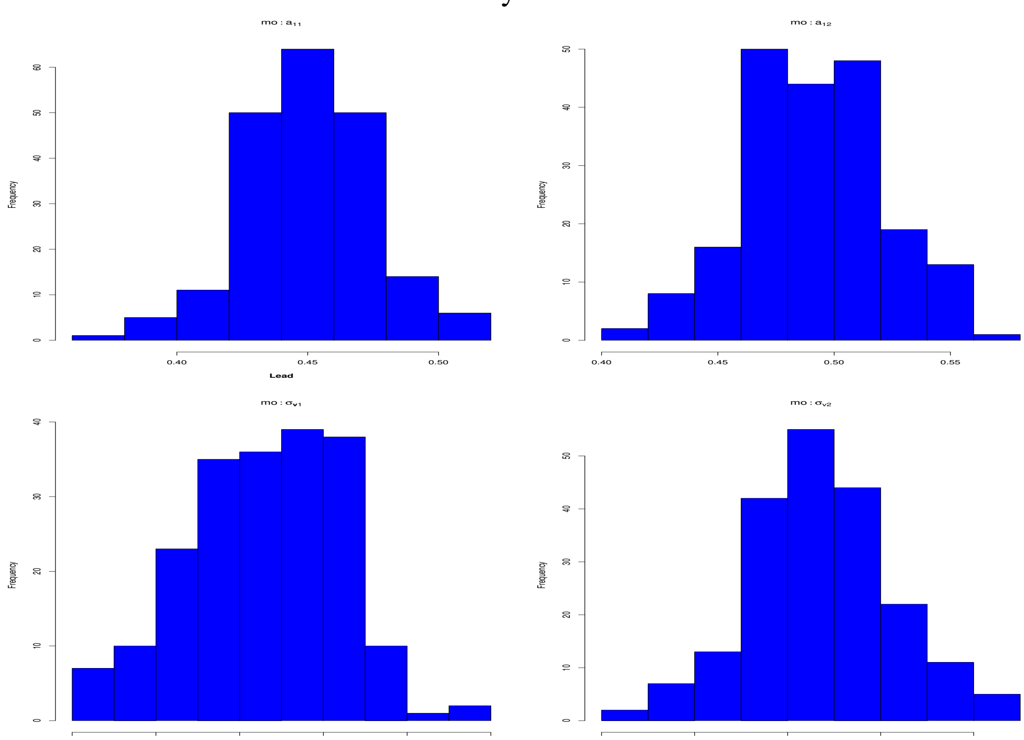
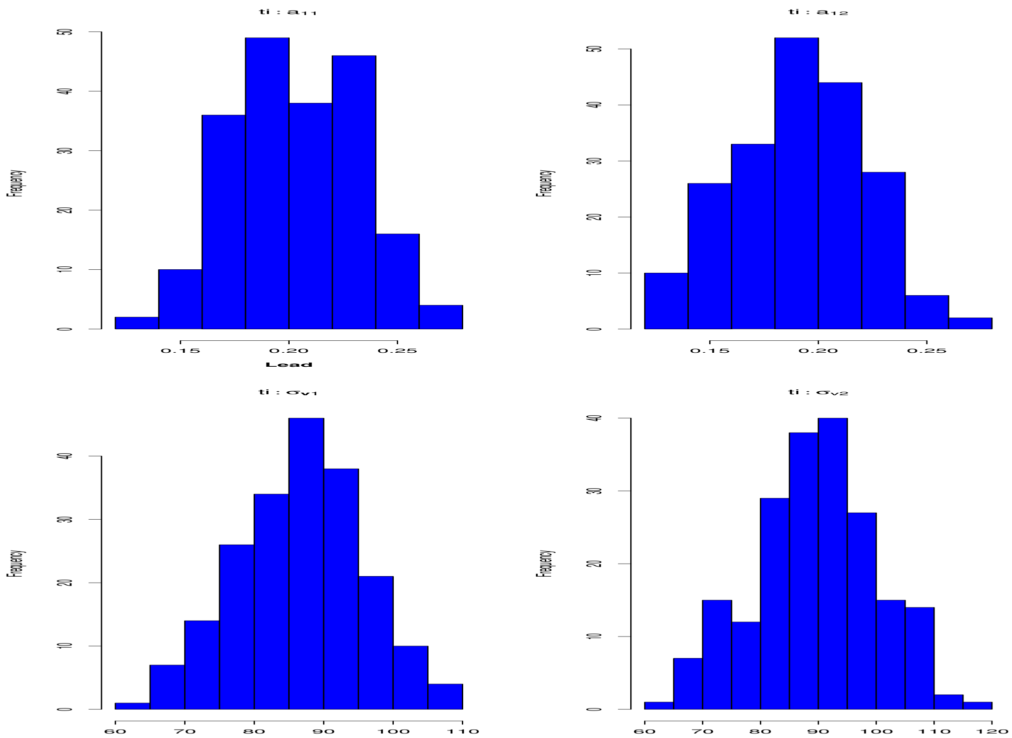


Figure 4.11: Histograms of the parameters (b)

Tin



Lead

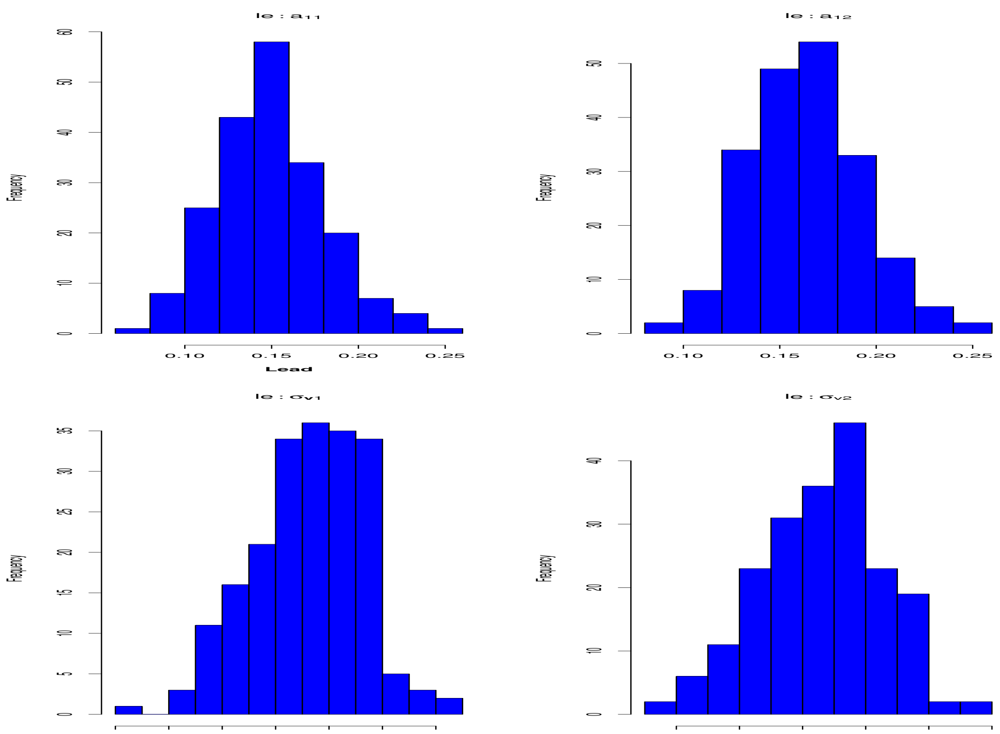


Figure 4.12: Euro Stoxx Index and June 2017 Futures prices

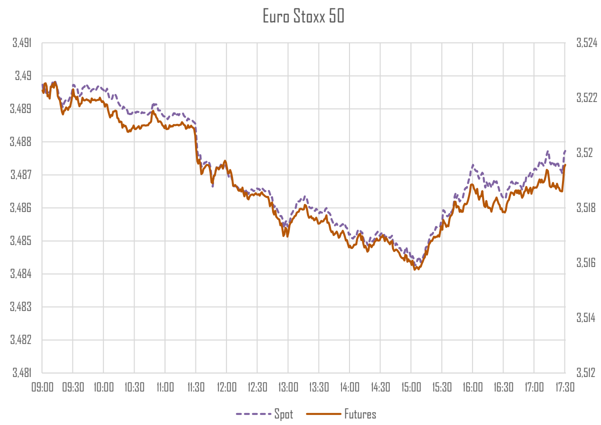


Figure 4.13: ACF of Conditionnal variances from the VECM ASV

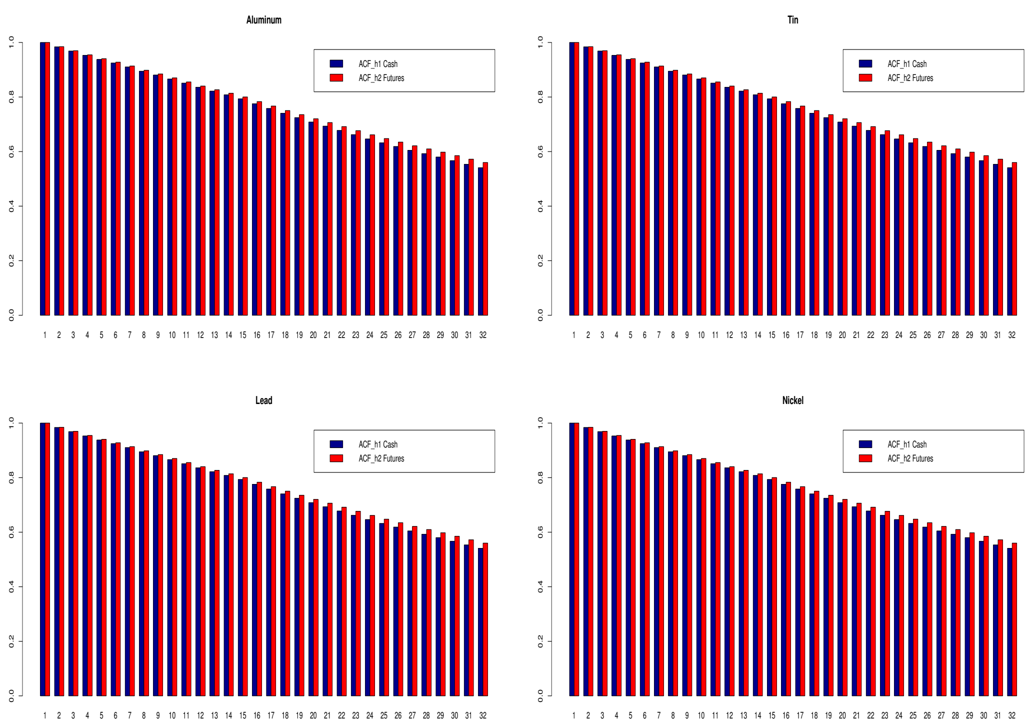
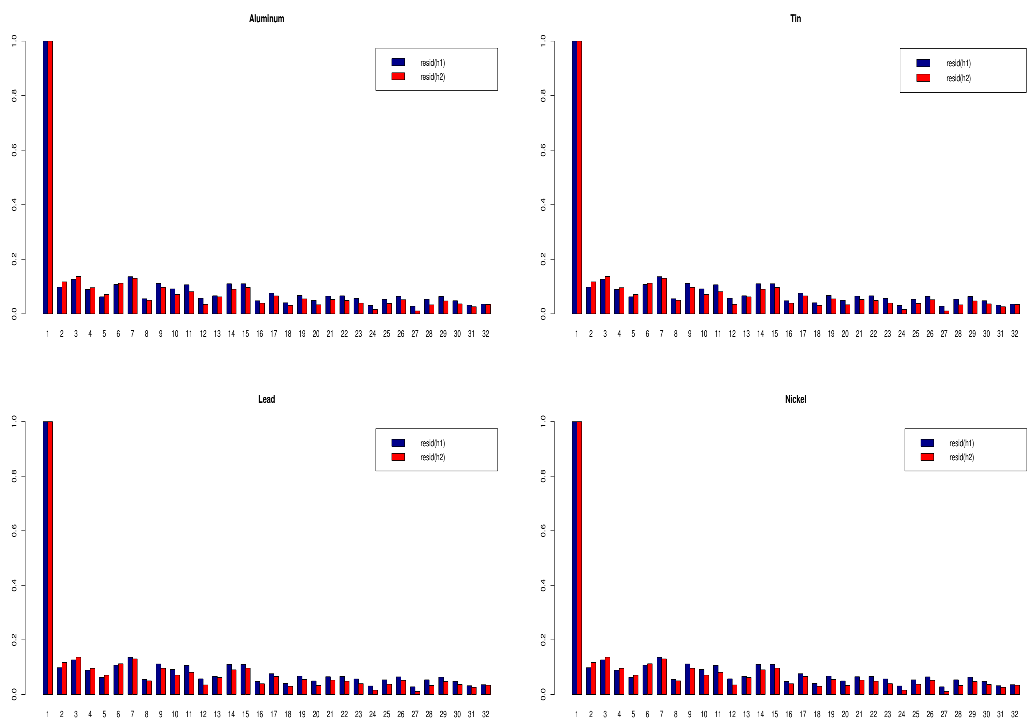


Figure 4.14: Residuals autocorrelation of the VECM-ASV



List of Figures

List of Tables

Bibliography

Andersen, T. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*(2-3), 115–158. 129

Andersen, T., T. Bollerslev, F. Diebold, and P. Labys (2000). Great realizations. *RISK 13*, 105–108. 3, 4, 21, 92, 135

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *The American Economic Review 93*(1), 38–62. 27

Antoniou, A. and P. Holmes (1995). Futures trading, information and spot price volatility: evidence for the ftse-100 stock index futures contract using garch. *Journal of Banking & Finance 19*(1), 117 – 129. 113

Baillie, R. T., G. G. Booth, Y. Tse, and T. Zabotina (2002). Price discovery and common factor models. *Journal of Financial Markets 5*(3), 309 – 321. 2, 9, 59, 64, 67, 92

Barclay, M. J., T. Hendershott, and D. T. M. N (2003). Competition among trading venues: Information and trading on electronic communications networks. *Journal of Finance 58*, 2637–2665. 4, 92

Baule, R., B. Frijns, and M. E. Tieves (2017). Volatility discovery and volatility quoting on markets for options and warrants. *Available at SSRN: https://ssrn.com/abstract=3045260*. 113

Bhar, R. (2001). Return and volatility dynamics in the spot and futures markets in australia: An intervention analysis in a bivariate egarch-x framework. *Journal of Futures Markets 21*(9), 833–850. 113

Blume, M. E. and M. A. Goldstein (1997). Quotes, order flow, and price discovery. *The Journal of Finance 52*(1), 221–244. 92

Cavaliere, G., A. Rahbek, and A. R. Taylor (2010). Testing for co-integration in vector autoregressions with non-stationary volatility. *Journal of Econometrics 158*(1), 7 – 24. 116

Chakravarty, S., H. Gulen, and S. Mayhew (2004). Informed trading in stock and option markets. *The Journal of Finance 59*(3), 1235–1257. 4, 7, 92, 112

Chambers, M. J. (1999). Discrete time representation of stationary and non-stationary continuous time systems. *Journal of Economic Dynamics and Control 23*(4), 619 – 639. 100

Chambers, M. J. (2011). Cointegration and sampling frequency. *The Econometrics Journal 14*(2), 156–185. 100

Chan, K., K. C. Chan, and G. A. Karolyi (1991). Intraday volatility in the stock index and stock index futures markets. *The Review of Financial Studies 4*(4), 657–684. 112, 113

Christensen, K., S. Kinnebrock, and M. Podolskij (2010). Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data. *Journal of econometrics 159*(1), 116–133. 95, 100

Cochrane, J. H. (2012). Continuous-time linear models. *Foundations and Trends in Finance 6*(3), 165–219. 95

Comte, F. (1999). Discrete and continuous time cointegration. *Journal of Econometrics 88*(2), 207 – 226. 100

De Jong, F. (2002). Measures of contributions to price discovery: a comparison. *Journal of Financial Markets 5*(3), 323 – 327. 2, 8, 59, 92

Dias, G. F., C. M. Scherrer, and F. Papailias (2016). Volatility discovery. *CREATES Research Papers*. 113, 114

Epps, T. W. and M. L. Epps (1976). The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distributions hypothesis. *Econometrica 44*(2), 305–321. 114, 128

Eun, C. S. and S. Sabherwal (2003). Cross-border listings and price discovery: Evidence from u.s.-listed canadian stocks. *The Journal of Finance 58*(2), 549–575. 2, 4, 9, 64, 92, 112

Figuerola-Ferretti, I. and J. Gonzalo (2010). Modelling and measuring price discovery in commodity markets. *Journal of Econometrics 158*(1), 95 – 107. iii, v, 60, 72, 73, 77

Fricke, C. and L. Menkhoff (2011). Does the âĂIJbundâĂİ dominate price discovery in euro bond futures? examining information shares. *Journal of Banking & Finance 35*(5), 1057 – 1072. 112

Frijns, B., I. Indriawan, and A. Tourani-Rad (2015). Macroeconomic news announcements and price discovery: Evidence from canadian-u.s. cross-listed firms. *Journal of Empirical Finance 32*, 35–48. 27

Garbade, K. D. and W. L. Silber (1983). Price movements and price discovery in futures and cash markets. *The Review of Economics and Statistics 65*(2), 289–297. iii, v, 2, 58, 60, 72, 73

Gonzalo, J. and C. Granger (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics 13*(1), 27–35. 2, 8, 9, 59, 63, 64

Grammig, J. and F. J. Peter (2013). Telltale tails: A new approach to estimating unique market information shares. *Journal of Financial and Quantitative Analysis 48*, 459–488. 7, 59, 63

Harris, F. H., T. H. McInish, and R. A. Wood (2002a). Common factor components versus information shares: a reply. *Journal of Financial Markets 5*(3), 341 – 348. 8, 9, 64, 70, 125

Harris, F. H., T. H. McInish, and R. A. Wood (2002b). Security price adjustment across exchanges: an investigation of common factor components for dow stocks. *Journal of Financial Markets 5*(3), 277 – 308. 2, 8, 30, 59, 92, 112

Harris, F. H., T. H. McInish, and R. A. Wood (2002c). Security price adjustment across exchanges: an investigation of common factor components for dow stocks. *Journal of Financial Markets 5*(3), 277 – 308. 114, 131

Hasbrouck, J. (1995). One security, many markets: Determining the contributions to price discovery. *The Journal of Finance 50*(4), 1175–1199. ii, iv, 2, 3, 4, 6, 7, 8, 27, 30, 58, 59, 60, 61, 62, 65, 67, 68, 92, 93, 97, 112, 113, 114, 125, 131

Hasbrouck, J. (2002). Stalking the efficient price in market microstructure specifications: an overview. *Journal of Financial Markets 5*(3), 329 – 339. 2, 9, 15, 59, 70, 71, 92

Huang, R. D. (2002). The quality of ecn and nasdaq market maker quotes. *The Journal of Finance 57*(3), 1285–1319. 4, 92

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications 119*(7), 2249 – 2276. 3, 21, 92, 110

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica 59*(6), 1551–1580. 5, 60, 115

Kim, H. (2013). Generalized impulse response analysis: General or extreme? *EconoQuantum 10(2), 135-141*. 68

Koop, G., M. Pesaran, and S. Potter (1996). Impulse response analysis in nonlinear multivariate models. *Journal of Econometrics 74*(1), 119–147. 65

Lanne, M. and H. Lütkepohl (2010). Structural vector autoregressions with nonnormal residuals. *Journal of Business & Economic Statistics 28*(1), 159–168. 8, 59

Lehmann, B. N. (2002). Some desiderata for the measurement of price discovery across markets. *Journal of Financial Markets 5*(3), 259 – 276. 2, 9, 59, 92, 113

Lien, D. and K. Shrestha (2009). A new information share measure. *Journal of Futures Markets 29*(4), 377–395. 63, 112

Lien, D. and K. Shrestha (2014). Price discovery in interrelated markets. *Journal of Futures Markets 34*(3), 203–219. 7, 59, 62, 92

Lütkepohl, H. (2007). *New Introduction to Multiple Time Series Analysis*. Springer Publishing Company, Incorporated. 85

Menkveld, A. J., S. J. Koopman, and A. Lucas (2007). Modeling around-the-clock price discovery for cross-listed stocks using state space methods. *Journal of Business & Economic Statistics 25*(2), 213–225. 6, 92, 94

Ni, S., J. PAN, and A. M. POTESHMAN (2008). Volatility information trading in the option market. *The Journal of Finance 63*(3), 1059–1091. 128

Nuria, A., V. Arago, and E. Salvador (2017). The influence of intraday seasonality on volatility transmission pattern. *SSRN: https://ssrn.com/abstract=3043130*. 129

Pesaran, H. and Y. Shin (1998). Generalized impulse response analysis in linear multivariate models. *Economics Letters 58*(1), 17 – 29. 59, 65, 78, 85, 86, 93

Phillips, P. (1991). Error correction and long-run equilibrium in continuous time. *Econometrica 59*(4), 967–980. 100

Putniņš, T. J. (2013). What do price discovery metrics really measure? *Journal of Empirical Finance 23*, 68 – 83. 3, 7

Ranaldo, A. (2004). Order aggressiveness in limit order book markets. *Journal of Financial Markets 7*(1), 53 – 74. 128

Rigobon, R. (2003). Identification through heteroskedasticity. *The Review of Economics and Statistics 85*(4), 777–792. 8, 59

Ross, S. A. (1989). Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. *The Journal of Finance 44*(1), 1–17. 112

Tauchen, G. and M. Pitts (1983). The price variability-volume relationship on speculative markets. *Econometrica 51*(2), 485–505. 114, 128

Tsay, R. S. (2005). *Analysis of financial time series*. Wiley series in probability and statistics. Wiley-Interscience. 113, 119, 120

Vlaar, P. J. G. (2004). On the asymptotic distribution of impulse response functions with long-run restrictions. *Econometric Theory 20*(5), 891–903. 86

Wang, Q. (2014). Volatility discovery across stock limit order book and options markets. *Journal of Futures Markets 34*(10), 934–956. 128

Yan, B. and E. Zivot (2010). A structural analysis of price discovery measures. *Journal of Financial Markets 13*(1), 1 – 19. 2, 3, 20, 59, 65, 92

Yang, J., Z. Yang, and Y. Zhou (2012). Intraday price discovery and volatility transmission in stock index and stock index futures markets: Evidence from china. *Journal of Futures Markets 32*(2), 99–121. 113

Zhang, F. (2010). High-frequency trading, stock volatility, and price discovery. *Yale University Working Papers*. 3, 21

Zhang, L. (2006, 12). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli 12*(6), 1019–1043. 92

Zhong, M., A. F. Darrat, and R. Otero (2004). Price discovery and volatility spillovers in index futures markets: Some evidence from mexico. *Journal of Banking & Finance 28*(12), 3037–3054. 113