Price discovery and instantaneous effects among cross listed stocks

Cristina Mabel Scherrer*
Queen Mary University of London

This version: November 20, 2012

Job Market Paper

Abstract

This paper investigates instantaneous and long-run linkages between common and preferred shares traded at both domestic and foreign markets. I develop a market microstructure model in which the dynamics of the different share prices react to three common factors, namely, the efficient price, the efficient exchange rate, and the efficient voting premium. I show how to identify the structural innovations so as to differentiate instantaneous and long-run effects. First, I obtain dynamic measures of price discovery that quantify how prices traded at different venues respond to shocks on the common factors. Second, I am able to test whether shocks in the efficient exchange rate change the value of the firm. Third, I test whether shocks on the efficient voting premium have a permanent effect on preferred shares. I implement an empirical application using high-frequency data on six Brazilian large companies. I find that, in the long-run, a depreciation of the Brazilian currency leads to a depreciation of the value of the firm that exceeds the expected arbitrage adjustment. In addition, a positive shock on the voting premium yields a positive impact on the value of the firm. My price discovery analysis also reveals that one trading day suffices to impound new information on all share prices, regardless of the venue they trade at.

JEL classification numbers: G15, G12, G32, C32

Keywords: price discovery, structural VECM, high frequency data, market microstructure

Acknowledgments: I am especially grateful to Marcelo Fernandes for his comments and suggestions in this project. I am also indebted to Erik Hjalmarsson, Tim Bollerslev, Daniel Ferreira, Walter Novaes, Samy Yoshima, Pedro Saffi, Gabriela Bertol Domingues and Andrea Vedolin for valuable comments as well as the seminar participants at the Queen Mary Econometrics Reading Group. The latest version of this paper can be downloaded from http://webspace.qmul.ac.uk/cscherrer/Scherrer,Cristina_2012_Job_Market_Paper.pdf. The usual disclaimer applies.

*School of Economics and Finance, Queen Mary University of London, London, E1 4NS, UK. E-mail: c.scherrer@qmul.ac.uk. Tel: +44 (0)20 7882 8846. Fax: +44 (0)20 8983 3580
1 Introduction

Financial markets can be very informative of the economic situation of a business, an industry, a country. The interaction among players for allocation of capital ownership defines financial prices. Assets price changes reflect how new information is incorporated into perceptions, beliefs and assumptions. Prices at different markets may react differently to news, as a consequence of market’s structure, design, liquidity, efficiency, behaviour. I study how assets driven by the same fundamentals and traded at various markets react to news. I show that price’s fundamentals may have cross linkages, meaning that an innovation on exchange rate may have a permanent effect on the asset latent price, for instance. This is a major contribution to the price discovery literature, since no study has considered cross linkages among common factors before.

I compute a dynamic price discovery measure on common and preferred shares traded at both domestic and foreign markets. I set up a novel market microstructure model allowing for cross linkages among common factors. In this model, I isolate instantaneous and long-run effects on prices given structural innovations associated to common factors. My flexible econometric specification, differently from previous dynamics measures in price discovery, allows me to look at cross linkages in the same way my theoretical model does. I show how to identify the structural innovations with minimal restrictions. My theoretical model and econometric methodology yield three developments: First, I obtain dynamic measures of price discovery that quantify how prices traded at different venues respond to shocks on common factors over time. Second, I am able to test whether a devaluation of a currency leads to a devaluation of the asset exceeding the expected arbitrage adjustment. Third, I test whether changes on the latent voting premium have effect not only on common shares, but also on preferred ones. These results in a novel conclusion on the dynamic price discovery analyses: an innovation of one unit on common factors may have a larger impact on observed prices, given cross linkages among common factors.

Static price discovery has been a subject widely studied. The two most prominent static measures of price discovery are the information share (IS) and component share (CS) frameworks of Hasbrouck (1995) and Gonzalo and Granger (1995) respectively. These two methodologies and their numerous variations were broadly applied to different markets, assets and financial instruments. These studies have mainly focused on identifying either the market or the financial instrument that is the fastest on impounding new information. Harris, McInish, and Wood (2002) measure price discovery for stocks in the Dow Jones Industrial Average (DJIA) index traded at different exchanges and find that NYSE is “information dominant”. Chu, Hsieh, and Tse (1999) and Booth, So, and Tseh (1999) find that future markets are the leaders in impounding new information.

1Preferred shares have preference on receiving dividends and do not hold voting rights.
Hasbrouck (2003) shows that E-mini contracts and exchange-traded index mutual funds (ETFs) are the fastest ones in impounding new information. More recently, Grammig, Melvin, and Schlag (2005) introduced the exchange rate on the price discovery analysis using the IS framework. They focus on three German firms cross listed in the US market and their results strongly suggest that the home market is the one playing the main role on the price discovery process. By adopting a variant of the framework suggested by Grammig, Melvin, and Schlag (2005) and introducing a unique IS measure, Fernandes and Scherrer (2011) find that the relative importance of the home and foreign markets may change over time, specially given financial crises events.

The main drawback associated with the two static methodologies is that both are based on market innovations in their reduced form. Lehmann (2002) points out that price discovery is a dynamic process and that both IS and CS measures are not able to capture the price dynamics. Fundamentally, shocks on observed prices may be correlated. Specifically to measure price discovery, one would be interested in looking at instantaneous and total effects of uncorrelated innovations. Therefore, in order to appropriately address the price discovery dynamics it is paramount to adopt a structural methodology such that changes in the structural innovations can be correctly assigned to markets. Yan and Zivot (2007) introduce the first structural measure of price discovery. Their framework considers only one common factor, and they find that price discovery of the yen/euro exchange rate happens mainly through the US dollar market.

In my market microstructure model, I include latent and observed prices. Observed prices are allowed to share not only one common factor, but either two or three, namely, the efficient price, the efficient exchange rate, and the efficient voting premium. The innovations attached to each one of the three common factors are allowed to be contemporaneously correlated. This provides a considerable advantage to my theoretical approach, since it encompasses results such as cross linkages among the common factors in the short- and long-run, as well as on the observed prices. By using the structural framework, I am able to retrieve the structural counterpart associated to each common factors innovation. These structural innovations have a diagonal covariance matrix, allowing to explicitly obtain the effects of shocks on each of the observed prices across every period of time. I present in detail the short- and long-run solutions for a five-variable model.

With regard to the estimation process, I merge the two-step procedure proposed by Gonzalo and Ng (2001) and the methodology suggested by Warne (1993) in order to recover the permanent and transitory shocks in their reduced form. As a result, I do not have to face normalization issues, regarding the cointegrating vectors. To obtain the structural counterparts, I decompose the covariance matrix of the reduced form innovations allowing for the structural innovations variance to be different than one. Moreover, I chose a decomposition flexible enough to deliver any shape of relation between reduced
and structural innovations, differently than the decompositions proposed in the price discovery literature. Combining both methodological contributions, I find that my measure is order invariant. Moreover, I present a Monte Carlo exercise showing that my measure outperforms, in terms of relative mean squared error (RMSE), the measures proposed in Gonzalo and Ng (2001) and Yan and Zivot (2007). Yan and Zivot (2007) apply the Gonzalo and Ng (2001) approach to the price discovery study for a two-variable model, also allowing for the structural variance matrix to be different from the identity matrix. Kim (2010a) expands the work of Yan and Zivot (2007) to a three-variable model sharing two common factors (only one cointegrating vector). My work differs from theirs in three ways: first, I introduce a theoretical model where I allow for contemporaneous correlation among common factors, and hence, short- and long-run solutions are identified as functions of the parameters that relate correlated innovations to structural innovations. Second, I decompose the covariance matrix of the reduced innovations in such a way that common factors may impact all observed prices in the long-run. This is an important advantage of my approach when compared to the standard triangular decomposition adopted in Yan and Zivot (2007) and Kim (2010a). Under their approach the sub-matrix containing the long-run effects of changes in the structural permanent innovations is restricted to have ones in its diagonal and zeros on the upper block. Third, as discussed in the Monte Carlo section, my methodology contribution overcomes the estimation problem that arises in their work when models have more than one cointegrating vector (the number of cointegrating vectors increase as one adds more variables sharing the same common factor to my model, since there are additional combinations of variables delivering a stationary process). Furthermore, I show that my theoretical results and methodology holds for a seven-variable model with three common factors.

I study the instantaneous and long-run price effects for cross listed Brazilian firms traded at Brazil and US. I work with a tick by tick data set from BM&FBovespa (the Brazilian stock exchange), NYSE and ARCA incorporating common and preferred shares. Brazilian firms are traded at the US markets as American Depositary Receipts (ADR). The data ranges from December 2007 to November 2009. There are important characteristics that make this data set relevant for this study. The first one is regarding core trading hours. There is a large trade intersection period in the entire year between Brazil and US, varying from six hours and thirty minutes to five hours and thirty minutes. This allows us to have much more information regarding the price discovery process, when compared to studies among companies traded at the European and the US markets, where the intersection period is merely two to three hours. Second, I work with companies that possess distinctive characteristics such as: core business, ownership structure, global insertion, and strategic and political relevance. Brazilian firms are Ambev (beverage, private owned, global market), BR Telecom (telecommunication, private owned, domestic market), Bradesco (finance, private owned, domestic market), Gerdau (steel, private
owned, global company), Vale (mining, private owned with governmental influence, global company) and Petrobras (oil, state owned, global company). Apart from BR Telecom, they all are part of Ibovespa, the main index of the Brazilian stock exchange. Vale and Petrobras are the largest Brazilian companies and this is reflected in their weights on the Ibovespa index.

Liquidity plays an important role on defining whether both common and preferred shares traded at all exchanges can be considered in my baseline model. I require stocks to be highly liquid as a condition to be part of the analysis. This implies that the number of variables varies from company to company. I report results considering four-, five- and seven-variable models. In all different specifications, I am able to understand which markets are the main drivers for the price determination of these Brazilian cross listed firms. Moreover, by including the exchange rate in all models, I test whether shocks on the innovations associated with the efficient price and efficient exchange rate have a long-run impact on the exchange rate and firm’s value, respectively. The five- and seven-variable models include the common shares traded at both domestic and foreign markets. This adds the voting premium, as a further common factor, to the price dynamics analysis. Under this specification, I test whether shocks on the voting premium lead to a permanent effect on the preferred shares.

The empirical results corroborate the solutions of my market microstructure model, implying that innovations associated with the latent processes are contemporaneously correlated, leading to cross linkages among common factors. This shows that measuring price discovery independently of exchange rate or other common factors may lead to misleading results. I document that, in the long-run, a depreciation of the Brazilian currency leads to a depreciation of the value of the firm that exceeds the expected arbitrage adjustment. In addition, a positive shock on the voting premium yields a positive impact on the value of the firm. In general, ARCA is faster than NYSE in the short run, but they are equally important in the long-run. These results are consistent across all the six companies, as well as at different sampling frequencies. My price discovery analysis also reveals that one trading day suffices to impound new information on all share prices, regardless of the venue they trade at.

The remaining of the paper proceeds as follows: Section 2 presents the institutional background of the trading venues and the main data features. Section 3 introduces my theoretical model. Section 4 describes the estimation procedure and show my identification strategy. Section 5 documents the empirical results for Brazilian firms, whereas Section 6 offers some concluding remarks. The appendix presents additional results regarding my identification strategy and the Monte Carlo study addressing the performance of my estimation methodology.
2  Institutional background and data handling

The use of data from the Brazilian stock exchange is particularly interesting. The large
time intersection between Brazil and the US provides just a very small daily period of time
where information is coming to only one market, because of opening and closure hours.
The two markets have a time overlap period of six hours and a half during the majority of
the year, from mid February to mid November, with the Brazilian exchange being open
only for thirty minutes while NYSE is closed. The least intersection between the two
markets occurs from mid November to mid February, where they are still jointly open for
five hours and a half. This is of great value to analyse the price discovery mechanism,
since markets are less likely to lose their importance on price discovery because of trading
hours. Hence, it becomes easier to isolate the different aspects driving the price dynamics
in all markets.

I use a tick by tick data set of Brazilian blue-chip companies traded at three different
venues, Bovespa, NYSE and ARCA. The sample period is beneficial (December, 2007 to
November, 2009) since it is large enough to englobe a variety of movements in the stock
markets, including the 2008/2009 financial crises. With this data set, I measure price
discovery considering stable and highly instable periods. The use of a high frequency
data set provides timely incorporation of new information in each different market. A
daily data set of these markets would not provide the information needed to measure
price discovery, since at a day to day level, all markets would have incorporated new
information.

The BM&FBovespa is the only stock exchange in Brazil and the leading exchange in
Latin America in terms of number of contracts traded. It is a very active market, with
2011 average daily trading volume of USD 3.5 billion (Bovespa segment), being one of
the top 10 stock exchanges in the world (in market capitalization). In 2002, Bovespa
bought equity membership of the Rio de Janeiro stock exchange (BVRJ) and in 2008,
Bovespa merged with the Brazilian Mercantile & Futures Exchange (BM&F), forming
the BM&FBovespa. It is a fully electronic exchange (end of open outcry transactions at
Bovespa took place in 2005 and derivatives transactions was in 2009) and operates under
supervision of the CVM (Brazilian Securities Commission). BM&FBovespa markets in-
clude equity, commodities and futures, foreign exchange, securities and ETF’s (exchange
traded funds). Brazil achieved the investment grade rating from Standard & Poor’s in
April 2008. Fitch and Moody’s increased Brazilian ratings in May 2008 and September
2009, respectively. IBOVESPA is the most important index at BM&FBovespa, the index
and the exchange as a whole have reflected the improvement in these ratings in the past
five years with a trading volume increase of 7.3% (compound annual growth rate).

Foreign companies usually are traded and listed in the US market through American
Depositary Receipts (ADR). An ADR is a physical certificate evidencing ownership of
a US dollar denominated form of equity in a foreign company. It represents the shares of the company held on deposit by a custodian bank in the company’s home country and carries the corporate and economic rights of the foreign shares, subject to the terms specified on the ADR certificate.

Brazilian companies are very liquid in the US market, sometimes having more trading activity there than in Brazil. I choose to use very liquid companies in the three markets, in order not to lose information once aggregating a very illiquid firm with one presenting a much larger number of trades. I work with the following firms: Ambev (beverage), BR Telecom (telecommunication), Bradesco (finance), Gerdau (steel), Vale (mining) and Petrobras (oil). Apart from BR Telecom, they all belong to IBOVESPA. Preferred shares of Vale and Petrobas are the two most heavily traded shares at the Brazilian market, with Gerdau and Bradesco coming in the top 15. The number of trades at Bovespa for preferred Petrobras is around 9 million for the 2-year data set. For Vale is 6.4 million, Gerdau 3.2 million, Bradesco 3 million, Ambev 0.7 million and 0.6 for BR Telecom. At Arca and Nyse, these stocks have also a significant number of trades, for some shares larger than at the Brazilian exchange (see table 1).

As I start from a tick-by-tick data base, I had to follow two important steps before I could actually estimate measures for price discovery. The first step relies on cleaning the data, once it may present microstructure effects, such as bid ask bounces, discreetness of price changes, etc. I use the algorithm proposed by Brownlees and Gallo (2006) to exclude the entries not plausible to normal market activity. Table 1 presents the number of raw observations, the identified outliers and the number of trades after the implementation of the cleaning filter.

Secondly, I need to deal with the non synchronization issue of the data set. Some stocks are more intensively traded than others at certain periods and/or overall the time span. I aggregate these series, based on a defined time interval ranging from 30s to 300s. I allow this range because some stocks do not present enough trades to aggregate at a higher frequency, for instance 30s. This is the case for BR Telecom. Aggregating at a higher frequency would result in a large amount of missing observations (since some 30s-intervals would not have any trades), which may lead to serial correlation. The benefits for the price discovery measure would not be that high, since no trades are happening. Even though I are careful on choosing the intervals for each firm, I estimate the covariance matrix using the Newey-West estimator in order to control for serial autocorrelation. To aggregate the series I use the methodology proposed by Harris, McInish, L.Shoesmith, and Wood (1995). For each interval, I identify the last market to have the first trade, and acquire the most recent trade from the other markets, forming the first time tuple, and so on (called in their paper as the ‘replace all’ method). Table 2 has the initial and final number of observations given the aggregation process.

Figure 1 shows the price evolution for the shares used in this study. There are periods
of large price instability, characterized specially by the effects of the 2008/2009 financial crises. A significant drop in prices is observed for the majority of stocks, jointly with an increase in volatility.

3 A simple model for price discovery

I present a simple market microstructure model, in order to guide the understanding of the empirical results. I consider a firm traded at four markets and the exchange rate. There is a common and preferred share in both the home market and the foreign market. This model setup can be easily extended to the case with six markets plus the exchange rate and also reduced to the case of three markets plus exchange rate case (as is the instance for some companies analyzed in the empirical section).

The main target of this model is to have price variations in the short run (instantaneous effects) and in the long-run as a function of permanent and transitory uncorrelated innovations, as presented in Gonzalo and Granger (1995) and Gonzalo and Ng (2001) and explained in detail in the next section. By implementing this breakdown, I am able to isolate permanent innovations coming from different sources. I have instantaneous effects when $L = 0$ ($L$ being the lag operator). This is the effect at the same period in time of the innovation. This comes from a Vector Moving Average (VMA) Model, where the matrix giving instantaneous effects into prices is different than the identity matrix (structural form). As long-run, I refer to $L = 1$, which gives the sum of impacts from innovations across time in a VMA model. This measure gives the total effect on prices of an innovation.

I write the efficient price and efficient exchange rate as random walk processes being affected by permanent innovations. I insert a third efficient factor that is common to the observed prices, the efficient voting premium. I also model it as a random walk process, since from the empirical results, I find systems sharing three common factors, leading to the conclusion that the voting premium is a non-stationary random walk process. Additionally, I could not find any plausible reason for the voting premium not be a random walk. It is a financial asset at last, and as so, it follows the unpredictability characteristic on its returns. These three random walks are non observable prices. Each permanent innovation related to a particular common factor may affect other non observable prices. This effect is given by $\lambda, \rho, \pi$ and $\kappa$. In other words, innovations on common factors are correlated. I aim to isolate them and quantify their impacts in each of the observed prices. The intuition on the allowance for this correlation comes from empirical analysis of the date. For instance, the correlation between the exchange rate (Brazilian currency over US dollars) and the main index of the Brazilian stock exchange (Ibovespa) is equal to -0.60 during December 2007 and November 2009. These are observed prices, but brings the questions if the same is true for the latent process of these prices. If there were no cross
linkages among the common factors, one would expect to find these parameters equal to zero in the empirical results. If markets are instantaneously efficient, the difference between each of these prices and their respective observed counterparts are the transitory effects at each point in time. For instance, the observed price is equal to the efficient price plus transitory effects, such as bid ask bounces, price discreetness, inventory effects, etc. In the same way that the observed voting premium (the difference between the common and the preferred share, considering efficiency of these observable prices on incorporating news on the efficient price) is equal to the efficient voting premium plus transitory innovations, such as liquidity effects.

The permanent innovations, \( \eta_t \), are defined with \( \mathbb{E}(\eta_t) = 0 \) in their structural form, implying that \( \text{Var}(\eta_t) \) is a diagonal matrix. I define the logarithm function of the efficient price of the asset \( m_t \), of the efficient exchange rate \( \dot{e}_t \) and of the voting premium \( v_t \), such that:

\[
\begin{align*}
\dot{e}_t &= \dot{e}_{t-1} + \eta^e_t + \lambda \eta^m_t \\
m_t &= m_{t-1} + \eta^m_t + \rho \eta^e_t + \pi \eta^v_t \\
v_t &= v_{t-1} + \eta^v_t + \kappa \eta^m_t
\end{align*}
\]

where \( \eta^e_t, \eta^m_t \) and \( \eta^v_t \) are the permanent innovations associated to exchange rate, efficient price and voting premium, respectively; \( \eta_t = (\eta^e_t, \eta^m_t, \eta^v_t)^\prime \); and \( \dot{e}_t \) is defined in terms of home currency. Note that, from the structure imposed in (1), (2) and (3), the efficient price, exchange rate and voting premium are random walk processes, implying that their first difference are \( I(0) \) processes with covariance matrix given by:

\[
\text{Var}([\Delta \dot{e}_t, \Delta m_t, \Delta v_t]) = \\
\begin{pmatrix}
\varsigma^2_e + \lambda^2 \varsigma^2_m & \rho \varsigma^2_e + \lambda \varsigma^2_m & \lambda \kappa \varsigma^2_v \\
\rho \varsigma^2_e + \lambda \varsigma^2_m & \varsigma^2_m + \rho^2 \varsigma^2_e + \pi^2 \varsigma^2_v & \kappa \varsigma^2_e + \pi \varsigma^2_v \\
\lambda \kappa \varsigma^2_v & \kappa \varsigma^2_e + \pi \varsigma^2_v & \varsigma^2_v + \kappa^2 \varsigma^2_m
\end{pmatrix}
\]

where \( \text{Var}(\eta^e_t) = \varsigma^2_e, \text{Var}(\eta^m_t) = \varsigma^2_m \) and \( \text{Var}(\eta^v_t) = \varsigma^2_v \).

Denote \( Y_t \) the vector containing the logarithm function of the exchange rate and observed prices on different venues. I want a high frequency trading model that reflects price adjustments in a partial way, such that innovations are not completely incorporated by all market in each \( t \), i.e. in each microsecond. If markets were efficient, I could have the parameters giving this partial adjustment set to unit, making observed prices equal to efficient price plus some transitory effects, such as bid-ask bounce, price discreetness, liquidity effects, etc. The model below is a modified and extended version of other models used in the literature (see Amihud and Mendelson (1987), Hasbrouck and Ho (1987) and Yan and Zivot (2010)).
I present the observed exchange rate as a function of the efficient exchange rate, past observed exchange rate and transitory effects (denoted by the $2 \times 1$ vector $\eta^*_{t}$ and the $1 \times 2$ vector of parameters $b_1$). I impose two transitory innovations because I need the number of innovations to be equal to the number of markets. This goes in line with Gonzalo and Ng (2001) and Yan and Zivot (2007), since there is the need to invert a decomposition of the covariance matrix, as the next section explains in detail. Hence, if one has $n$ number of markets, it will be necessary to have $n - p$ number of transitory effects, and $p$ number of permanent effects. In this model, I have three permanent innovations ($\eta^*_i, \eta^{**}_i$ and $\eta^{**}_i$), two transitory ones and five variables, allowing invertibility of some specific matrices in my identification strategy. Adjustments to permanent innovations are allowed to happen in a partial way, by inserting $\gamma_i$. Hence, all observed prices are allowed to adjust to the three latent prices. I distinguish prices traded at different currencies, such that $y^*_{i,t}$ and $y^*_n,t$ entail prices in foreign currency (currency that they are actually traded), whereas $y_{i,t}$ and $y_n,t$ are expressed in home currency.

I write the price process for each $y_{i,t}$ as below.

$$ e_t = e_{t-1} + \gamma_1 (m_t - m_{t-1}) + \gamma_2 (e_t - e_{t-1}) + \gamma_3 (\bar{v}_t - v_{t-1}) + b_1 \eta^*_{t} \tag{5} $$

$$ y_{2,t} = y_{2,t-1} + \gamma_4 (m_t - y_{2,t-1}) + \gamma_5 (e_t - e_{t-1}) + \gamma_6 (v_t - v_{t-1}) + b_2 \eta^*_{t} \tag{6} $$

$$ y_{3,t} = y_{3,t-1} + \gamma_4 (m_t - y_{3,t-1} + v_{t-1}) + \gamma_5 (e_t - e_{t-1}) + \gamma_6 (v_t - v_{t-1}) + b_3 \eta^*_{t} \tag{7} $$

$$ y^*_{i,t} = y^*_{i,t-1} + \gamma_4 (m_t - y^*_{i,t-1}) + \gamma_5 (e_t - e_{t-1}) + \gamma_6 (\bar{v}_t - v_{t-1}) + b_4 \eta^*_{t} \tag{8} $$

$$ y^*_{n,t} = y^*_{n,t-1} + \gamma_4 (m_t - y^*_{n,t-1} + v_{t-1}) + \gamma_5 (e_t - e_{t-1}) + \gamma_6 (\bar{v}_t - v_{t-1}) + b_5 \eta^*_{t} \tag{9} $$

Where $b_1, b_2, b_3, b_4$ and $b_5$ are $1 \times 2$ vectors and $\eta^*_{t}$ is a $2 \times 1$ vector. I show the steps to obtain $\Delta y_{i,t}$ only for the preferred share traded at the domestic market ($\Delta y_{2,t}$). The remaining equations ((5),(7), (8) and (9)) are obtained in a similar manner.

$$ y_{2,t} - y_{2,t-1} = y_{2,t-1} - y_{2,t-2} + \gamma_2 (m_t - m_{t-1} - y_{2,t-1} + y_{2,t-2}) + b_2 (\eta^*_i - \eta^*_{i-1}) $$

$$ (1 - L + L \gamma_2) \Delta y_{2,t} = \gamma_2 (\eta^{**}_i + \rho \eta^*_i + \pi \eta^*_i) + b_2 (\eta^*_i - L \eta^*_i) $$

$$ \Delta y_{2,t} = (1 - L + L \gamma_2)^{-1} [\gamma_2 (\eta^{**}_i + \rho \eta^*_i + \pi \eta^*_i) + b_2 (\eta^*_i - L \eta^*_i)] $$

By setting the lag operator equal to zero, I have the instantaneous effects of permanent
and transitory innovations for each price series.

\[
\begin{pmatrix}
\Delta e_t \\
\Delta y_{2,t} \\
\Delta y_{3,t} \\
\Delta y_{4,t} \\
\Delta y_{5,t}
\end{pmatrix}
= 
\begin{pmatrix}
\hat{\gamma}_1 + \gamma_1 \rho & \hat{\gamma}_1 \lambda + \gamma_1 + \hat{\gamma}_1 \kappa & \hat{\gamma}_1 + \gamma_1 \pi & b_1 \\
\hat{\gamma}_2 + \gamma_2 \rho & \hat{\gamma}_2 \lambda + \gamma_2 + \hat{\gamma}_2 \kappa & \hat{\gamma}_2 + \gamma_2 \pi & b_2 \\
\hat{\gamma}_3 + \gamma_3 \rho & \hat{\gamma}_3 \lambda + \gamma_3 + \hat{\gamma}_3 \kappa & \hat{\gamma}_3 + \gamma_3 \pi & b_3 \\
\hat{\gamma}_4 + \gamma_4 \rho & \hat{\gamma}_4 \lambda + \gamma_4 + \hat{\gamma}_4 \kappa & \hat{\gamma}_4 + \gamma_4 \pi & b_4 \\
\hat{\gamma}_5 + \gamma_5 \rho & \hat{\gamma}_5 \lambda + \gamma_5 + \hat{\gamma}_5 \kappa & \hat{\gamma}_5 + \gamma_5 \pi & b_5
\end{pmatrix}
\begin{pmatrix}
\eta_t \\
\eta_{T_t}^m \\
\eta_t^c \\
\eta_{T_t}^T
\end{pmatrix}
\]  
(10)

Where \( b_1, b_2, b_3, b_4 \) and \( b_5 \) are \( 1 \times 2 \) vectors and \( \eta_{T_t}^T \) is a \( 2 \times 1 \) vector. By making the lag operator equal to unit, I get the long-run effect on prices, as below:

\[
\begin{pmatrix}
\Delta e_t \\
\Delta y_{2,t} \\
\Delta y_{3,t} \\
\Delta y_{4,t} \\
\Delta y_{5,t}
\end{pmatrix}
= 
\begin{pmatrix}
1 & \lambda & 0 & 0 \\
\rho & 1 & \pi & 0 \\
\rho & \kappa + 1 & \pi + 1 & 0 \\
\rho - 1 & 1 - \lambda & \pi & 0 \\
\rho - 1 & 1 + \kappa - \lambda & \pi + 1 & 0
\end{pmatrix}
\begin{pmatrix}
\eta_t \\
\eta_{T_t}^m \\
\eta_t^c \\
\eta_{T_t}^T
\end{pmatrix}
\]  
(11)

where \( 0 \) are \( 1 \times 2 \) vectors and \( \eta_{T_t}^T \) is a \( 2 \times 1 \) vector.

To evaluate which market is more important in the price discovery process, I need to look at the combination of parameters from each equation in the short-run (\( L = 0 \)) and long-run (\( L = 1 \)). When I refer to long-run this is about a day or even less in the high frequency context, differently than in macroeconomic scenarios of years or decades.

The combination of parameters is exactly what I estimate and quantify as price discovery measures in the simulations subsection (for a simpler model) and in the empirical results.

Let us call the matrix with the combination of parameters that gives the instantaneous effects as \( D_0 \), where on the rows there is each market’s effect given an innovation on the permanent and transitory shocks. Similarly, I define \( D(1) \) as the effects when \( L = 1 \).

Therefore, I can gauge price discovery by two measures: importance compared to other markets, considering each permanent innovation and fastness, as below:

Importance = Max \( d_{0,ii} \) and Max \( d(1)_{ii} \)  
Fastness = \( d_{0,ii}/d(1)_{ii} \)  
(12)
(13)

where \( d_{0,ii} \) and \( d(1)_{ii} \) are elements of \( D_0 \) and \( D(1) \) matrices, respectively.

The matrices \( D_0 \) and \( D(1) \) have dimension \( K \times K \), and they give the instantaneous and long-run impacts on market prices from shocks on the permanent and transitory innovations. I am only interested in the parameters accompanying the permanent innovations, because these are the relevant parameters for price discovery. Therefore, I do not look at the part of \( D_0 \) or \( D(1) \) related to the transitory shocks.

I look at the overall importance when I compare the parameters of \( D_0 \) and \( D(1) \) for all
the markets. The fastness measure may be understood as the proportion of permanent shocks incorporated instantaneously compared to what is incorporated in the long-run. It is important to point out that both measures depend on the positivity of the elements in $D_0$ and $D(1)$.

Apart from the price discovery analyses, a question now lies on what to expect from the parameters defined in (1), (2) and (3). I expect to find $\pi > 0$, meaning that a positive innovation on the efficient voting premium delivers a positive impact on the efficient price. If the voting rights of a given firm turn to be more valuable, this would deliver an increase in the value of the asset itself, i.e., the efficient price. An increase in the voting premium is the same as an increase in the price for a vote. Hence, if a vote turns to be more expensive, it is possible that better decisions will be taken by the ones owning these rights, once they have paid more for them. Rational individuals should do a better use of something they start paying more for. Therefore, the efficient price is affected positively by the expectations and realizations of better votes.

I conjecture to observe $\rho < 0$, where a positive innovation on the exchange rate (depreciation of the home currency) leads to a negative effect on the efficient price. Ignoring firm specificities regarding imports and exports, a depreciation of the currency where the business is situated and has its main operations would result in a decrease in its value. Exchange rate may affect a firm business in many different fronts: transaction (imports and exports), competitors, suppliers, suppliers competitors, access to international capitals, and so on. This last one, particularly, may impact considerably the cost of firms that aim to finance investments with external resources. This may affect Brazilian companies, as they are located at an emerging country, they do search for external capital resources. A useful literature review on effects of exchange rate on firm value can be found at Muller and Verschoor (2006).

I would expect $\lambda$ to be in general equal to zero, since innovations on the efficient price should not lead to effects on the exchange rate. However, depending on how related the firm activities can be to the exchange rate, or how the overall movements of the stock exchange in the home country can be correlated to exchange rate, I might see $\lambda$ different than zero. I would expect anything different than zero in $\lambda$ to be related to correlation between observed exchange rate and observed prices.

The observed voting premium (observed difference between common and preferred shares) can be defined as the efficient voting premium plus some transitory effects. These effects can be the result of liquidity issues, such as that either the common or the preferred share has a more liquid market than the other, allowing investors to price this difference. The efficient voting premium is a function of private benefits an investor could get from holding voting rights as well as function of a possible premium over the preferred share, in the instance of a merger or an acquisition (see Zingales (1994) and Zingales (1995) for explanations on private benefits and merger premium, and for reasons on why voting
premium vary across countries). Given that, if an increase in the firm’s value generates an increase in private benefits or a potential acquisition premium, \( \kappa \) should be positive, leading to a positive impact on the efficient price. This yields a positive effect on the voting premium. However, if I consider that appropriation of private benefits is more related to the culture of the company, and how strong or weak the country institutions are, they should not be affected by the efficient price, unless this innovation on the efficient price is coming from a change in the company regarding the appropriation of private benefits per se.

4 Getting structural parameters from reduced-form VECM

My data set is composed by a single security being traded at different markets, namely Brazil and USA. As they share at least one common factor among them, they are cointegrated. Hence, I use a vector error correction model (VECM) to estimate the price discovery parameters. As I aim to have a structural measure, I want to recover the structural innovations from the VECM residuals. I use Gonzalo and Granger (1995) and Gonzalo and Ng (2001) methodology to retrieve reduced form permanent and transitory innovations from market residuals in their reduced form. I merge their methodology with the work of Warne (1993), making my price discovery measures more accurate. In order to transform reduced form permanent innovations into structural innovations, I modify the standard procedure on Gonzalo and Ng (2001), allowing the variance of the structural innovations to be different than the identity matrix. I show that this methodology works well for models with one, two or three common factors.

Hence, the first step is to estimate a reduced-form VEC model as,

\[
\Delta y_t = \xi_1 \Delta y_{t-1} + \xi_2 \Delta y_{t-2} + \ldots + \xi_p \Delta y_{t-p} + \zeta + \xi_0 y_{t-1} + \epsilon_t, \tag{14}
\]

where \( y_t \) is a vector of price series in different markets, \( \xi_0 = \alpha/\beta' \) and \( \epsilon_t \) is a zero mean white noise process with a non-diagonal covariance matrix \( \Omega \). I impose restrictions on the constant term for the absence of deterministic time trends.

Gonzalo and Granger (1995) propose a way to estimate the common factors from a reduced form model (VECM) and Gonzalo and Ng (2001) show a two-step procedure on how to obtain permanent and transitory structural innovations from the reduced-form errors. To this purpose, I first estimate (14) using a the full-information maximum likelihood (FIML) approach proposed by Johansen (1988) and Johansen (1991) and discussed in Hamilton (1994) in order to avoid any possible misspecification in the model derived from setting normalization conditions on the cointegrating vector.
Once I estimate the VEC parameters, I am in position where I can back out the vector moving average (VMA) coefficients through dynamic simulation (see Hamilton (1994)). Note that the VMA equation in (15) is driven by the reduced form errors.

\[ \Delta y_t = \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \ldots = \Psi(L)\epsilon_t, \]  

(15)

In the above equation, change in prices are given by the market reduced form contemporaneous innovations \( \epsilon_t \) and lagged values. The problem lies on the likely contemporaneous correlation of \( \epsilon_t \). Innovations in the home market might be contemporaneously correlated with innovations in the foreign market, making the price discovery analysis problematic. My target is to have a VMA expression with relation to structural innovations, where the contemporaneous correlation among these innovations will be null.

If I were to estimate VMA parameters with relation to market structural innovations, I would have impose a series of restrictions in order to identify these parameters. These restrictions would require prior knowledge from the markets, which turns to be difficult to be justified. It is less harmful, therefore, to consider assumptions on permanent and transitory structural innovations since they are easier to implement, not requiring any prior judgement about the markets.

Hence, I want to obtain a VMA expression as in (15), that now is driven by the structural permanent and transitory shocks.

\[ \Delta y_t = d_0 \eta_t + d_1 \eta_{t-1} + d_2 \eta_{t-2} + \ldots = D(L)\eta_t. \]  

(16)

The covariance matrix of \( \eta_t = (\eta_t^p, \eta_t^T)' \) is diagonal, where \( \eta_t^p \) entails the permanent effect in \( y_t \) and \( \eta_t^T \) consists only of transitory effects. Gonzalo and Granger (1995) define \( \lim_{h \to \infty} \partial E_t(\partial y_{t+h}/\partial \eta_t^p) \neq 0 \) and \( \lim_{h \to \infty} \partial E_t(\partial y_{t+h}/\partial \eta_t^T) = 0 \), where \( E_t \) denotes the conditional expectation with relation to past information up to time \( t \). Therefore the long-run forecast of change in prices can only come permanent innovations. Gonzalo and Ng (2001) rotate \( \epsilon_t \) and split it into permanent and transitory innovation using matrix \( G \) (their first step on finding structural innovations), as below.

\[ G = [\alpha'_\perp, \beta']', \]  

(17)

where \( \alpha_\perp \) is a \((k-r) \times k\) matrix and \( \beta \) is a \(r \times k\) matrix.

To rotate \( \epsilon_t \) and find the reduced form permanent and transitory innovations, I need to multiply matrix \( G \) by \( \epsilon_t \) as in (18),

\[ \epsilon_t = G\epsilon_t, \]  

(18)

where \( \epsilon_t = (\epsilon_t^p, \epsilon_t^T)' \). The permanent shock in the reduced form is given by \( \alpha'_\perp \epsilon_t \), whereas
the transitory shock is \( \beta' \epsilon_t \). The variance (\( \Xi \)) of the unorthogonalized shocks \( \epsilon_t \) is still non diagonal, which opens the way for the second step: decompose \( \Xi \) in such a way that I find a relation between \( \Xi \) and the variance of the structural innovations (\( \text{Var}(\eta_t) \)). Once I have this, I can extend it to find the relation between \( \epsilon_t \) and \( \eta_t \), as well as between the parameters of the VMA in its reduced form and of the structural VMA. Gonzalo and Ng (2001) decompose \( \Xi \) using the well known Cholesky decomposition, as below:

\[
\Xi = G\Omega G' = CC' = CIC',
\]

(19)

where \( C \) is the Cholesky decomposition of the variance matrix of \( \epsilon_t \).

Setting \( \text{Var}(\eta_t) \) equal to an identity matrix gives the easy relation between the variance matrices, with \( \Xi = CIC' = CVar(\eta_t)C' \). Now, I am in the position to find the relation between \( \epsilon_t \) and \( \eta_t \), given by

\[
\eta_t = C^{-1}\epsilon_t = C^{-1}G\epsilon_t = D_0^{-1}\epsilon_t,
\]

(20)

where \( D_0 = G^{-1}C \).

Change in price series are given by the VMA parameters as in (15), as well as (16). The way to recover the latter one using the estimated parameters from the former is given below.

\[
\Delta y_t = \Psi(L)G^{-1}CC^{-1}G\epsilon_t = D(L)\eta_t,
\]

(21)

which delivers \( D(L) = \Psi(L)G^{-1}C = \Psi(L)D_0 \). Yan and Zivot (2007) use this approach in the price discovery analysis. They implement a modification that allows for the variance matrix of \( \eta_t \) to be different than the identity matrix, but obviously still diagonal. In particular, they implement a slightly different decomposition than the Cholesky one, as below:

\[
\Xi = G\Omega G' = LDL',
\]

(22)

where \( L \) is a unique lower triangular matrix with ones in its main diagonal and \( D \) is a diagonal matrix with positive entries. I call this decomposition LDL from now onwards. Kim (2010b) extends Yan and Zivot (2007)’s model to a 3-variable model with two common factors, using the same LDL decomposition as above.

My target in the paper is to measure price across Brazilian and US market, considering two trading venues in the US. I include exchange rate, in order to split exchange rate shocks and price fundamental’s change. I also add common shares for the companies that present enough liquidity for both common and preferred shares in all markets. This leads to a minimum of four-variable model to a maximum of seven-variable model. As
I work with high frequency data, the variance of the innovations in its reduced form is much smaller than one. I do not want to force the variance of the structural innovations \( \text{Var}(\eta_t) \) to be equal to one, hence I need a methodology that allows it to be different than the identity matrix, nevertheless still diagonal. My first choice is to follow Yan and Zivot (2007). Applying LDL decomposition on matrix \( G \), as in (17), brings a bias to the estimates in models with more than one cointegrating vector (which is always the case in this paper). This is illustrated with the Monte Carlo simulations in the appendix. I try to correct this problem by constructing matrix \( G \) in an alternative way, in particular, on the way to recover the transitory innovations.

My proposed methodology is to recover the reduced form permanent and transitory innovations using matrix \( G^* \). \( G^* \) is built, as in Warne (1993), with \( \alpha' \Omega^{-1} \) instead of \( \beta' \), viz.

\[
G^* = [\alpha'_1, \alpha' \Omega^{-1}]'
\]  

(23)

The intuition here is the same as the one when using \( \beta' \), i.e. to get everything that is transitory and therefore vanishes away. The drawback of \( \beta' \) is that when you have more than one cointegrating vector, it does not work properly. The Monte Carlo exercises show that this change solves the problem of bias. A similar result could be obtained by implementing the Cholesky decomposition as in Gonzalo and Ng (2001).

The second methodological issue that I address is with relation to correlation between common factors. To go from innovations in their reduced form to structural innovations, I need to decompose the covariance matrix. By using the well known Cholesky decomposition, there is the imposition that the variance of the structural innovations are equal to one and most important, that the first common factor can only have impacts from the first structural innovation, the second can only have impacts from the first and the second and so on. Looking at equations (1), (2) and (3), these restrictions mean that \( \lambda \) and \( \phi \) are equal to zero. With the slightly modified version of the Cholesky decomposition proposed by Yan and Zivot (2007), they impose that not only \( \lambda \) and \( \phi \) are equal to zero, but also \( \pi \) and \( \kappa \). This happens because they force the innovations associated with the common factors to have impact equal to one in the long-run, which makes them uncorrelated between each other and hence, equal to the structural innovations.

What I address differently in this paper is that shocks on common factors might be correlated, hence I need to recover innovations associated to each common factor allowing for this correlation. I construct that by saying that each common factor may respond to three different structural innovations (it could be more than that, but I use three following equations (1), (2) and (3). In order to allow the parameters in these equations to be different than zero, I use the spectral decomposition on a normalized covariance matrix. By normalizing the matrix, I impose only one restriction: the variance of the structural
permanent and transitory innovations must be equal to the variance of the permanent and transitory innovations in their reduced form. With the normalized matrix, I gain \((k^2 - k)/2\) equations (as I show in the appendix), which allows me not to impose the restrictions Cholesky and LDL decomposition impose. In summary, I just implement a decomposition that does not give me any predefined shape on the decomposed matrices, allowing for the parameters of equations (1), (2) and (3) to have any value.

Hence, I decompose the variance matrix of \(\varepsilon_t\) using the spectral decomposition. With this change, I allow for the variance of \(\eta_t\) to be different than the identify matrix (opposed to Gonzalo and Ng (2001)’s methodology) and at the same time I do not have the need to restrict the long-run impact of permanent innovations to be equal to either one or zero (as it is the case on Yan and Zivot (2007)). This brings a significant benefit, specially when one realizes that the long-run impact can be different than one and zero, as my empirical results prove.

The first step is to normalize \(\Xi\) such that \(\tilde{\Xi} = \Xi \Theta^{-1}\), with \(\Theta\) being a diagonal matrix constructed with the diagonal elements of \(\Xi\). In order to identify the model, I need to impose \(\text{Var}(\eta_t) = \Theta\), as shown in the appendix. When I implement the spectral decomposition on \(\tilde{\Xi}\), I have the following:

\[
\tilde{\Xi} = \tilde{S} \tilde{S}'
\] (24)

where \(\tilde{S}\) is the squared root of \(\tilde{\Xi}\) obtained from an eigenvalue decomposition.

\[
\tilde{\Xi} = V \Lambda V^{-1} \Rightarrow \Xi^{1/2} = V \Lambda^{1/2} V^{-1},
\] (25)

where the columns of \(V\) are the eigenvectors of \(\Xi\) and \(\Lambda\) is a diagonal matrix with the corresponding eigenvalues. Hence, I can recover \(\Xi\) just by multiplying back \(\Theta\).

\[
\Xi = \tilde{\Xi} \Theta = \tilde{S} \tilde{S}' \Theta
\] (26)

If I prove that \(\tilde{S} \tilde{S}' \Theta = \tilde{S} \Theta \tilde{S}'\), I can use exactly the same steps defined above to recover \(\eta_t\). To this purpose:

\[
\tilde{S} \Theta \tilde{S}' = \Xi
\] (27)

\[
\text{Var}(\eta_t) = \Theta = \tilde{S}^{-1} \Xi \tilde{S}^{-1'}
\]

\[
\eta_t = \tilde{S}^{-1} \varepsilon_t.
\]

In the appendix I present the proof of \(\tilde{S} \tilde{S}' \Theta = \tilde{S} \Theta \tilde{S}'\). This methodology can also be applied when one desires to set \(\text{Var}(\eta_t) = I\). I then decompose \(\Xi\), as below:

\[
\Xi = G^* \Omega G^* = SS'
\] (28)
where $S$ is the squared root of $\Xi$ obtained from the eigenvalue decomposition of matrix $\Xi$ as stated in (25). I show in the appendix the identification steps for this case.

The benefit of using $\alpha'\Omega^{-1}$ instead of $\beta'$ is mainly on normalization. The use of $\beta'$ carries the question on how to normalize the cointegrating vectors. The standard way is to implement the triangular normalization, however this makes the uniqueness of parameters slightly hard to compute, since for each different order, one would also need to change the cointegrating vectors, not keeping the original triangular normalization. With $\alpha'\Omega^{-1}$ there is no question mark here. The construction of $G$ is straightforward from the results of the VEC model, delivering uniqueness on the parameters.

Regarding the use of spectral decomposition, the main contribution is that common factors might have innovations that are correlated, and so, these innovations are not in their structural form yet. Hence, I choose a methodology that allows to retrieve them in their structural form. My empirical results show that the parameters on equations (1), (2) and (3) are not zero, indicating that innovations associated with the common factors are indeed correlated. In fact, I find that an innovation related to the efficient exchange rate has a long-run (by long-run I mean one trading day) impact on the observed prices higher than the expect arbitrage adjustment. The main benefit is that I allow long-run effects to be different than one or zero. By allowing this, I include in my model long-run behavior of innovations on common factors derived from other common factors, extracted from their contemporaneous correlation. This permits a dynamic process not only on observed prices, but also on the common factors themselves. I show on the Monte Carlo simulations that this methodology achieves the best results in finite sample for models with one and two common factors. Extension to the case of more common factors can be easily implemented.

5 Price discovery for Brazilian cross listed stocks

I have two targets in this section. The first one is related to price discovery analysis, where I am keen on finding what are the roles of each of the markets in the price discovery dynamics. Secondly, I want to check whether the model presented in Section 3 is a valid model for my data set.

Starting with the latter one, I find a significant difference between instantaneous and long-run effects, averring that $\gamma_i$ (in equations (5) to (9)) is different than unit and there is a partial adjustment process given an a shock on the permanent innovations. Hence, the model stated in Section 3 seems to fit well in terms of the partial adjustments. Important to mention that, as in Section 3, by long-run I mean hours within a day, since I am dealing with high frequency data.

As the market microstructure model shows, I believe the efficient price does incorporate part of shocks on the permanent innovations associated with both the exchange rate
and voting premium. The same is true for the efficient voting premium and the efficient exchange rate (both being affected by shocks on the permanent innovation associated with the efficient price). These would firstly imply that the parameters $\lambda$, $\kappa$, $\pi$ and $\rho$ to be different than zero. Furthermore, if the direction of the cross linkages among the assets are the ones discussed in section 3, I should expect the parameters signs to reflect that.

Indeed, I do find the majority of the elements of $D(1)$ being statistically different than zero across all companies. Therefore, the conclusions on that are twofold: in the long-run, a depreciation of the Brazilian currency leads to a decrease on the value of the firm that exceeds the expected arbitrage adjustment. Second, a positive innovation in the efficient voting premium leads to an increase in the asset’s value. These are exactly what I infer when analysing the parameters in Section 3: $\pi > 0$ and $\rho < 0$ respectively. As I am not able to identify all the parameters in some equations (specially in foreign common shares, where the number of parameters is higher), I do not find values of $\lambda$ and $\kappa$ as high as I find for $\pi$ and $\rho$. When I was able to identify them, I find $\lambda$ very close to zero. Although I expected to find a positive $\kappa$, I do not find strong evidences on that, finding it to be closer to zero. I am also not able to identify all the $\gamma$ parameters, however, it is not so much of interest to look at them individually, since the effects of innovations on the efficient prices are given by the combination of parameters, which is what I estimate and analyse below.

Tables 5, 2, 5, 6, 7 and 8 report results for the six companies considered in this paper. For the first four companies (Gerdau, BR Telecom, Bradesco and Ambev) there are four markets: exchange rate (Brazilian Reais/USD dollars), preferred shares traded at the Brazilian market, at NYSE and at ARCA. The shares traded at Brazil are quoted in Brazilian Reais (R$) and the shares traded at the US market are expressed in US dollars. For these four companies, I find two cointegrating vectors ($\beta$), which are in the last two columns of their tables. This leads us to two common factors, seen as the efficient exchange rate and the efficient price. Hence, I analyse the first two columns of $D_0$ and $D(1)$, since these are the ones related to the two permanent innovations. I call these permanent innovations $p^e$ and $p^m$, permanent innovation on efficient exchange rate and efficient price respectively. $D_0$ has the instantaneous effect of a permanent innovation, whereas $D(1)$ has the long-term effect, as defined and computed in Section 4.

For the last two companies (Petrobras and Vale) there are seven and five markets respectively: exchange rate, preferred and common shares traded at the Brazilian market and at NYSE. Petrobras has also common and preferred shares at ARCA. For these two companies I find four and two cointegrating vectors ($\beta$), respectively. This leads us to three common factors for both companies, seen as the efficient exchange rate, the efficient price and the efficient voting premium. The observed voting premium is the difference between the common and preferred shares. The efficient voting premium is rid
of transitory effects, such as differences in liquidity between the two stocks. I analyse the first three columns of $D_0$ and $D(1)$, since these are the ones related to the three permanent innovations. I call these permanent innovations $p^e$, $p^m$ and $p^v$, as in my theoretical model, where they stand for permanent innovations on the efficient exchange rate, efficient price and efficient voting premium, respectively.

In general, I find the US market as the most important for the price discovery process. In particular, ARCA impounds more information than NYSE which can be explained by the fact that ARCA has a smarter router system. This router is able to check among other exchanges if there is a better quote than the one at ARCA. If this is the case, it executes the order at the venue where the best quote is available. This special characteristic seems to give a more important role for ARCA in the price discovery process when compared to NYSE. In addition to that, it is important to give attention to how liquid (in terms of number of trades) the stocks are in each exchange. ARCA has a similar or higher number of trades than NYSE for the majority of stocks, apart from BR Telecom and Ambev. For these two companies, ARCA has half the number of trades than NYSE, which might affect the importance of ARCA. In the long-run, ARCA and NYSE are equally important.

The highest importance of the US market may be explained by the characteristics of investors in the two markets. Brazilian investors do trade at the US market (reasons might include the fact that US is a much bigger market with a higher potential for diversification, exchange rate issues, etc). It is less likely, however, for an US investor to trade at the Brazilian market, when the stock is available in the US. Thus I would have two sources of information in the US market, whereas only one in Brazil.

Since the US market is the most important, one would expect this market to incorporate not only innovations on the unobserved efficient price, but also innovations on the exchange rate. If the US market is faster on getting permanent news regarding the intrinsic value of the company, it also shows to be faster on adjusting the share price given a permanent innovation on exchange rate.

Regarding the effect of innovations on the efficient exchange rate on prices, there is an instantaneous overshooting. Given a unit shock on the efficient exchange rate (R$/USD), i.e. a depreciation of the Brazilian currency, there is a higher depreciation instantaneously than in the long-run. I observe this behavior for all stocks. For the BR Telecom case, I find that exchange rate overshooting is significantly smaller than the ones found for the other companies. This might be explained by the fact that I use 300s interval, leading artificial longer period assigned as short-run. Looking at the theoretical model in Section 3, this would be explained by parameter $\gamma_1$ being higher than the unit. Intuitively, it could be a signal of herd behavior during turbulent periods, specially if one considers that the Brazilian currency devaluated 49% over 90 days during mid July 2008 and beginning

The US market is the one that adjusts the price instantaneously, given a change in the exchange rate. In the long-run, I find that a depreciation of the Brazilian currency actually devalues Brazilian assets, since the parameters I find are negative for the Brazilian market and higher than one in absolute value for the US market (same amount as the Brazilian plus one unit, all negative). This comes as a strong finding, since it associates exchange rate shocks to value of assets using high frequency data. I find this result for all stocks, although some appear to have a stronger depreciation than others. Gerdau, Vale and Petroras range from 0.50 to 0.60, while BR Telecom and Bradesco are in 0.40’s and Ambev is the one to have the least depreciation of asset, with 0.25. This goes in line with the finding of $\rho < 0$ in my theoretical model.

Regarding innovations on the efficient price, the US market has the highest parameters for all stocks, but Ambev. When I compare the importance of NYSE and ARCA, ARCA seems to be faster in incorporating news. I do not find this result only for BR Telecom and Ambev, explained by the fact the these two stocks are so traded at ARCA as they are at NYSE. Ambev seems to be less affected by exchange rate innovations and is the only one where the Brazilian market is more important than the US market.

I now move to the voting premium aspect looking at Petrobras and Vale, since they are the companies presenting enough trades at common and preferred shares that allowed this analysis. Given a shock on the innovation associated to the efficient voting premium, the common stocks suffer an instantaneous overshooting, while the preferred shares have a negative impact, adjusting for this overshooting. In the long-run, preferred and common shares have a positive impact from a shock on the voting premium. Hence, an increase in the voting premium of a company increase the value of its asset. Again, this is the same result as the one in the theoretical model, where $\pi$ is higher than zero.

### 5.1 Robustness

To check the validity of my main results, I perform two robustness checks. The first one is with relation to time stamp, checking whether I see a difference in terms of market leadership across different periods of time. The second one checks if the way I aggregate the data may impact my final results, specially for the stocks where I have to aggregate at a lower frequency. This check is particularly important for the long-run effects, since I do expect to have differences on the short-run effects given distinct frequencies.

#### 5.1.1 Rolling Window

I perform a rolling window exercise as robustness check for my main results on instantaneous effects and long-run impact. I estimate my model considering a smaller sample

---

2http://www.oanda.com/currency/historical-rates
size, such that each regression accounts for approximately two months. I do not want to have a very small period and not be able to capture the price dynamics, but at the same time, I look for a reasonable number of windows. The shift window size is set to have the number of observations closely resembling two weeks.

Figures 2 to 7 displays the price dynamics over time. A few conclusions arise: although the majority of measures is considerably stable over time (specially if I consider the bootstrap intervals also graphed), there are changes in market’s importance, specially during the second half of 2008 and first half of 2009. I claim this change of behavior comes from uncertainties derived from the 2008/2009 crises, since the stability is recovered by the end of 2009. This period presents changes in the number of cointegrating vectors for Vale and Petrobras. This is the reason why I do not report the impact on prices derived from shocks on innovations associated with the voting premium, given that I would have missing estimates. I claim the change from three to two common factors comes from the elimination of the voting premium as a common factors in certain periods of the data set. As Fernandes and Scherrer (2011) find with the same data set, there is a clear change in investor’s behavior during the crises period regarding voting premium. They point out this may come from the fact that during turbulent periods, financial assets tend to be more correlated and even share a single common factor among them.

5.1.2 Interval Frequency

In Section 2, I briefly explain how I deal with non synchronous trading. Depending on the number of observations each share presents, I accommodate the size of the interval to aggregate the series. For instance, if share A has ten trades for each 30-second interval, and share B had ten trades for each 3-minute interval, I can not aggregate them in 30 seconds, since I would be incurring in a high risk of serial correlation for share B. At the same time, I do not want to aggregate at 3 minutes and loose some important information from share A. I need to find a situation in between. I estimate the covariance matrix using the Newey-West estimator as a way to overcome the serial correlation issue. Additionally, I sample the data at different frequencies, checking wether my main conclusion on market leadership change. This is specially important for stocks where I had to sample at a much lower frequency, for instance 240 seconds. The results regarding long-run impacts do not alter with this change, implying that my results are robust to different sampling frequencies.

6 Conclusion

In this paper, I investigate the price discovery for cross listed Brazilian companies. I am interested on measuring how fast permanent innovations are impounded by the different
platforms, as well as what markets are the most important on incorporating this new information.

I present a simple market microstructure model, that guides the understanding of the empirical results. My model allows the observed prices to depend on three different common factors: the efficient exchange rate, the efficient asset price and efficient voting premium. Moreover, I allow the common factors to be contemporaneously correlated, yielding the necessary conditions for cross linkages among the common factors. I provide short-run and long-run solutions as function of the structural parameters, as well as price discovery dynamic measures.

I propose an alternative methodology to measure instantaneous effects of permanent shocks on prices in spirit of Yan and Zivot (2007). My methodology does not present any issues regarding normalization of the cointegrating vectors. It is order invariant and works properly even for a large number of variables and cointegrating vectors. By using the structural framework, I am able to assess whether a permanent shock on the exchange rate changes the company’s value more than the expected arbitrage adjustment. This is an interesting point which so far has not been analyzed using the price discovery framework. I find through a Monte Carlo exercise that my measure presents better performance in finite sample, when compared to all competitors.

On the empirical results, I find that the trading platform ARACA is the most efficient market on incorporating shocks instantaneously. US market is the one that adjust for exchange rate shocks. I also observe that liquidity plays an important role on how fast markets impound new information.

Finally, the theoretical model proposed lead to interesting results regarding the price process. I find that in real terms, Brazilian companies lose value in the “long-run”, following a depreciation of the Brazilian currency and a positive innovation in the efficient voting premium leads to an increase in the asset’s value.

23
References


7 Appendix

7.1 Identification Issues

Regarding the identification strategy two issues have to be addressed to retrieve the price discovery measures. The first one refers to the implementation of the spectral decomposition, instead of the ones previously adopted in the literature: LDL and Cholesky. The second issue relates to the computation of matrix G, since I allow for specifications containing up to three common factors.

To show that the spectral decomposition can replace either the LDL or the Cholesky decompositions, I need to show that it carries the same number of restrictions as both alternative decompositions. To this purpose, I consider two different scenarios: \( \text{Var}(\eta_t) = I \) and \( \text{Var}(\eta_t) \neq I \). The reason why I consider these two situations is because in the empirical exercises I do allow for the \( \text{Var}(\eta_t) \) to be different than an identity matrix. As I am dealing with high frequency data, the variance of the reduced form error terms, \( \Omega \), is very small. Hence, if \( \text{Var}(\eta_t) \) is set to be equal to an identity matrix, the resultant parameters in the \( G \) matrix would be unrealistic small.

Let us start with the simpler case: \( \text{Var}(\eta_t) = I \). The matrix \( \Xi = S \text{Var}(\eta_t) S' \) has \( (K^2 - K)/2 + K \) equations and \( K^2 + (K^2 - K)/2 + K \) unknown variables. Hence, in order to completely identify the model, I need to add further \( K^2 \) restrictions. These identifies \( S \) and \( \text{Var}(\eta_t) \). The first set of restrictions comes from the assumption governing the variance of permanent and transitory errors. I assume these innovations have unit variance, which adds \( K \) restrictions to my model. The second set of restrictions arises from the use of the structural framework, where \( \eta_t = (\eta_t^e, \eta_t^w, \eta_t^v, \eta_T^T)' \) is the vector with all innovations on their structural form. This implies that \( \text{Var}(\eta_t) \) is a diagonal matrix, i.e, permanent and transitory shocks are uncorrelated. This adds \( (K^2 - K)/2 \) restrictions. Finally, when the spectral decomposition is applied to a symmetric matrix, it decomposes symmetric matrices. Hence, as \( \Xi \) is a symmetric matrix, likewise will be \( S \), adding the final \( (K^2 - K)/2 \) restrictions needed.

The second scenario addresses the case which \( \text{Var}(\eta_t) \neq I \). To show that the spectral decomposition also holds in this case, I need to change the first and third set of identification restrictions from the previous scenario. I replace the first set of restrictions by imposing \( \text{Var}(\eta_t) = \Theta \). This yields the same \( K \) restrictions I consider in the previous example. The third set of conditions arises by applying the spectral decomposition to a nonsymmetric matrix \( \bar{\Xi} \), implying that the resulting decomposed matrix is no longer a symmetric matrix, adding additional \( (K^2 - K)/2 \) equations which completely identify the model. I show in Section 4 that if I am able to prove that \( \bar{S}S\Theta = \bar{S}\Theta\bar{S}' \) holds, I can
To recover $\eta$, I show the proof for a $2 \times 2$ matrix. Define $\Xi$ as:

$$\Xi = \begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

(29)

Define $\Theta$ as a diagonal matrix containing the vector $\theta = (a, c)'$ on its diagonal. Hence, I compute $\tilde{\Xi}$ as

$$\tilde{\Xi} = \Xi \Theta^{-1} = \begin{pmatrix} 1 & b \\ b & c \end{pmatrix}.\quad (30)$$

Define $V$ as the matrix containing the eigenvectors associated with $\tilde{\Xi}$ and $\Lambda$ the diagonal matrix with the eigenvalues of $\tilde{\Xi}$ on its diagonal, such that:

$$V = \begin{pmatrix} -\frac{\sqrt{a}}{\sqrt{ac}} & \frac{\sqrt{a}}{\sqrt{ac}} \\ 1 & 1 \end{pmatrix}\quad (31)$$

$$\Lambda = \begin{pmatrix} \frac{\sqrt{ac} - b}{\sqrt{ac}} & 0 \\ 0 & \frac{b + \sqrt{ac}}{\sqrt{ac}} \end{pmatrix}\quad (32)$$

By applying the spectral decomposition, I have that $\tilde{\Xi} = \tilde{S}S$ and $\Xi = \tilde{S}S\Theta$, with $\tilde{S}$ given by:

$$\tilde{S} = VA^{1/2}V^{-1} = \begin{pmatrix} \frac{1}{2} \left[ (\frac{\sqrt{ac} - b}{\sqrt{ac}})^{1/2} + \frac{b + \sqrt{ac}}{\sqrt{ac}} \right] \frac{\sqrt{a}}{\sqrt{ac}} \left[ (\frac{b + \sqrt{ac}}{\sqrt{ac}})^{1/2} - (\frac{\sqrt{ac} - b}{\sqrt{ac}})^{1/2} \right] \\ \frac{\sqrt{a}}{\sqrt{ac}} \left[ (\frac{b + \sqrt{ac}}{\sqrt{ac}})^{1/2} - (\frac{\sqrt{ac} - b}{\sqrt{ac}})^{1/2} \right] \frac{1}{2} \left[ (\frac{b + \sqrt{ac}}{\sqrt{ac}})^{1/2} + (\frac{\sqrt{ac} - b}{\sqrt{ac}})^{1/2} \right] \end{pmatrix} \quad (33)$$

By computing $\tilde{S}S\Theta$ and $\tilde{S}\Theta\tilde{S}'$, as in (34), I show that these two quantities are equal to each other, proving that the normalization holds $^3$.

$$\tilde{S}S\Theta \approx \tilde{S}\Theta\tilde{S}' = \begin{pmatrix} a \frac{\sqrt{ac}}{2} \left[ (1 - \frac{b}{\sqrt{ac}}) + \left( \frac{b}{\sqrt{ac}} + 1 \right) \right] \sqrt{\frac{ac}{2}} \left[ \left( \frac{b}{\sqrt{ac}} + 1 \right) - \left( 1 - \frac{b}{\sqrt{ac}} \right) \right] \quad (34) \end{pmatrix}$$

I solve the second identification issue by showing that I can use some of the rows of $\Psi(1)$ in the place of $\alpha'_\perp$, following Yan and Zivot (2007). Using the Johansen’s Factorization as in Johansen (1991), the matrix $\Psi(1)$ can be decomposed as:

$$\Psi(1) = \beta'_\perp (\xi(1)\beta'_\perp)^{-1}\alpha'_\perp = \Gamma\alpha'_\perp\quad (35)$$

If I multiply both sides by the error term obtained from the reduced form VEC model, I

$^3$ A numerical exercise showing that (34) holds for matrix with dimensions greater than two is available upon request.
have
\[ \Psi(1)\epsilon_t = \Gamma \alpha'_{\perp} \epsilon_t \]  

(36)

Matrix $G$ is built in such a way that the right-hand side of (36) contains the portion of $\epsilon_t$ related to permanent innovations, $\epsilon^p_t = (\epsilon^p_t, \epsilon^m_t, \epsilon^v_t)'$, since $\epsilon_t$ is multiplied by the upper part of matrix $G$.

\[ \Psi(1)\epsilon_t = \Gamma \epsilon^p_t \]  

(37)

From (37), the long-run impact of changes in $\epsilon^p_t$ on the market prices is given by $\Gamma$. Considering the a model that accounts for exchange rate, preferred and common shares traded at both domestic and foreign markets as the one discussed in Section 3, I need to add assumptions that allow us to identify $\alpha'_{\perp}$ using the rows of $\Psi(1)$. This is a modification of the original identification strategy proposed by Gonzalo and Granger (1995), where they assume that permanent innovations present a long-run effect different than zero, whereas transitory shocks vanish away in the long-run. Assuming a simpler model than mine, Kim (2010a) and Yan and Zivot (2010) impose long-run restrictions on the permanent innovations in their reduced form to justify the use of common rows in $\Psi(1)$ to identify $\alpha'_{\perp}$. My identification strategy follows along these lines, but I recover matrix $\Gamma$ in (36) using the parameters that drive the common factors dynamics. This covers the case where $\Psi(1)$ does not have clear common rows. To this purpose, from (1), (2) and (3), I construct a matrix $\Phi$ such that:

\[ \Phi = \begin{pmatrix} 1 & \lambda & 0 \\ \rho & 1 & \pi \\ 0 & \kappa & 1 \end{pmatrix} \]  

(38)

Equation (11) gives the long-run dynamics of prices as function of the permanent and transitory innovations on their structured form. Define $D_p(1)$ as the sub-matrix containing all rows and the first three columns of the $D(1)$ matrix. Hence, I want to impose restrictions on $\Gamma$ in the right-hand side of (37) such that the long-run dynamics depicted in (11) holds. To this purpose, it is sufficient to find $\Gamma$ that makes $D_p(1) \eta_t = \Gamma \epsilon^p_t$, provided that $\epsilon^p_t = \Phi \eta_t$ holds.

\[ \Gamma \epsilon^p_t = D_p(1) \eta_t \]
\[ \Gamma \Phi \eta_t = D_p(1) \eta_t \]
\[ \Gamma \Phi = D_p(1) \]
\[ \Gamma = D_p(1) \Phi^{-1} \]  

(39)
Hence, the left-hand side of (39) resumes to:

\[
\Gamma = \begin{pmatrix}
1 & \lambda & 0 \\
\rho & 1 & \pi \\
\rho & \kappa + 1 & \pi + 1 \\
\rho - 1 & 1 - \lambda & \pi \\
\rho - 1 & \kappa - \lambda + 1 & \pi + 1
\end{pmatrix}
\]

Combining (37) with (40), I have that the first two rows of \(\Psi(1)\) can be used in place of the first two rows of \(\alpha'_\perp\) and the third minus the second row of \(\Psi(1)\) as the third row of \(\alpha'_\perp\). These imply that the reduced form innovations associated with the efficient price do not have a permanent impact on the exchange rate in the long-run. This is not harmful in my analysis for two reasons: first, there is no reason to imagine that an innovation on the efficient price should have an effect on exchange rate, apart from correlation aspects governing the structural innovations associated to the common factors. Second, if there are such correlation among the structural errors, this still can be captured by the model, since the restriction is on \(\varepsilon_{\text{m}}^\pi\) and not on \(\eta_{\text{m}}^\pi\). Moreover, \(\Gamma\) imposes that changes in the reduced innovation associated with the exchange rate affects only the foreign market. Again, this is not harmful, since both restrictions are constructed in terms of \(\varepsilon^\pi\).

### 7.2 Simulations

This section illustrates my proposed estimation methodology by comparing it with the alternative frameworks available in the literature. I focus my analysis on computing the instantaneous and long-run measures for price discovery. As pointed out in Section 4, I propose two changes in the methodology. The first one refers to the computation of matrix \(G\), using \(\alpha'\Omega^{-1}\) instead of \(\beta'\). The second one adopts the spectral decomposition rather than the LDL or Cholesky decompositions.

The model I use here is a simplified version of the one presented in Section 3. I work with two common factors, but the extension to the case with more common factors is straightforward. I also assume that the parameters \(\lambda, \rho, \pi, \kappa\) are all equal to zero, which simplifies my results. Given these restrictions, the elements of \(D_0\) are the parameters giving the partial adjustment between efficient and observed prices. I additionally assume that the efficient exchange rate is an observed process. Therefore, the
data generation process is given by:

\[ e_t = e_{t-1} + \eta^e_t \]
\[ m_t = m_{t-1} + \eta^m_t \]  

where \( e_t \) is the efficient exchange rate and \( m_t \) is the asset efficient price. The structural innovations \( \eta^e_t \) and \( \eta^m_t \) are random normal processes generated with a diagonal covariance matrix. The transitory innovations \( \eta^T_t \) and \( \eta^\tau_t \) are also normally distributed. The observed prices are given by:

\[ \Delta y_{2,t} = \gamma_2 (m_t - y_{2,t-1}) + b_2 \eta^T_t \]
\[ \Delta y_{3,t} = \gamma_3 (m_t - y_{3,t-1}) + b_3 \eta^T_t \]
\[ \Delta y^*_{4,t} = \gamma_4 (m_t - y^*_4,t-1) - \dot{\gamma}_4 (e_t - e_{t-1}) + b_4 \eta^\tau_t \]
\[ \Delta y^*_{5,t} = \gamma_5 (m_t - y^*_5,t-1) - \dot{\gamma}_5 (e_t - e_{t-1}) + b_5 \eta^\tau_t \]

where \( y_{2,t} \) and \( y_{3,t} \) are transactions prices observed in the domestic market, whereas and \( y^*_{4,t} \) and \( y^*_{5,t} \) are prices observed in the foreign market expressed in foreign currency. The \( 1 \times 3 \) vector \( b_i \) has the parameters accompanying the transitory innovations.

\[
D_0 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & \gamma_2 & b_{1i} & b_{2i} & b_{3i} \\
0 & \gamma_3 & b_{1i} & b_{2i} & b_{3i} \\
\dot{\gamma}_4 & \gamma_2 & b_{1i} & b_{2i} & b_{3i} \\
\dot{\gamma}_5 & \gamma_5 & b_{1i} & b_{2i} & b_{3i}
\end{pmatrix}
\]

(43)

Table 9 reports results based on the four different comparisons. Firstly, I want to measure the benefit of computing \( D_0 \) using the matrix \( G \) constructed with \( \alpha' \Omega^{-1} \). Therefore, I compare \( \tilde{D}_0 \) versus \( \hat{D}_0 \), where \( \hat{D}_0 \) stands for \( D_0 \) computed using \( \alpha' \Omega^{-1} \) and decomposed with LDL, whereas \( \tilde{D}_0 \) stands for \( D_0 \) calculated with the matrix \( G \) computed using \( \beta' \) and LDL decomposition. The second comparison assesses the benefit of using only the spectral decomposition. Hence, I compute two estimates of \( D_0 \): the first one uses the \( \alpha' \Omega^{-1} \) expression in \( G \) and the spectral decomposition (denoted as \( \hat{D}_0 \)), whereas the second measure uses \( \alpha' \Omega^{-1} \) and the LDL decomposition (denoted as \( \tilde{D}_0 \)). The third comparison addresses the benefits of combining my two methodological suggestions. I compute \( D_0 \) using both the \( \alpha' \Omega^{-1} \) expression and the spectral decomposition (denoted as \( \hat{D}_0 \)) and I denote \( \hat{D}_0 \) as the estimates computed using with \( \beta' \) and the LDL decomposition. Finally, I also want to compare \( \hat{D}_0 \) with the methodology suggested by Gonzalo and Ng (2001). I denote it as \( \overline{D}_0 \) and I compute it using \( \beta' \) and the Cholesky decomposition.

I report results in terms of the mean, relative mean squared errors (RelMSE) and
relative root mean squared error (RelRMSE). I display the ratio of the $D_0$ measures to indicate the way the relative measures are computed. For instance, $\hat{D}_0/D_0$ implies that the relative measures are computing having $D_0$ in the denominator and $\hat{D}_0$ in the numerator. Thus, relative measures smaller than one indicates that the $\hat{D}_0$ outperforms $D_0$.

The results show that $\tilde{D}_0$ is biased for systems with more than one cointegrating vector (I did compute $\tilde{D}_0$ for a smaller system with only one cointegrating vector and the biased is eliminated). By inserting $\alpha'\Omega^{-1}$ I am able to eliminate all the bias, and I could even continue to use LDL decomposition, as I see on the results considering $\hat{D}_0$. Hence, $\hat{D}_0$, $\bar{D}_0$ and $\tilde{D}_0$ are not biased. By analyzing the relative measures, I show that $\hat{D}_0$ presents massive gains when compared to the $\tilde{D}_0$ measures. Similar results are obtained when $\hat{D}_0$ is compared $\bar{D}_0$, indicating the by using $\alpha'\Omega^{-1}$ instead of $\beta'$, I am able to improve considerably my price discovery estimates. In summary, my proposed measure outperforms all competitors.
### Table 1: Data Cleaning Details

<table>
<thead>
<tr>
<th>Stock</th>
<th>Raw obs (M)</th>
<th>outliers (K)</th>
<th>final obs (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRLUSD</td>
<td>4.088</td>
<td>0.600</td>
<td>4.087</td>
</tr>
<tr>
<td>Brazil</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pref</td>
<td>PETR4</td>
<td>9.071</td>
<td>7.353</td>
</tr>
<tr>
<td>Com</td>
<td>PETR3</td>
<td>2.109</td>
<td>4.812</td>
</tr>
<tr>
<td>Pref</td>
<td>VALE5</td>
<td>6.385</td>
<td>5.236</td>
</tr>
<tr>
<td>Com</td>
<td>VALE3</td>
<td>2.066</td>
<td>8.139</td>
</tr>
<tr>
<td>Pref</td>
<td>AMBV4</td>
<td>0.720</td>
<td>4.109</td>
</tr>
<tr>
<td>Pref</td>
<td>BRTO4</td>
<td>0.555</td>
<td>1.564</td>
</tr>
<tr>
<td>Pref</td>
<td>GGBR4</td>
<td>3.237</td>
<td>3.000</td>
</tr>
<tr>
<td>Pref</td>
<td>BBDC4</td>
<td>2.958</td>
<td>3.909</td>
</tr>
<tr>
<td>Nyse</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pref</td>
<td>PBRa.N</td>
<td>5.021</td>
<td>4.485</td>
</tr>
<tr>
<td>Com</td>
<td>PBR.N</td>
<td>7.914</td>
<td>3.318</td>
</tr>
<tr>
<td>Pref</td>
<td>RIOp.N</td>
<td>3.577</td>
<td>1.823</td>
</tr>
<tr>
<td>Com</td>
<td>RIO.N</td>
<td>6.930</td>
<td>1.159</td>
</tr>
<tr>
<td>Pref</td>
<td>ABV_N</td>
<td>1.119</td>
<td>1.645</td>
</tr>
<tr>
<td>Pref</td>
<td>BTM_N</td>
<td>0.200</td>
<td>0.521</td>
</tr>
<tr>
<td>Pref</td>
<td>GGB_N</td>
<td>2.829</td>
<td>0.723</td>
</tr>
<tr>
<td>Pref</td>
<td>BBD_N</td>
<td>3.599</td>
<td>0.959</td>
</tr>
<tr>
<td>Arca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pref</td>
<td>PBRa.P</td>
<td>4.873</td>
<td>2.501</td>
</tr>
<tr>
<td>Com</td>
<td>PBR.P</td>
<td>11.821</td>
<td>3.460</td>
</tr>
<tr>
<td>Pref</td>
<td>ABV_P</td>
<td>0.506</td>
<td>1.190</td>
</tr>
<tr>
<td>Pref</td>
<td>BTM_P</td>
<td>0.106</td>
<td>0.391</td>
</tr>
<tr>
<td>Pref</td>
<td>GGB_P</td>
<td>4.164</td>
<td>0.871</td>
</tr>
<tr>
<td>Pref</td>
<td>BBD_P</td>
<td>6.091</td>
<td>1.038</td>
</tr>
<tr>
<td>Total</td>
<td>89.939</td>
<td>62.806</td>
<td>89.876</td>
</tr>
</tbody>
</table>

Number of observations (transaction prices) as raw observations in million, number of outliers in thousands; and final number of observations in million. PETR and PBR stands for Petrobras shares, VALE and RIO for Vale shares, AMBV and ABV for Ambev, BRTO and BTM for Brasil Telecom, GGBR and GGB for Gerdau and BBDC and BBD for Bradesco. I identify the outliers using the filter proposed by Brownlees and Gallo (2006).
### Table 2: Data Aggregation

<table>
<thead>
<tr>
<th>Freq.</th>
<th>Stock</th>
<th>Initial obs (M)</th>
<th>Missing obs (k)</th>
<th>Agg. Obs (k)</th>
<th>% Missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>30&quot;</td>
<td>PETR4</td>
<td>7.79</td>
<td>4.53</td>
<td>352.68</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>PETR3</td>
<td>1.94</td>
<td>100.07</td>
<td>352.68</td>
<td>28%</td>
</tr>
<tr>
<td></td>
<td>PBRaP</td>
<td>4.63</td>
<td>44.20</td>
<td>352.68</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>PBRp</td>
<td>11.19</td>
<td>9.95</td>
<td>352.68</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>PBRaN</td>
<td>4.80</td>
<td>16.66</td>
<td>352.68</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>PBRaN</td>
<td>7.42</td>
<td>4.86</td>
<td>352.68</td>
<td>1%</td>
</tr>
<tr>
<td></td>
<td>BRLUSD</td>
<td>2.84</td>
<td>34.47</td>
<td>352.35</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>VALE5</td>
<td>6.38</td>
<td>6.28</td>
<td>352.35</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>VALE3</td>
<td>2.06</td>
<td>88.46</td>
<td>352.35</td>
<td>25%</td>
</tr>
<tr>
<td></td>
<td>RIO5</td>
<td>6.92</td>
<td>8.76</td>
<td>352.35</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>RIOpN</td>
<td>3.57</td>
<td>36.89</td>
<td>352.35</td>
<td>13%</td>
</tr>
<tr>
<td></td>
<td>BBDC4</td>
<td>2.60</td>
<td>123.32</td>
<td>352.35</td>
<td>35%</td>
</tr>
<tr>
<td></td>
<td>BBpN</td>
<td>3.21</td>
<td>29.21</td>
<td>352.35</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>GGBR4</td>
<td>5.53</td>
<td>68.91</td>
<td>352.35</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td>GGBP</td>
<td>2.78</td>
<td>38.5</td>
<td>352.35</td>
<td>11%</td>
</tr>
<tr>
<td></td>
<td>BRLUSD</td>
<td>2.84</td>
<td>5.72</td>
<td>117.49</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>AMBV4</td>
<td>0.65</td>
<td>21.19</td>
<td>117.49</td>
<td>18%</td>
</tr>
<tr>
<td></td>
<td>ABVN</td>
<td>1.01</td>
<td>10.41</td>
<td>117.49</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>ABVP</td>
<td>0.46</td>
<td>31.39</td>
<td>117.49</td>
<td>27%</td>
</tr>
<tr>
<td>90&quot;</td>
<td>BRLUSD</td>
<td>2.84</td>
<td>0.89</td>
<td>35.23</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>BRTO4</td>
<td>0.49</td>
<td>1.63</td>
<td>35.23</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>BMN</td>
<td>0.18</td>
<td>5.97</td>
<td>35.23</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>BPM</td>
<td>0.10</td>
<td>12.70</td>
<td>35.23</td>
<td>36%</td>
</tr>
<tr>
<td>300&quot;</td>
<td>BRLUSD</td>
<td>2.84</td>
<td>0.09</td>
<td>35.23</td>
<td>3%</td>
</tr>
<tr>
<td></td>
<td>BRTO4</td>
<td>0.49</td>
<td>1.63</td>
<td>35.23</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>BTMn</td>
<td>0.18</td>
<td>5.97</td>
<td>35.23</td>
<td>17%</td>
</tr>
<tr>
<td></td>
<td>BTMp</td>
<td>0.10</td>
<td>12.70</td>
<td>35.23</td>
<td>36%</td>
</tr>
</tbody>
</table>

Number of initial observations in million, number of missing observations and aggregated observations in thousands. PETR and PBR stands for Petrobras shares (3: common, 4 and a: preferred, N:NYSE, P: ARCA), VALE and RIO for Vale shares (3: common, 5 and p: preferred, N:NYSE), AMBV and ABV for Ambev (all preferred, N:NYSE and P:ARCA), BRTO and BTM for Brasil Telecom (all preferred, N:NYSE and P:ARCA), GGBR and GGBP for Gerdau (all preferred, N:NYSE and P:ARCA), BBDC and BBD for Bradesco (all preferred, N:NYSE and P:ARCA) and BRLUSD for the Brazilian Reais/US dollar exchange rate. I aggregate the data into time tuples using the methodology proposed by Harris, McInish, L. Shoesmith, and Wood (1995).

### Table 3: Price Discovery BR Telecom

<table>
<thead>
<tr>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>Fastness</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRLUSD</td>
<td>$\eta_i^e$</td>
<td>$\eta_i^m$</td>
</tr>
<tr>
<td>BR</td>
<td>0.17</td>
<td>-0.06</td>
</tr>
<tr>
<td>BTMn</td>
<td>-0.88</td>
<td>1.00</td>
</tr>
<tr>
<td>ARCA</td>
<td>-0.83</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Exchange rate (ExRate) is expressed in R$ per US dollars. Permanent shocks are denoted as $\eta_i^e$ and $\eta_i^m$, where the former one is related to the efficient exchange rate, whereas the latter one refers to the efficient price of the underlying security. I determine lag length in the VEC model through Schwarz criteria. BR Telecom prices are sampled at 240 and 300 seconds frequency ($T = 44, 103$ and $T = 35, 229$). The bootstrap standard errors are in the parenthesis.
Exchange rate (ExRate) is expressed in R$ per US dollars. Permanent shocks are denoted as $\eta^e$ and $\eta^m$, where the former one is related to the efficient exchange rate, whereas the latter one refers to the efficient price of the underlying security. I determine lag length in the VEC model through Schwarz criteria. Gerdau prices are sampled at 30 seconds frequency and 60 seconds frequency (T = 352, 183). The bootstrap standard errors are in the parenthesis.

### Table 4: Price Discovery Bradesco

<table>
<thead>
<tr>
<th></th>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>FastnessSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta^e$</td>
<td>$\eta^m$</td>
<td>$\eta^e$</td>
</tr>
<tr>
<td>BRLUSD</td>
<td>1.30</td>
<td>0.01</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>(0.019)</td>
<td>(0.009)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>BR</td>
<td>0.31</td>
<td>0.78</td>
<td>−0.47</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.028)</td>
<td>(0.024)</td>
</tr>
<tr>
<td>NYSE</td>
<td>−1.10</td>
<td>0.81</td>
<td>−1.45</td>
</tr>
<tr>
<td></td>
<td>(0.061)</td>
<td>(0.025)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>ARCA</td>
<td>−1.29</td>
<td>0.97</td>
<td>−1.45</td>
</tr>
<tr>
<td></td>
<td>(0.051)</td>
<td>(0.02)</td>
<td>(0.022)</td>
</tr>
</tbody>
</table>

Exchange rate (ExRate) is expressed in R$ per US dollars. Permanent shocks are denoted as $\eta^e$ and $\eta^m$, where the former one is related to the efficient exchange rate, whereas the latter one refers to the efficient price of the underlying security. I determine lag length in the VEC model through Schwarz criteria. Bradesco prices are sampled at 30 seconds frequency (T = 352, 183). The bootstrap standard errors are in the parenthesis.

### Table 5: Price Discovery Gerdau

<table>
<thead>
<tr>
<th></th>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>Fastness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta^e$</td>
<td>$\eta^m$</td>
<td>$\eta^e$</td>
</tr>
<tr>
<td>BRLUSD</td>
<td>1.34</td>
<td>0.00</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.007)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>BR</td>
<td>0.19</td>
<td>0.84</td>
<td>−0.62</td>
</tr>
<tr>
<td></td>
<td>(0.047)</td>
<td>(0.029)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>NYSE</td>
<td>−0.82</td>
<td>0.74</td>
<td>−1.58</td>
</tr>
<tr>
<td></td>
<td>(0.094)</td>
<td>(0.033)</td>
<td>(0.025)</td>
</tr>
<tr>
<td>ARCA</td>
<td>−1.34</td>
<td>0.90</td>
<td>−1.58</td>
</tr>
<tr>
<td></td>
<td>(0.087)</td>
<td>(0.029)</td>
<td>(0.025)</td>
</tr>
</tbody>
</table>

Exchange rate (ExRate) is expressed in R$ per US dollars. Permanent shocks are denoted as $\eta^e$ and $\eta^m$, where the former one is related to the efficient exchange rate, whereas the latter one refers to the efficient price of the underlying security. I determine lag length in the VEC model through Schwarz criteria. Gerdau prices are sampled at 30 and 60 seconds frequency (T = 352, 159 and T = 176, 313). The bootstrap standard errors are in the parenthesis.

### Table 6: Price Discovery Ambev

<table>
<thead>
<tr>
<th></th>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>FastnessSQ</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\eta^e$</td>
<td>$\eta^m$</td>
<td>$\eta^e$</td>
</tr>
<tr>
<td>BRLUSD</td>
<td>1.14</td>
<td>0.03</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.011)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>BR</td>
<td>0.17</td>
<td>0.95</td>
<td>−0.26</td>
</tr>
<tr>
<td></td>
<td>(0.034)</td>
<td>(0.032)</td>
<td>(0.022)</td>
</tr>
<tr>
<td>NYSE</td>
<td>−0.76</td>
<td>0.85</td>
<td>−1.25</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.036)</td>
<td>(0.02)</td>
</tr>
<tr>
<td>ARCA</td>
<td>−0.91</td>
<td>0.96</td>
<td>−1.25</td>
</tr>
<tr>
<td></td>
<td>(0.072)</td>
<td>(0.033)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

Exchange rate (ExRate) is expressed in R$ per US dollars. Permanent shocks are denoted as $\eta^e$ and $\eta^m$, where the former one is related to the efficient exchange rate, whereas the latter one refers to the efficient price of the underlying security. I determine lag length in the VEC model through Schwarz criteria. Ambev prices are sampled at 90 seconds frequency (T = 117, 087). The bootstrap standard errors are in the parenthesis.
Table 7: Price Discovery Petrobras

<table>
<thead>
<tr>
<th></th>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>Fastness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \eta^*_t )</td>
<td>( \eta^*_\infty )</td>
<td>( \eta^*_t )</td>
</tr>
<tr>
<td><strong>BRLUSD</strong></td>
<td>1.22</td>
<td>0.05</td>
<td>0.20</td>
</tr>
<tr>
<td><strong>PETRA</strong></td>
<td>-0.18</td>
<td>0.86</td>
<td>-0.41</td>
</tr>
<tr>
<td><strong>PETR3</strong></td>
<td>0.15</td>
<td>0.46</td>
<td>1.45</td>
</tr>
<tr>
<td><strong>NYSE</strong> PBRaN</td>
<td>-1.14</td>
<td>0.90</td>
<td>-0.65</td>
</tr>
<tr>
<td><strong>NYSE</strong> PBRaN</td>
<td>-1.23</td>
<td>0.95</td>
<td>1.92</td>
</tr>
<tr>
<td><strong>ARCA</strong> PBRaP</td>
<td>-1.46</td>
<td>1.00</td>
<td>-0.49</td>
</tr>
<tr>
<td><strong>ARCA</strong> PBRaP</td>
<td>-1.48</td>
<td>1.04</td>
<td>1.79</td>
</tr>
</tbody>
</table>

Table 8: Price Discovery Vale

<table>
<thead>
<tr>
<th></th>
<th>Inst. Effect</th>
<th>Long-Run</th>
<th>Fastness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \eta^*_t )</td>
<td>( \eta^*_\infty )</td>
<td>( \eta^*_t )</td>
</tr>
<tr>
<td><strong>BRLUSD</strong></td>
<td>1.16</td>
<td>-0.05</td>
<td>0.03</td>
</tr>
<tr>
<td><strong>VALE5</strong></td>
<td>-0.47</td>
<td>0.99</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>VALE3</strong></td>
<td>-0.53</td>
<td>0.93</td>
<td>2.25</td>
</tr>
<tr>
<td><strong>NYSE</strong> RIOaN</td>
<td>-1.58</td>
<td>1.12</td>
<td>0.17</td>
</tr>
<tr>
<td><strong>NYSE</strong> RIOaN</td>
<td>-1.61</td>
<td>1.14</td>
<td>2.31</td>
</tr>
</tbody>
</table>

Table 9: Monte Carlo Simulations

<table>
<thead>
<tr>
<th>True value</th>
<th>Mean</th>
<th>RMSE</th>
<th>RRMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_0 )</td>
<td>( \alpha_{LDL} )</td>
<td>( \beta_{LDL} )</td>
<td>( \alpha_S )</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>1.0</td>
<td>0.99</td>
<td>0.99</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.0</td>
<td>0.46</td>
<td>0.00</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.0</td>
<td>0.30</td>
<td>0.00</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.2</td>
<td>1.15</td>
<td>0.20</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.5</td>
<td>0.95</td>
<td>0.50</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.0</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.8</td>
<td>0.87</td>
<td>0.79</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.4</td>
<td>0.53</td>
<td>0.40</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.2</td>
<td>0.45</td>
<td>0.20</td>
</tr>
<tr>
<td>( d_{ij} )</td>
<td>0.5</td>
<td>0.54</td>
<td>0.50</td>
</tr>
</tbody>
</table>

Results are expressed in terms of Relative Mean Squared Error (RMSE) and Relative Root Mean Squared Error (RRMSE). Sample size and replication number are fixed at 10,000 and 1,000 respectively. The variable \( d_{ij} \) denotes the \( ij \)th element of the \( D_0 \) matrix.
Figure 1: Price Evolution

Displays price evolution of Ambev, Gerdau, Bradesco, BR Telecom, Petrobras and Vale stocks traded at BOVESPA, NYSE and ARCA. Sampling frequency are fixes as follows: Ambev, 90"; Gerdau, 60"; Bradesco, 30"; BR Telecom, 300"; Petrobras 30"; and Vale 30". Prices are aggregated following Harris, McNish, L.Shoesmith, and Wood (1995) and free of non plausible values.
Figure 2: Bradesco

Displays rolling window estimates of the short and long-run structural polynomials, $D(0)$ and $D(1)$ respectively. Sample size and shift window size are fixed equal to 30,547 and 7,637 observations respectively, resulting in 43 windows. Prices are aggregated at 30 seconds. BOVESPA accounts for preferred shares traded at BOVESPA (Brazil), NYSE for the ADR’s on preferred shares traded at NYSE and ARCA for preferred shares traded at ARCA. Exchange rate (ExRate) is expressed in R$ per US dollars. 95% confidence intervals are obtained using the bootstrap empirical distribution.
Figure 3: Gerdau

Displays rolling window estimates of the short and long-run structural polynomials, \( D(0) \) and \( D(1) \) respectively. Sample size and window size are fixed equal to 30,547 and 7,637 observations respectively, resulting in 42 windows. Prices are aggregated at 30 seconds. BOVESPA\( _p \) accounts for preferred shares traded at BOVESPA (Brazil), NYSE\( _p \) for the ADR’s on preferred shares traded at NYSE and ARCA\( _p \) for preferred shares traded at ARCA. Exchange rate (ExRate) is expressed in R\$ per US dollars. 95% confidence intervals are obtained using the bootstrap empirical distribution.
Figure 4: Ambev

Displays rolling window estimates of the short and long-run structural polynomials, $D(0)$ and $D(1)$ respectively. Sample size and window size are fixed equal to 10,000 and 2,000 observations respectively, resulting in 54 windows. Prices are aggregated at 90 seconds. $BOVESPA_p$ accounts for preferred shares traded at BOVESPA (Brazil), $NYSE_p$ for the ADR’s on preferred shares traded at NYSE and $ARCA_p$ for preferred shares traded at ARCA. Exchange rate (ExRate) is expressed in R$ per US dollars. 95% confidence intervals are obtained using the bootstrap empirical distribution.
Displays rolling window estimates of the short and long-run structural polynomials, $D(0)$ and $D(1)$ respectively. Sample size and window size are fixed equal to 5,000 and 1,000 observations respectively, resulting in 39 windows. Prices are aggregated at 240 seconds. Prices are aggregated at 60 seconds. BOVESPA accounts for preferred shares traded at BOVESPA (Brazil), NYSE for the ADR's on preferred shares traded at NYSE and ARCA for preferred shares traded at ARCA. Exchange rate (ExRate) is expressed in R$ per US dollars. 95% confidence intervals are obtained using the bootstrap empirical distribution.
Displays rolling window estimates of the short and long-run structural polynomials, $D(0)$ and $D(1)$ respectively. Sample size and window size are fixed equal to 30,547 and 7,637 observations respectively, resulting in 43 windows. Prices are aggregated at 43 seconds. Prices are aggregated at 60 seconds. BOVESPA$_p$ and BOVESPA$_c$ accounts for preferred and common shares traded at BOVESPA (Brazil) and NYSE$_p$ and NYSE$_c$ for the ADR’s on preferred and common shares traded at NYSE. Exchange rate (ExRate) is expressed in R$ per US dollar. 95% confidence intervals are obtained using the bootstrap empirical distribution.
Figure 7: Petrobras

Displays rolling window estimates of the short and long-run structural polynomials, $D(0)$ and $D(1)$ respectively. Sample size and window size are fixed equal to 30,547 and 7,637 observations respectively, resulting in 43 windows. Prices are aggregated at 60 seconds. $BOVESPA_p$ and $BOVESPA_c$ accounts for preferred and common shares traded at BOVESPA (Brazil), $NYSE_p$ and $NYSE_c$ for the ADR’s on preferred and common shares traded at NYSE and $ARCA_p$ and $ARCA_c$ for preferred and common traded at ARCA. Exchange rate (ExRate) is expressed in R$ per US dollars. 95% confidence intervals are obtained using the bootstrap empirical distribution.