The price formation of substitute markets: Theory and application to twin-board China firms

Michael T. Chng* and Aihua Xia†

Current draft 28th March, 2006

Abstract

Understanding price formation to measure the price discovery of substitute markets (cross listing, contract proliferation, spot-futures-option) is relevant not only to investors and regulators, but also to the financial exchanges that host such markets. Price formation models vary in the trading parameters that each model highlights to depict the trading process. They include trade size in Hasbrouck (1991), cross-market prices in Hasbrouck (1995), order size in Al-Suhaibani and Kryzanowski (2000), time between trade in Dufour and Engle (2000) and trade and order sizes in Chng (2005). In this paper, we propose a theoretical model that considers the joint trade directions of a pair of substitute markets. Our model extends that in Madhavan, Richardson and Roomans (1997). We apply the model to analyze the price formation of A-B and A-H Chinese twin shares. Results from the A-B group does not indicate evident price leadership by either boards. However, results from the A-H group support our proposition that the H-board provides price leadership for the A-board.

JEL classification: G14, G15.

Key words and phrases: price discovery, trading dynamics, orders, trades, cross-listing.

*Corresponding author's contacts: Department of Accounting and Finance, Monash University, VIC 3800, Australia. Tel: 613-99059799. Email: michael.chng@buseco.monash.edu.au. I retain full property rights to all existing errors. Funding from both my department and the Melbourne Center for Financial Studies are gratefully acknowledged.

†Department of Mathematics and Statistics, University of Melbourne, VIC 3010, Australia. Tel: 613-83444247. Email: A.Xia@ms.unimelb.edu.au.

The price formation of substitute markets: Theory and application to twin-board China firms

Abstract

Understanding price formation to measure the price discovery of substitute markets (cross listing, contract proliferation, spot-futures-option) is relevant not only to investors and regulators, but also to the financial exchanges that host such markets. Price formation models vary in the trading parameters that each model highlights to depict the trading process. They include trade size in Hasbrouck (1991), cross-market prices in Hasbrouck (1995), order size in Al-Suhaibani and Kryzanowski (2000), time between trade in Dufour and Engle (2000) and trade and order sizes in Chng (2005). In this paper, we propose a theoretical model that considers the joint trade directions of a pair of substitute markets. Our model extends that in Madhavan, Richardson and Roomans (1997). We apply the model to analyze the price formation of A-B and A-H Chinese twin shares. Results from the A-B group does not indicate evident price leadership by either boards. However, results from the A-H group support our proposition that the H-board provides price leadership for the A-board.

JEL classification: G14, G15.

Key words and phrases: price discovery, trading dynamics, orders, trades, cross-listing.

1 Introduction

The price of a security moves when investors react simultaneously and similarly to new public information, inducing price adjustments without trading. Trading among asymmetrically informed investors also leads to price change. The trading process itself induces price movement in the presence of an array of microstructure features that govern trading. A better understanding of intraday price formation in a given market necessarily involves the preceding considerations. In doing so, we can begin to gauge a market's price discovery i.e. how private information is revealed through trading. Additionally, if the same security is traded across two markets, a more comprehensive depiction of intraday cross market price formation would necessarily consider the trading process across both markets.

Indeed, a market's trading process can be described by a set of observable trading parameters. This includes price, return, volatility, trade size, number/size of best bid and ask, number/size of trades within a given time interval, time between trades and trade direction i.e. whether a trade is buyer or seller-initiated. Some studies include Hasbrouck (1991) on trade size, Hasbrouck (1995) on cross-market prices, Al-Suhaibani and Kryzanowski (2000) on order size, Dufour and Engle (2000) on time between trade and Chng (2005) on trade and order sizes. For a pair of substitute markets, incorporating the joint trading process of both markets potentially imply twice as many parameters to consider. We regard Market X and Market Y as substitute markets when there exists a clear fundamental link between X and Y. This could refer to cross-listing, contract proliferation[1] or spot-futures-option markets.[2] It could also apply to a firm that is listed on multiple boards on a stock exchange.[3] To a

[1]E.g. The Nikkei 225 index futures markets of Osaka and Singapore.

[2]E.g. Standard and Poors 500 futures and futures option markets of the Chicago Mercantile Exchange.

[3]E.g. On the Shanghai Stock Exchange, other than the main A-Board, firms have the option to list on

lesser extent, it could even refer to pairwise firms whose return generating processes display very similar factor structures.[4] The preceding discussion gives a sense of the potential comprehensiveness of our model's application in examining the price discovery between two markets.

The majority of trading parameters are either magnitude-only measures e.g. price, volatility, number/size of best bid/ask, time between trades, number of trades per fixed time interval etc, or magnitude and direction-based measures e.g. returns and signed trade size. While there is abundant theoretical and empirical justification to consider a given trading parameter in single market price formation e.g. signed trade size in Easley and O'Hara (1985), it is less clear if the same parameter is just as readily applicable in a cross-market setting. For example, the presence of stealth trading may suggest that trade direction is more relevant than trade size. Also, if trading in one market is dominated by uninformed institutional hedging, then size does not necessarily matter. A good example surrounding the Nikkei 225 futures trading between the Singapore Exchange (SGX) and Osaka Securities Exchange (OSE) is discussed in Chng (2004a). Although OSE hosts the lion share of trading volume, SGX performs the majority of price discovery.

We argue that a direct comparison of a magnitude-based parameter between two markets is somewhat awkward. To elaborate, consider a cross-market analysis on the information content of realized volatility i.e. permanent price impact. The results suggest that (say) Market X contributes 60% of price discovery and Market Y 40%. Does that imply that Market Y is necessarily inferior to Market X? Shouldn't Market Y be deemed inferior only if it does not contribute to price discovery at all? Does it mean that Market X is 50%

either the B-Board or the H-Board.

[4]E.g. Google and Yahoo; Dell and IBM, Ford and GM

more efficient than Market Y? Similar questions can be raised on other magnitude-based parameters[5].

Our proposed model is based on the single market Madhavan, Richardson and Roomans (1997), or MRR, model. The latter highlights the one period autocorrelation in trade direction ρX_{t-1} as an explanatory variable for observed price change, where $X_t = \{1, 0, -1\}$ is an indicator variable that is assumed to follow a general Markov process. If a trade is buyer (seller) initiated, $X_t = 1(-1)$. When a trade is transacted at a price straddled by the prevailing bid ask spread, it is indicated with $X_t = 0$. This mixed-signal state of $X_t = 0$ is conceptually trivial in measuring price discovery.

In our model, the focus is on the joint trade direction of Market X and Market Y. We denote each of X_t and Y_t as a dichotomy by removing the third state of zero. Put differently, all trades are assumed clearly classifiable as either buyer or seller initiated. We will elaborate on this point in the next section. In our model, there are two mixed signal states in $(X_t = 1, Y_t = -1)$ and $(X_t = -1, Y_t = 1)$. Unlike the conceptually trivial mixed signal state in the MRR model, the mixed-signal states in our model offer a channel by which to gauge the price discovery between two substitute markets. The emphasis of our model is to measure the likelihood of one market following the other at time t when the two markets generate conflicting trade direction signals at time $t - 1$.

Our theoretical contribution is to formalize a potential extension suggested by Madhavan,

[5]E.g. 1 If Markets X and Y correspondingly handle 20 and 10 trades on average within a given 10 minute interval, or, if a trade occurs once every 45 seconds on average for Market X and for Market Y, once every 90 seconds, does that imply Market X is more, equal or less than twice as liquid as Market Y? E.g. 2 If lag 1 futures returns R^f_{t-1} is significant in the spot return R^s_t process, but not vice versa, it indicates that futures returns leads spot returns change. But if the price data used is observed at five minutes intervals, does the result imply that the spot price reaction is slow, or it is fast but not as fast as the futures? Is the significance of R^f_{t-1} and the non-significance of R^s_{t-1} trivial?

Richardson and Roomans (1997) to compare and contrast price formation across assets and/or trading systems. The one-dimensional MRR model is unlikely to be appropriate if applied directly on a cross-section of stocks/futures contracts for a given exchange, or, a cross-section of trading systems for a given stock/futures contract. What is required is a multi-dimensional structural model that allows for dynamic interactions. In our model, a discrete joint trade direction variable (X_t, Y_t) is introduced into a bivariate structural system of cross-market price formation to gauge the price discovery between pairwise substitute markets. Since trade direction is a dichotomy, it can be clearly observed if Market X conforms or contradicts Market Y. From there, cross-market price discovery can be measured based on the coefficient of $(X_t + Y_t)$

A firm listed on either the Shanghai or Shenzhen stock exchange (SHSE and SZSE) main A-board has an option to list on either the B- or H- board. B-shares are traded mainly by foreign investors in USD on the SHSE and HKD on the SZSE. H-shares are traded in HKD on the Hong Kong Exchange (HKEx), which possesses a significant foreign institutional investor clientele. Either board offers a Chinese firm exposure to the foreign investment community. We apply our extended model to analyze the price formation of A- and B/H boards for a given firm. We also analyze comparable pairwise A-B and A-H firms to examine potential differences in the relative pricing between A-B and A-H firms. If one channel of internationalization provides better price discovery, this has policy implications concerning the dissimilar foreign investor clientele present on the B and H boards.

Our choice of application using Chinese stock market data is motivated by a number of factors. First, as China's economy grows, her financial markets are also gaining importance among institutional investors. Contemporaneously, the Chinese stock market, with its multiple listing boards (A-B-H) on multiple exchanges SHSE, SZSE and HKSE, is generating interest among market microstructure researchers as well. Second, examining cross-listings between exchanges necessarily restricts the number of firms and the extent of simultaneous trading between the two markets. In stark contrast, the Chinese market offers a unique opportunity to examine a large number of twin-share firms listed on the same exchange and/or time zone. Third, we are interested to contrast the relative pricing between A-B and A-H shares, given some degree of substitutability between the B and H boards.

Fourth, we argue that our empirical application alone suffices a reasonable contribution to existing studies on the China stock market. Most prior studies examine either A-B firms or A-H firms, but seldom both. Yeh, Lee and Pen (2002) analyze the role of lagged pricing gap of A-share price over B-share price in affecting subsequent returns. Wang and Jiang (2004) examine factors that influence the A-share and H-share pricing gap. Lee and Rui (2000) and Sun and Tong (2000) are two studies that perform a more comprehensive analysis across all three listing boards, and both papers consider potential substitutability between the B- and H-boards. Lee and Rui (2000) find little evidence of either US or Hong Kong trading affecting trading on the mainland markets. In contrast, Sun and Tong (2000) find that increases in the number of listings on the H-board expands the A-B pricing gap.

While Sun and Tong (2000) examine relative pricing of both A-B and A-H sample groups, the sample period did not cover reforms that suggest a partial merging between the A and B boards. From Feb 2001, mainlanders are allowed to use foreign currency accounts to trade B-shares. From Dec 2002, Qualified Foreign Institutional Investors (QFII) are allowed to trade A-shares. More importantly, the fact that the paper uses monthly prices will subject their results to a non-trivial non-synchronous trading problem. Wang and Jiang (2004) examine a more recent daily database from Jun 1995 to Dec 2001. While their sample covers the Feb 2001 reform, they did not address this issue. That mainlanders are allowed to trade B-shares may seem trivial to Wang and Jiang (2004), whose focus is on A-H relative pricing. But local investors migrating to and from the A-board, which has been widely documented in the early months post reform, suggests that their A-board sample contains a potential structural shift that has not been accounted for. By utilizing minute-by-minute data for the period 4th Jan to 30th Sep 2005, we address the preceding critiques in our empirical application.

The theoretical model is presented in section 2. Empirical application of our model, data and sampling methodology are discussed in section 3. Results are reported and discussed in section 4. Section 5 concludes.

2 Theoretical Model

We develop a cross-market price formation model based on Madhavan, Richardson and Roomans (1997). The MRR model is a single market structural model that incorporates public information shocks and microstructure effects. It highlights trade direction as the trading parameter to explain intraday price movements. Specifically, the model focuses on both the change Δ in trade direction $\{X_t\}$ and the one period autocorrelation ρ in trade direction $\{X_t\}$ as the key parameters to explain intraday price formation. The conceptual focus is on the values $\{+1, -1\}$ that X_{t-1} takes. For completeness, the model also considers a third possible state of $X_{t-1} = 0$. This mixed outcome is indicative that the trade at time $t-1$ has occurred within the prevailing spread. This occurrence causes a break in the

Markov chain such that the expected trade direction at time t is also zero. Conceptually, limited inference can be extracted from a mixed direction. Statistically, limited inference can be drawn from zero. The progression of X_t over time is illustrated in Figure One.

INSERT FIGURE ONE

The extended model examines the price formation of a risky security traded across substitute markets X and Y by considering the joint trade variable (X_t, Y_t) observed at time t. Since each coordinate is a dichotomy taking the value of $+1$ or -1, (X_t, Y_t) has four possible states: $\{(1,1),(1,-1),(-1,1),(-1,-1)\} := \mathcal{S}$. In contrast to the MRR model, the focus of the extended model is on the two prior states where trade directions are mixed i.e. $(X_{t-1}, Y_{t-1}) = (1,-1)$ or $(-1,1)$. This is because modeling (X_t, Y_t) conditional on substitute markets giving mixed signals $X_{t-1} \neq Y_{t-1}$ facilitates the evaluation as to which market exerts more influential on (X_t, Y_t). This forms the basis of a cross-market price discovery measurement approach. This departure from the original model is what drives the motivation of our paper and the extended model's contribution to the existing literature.

For quote driven and/or floor traded markets, there are various reasons why a trade is observed to be transacted within the prevailing spread[6]. The literature offers various approaches for inferring the direction of such trades, including the commonly used quote

[6]One example is on a pre-1997 quote-driven NASDAQ, where direct negotiation between dealer and a client with a large buy order could lead to a trade occurring at a price marginally below the dealer's quoted ask price. For another example, note that the time stamps on floor traded futures tick data refer to the time of recording, not time of occurrence. Consider two locals shouting a best bid of \$10 and best ask of \$10.20. When the bidding local's order is filled by a floor broker, their respective clerks will gather corresponding details for confirmation in the upstairs offices. Once confirmed, the trade is recorded at a price of \$10 with a time delay. While this is happening, another local could be announcing the next best bid at \$9.90. If this new bid is recorded ahead of the latest trade, then when the data is consolidated and sorted according to the time stamps, the following sequence is observed: Spread is 10-10.2; spread widens to 9.90-10.20; a trade occurred within the prevailing spread at \$10. The reason why a given spread could actually widen in a floor environment is because prevailing quotes have very short term to maturity as they can easily be withdrawn or cancelled.

and tick rules outlined in Lee and Ready (1991) and hybrid rules proposed in Ellis, Michaely and O'Hara (2000) and Chng (2004b). To reiterate, the focus of our model is on states of the world where substitute markets produce either consistent or conflicting trade direction signals. It is necessary for the direction of all trades to be clearly identified without relying on appropriate trade-signing algorithms, which is suitably a separate issue[7]. Ideally, the model is applied to examine substitute markets whereby trading is governed by a central limit order book. This is because in the presence of a limit order queue, all trades can easily be classified as either buyer or seller initiated.

The extended model considers three price processes μ_t, p_t^X and p_t^Y. Outlined in equation (1), μ_t is defined as the post-trade expected value of a risky security incorporating all public information, including the joint trade variable $(X_t + Y_t)$. It takes on a random walk process with the residual expressed as the sum of two terms. This is indicative that the update of public beliefs and the adjustment of μ_t could arise from one of two possible sources.

$$\mu_t = \mu_{t-1} + \theta \left(X_t + Y_t - \mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1}) \right) + \epsilon_t. \tag{1}$$

First, non-trade new public information released between $t-1$ and t generates a revision in beliefs. Denote this as ϵ_t, where $\{\epsilon_t, \ t \geq 0\}$ are independent and identically distributed random variables with mean 0 and variance σ_ϵ^2. Second, adjustment in μ_t can also be triggered by the joint trade innovation $(X_t + Y_t) - \mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1})$. When traders are asymmetrically informed, the joint trade innovation represents a surprise in the joint trade direction observed at time t. The revision in belief resulted from that surprise elicits an adjustment in μ_t based on the coefficient θ, which indicates the degree of information asymmetry. Fol-

[7]The restriction is also imposed for simplicity. The one-period transition by two variables, each taking only one of two possible values gives a total of 16 possible joint states. The one-period transition by two variables, each taking one of three possible values raises the total number of possible joint states to 81.

lowing Glosten and Milgrom (1985), assume revision in beliefs is positively correlated with the joint trade innovation, such that $\theta \geq 0$. A larger coefficient θ implies greater information asymmetry since it generates a larger revision in belief for a given joint trade innovation.

$$
\begin{aligned}
p_t^X &= \mu_t + \phi^X X_t + \xi_t^X \\
p_t^Y &= \mu_t + \phi^Y Y_t + \xi_t^Y.
\end{aligned} \tag{2}
$$

A pair of transaction prices p_t^X and p_t^Y observed on Market X and Market Y correspondingly at time t are presented in equation (2)[8]. Since the same security is traded on two different markets, it is assumed that p_t^X and p_t^Y contain the common factor μ_t while displaying idiosyncratic pricing errors ξ_t^X and ξ_t^Y. For Market X, the pricing error $p_t^X - \mu_t$ is the sum of two components. First, the term $\phi^X X_t$ represents the transitory price impact by X_t i.e. a buyer (seller) initiated trade has a positive (negative) effect on p_t^X. The latter incorporates the impact of the current trade direction $\phi^X X_t$ since it is a regret-free price. The fact that a trade at the prevailing ask (bid) is expected to lift (lower) subsequent prices is already impounded into the current price. Second, all microstructure effects, e.g. tick size, price rounding, are manifested in the residual ξ_t^X, where $\{\xi_t^X, t \geq 0\}$ are independent and identically distributed random variables with mean 0 and standard deviation σ_{ξ^X}. The price process p_t^Y is described in a similar fashion.

Define Δ as the backward difference operator, e.g., $\Delta p_t = p_t - p_{t-1}$ and let $r_t = \Delta p_t$, then equations (1) and (2) can be used to derive

$$
\begin{aligned}
r_t^X &= \phi^X \Delta X_t + \theta(X_t + Y_t - \mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1})) + \Delta\xi^X + \epsilon_t, \\
r_t^Y &= \phi^Y \Delta Y_t + \theta(X_t + Y_t - \mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1})) + \Delta\xi^Y + \epsilon_t.
\end{aligned} \tag{3}
$$

[8]This is a general approach for expressing a non-stationary series into a random walk and a residual stationary component. Stock and Watson (1988) provides a comprehensive review

Equation (3) depicts the price change process r_t. But estimating (3) requires additional structure to be specified on the joint trade innovation, in particular, how it behaves over time. Assume the joint trade variable $\{(X_t, Y_t)\}$ follow a Markov process. Based on the current set up, the progression of $\{(X_t, Y_t)\}$ over time can be illustrated in Figure Two with a set of coordinates. At any given time, only four sets of coordinates in $\mathcal{S}$ are possible.

INSERT FIGURE TWO

Since the outcome of the joint trade direction at time t depends only on the outcome exhibited at time $t-1$, the progression of $\{(X_t, Y_t)\}$ over time can be categorized into one of four possible scenarios:

1. Full continuation: $\{X_t = X_{t-1}\} \cap \{Y_t = Y_{t-1}\}$
2. X-continuation: $\{X_t = X_{t-1}\} \cap \{Y_t = -Y_{t-1}\}$
3. Y-continuation: $\{X_t = -X_{t-1}\} \cap \{Y_t = Y_{t-1}\}$
4. Full reversal: $\{X_t = -X_{t-1}\} \cap \{Y_t = -Y_{t-1}\}$

From time $t-1$ to t, full continuation could refer to one of the four cases: $(1,1) \rightarrow (1,1)$, $(-1,1) \rightarrow (-1,1)$, $(1,-1) \rightarrow (1,-1)$ and $(-1,-1) \rightarrow (-1,-1)$. The same can be said for each of the three remaining scenarios. This provides a good perception of the sixteen possible joint states of the world depicting the progression of $\{(X_t, Y_t)\}$ over time. Computing the conditional expectation of $(X_t + Y_t)$ requires the consideration of conditional probabilities since each of the four scenarios relate to how (X_t, Y_t) is transitioned from (X_{t-1}, Y_{t-1}). Denote the probability of full continuation as $\mathbb{P}(X_t = X_{t-1}, Y_t = Y_{t-1}) = \gamma$. Along similar lines, denote the the probability of X-continuation as $\mathbb{P}(X_t = X_{t-1}, Y_t = -Y_{t-1}) = \lambda_X$ and the probability of Y-continuation as $\mathbb{P}(X_t = -X_{t-1}, Y_t = Y_{t-1}) = \lambda_Y$. It follows that the probability of full reversal is $\mathbb{P}(X_t = -X_{t-1}, Y_t = -Y_{t-1}) = 1 - \gamma - \lambda_X - \lambda_Y$.

The 4×4 transition probability matrix P is presented in Equation (4). Note that in the current model specification, P is a doubly stochastic matrix, namely, both row and column values of P sum to 1. This arises from the manner in which the conditional probabilities are defined. An event that occurs at time t is defined in terms of whether it is exactly the same, partially the same or totally difference from an event being observed at time $t-1$. As such, no distinction is made between (say) $\mathbb{P}\left((X_t, Y_t) = (1,1) | (X_{t-1}, Y_{t-1}) = (1,1)\right)$ and $\mathbb{P}\left((X_t, Y_t) = (-1,-1) | (X_{t-1}, Y_{t-1}) = (-1,-1)\right)$.

$$P = \begin{array}{r} \\ (1,1) \\ (1,-1) \\ (-1,1) \\ (-1,-1) \end{array} \begin{array}{c} \begin{array}{cccc} (1,1) & (1,-1) & (-1,1) & (-1,-1) \end{array} \\ \begin{pmatrix} \gamma & \lambda_X & \lambda_Y & 1-\gamma-\lambda_X-\lambda_Y \\ \lambda_X & \gamma & 1-\gamma-\lambda_X-\lambda_Y & \lambda_Y \\ \lambda_Y & 1-\gamma-\lambda_X-\lambda_Y & \gamma & \lambda_X \\ 1-\gamma-\lambda_X-\lambda_Y & \lambda_Y & \lambda_X & \gamma \end{pmatrix} \end{array} \tag{4}$$

The following properties of $\{(X_t, Y_t)\}$ under the steady state regime are proved in the appendix. It is shown that $\mathbb{E}(X_t) = \mathbb{E}(Y_t) = \mathbb{E}(X_t Y_t) = 0$, $\operatorname{Var}(X_t) = \operatorname{Var}(Y_t) = 1$, $\operatorname{Cov}(X_t, X_{t-1}) = 2(\gamma + \lambda_X) - 1 = \rho$ and $\operatorname{Cov}(Y_t, Y_{t-1}) = 2(\gamma + \lambda_Y) - 1 = \delta$. It is also shown that the current setup implies that the stationary distribution of the Markov chain is $\mathbb{P}((X_t, Y_t) = (i,j)) = 0.25$ for $(i,j) \in \mathcal{S}$. The implication is twofold. First, this suggests there is no reason to believe buyer-initiated trades are more likely than seller-initiated trades. Second, if all trades are information motivated, substitute markets should generate consistent signals more often than conflicting signals in the long run, such that

$\mathbb{P}\left((X_t, Y_t) = (1,1) \text{ or } (-1,-1)\right) > \mathbb{P}\left((X_t, Y_t) = (1,-1) \text{ or } (-1,1)\right)$. But since our price formation model allows for noise trading, there is no reason why substitute markets are more likely to give conforming than conflicting signals.

It is shown in the appendix that the coefficient matrix Σ in the linear predictor $\Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}$ for $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ that is set to

$$\min \mathbb{E}\left(\left(\begin{pmatrix} X_t \\ Y_t \end{pmatrix} - \Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}\right)^T \times \left(\begin{pmatrix} X_t \\ Y_t \end{pmatrix} - \Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}\right)\right),$$

has to satisfy: $\Sigma = \begin{pmatrix} \mathrm{Cov}\left(X_t, X_{t-1}\right) & \mathrm{Cov}\left(X_t, Y_{t-1}\right) \\ \mathrm{Cov}\left(Y_t, X_{t-1}\right) & \mathrm{Cov}\left(Y_t, Y_{t-1}\right) \end{pmatrix} = \begin{pmatrix} 2(\gamma + \lambda_X) - 1 & 0 \\ 0 & 2(\gamma + \lambda_Y) - 1 \end{pmatrix}$.

As it turns out, Σ is an auto-cross correlation matrix. Solving for the expected joint trade direction conditional on the joint trade direction variable observed in the prior state between substitute markets gives

$$\mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1}) = \rho(X_{t-1}) + \delta(Y_{t-1}) \tag{5}$$

Equation (5) outlines the conditional expectation of $(X_t + Y_t)$. Let the residuals $\Delta\xi^X + \epsilon_t$ and $\Delta\xi^Y + \epsilon_t$ be denoted as ω_t^X and ω_t^Y respectively, then combining equations (3) and (5) gives

$$\begin{aligned} r_t^X &= (\theta + \phi^X)\Delta X_t + \theta\Delta Y_t + \theta(1-\rho)X_{t-1} + \theta(1-\delta)Y_{t-1} + \omega_t^X, \\ r_t^Y &= (\theta + \phi^Y)\Delta Y_t + \theta\Delta X_t + \theta(1-\rho)X_{t-1} + \theta(1-\delta)Y_{t-1} + \omega_t^Y. \end{aligned} \tag{6}$$

Equation (6) displays a pair of transaction price change equations that can be estimated using transaction data for any pair of substitute markets. Each price formation is described by both change in trade directions ΔX_t and ΔY_t as well as prior trade direction ΔX_{t-1} and

ΔY_{t-1}. Note that if there is no information asymmetry i.e. $\theta = 0$, no market frictions i.e. $\phi^X = \phi^Y = 0$ and no trading frictions i.e. $\xi_t^X = \xi_t^Y = 0$, equation (6) reduces to a traditional efficient price random walk process $r_t^X = r_t^Y = \epsilon_t$

.

3 Empirical Application

Here we demonstrate the use of our model. To reiterate, unlike the MRR model where the mixed state of $X_t = 0$ is of secondary focus, the two mixed scenarios of X-continuation and Y-continuation in our model provide a framework for which to gauge the price discovery between two substitute markets. The extended model's potential empirical application is comprehensive, ranging from cross-listed stocks, competing futures contracts, spot-futures-option markets and different listing boards of a given stock exchange e.g. Stock Exchange of Thailand and the Shanghai Stock Exchange (SHSE).

Model estimation

In our model, there are five trading parameters to describe intraday cross-market price formation. They are θ (overall asymmetric information), ϕ_X and ϕ_Y (trading cost of markets X and Y), and lastly, ρ and δ (trade autocorrelation of markets X and Y). Substituting equation (5), which outlines observable representation of $\mathbb{E}(X_t+Y_t|X_{t-1}, Y_{t-1})$, into equation (3) yields a bivariate price formation system in equation (6). Let $\beta = (\theta, \phi^X, \phi^Y, \rho, \delta)$ denote the vector of model parameters to be derived from the estimation of equation (6). Each of r_t^X and r_t^Y is expressed as a linear sum of contemporaneous and lagged trading variables. By imposing cross equation coefficient restrictions, we estimate equation (6) using generalized method of moments (GMM). This procedure requires weak assumptions about the stochastic data generating process, although it does require a large sample size to induce consistency in estimates.[9]

The GMM approach selects values for β in order to minimize a criterion function based on a set of theoretical moment conditions that β should satisfy. In our model, the theoretical moment conditions are:

$$E\begin{pmatrix} (X_t - \rho X_{t-1})X_{t-1} \\ (Y_t - \delta Y_{t-1})Y_{t-1} \\ \theta(X_t + Y_t - (\rho X_{t-1} + \delta Y_{t-1})) \\ (w_t^X - c)X_{t-1} \\ (w_t^Y - c)Y_{t-1} \end{pmatrix} = 0$$

The first two moment conditions depict the trade autocorrelation of markets X and Y. The third is the joint trade innovation described in equations (1) and (5). The last two moment conditions describe the orthogonality between the residual w_t^X and X_{t-1}, and between w_t^Y and Y_{t-1}. Since $\rho = 2(\gamma + \lambda_X) - 1$, $\delta = 2(\gamma + \lambda_Y) - 1$ and γ, λ_X and λ_Y are probabilities, this implies $-1 \leqslant \rho \leqslant 1$ and $-1 \leqslant \delta \leqslant 1$. This is already shown in the appendix by the fact that ρ and δ are autocorrelation coefficients of X_t and Y_t respectively.

To note, $\rho - \delta = 2(\lambda_X - \lambda_Y)$, such that $\rho > \delta$ implies $\lambda_X > \lambda_Y$ and vice versa. With this, the estimation of β allows the conditional probabilities of X and Y continuation to be inferred. For a given firm, $\lambda_X < \lambda_Y$ suggests that whenever the two listing boards are emitting conflicting signals at time $t-1$, then it is more likely for Board Y to persist in the same direction and for Board X to reverse its direction and follow Board Y at time t.

[9]This should not pose a major hinderance since there are at least 15,000 observations per variable per firm per board.

Whether if λ_X being greater or less than λ_Y is suitably an empirical question depends on the actual pair of markets that the model is applied to. For instance, if we apply the model to examine the joint trade direction of (say) CBOE's S&P 500 index option versus CME's S&P 500 futures options, then whether $\lambda_{ind-opt}$ is greater than $\lambda_{fut-opt}$ is suitably an empirical question. But if we apply it to examine IBM stock versus IBM futures trade, then we would expect $\lambda_f^{IBM} > \lambda_s^{IBM}$

Institutional background

If the birth of a stock market goes contemporaneously with the birth of a stock exchange, then the birth of the China stock market can be traced back to the establishment of its twin exchanges: SHSE on 26th Nov 1990 and SZSE on 11th Apr 1991. As at 31st Dec 2005, there are 878 and 586 listed firms on the SHSE and SZSE respectively. Understandably, most of the large Chinese firms are from predominant sectors of China's economy like telecommunications, textile, transport, airline, infrastructure, resource and basic materials (steel, cement, aluminium, petrochemicals etc).

Each firm incorporated in mainland China not only has to decide which exchange to list on, but also which board to issue shares on. There are three main listing boards: A-, B- and/or H-boards. First, shares on the main A-board are traded in Renminbi (RMB). Only mainlanders are allowed to trade A-shares, but from Dec 2002, QFII, mainly global investment banks[10] are allowed to trade A-shares. As at Dec 2003, there are twelve licensed QFII with an aggregate investment quota of USD 1.7 billion.

[10]e.g. HSBC, JP Morgan Chase, Goldman Sachs, Standard Charter etc.

Second, China incorporated firms can also choose to issue B-shares, which entitle holders to the same voting rights and revenue-streams as A-share holders. But while their face values are also denominated in RMB and traded on domestic exchanges, they are actually subscribed, quoted and traded in a foreign currency (USD on SHSE; HKD on SZSE). As at 31st Dec 2005, there are 54 and 55 B-board firms listed on the SHSE and SZSE respectively. Although the B-board was originally set up for foreign investors, but from Feb 2001, mainlanders with legal foreign currency accounts (either USD or HKD) are also allowed to trade B-shares. The China Securities Regulatory Body (CSRB) introduced these two reforms as a first step towards integrating the two domestic boards.

Third, other than B-shares, a China-incorporated firm can instead choose to issue H-shares on the Hong Kong Exchange (HKEx). As at 31st Dec 2005, there are 80 H-board firms listed on the main board, and another 40 firms listed on the Growth Enterprize Market (GEM) board. Listing on the H-board facilitates potentially better access to foreign investors since the HKEx has a larger proportion of foreign investor participation relative to either SHSE or SZSE. Understandably, firms need to comply with relatively more stringent listing and disclosure requirements as well. Their H-shares are denominated, subscribed and traded in HKD, as with all HKEx listed firm.

INSERT TABLE ONE

The number of listed firms on the various boards as at 31st Dec 2005 are reported in Table 1. Note that a firm can choose whether to list A-, B-, H-, AB- or AH- shares. However, a firm is not allowed to list both B- and H-shares. They are also disallowed, and in any case, do not see the need to cross-list on SHSE and SZSE. The majority of firms that issue either

B- or H-shares would have already issued A-shares, but this is not necessarily the case. The number of firms issuing various classes of shares from 1994 to 2003 are reported in Table 2 Panel A.

INSERT TABLE TWO

In addition to publicly traded A-, B- or H-shares, it is important to note that many China-incorporated firms were former state-owned-enterprisers (SOE). Although partly privatized, the portion of non-tradable shares remains non-trivial, the majority of which are state-owned shares i.e. held by the central government. Another class of non-tradable shares, known as staff shares, are offered to and/or subscribed by employees of the firm. Lastly, a firm could be set up as a joint venture between the government, who provides the land, and a foreign company, who provides capital in return for foreign legal-persons shares. Along similar lines, other variations include sponsors and private placement legal-persons shares. The number (millions) and proportion of various classes of tradable and non-tradable shares issued by Chinese firms from 1994 to 2003 are outlined in Table 2 Panel B.

INSERT TABLE THREE

Trading on all three exchanges are divided into a morning and an afternoon session. For SHSE and SZSE, the pre-opening period is between 9:15am and 9:25am. The morning session is from 9:30am to 11:30am. Afternoon session trading is between 1:00pm and 3:00pm. On the HKEx, the pre-opening period is between 9:30am and 10:00am. Trading starts at 10:00am and finishes at 12:30pm. The market reopens at 2:30pm and closes for the day at 4:00pm. Table 3 provides some basic listing data across the three exchanges over a four year period. While a foreign investor can trade either SHSE/SZSE B-shares or comparable HKEx H-shares, the trading environment in China is quite different from Hong Kong.

First, the HKEx has a longer history hence is more well-established. It imposes more rigorous listing requirements on potential new firms and information disclosure requirements on listed firms. Second, the HKEx trading environment is more 'foreign-user friendly' in terms of cultural and language. For example, the SHSE/SZSE English-version website contain only very basic information, and combined to equal roughly a quarter of the amount of information provided on the HKEx website. In addition, English versions of financial reports and basic media documents are not readily available and/or are not translated in detail. Lastly, as displayed in Table 4, the proportional mix among investor clienteles between China and Hong Kong is also disparate. In terms of the local/foreign clientele, on the HKEx, although foreign investor trading is less than local investor trading, it still averages slightly less than 40% between 2002 and 2004. For the SHSE/SZSE, we use A- and B- share turnover to approximate local and foreign investor participation. In contrast, the mainland exchanges are dominated by local trading, stable at 97% for 2003 and 2004 after a jump from an already high level of 87% in 2002. To note, the QFII was implemented in Dec 2002.

INSERT TABLE FOUR

In terms of the institutional/retail clientele, on the HKEx, institutional trading constitutes an average of about 60%. This is not surprising given that foreign investor participation on the HKEx is non-trivial and that the majority of those are actually institutional investors. Contrastingly, on the mainland exchanges, institutional trading, which we approximate using the percentage of institutional accounts, is only around half a percent of total accounts.

Propositions, data and sampling methodology

If a firm has an option to issue either B- or H-shares, the question becomes, does it matter whether a firm chooses to issue (say) HKD traded B-shares on the SZSE rather than HKD traded H-shares on the HKEx. We attempt to shed some light on this issue. Please note that in discussing the empirical sections, we sometimes use X denotes A-board parameters and Y denotes B- or H-board parameters.

Proposition 1: Information asymmetry

Given that there is more institutional trading on the H-board relative to the B-board, we propose that there is relatively more information asymmetry for stocks listed solely on the mainland boards i.e $\hat{\theta}_{AB} > \hat{\theta}_{AH}$.

Proposition 2a: Price impact of trading

Given that the A-board encompasses more retail trading and the B- and H- boards encompass more institutional trading, we propose that the transitory price impact of order flow is greater on the A-board i.e. $\hat{\phi}_A > \hat{\phi}_B$ *and* $\hat{\phi}_A > \hat{\phi}_H$.

Proposition 2b

Given that there is more institutional trading on the H-board relative to the B-board, we propose that the discrepancy in transitory price impact between A- and B- board is larger than between A- and H- board i.e. $|\hat{\phi}_A - \hat{\phi}_B| > |\hat{\phi}_A - \hat{\phi}_H|$.

Proposition 3

Given a potential convergence of the investor pool between the A- and B-boards, we propose that the presence of price leadership is less (more) evident between A-B (A-H) shares. If this is the case, sub-sample and/or cross-sectional results on the relation between λ_A and λ_B is less robust than between λ_A and λ_H i.e. $\lambda_A <> \lambda_B; \lambda_A < \lambda_H$.

Taken collectively, the results relating to the preceding propositions would allow us to contrast the trading quality of B-shares and H-shares. This in turn offers implications pertaining to the dissimilar investor clientele participating on the two foreign boards. In this paper, we do not differential between Shanghai and Shenzhen listed A-B firms.[11].

We use minute-by-minute data from Reuters for pair wise A-B and A-H firms over the sample period 4th Jan to 30th Sep 2005. For A-H shares, we consider only data generated from when both boards are trading simultaneously.[12] As such, the A-H sample contains two hours worth of data per day. This is between 10:00am and 11:30am in the morning and between 2:30pm and 3:00pm in the afternoon. To note, trading on the HKEx commences at 10:00am and recommences at 2:30pm. Conversely, SHSE/SZSE markets break at 11:30am and closes at 3:00pm. To reduce potential market opening/closing effects, we exclude the first and last five minutes from both trading sessions i.e. our sample covers from 10:05am to 11:24am and from 2:35pm to 2:54pm. This gives a total of 100 observations per trading day (80 in the morning; 20 in the afternoon)

[11] This issue will be investigated in a separate paper on the Chinese market, which has a more empirical focus. The empirical section in the current paper is mainly to demonstrate the use of our model.

[12] There is no time zone difference between China and Hong Kong.

Our propositions involve comparing results between A-B and A-H pair-wise firms. Hence for consistency, we restrict our analysis of A-B firms to the same trading hours as their A-H counterpart. Trading days considered in the A-B group differ slightly from the A-H group since both A and B boards are hosted by the same mainland exchange. Furthermore, when selecting firms into our A-B and A-H group, we need to address the issue of compatibility between the two sample groups. We attempt to promote compatibility by addressing two issues. First is the issue of tradability. Despite significant developments in the privatization of former SOEs, many Chinese firms remain majority state-owned with an average of 30%-40% of issued shares still being held by the state. The proportion of non-tradable shares, which also include legal-person and staff-held shares, is even higher at around 60%-70% on average. For a study that examines intraday price formation, we argue that there is little merit in analyzing a firm where only (say) 15% of its issued capital is tradable.

Second, we address the issue of obtaining appropriate pair-wise A-B and A-H firms. Firm attributes to base the matching criteria on is wide-ranging e.g. size, industry/sector group and leverage etc. Our earlier attempts to match firms based on two or more attributes has led to either or both A-B and A-H groups containing only one firm. While the number of firms can be increased by matching based on one firm attribute, the paper becomes expose to the risk that results may be driven by size, industry or leverage effects. Since our model examines cross-market price formation between main A-board and its corresponding subsidiary board, we ascertain that it is a futile exercise if we examine a firm that has (say) 90% of its tradable shares listed on one board and only 10% on the other.

INSERT TABLE FIVE

Table 5 provides some basic details of the firms included in our two sample groups. We impose three selection criteria. First, we consider only firms whose non-tradable share is less than 70% of total issued capital, which is the overall average according to the 2005 CSRC report. Second, we match A-B and A-H firms based on industry/sector classification, but also attempt to include firms from various sectors of the Chinese economy. Third, to be included in either of our sample groups, a firm has to possess a 'reasonable' spread of tradable shares across its two relevant boards. In this paper, we consider a firm as having a reasonable spread if it has at least 10% of issued capital allocated to each board. In addition, the proportion of issued capital on the relatively smaller board has to be at least one-fifth that of the larger board.[13]

In sum, our overall sample covers a pair of firms in each of ten sectors of the Chinese economy. Each firm has an A-board sample and a B (or H)-board sample. Each sample contains 100 minute-by-minute observations per day. The sample period covers around 170 trading days.

4 Empirical Results

The percentage of buyer and seller initiated trades for each firm is reported in Table 6. If no trade occurs at time t for Market X, then $X_t = 0$ and $p_t^X = p_{t-1}^X$ i.e. $r_t^X = 0$. The same applies for Market Y. The proportion of buy, sell and no-trade averaged over ten firms of a given board is presented as a series of pie-charts in Figure Four. To note our sampling methodology necessarily imposes the same sample size for the two listing boards of a given

[13]Indeed, some of these threshold values may be ad-hoc. But if there is a concern, robustness checks can be performed by varying the minimum 10% issued capital and/or the one-fifth relative size tradable capital rules.

firm.

INSERT TABLE SIX

First, a similar proportion of buyer and seller initiated trades is being observed on a given board for a given firm. E.g. China Shipping Development A-share sample contains 48% and 46% buyer and seller initiated trades respectively. Similarly, its H-share sample consists of 26% buy and 22% sell trades. Inner Mongolia Cashmere has about the same proportion of buy and sell trades on its A-board and 20% buy and 16% sell trades on its B-board. The last two rows of Table 6 report the mean trade classifications for A-H and A-B firms. In the A-H group, the mean proportion of buy and sell trades on the A-board are fairly even at 40.1% and 38.4% respectively. The discrepancy is moderately wider on the H-board with the mean buyer and seller trades at 13.4% and 9.4% respectively. Similarly, in the A-B group, A-board generates an average of 36.9% buy and 35.4% sell trades. Again, the discrepancy is wider on the B-board, with an average of 26.5% buy and 18.9% sell trades being handled. That buy and sell trades are fairly evenly distributed is not surprising since there is no reason ex-ante to expect more buying than selling, or vice versa, for a given firm on a given board over a given time period.

INSERT FIGURE FOUR

Second, the absence of trades is more frequent on either the B- and H-board relative to the main A-board across all firms. Furthermore, the difference in the percentage of no-trades between the main and subsidiary boards is larger for an A-H firm relative to its corresponding A-B counterpart. E.g. Tsingtao Brewery A-board sample has a 45% no-trade classification, but the figure is higher at 86% on its H-board sample. Its counterpart, Yantai Changyu Winery, has a 55% and 60% no trade classification on its A- and B- board samples correspondingly. The average proportion of no-trades on the H-board is 77.2% compared with 54.6% no-trades being observed on the B-board.

INSERT FIGURE THREE

We plot the A- and B/H boards daily price and dollar volume time series over the sample period in Figure Three. For the purpose of comparison, those series denominated in either USD or HKD are converted to RMB using the corresponding average cross rate i.e. RMB/USD or RMB/HKD over the sample period. First, A-shares are trading at a premium over both B- and H-shares. Second, the premium differs from one firm to the next. Third, casual visual observation suggests that price fluctuation between boards is fairly similar, suggesting that information linkages do exist and warrants further analysis. Interestingly, H-share prices for Tsingtao Brewery, Shandong Pharmaceutical, Jingwei Textile and NE electric are fairly stable and do not seem to co-move with their A-share prices.

According to Table 6, this comes as no surprise since there exists a huge proportion of no-trade on the H-board for each of the four firms. Among them, Tsingtao Brewery possess the lowest at 86% no trades. The highest is Shandong Pharmaceutical with 97% no trades. Preliminary evidence thus far suggests that the H-board is likely to play a limited price leadership role, at least for each of these four firms. Fourth, the Chinese government adjusted the RMB upwards in early July 2005. Specifically, the RMB/HKD and RMB/USD cross rates were adjusted from 1.062 to 1.04 and from 8.27 to 8.1 respectively. This has been picked up by a number of firms in terms of investors adjusting the relative pricing between the two boards. Interestingly, the adjustment is (visually) more evident in firms with cross-border dealings e.g. China Shipping, Zhenhua Port, Southern and Hainan airlines, Inner Mongolia Cashmere and BOE Tech.

Table 7 provides descriptive statistics in Panel A and the correlation matrix of key variables in Panel B. While we expect the correlation of own-board parameters to be stronger than cross-board, the direction of the correlation is less clear ex-ante. Over a given sample period, if the 'market-wide' effect dominates e.g. a series of unexpected downgrades in earning forecast, then a positive correlation amongst parameters across boards is observed. But if the 'substitution effect' dominates e.g. arbitrage between boards, then negative correlations among parameters between boards.

INSERT TABLE SEVEN

For China Shipping, r_t^X and r_t^Y are negatively correlated, which is a preliminary indication of substitution between A- and H-board shares. Also, $r_t^X(r_t^Y)$ has a stronger and positive correlation with $X_t(Y_t)$ at 0.553 (0.445) and a weaker and negative correlation with $Y_t(X_t)$ at -0.0079 (-0.0019). Interestingly, the converse is true for Shandong Zhenhua Machinery, where r_t^X and r_t^Y) are positively correlated. This suggests market-wide trading across both A- and B-boards. Also, r_t^X and r_t^Y display positive correlations with both X_t and Y_t, although own board correlations are stronger as expected. Descriptive statistics for other firms from the remaining 9 sectors generated comments that are more or less similar with the preceding discussion on the heavy machinery sector. Hence they are excluded to reduce the length of our paper, but are available upon request.

Table 8 reports trading parameters generated from GMM estimation. First, we examine the first proposition by discussing the results of $\hat{\theta}_{AB}$ and $\hat{\theta}_{AH}$. Results from results from eight of the ten sectors indicate that $\hat{\theta}_{AB} > \hat{\theta}_{AH}$. This would support our first proposition that information asymmetry is relatively more evident in A-B firms due to the lack of participation by institutional investors. [14]

INSERT TABLE EIGHT

Second, $\hat{\phi}^{A} > \hat{\phi}^{B/H}$ in 18 out of the 20 firms in our overall sample. Here, our results strongly support Proposition 2a in suggesting that a higher proportion of institutional trading on the subsidiary boards relative to the A-board provides depth to the market, which lowers the transitory price impact of trading.

In contrast, the results on Proposition 2b indicate that the discrepancy $|\hat{\phi}^{X} - \hat{\phi}^{Y}|$ is not necessarily larger for A-B firms. In our sample, $|\hat{\phi}^{A} - \hat{\phi}^{B}| > |\hat{\phi}^{A} - \hat{\phi}^{H}|$ is observed in five of the ten sectors. They are liquor, transport, utility, shipping and electrical. The converse is true for the remaining five sectors. In light of the mixed outcome, we re-examine Proposition 2b. The original proposition is based on statistical figures on the overall level of institutional participation on the mainland and the Hong Kong share markets, which were taken from the CSRB and HKEx annual reports. While overall institutional trading on the mainland exchanges are comparatively lower, that does not per se imply that institutional trading on the B-board is definitely lower than that of the H-board. Hence we should not expect any robust relation between $|\hat{\phi}^{A} - \hat{\phi}^{B}|$ and $|\hat{\phi}^{A} - \hat{\phi}^{H}|$ ex-ante. Lastly, we find that both $\lambda_A < \lambda_H$ and $\lambda_A < \lambda_B$ in five of the ten A-H and A-B firms. Based on the A-B group, our mixed

[14]Note that the $\hat{\theta}_{AB}$ for textile and utility is actually negative. This is contrary to our expectation that $\theta \geq 0$

finding that price leadership is provided by the B-board in five firms, and by the A-board in five other firms supports our conjecture that the source of price leadership is less clear for a given A-B firm as a result of an quasi-merger between the two boards.

Based on the A-H group, the results are less supportive of our third proposition since we expect H-board to provide more evident price leadership over its A-board counterpart. Nevertheless, this is not totally unexpected since from our preliminary discussion, we made the point that four of the ten A-H firms display very little trading activity in that H-share prices display very low price volatility as a result of infrequent trading on the H-board. Indeed, these are also the firms (Tsingtao Brewery, Shandong Pharmaceutical, Jingwei Textile and NE Electric) where, according to our inference of λs, the A-board provided price leadership over the H-board. The other five firms still demonstrate price leadership despite non-trivial proportions of no-trades. As shown in Figure Four, the average proportion of no-trades on the H-board is 77%.

In sum, the results from full sample analysis are supportive of our three main propositions. Specifically, A-board shares display more trade reversals over time, while B- and H-board trading display more trade continuations. The results suggest that in general, whenever opposite trade directions are observed across two boards at a point in time, there is a tendency for the A-board to revert and follow its B- or H-board counterpart. If institutional investors conduct stealth trading, and are active on these subsidiary boards, then the stronger autocorrelation in B- and H- trade directions is not totally unexpected. Lastly, while our results show that price leadership is not necessarily more evident on the H-board than on the B-board, note that our H-sample has a 'handicap'. On average, 77% of the H sample consist of no trades as opposed to 55% in the B-sample. E.g. For the utility sector, Huadian

Power has a 78% no-trade compared with 60% by Heilongjiang Energy, yet $\lambda_A < \lambda_H$ and $\lambda_A > \lambda_B$. For the electrical sector, both λ_B for Guangdong Kelon and λ_H for BOE Tech are greater than their corresponding λ_A, yet the Guangdong Kelon H-sample contains 69% no-trade, compared with BOE Tech's 21% no-trade in its H-sample.

5 Concluding Remarks

In this paper we attempt to make one theoretical and one empirical contribution to the existing literature. In the research on price formation models, the one-dimensional MRR (1997) trade direction model considers the auto-correlation in trade direction ρX_{t-1} in subsequent price formation. Hence the mixed-signal state and its corresponding transition probability $\mathbb{P}(X_t = 0) = \frac{1-\lambda}{2}$ are less relevant in measuring price discovery.

In contrast, our extended model considers the joint trade direction of two markets, and since each of X_t and Y_t is a dichotomy, the mixed-signal states and their corresponding transition probabilities $\mathbb{P}(X_t = X_{t-1} \bigcap Y_t = -Y_{t-1}) = \lambda_X$ and $\mathbb{P}(X_t = -X_{t-1} \bigcap Y_t = Y_{t-1}) = \lambda_Y$ offer a new measure of price discovery between two substitute markets. The emphasis of the extended model is on the likelihood of one market following the other when the two markets generate conflicting signals in the previous state.

This application is demonstrated in our empirical analysis, which also constitute our empirical contribution to the literature that examines the trading dynamics of A-B and A-H twin shares issued by Chinese firms. We argue that our empirical analysis improves on both Sun and Tong (2000) and Jiang and Wang (2004) in terms of data frequency, sample period, sampling methodology. Most importantly, our testing methodology is based on a theoretical model of price formation.

Our results support the propositions in suggesting that for a A-H firm, the larger proportion of institutional investor clientele on the H-board provides price leadership over the A-board. In future research we expect results to be even stronger if we examine firms whose A- and H-board trading frequencies are more comparable. The source of price leadership is less evident for the A-B group. This is not surprising given that the clientele distinction is blurred on the two domestic boards, with mainlanders trading B-shares and foreigners trading A-shares. Our results do imply potential efficiency gains from order-flows consolidation by maintaining a single listing board.

Theoretical extension on our model is twofold. First, expand the state space for X_t and Y_t to formally acknowledge the no-trade scenario. Second, relax the zero cross-correlation restriction on Σ. This restriction arose from adopting a doubly stochastic transition matrix P that centers symmetrically around zero. Because of this, both cross-products $\mathrm{Cov}\,(X_t, Y_{t-1})$ and $\mathrm{Cov}\,(Y_t, X_{t-1})$ will always be zero. A direct remedy is to make P asymmetric by considering more probability parameters to distinguish between (say) $\mathbb{P}(X_t = Y_t = 1|X_{t-1} = Y_{t-1} = 1)$ and $\mathbb{P}(X_t = Y_t = -1|X_{t-1} = Y_{t-1} = -1)$. Under the current model, both are γ. We can accommodate the asymmetric price impact of buy versus sell trades by denoting $\mathbb{P}(X_t = Y_t = 1|X_{t-1} = Y_{t-1} = 1) = \gamma_b$ and $\mathbb{P}(X_t = Y_t = -1|X_{t-1} = Y_{t-1} = -1) = \gamma_s \neq \gamma_b$. As such, P becomes a regular transition matrix and Σ should contain non-zero off-diagonal cross-covariance terms. These will be presented as separate papers.

THE END

References

[1] Al-Suhaibani, M., and Kryzanowski, L., 2000. *The information content of orders on the Saudi stock market.* Journal of Financial Research, 23, 145-156.

[2] Chng, M., 2004a. *The trading dynamics of close-substitute futures markets: Evidence of margin policy spillover effects.* Journal of Multinational Financial Management, 14, 463-483.

[3] Chng, M., 2004b. *A model of price discovery and market design: Theory and empirical evidence.* Journal of Futures Markets, 24, 1107-1146.

[4] Chng, M., 2005. *Measuring the summary informativeness of orders and trades.* Review of Futures Markets, Fall, 245-281.

[5] Dufour, A., and Engle, R.F., 2000. *Time and the price impact of a trade.* Journal of Finance, LV, 2467-2498.

[6] Ellis, K., Michaely, R., and O'Hara, M., 2000. *The accuracy of trade classification rules.* Journal of Financial and Quantitative Analysis, 35, 529-551.

[7] Easley, D., and O'Hara, M., 1987. *Price, trade size, and information in securities markets.* Journal of Financial Economics, 19, 69-90.

[8] Glosten, L., and Milgrom, P., 1985. *Bid, ask and transaction prices in a specialist market with heterogeneously informed investors.* Journal of Financial Economics, 13, 71-100.

[9] Hamilton, J.D., 1994. *Time series analysis* (Princeton, NJ: University Press).

[10] Hasbrouck, J., 1991. *The summary informativeness of stock trades: An econometrics analysis.* Review of Financial Studies, 4, 571-595.

[11] Hasbrouck, J., 1995. *One security, many markets: Determining the contributions to price discovery.* Journal of Finance, 50, 1175-1199.

[12] Madhavan, A., Richardson, M., and Roomans, M., 1997. *Why do security prices change? A transaction-level analysis of NYSE stocks.* Review of Financial Studies, 10(4), pp 1035-1064.

[13] Lee, C.F., and Ready, M., 1991. *Inferring trade direction from intraday data.* Journal of Finance, 46, pp 733-746.

[14] Lee, C., and Rui, O., 2000. *Does trading volume contain information to predict stock returns? Evidence from China's stock markets.* Review of Quantitative Finance and Accounting, 14, pp 341-360.

[15] Stock, J.H., and Watson, M.W., 1988. *Variable trends in economic time series.* Journal of Economic Perspective, 2, 147-174.

[16] Sun, Q., and Tong, H.S., 2000. *The effect of market segmentation on stock prices: The China syndrome.* Journal of Banking and Finance, 24, p 1875-1902.

[17] Wang, S., and Jiang, L., 2004. *Location of trade, ownership restrictions, and market illiquidity: Examining Chinese A-H shares.* Journal of Banking and Finance, 28, p 1273-1297.

[18] Yeh, Y., Lee, T., and Pen, J., 2004. *Stock return and volatility under market segmentation: The case of Chinese A and B shares.* Review of Quantitative Finance and Accounting, 18, p 239-257.

Appendix

Notations:
Market i = Market X or Market Y
μ_t: Post-trade $\mathbb{E}(V_t|\Phi_t)$
p_t^i: Transaction price of Market i
ϕ^i: Collective market making cost parameter of Market i
θ: Asymmetric information parameter
ξ_t^i: Pricing error of Market i

$m_t = m_{t-1} + \theta(X_t + Y_t - \mathbb{E}(X_t + Y_t|X_{t-1}, Y_{t-1})) + \epsilon_t$
$p_t^X = m_t + \phi^X X_t + \xi_t^X$
$p_t^Y = m_t + \phi^Y Y_t + \xi_t^Y$

Derivation of equations:

$$\begin{aligned}\mathbb{E}(X_t + Y_t|X_{t-1}, Y_{t-1}) &= (2(\gamma + \lambda_X) - 1)X_{t-1} + (2(\gamma + \lambda_Y) - 1)Y_{t-1} \\ &= \rho(X_{t-1}) + \delta(Y_{t-1}) \text{ (derived below)}\end{aligned}$$

$$\begin{aligned}\Delta p_t^X &= \theta(X_t + Y_t) - \theta(\mathbb{E}(X_t + Y_t|X_{t-1}, Y_{t-1})) + \phi^X(\Delta X_t) + \Delta\xi_t^X + \epsilon_t \\ &= \theta(X_t + Y_t) - \theta(\rho X_{t-1} + \delta Y_{t-1}) + \phi^X(\Delta X_t) + \theta(X_{t-1}) - \theta(X_{t-1}) \\ &+ \theta(Y_{t-1}) - \theta(Y_{t-1}) + \Delta\xi_t^X + \epsilon_t \\ &= (\theta + \phi^X)(\Delta X_t) + \theta(\Delta Y_t) + \theta(1-\rho)X_{t-1} + \theta(1-\delta)Y_{t-1} + \Delta\xi_t^X + \epsilon_t\end{aligned}$$

$$\begin{aligned}\Delta p_t^Y &= \theta(X_t + Y_t) - \mathbb{E}(X_t + Y_t|X_{t-1}, Y_{t-1}) + \phi^Y(\Delta Y_t) + \Delta\xi_t^Y + \epsilon_t \\ &= \theta(X_t + Y_t) - \theta(\rho X_{t-1} + \delta Y_{t-1}) + \phi^Y(\Delta Y_t) + \theta(X_{t-1}) - \theta(X_{t-1}) \\ &+ \theta(Y_{t-1}) - \theta(Y_{t-1}) + \Delta\xi_t^Y + \epsilon_t \\ &= (\theta + \phi^Y)(\Delta Y_t) + \theta(\Delta X_t) + \theta(1-\rho)(X_{t-1}) + \theta(1-\delta)(Y_{t-1}) + \Delta\xi_t^Y + \epsilon_t\end{aligned}$$

To set up the transition matrix, denote the following states:

$$\begin{aligned}&\mathbb{P}(\text{Full-continuation}) = \mathbb{P}(X_t = X_{t-1}, Y_t = Y_{t-1}) = \gamma \\ &\mathbb{P}(\text{X-continuation}) = \mathbb{P}(X_t = X_{t-1}, Y_t \neq Y_{t-1}) = \lambda_X \\ &\mathbb{P}(\text{Y-continuation}) = \mathbb{P}(X_t \neq X_{t-1}, Y_t = Y_{t-1}) = \lambda_Y \\ &\mathbb{P}(\text{Full-reversal}) = \mathbb{P}(X_t \neq X_{t-1}, Y_t \neq Y_{t-1}) = (1 - \gamma - \lambda_X - \lambda_Y)\end{aligned}$$

$\mathbb{P}(X_t, Y_t) = (i, j) = \frac{1}{4}$
$\mathbb{P}(X_t = 1) = \mathbb{P}(X_t = -1) = \frac{1}{2} \longrightarrow \mathbb{E}(X_t) = 0, \mathrm{V}ar(X_t) = 1$

By symmetry, $\mathbb{E}(Y_t) = 0, \mathrm{V}ar(Y_t) = 1$ and

$$\mathrm{Cov}\,(X_t, Y_t) = \mathbb{E}(X_t Y_t) = \mathbb{P}(X_t = Y_t = 1) + \mathbb{P}(X_t = Y_t = -1) - 2\mathbb{P}(X_t \neq Y_t) = 0.$$

$$\begin{aligned}\mathbb{P}(X_t = X_{t-1} = 1) &= \mathbb{P}(X_t = X_{t-1} = Y_t = Y_{t-1} = 1) + \mathbb{P}(X_t = X_{t-1} = 1, Y_t = Y_{t-1} = -1)\\ &+ \mathbb{P}(X_t = X_{t-1} = 1, Y_t = 1, Y_{t-1} = -1) + \mathbb{P}(X_t = X_{t-1} = 1, Y_t = -1, Y_{t-1} = 1)\end{aligned}$$

$$\begin{aligned}=&\frac{1}{4}[\mathbb{P}(X_t = 1, Y_t = -1 | X_{t-1} = 1, Y_{t-1} = -1) + \mathbb{P}(X_t = 1, Y_t = -1 | X_{t-1} = 1, Y_{t-1} = 1)\\ &+ \mathbb{P}(X_t = 1, Y_t = 1 | X_{t-1} = 1, Y_{t-1} = -1) + \mathbb{P}(X_t = 1, Y_t = 1 | X_{t-1} = 1, Y_{t-1} = 1)]\end{aligned}$$

$$=\frac{1}{2}(\gamma + \lambda_X) = \mathbb{P}(X_t = X_{t-1} = -1) \rightarrow \mathbb{P}(X_t = X_{t-1}) = \gamma + \lambda_X.$$

Similarly, we have $\mathbb{P}(X_t = 1, X_{t-1} = -1) = \mathbb{P}(X_t = -1, X_{t-1} = 1) = \frac{1}{2}(1 - \gamma - \lambda_X)$ such that $\mathbb{P}(X_t \neq X_{t-1}) = 1 - \gamma - \lambda_X$

Hence $\mathrm{Cov}\,(X_t, X_{t-1}) = \mathbb{E}(X_t X_{t-1}) = \mathbb{P}(X_t = X_{t-1}) - \mathbb{P}(X_t \neq X_{t-1}) = 2(\gamma + \lambda_X) - 1 = \rho$

The proof for Y_t is similar, such that we obtain:

$\mathbb{P}(Y_t = Y_{t-1} = 1) = \mathbb{P}((Y_t = Y_{t-1} = -1) = \frac{1}{2}(\gamma + \lambda_Y)$

$\mathbb{P}(Y_t = 1, Y_{t-1} = -1) = \mathbb{P}(Y_t = -1, Y_{t-1} = 1) = \frac{1}{2}(1 - \gamma + \lambda_Y)$

$\mathrm{Cov}\,(Y_t, Y_{t-1}) = \mathbb{E}(Y_t Y_{t-1}) = 2(\gamma + \lambda_Y) - 1 = \delta$

Set the linear predictor $\Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}$ for $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ to minimize

$$\mathbb{E}\left(\left(\begin{pmatrix} X_t \\ Y_t \end{pmatrix} - \Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}\right)^T \times \left(\begin{pmatrix} X_t \\ Y_t \end{pmatrix} - \Sigma \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix}\right)\right),$$

where Σ satisfies:

$$\Sigma \begin{pmatrix} \mathrm{Cov}\,(X_t, X_{t-1}) & \mathrm{Cov}\,(X_t, Y_{t-1}) \\ \mathrm{Cov}\,(Y_t, X_{t-1}) & \mathrm{Cov}\,(Y_t, Y_{t-1}) \end{pmatrix} = \begin{pmatrix} 2(\gamma + \lambda_X) - 1 & 0 \\ 0 & 2(\gamma + \lambda_Y) - 1 \end{pmatrix}$$

,
which implies

$$\begin{aligned}\mathbb{E}(X_t + Y_t | X_{t-1}, Y_{t-1}) &= (2(\gamma + \lambda_X) - 1)X_{t-1} + (2(\gamma + \lambda_Y) - 1)Y_{t-1}\\ &= \rho(X_{t-1}) + \delta(Y_{t-1})\end{aligned}$$

The MRR (1997) model

$$
\begin{aligned}
\mu_t &= \mu_{t-1} + \theta(X_t - \mathbb{E}(X_t|X_{t-1}) + \epsilon_t \\
p_t &= \mu_t + \phi X_t + \xi_t \\
&= \mu_{t-1} + \theta(X_t - \mathbb{E}(X_t|X_{t-1}) + \phi X_t + \epsilon_t
\end{aligned}
$$

$$
X_t = \begin{cases} +1, & \text{if buyer-initiated trade;} \\ 0, & \text{if trade occurs within prevailing spread;} \\ -1, & \text{if seller-initiated trade.} \end{cases}
$$

$\mathbb{P}(X_t = 0) = \lambda$
$\mathbb{P}(X_t = X_{t-1}|X_{t-1} \neq 0) = \gamma$
$\mathbb{P}(X_t \neq X_{t-1}|X_{t-1} \neq 0) = (1 - \gamma - \lambda)$
$\mathbb{E}(X_t|X_{t-1} = 1) = \mathbb{P}(X_t = 1|X_{t-1} = 1) - \mathbb{P}(X_t = -1|X_{t-1} = 1)$
$= \gamma - (1 - \gamma - \lambda) = 2\gamma - (1 - \lambda) = \rho$

$\mathbb{E}(X_t|X_{t-1} = 0) = \mathbb{P}(X_t = 1|X_{t-1} = 0) - \mathbb{P}(X_t = -1|X_{t-1} = 0)$
$= \frac{1-\lambda}{2} - \frac{1-\lambda}{2} = 0$
$\mathbb{E}(X_t|X_{t-1} = -1) = \mathbb{P}(X_t = 1|X_{t-1} = -1) - \mathbb{P}(X_t = -1|X_{t-1} = -1)$
$= (1 - \gamma - \lambda) - \gamma = -\rho,$

hence $\mathbb{E}(X_t|X_{t-1}) = \rho X_{t-1}$.

$$
\begin{aligned}
\Delta p_t &= \theta(X_t - \rho X_{t-1}) + \phi \Delta X_t + \Delta \xi_t + \epsilon_t \\
&= (\phi + \theta)\Delta X_t + \theta(1 - \rho)X_{t-1} + \Delta \xi_t + \epsilon_t
\end{aligned}
$$

Let Cov $(X_t, X_{t-1}) = \frac{\sigma_{X_t, X_{t-1}}}{\sigma^2_{X_{t-1}}} = \frac{\mathbf{E}(X_t X_{t-1}) - \mathbf{E}(X_t)\mathbf{E}(X_{t-1})}{\mathbf{E}(X^2_{t-1}) - \mathbf{E}(X_{t-1})^2}$

$$
\begin{aligned}
\mathbb{E}(X_{t-1}) = \mathbb{E}(X_t) &= \mathbb{E}(X_t|X_{t-1} = 1)\mathbb{P}(X_{t-1} = 1) \\
&+ \mathbb{E}(X_t|X_{t-1} = 0)\mathbb{P}(X_{t-1} = 0) + \mathbb{E}(X_t|X_{t-1} = -1)\mathbb{P}(X_{t-1} = -1) \\
&= (\gamma - 1 + \gamma + \lambda) \times \frac{1-\lambda}{2} + (1 - \gamma - \lambda - \gamma) \times \frac{1-\lambda}{2} \\
&= 0
\end{aligned}
$$

On the other hand,

$$
\mathbb{E}(X^2_{t-1}) = \frac{1-\lambda}{2}(1)^2 + \frac{1-\lambda}{2}(-1)^2 = (1 - \lambda),
$$

and

$$\mathbb{E}(X_t X_{t-1}) = \mathbb{E}[\mathbb{E}(X_t X_{t-1}|X_{t-1})] = \mathbb{E}[X_{t-1}\mathbb{E}(X_t|X_{t-1})] = \rho\mathbb{E}(X_{t-1}^2) = (1-\lambda)\rho,$$

therefore,

$$\mathrm{Cov}\,(X_t, X_{t-1}) = \frac{(1-\lambda)\rho}{(1-\lambda)} = 2\gamma - (1-\lambda) = \rho.$$

Proposition: To recognize asymmetric price formation associated with buy and sell decisions, consider $\mathbb{P}(X_t = Y_t = 1|X_{t-1} = Y_{t-1} = 1) = \gamma_1$ and $\mathbb{P}(X_t = Y_t = -1|X_{t-1} = Y_{t-1} = -1) = \gamma_2$, where $\gamma_1 \neq \gamma_2$, which leads to the following asymmetric transition matrix:

$$P = \left(\begin{array}{cccc} \gamma_1 & \lambda_X & \lambda_Y & 1-\gamma_1-\lambda_X-\lambda_Y \\ \lambda_X & \gamma & 1-\gamma-\lambda_X-\lambda_Y & \lambda_Y \\ \lambda_Y & 1-\gamma-\lambda_X-\lambda_Y & \gamma & \lambda_X \\ 1-\gamma_2-\lambda_X-\lambda_Y & \lambda_Y & \lambda_X & \gamma_2 \end{array} \right)$$

Solving for Σ leads to:

$$\Sigma = \begin{pmatrix} \mathrm{Cov}\,(X_t, X_{t-1}) & \mathrm{Cov}\,(X_t, Y_{t-1}) \\ \mathrm{Cov}\,(Y_t, X_{t-1}) & \mathrm{Cov}\,(Y_t, Y_{t-1}) \end{pmatrix} = \begin{pmatrix} \gamma + \frac{1}{2}(\gamma_1+\gamma_2) + 2\lambda_X - 1 & \frac{1}{2}(\gamma_1+\gamma_2-2\gamma) \\ \frac{1}{2}(\gamma_1+\gamma_2-2\gamma) & \gamma + \frac{1}{2}(\gamma_1+\gamma_2) + 2\lambda_Y - 1 \end{pmatrix}$$

Imposing the restriction $\gamma_1 = \gamma_2 = \gamma$ yields the original

$$\Sigma = \begin{pmatrix} 2(\gamma+\lambda_X)-1 & 0 \\ 0 & 2(\gamma+\lambda_Y)-1 \end{pmatrix}$$

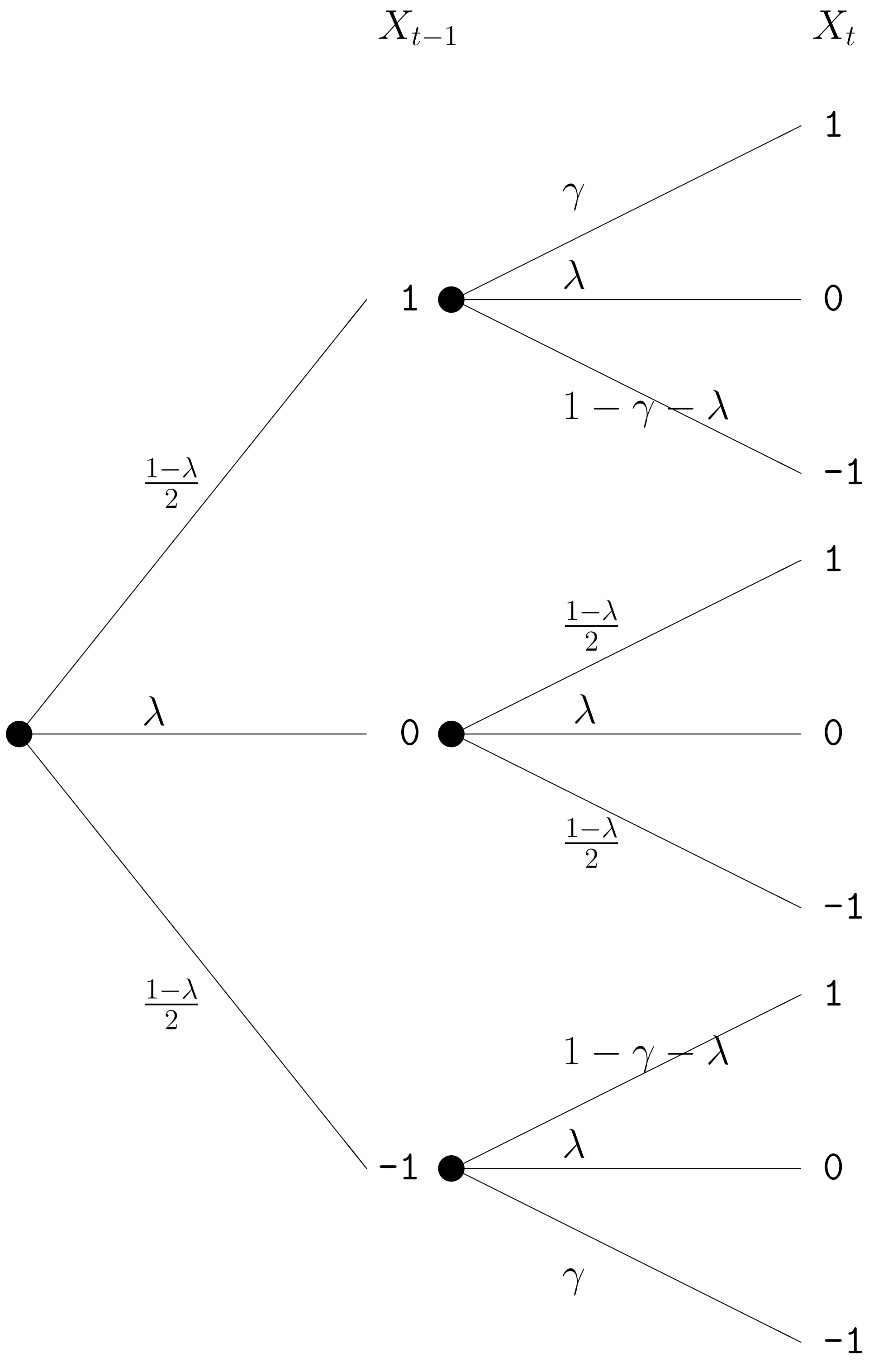

X_{t-1}
X_t
1
0
-1
γ
λ
$1-\gamma-\lambda$
$\frac{1-\lambda}{2}$
λ
$\frac{1-\lambda}{2}$
1
0
-1
$\frac{1-\lambda}{2}$
λ
$\frac{1-\lambda}{2}$
1
0
-1
$1-\gamma-\lambda$
λ
γ
1
0
-1

Figure 1:

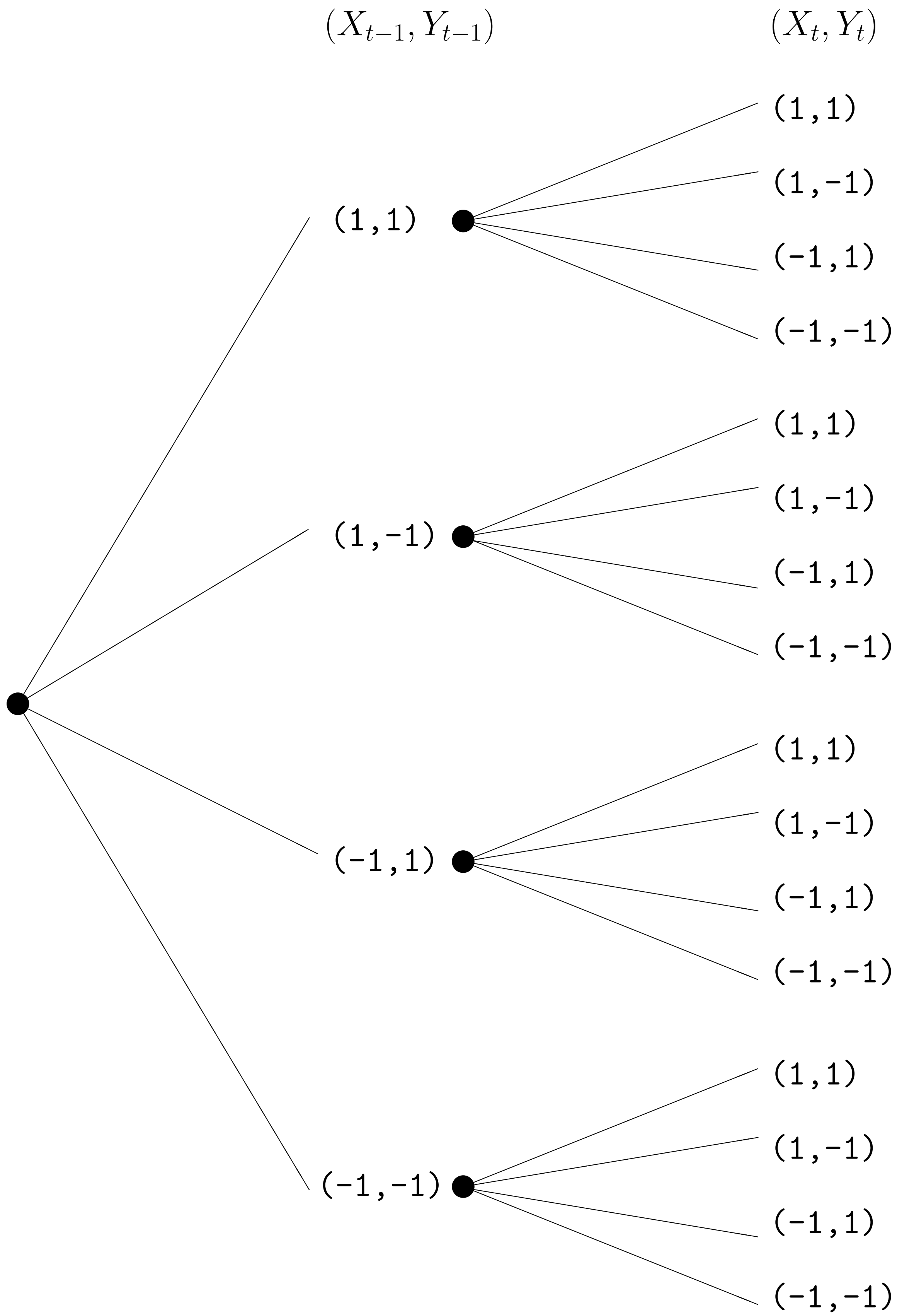

(X_{t-1}, Y_{t-1})
(X_t, Y_t)
(1,1)
(1,-1)
(-1,1)
(-1,-1)
(1,1)
(1,-1)
(-1,1)
(-1,-1)
(1,1)
(1,-1)
(-1,1)
(-1,-1)
(1,1)
(1,-1)
(-1,1)
(-1,-1)
(1,1)
(1,-1)
(-1,1)
(-1,-1)

Figure 2:

Table 1: Description of the main classes of shares issued by Chinese firms

Share class	Origin of incorporation	Currency traded on Exchange	Ownership restriction	Liquidity	No. of listed firms on Dec 2005
A	Mainland China	RMB on SHSE & SZSE	Domestic; from Dec 2002: QFII	High	1223 firms (694 on SHSE, 529 on SZSE)
B	Mainland China	USD on SHSE HKD on SZSE	Foreign; from Feb 2001: domestic with foreign currency accounts (USD or HKD)	Low	111 firms 54 on SHSE (includes 39 A-B firms) 57 on SZSE (includes 43 A-B firms)
H	Mainland China	HKD on HKEx	Foreign	High	121 firms (includes 31 A-H firms) 80 on main board & 40 on GEM
Red chip	Incorporated/managed in Hong Kong; possess substantial business interest/connections to Mainland China	HKD on HKEx	Foreign	High	80 listed firms
N	Mainland China	USD on NYSE	Foreign	Low	12 firms on HKEx traded as ADR. (includes 5 H-N & 7 A-H-N firms)

Table 2: Basic details of the main classes of shares issued by Chinese firms from 1994 to 2003

	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
Panel A: No. of listed firms in various share classes										
A-shares only firms	227	242	431	627	727	822	955	1025	1085	1146
A- & H-shares firms	6	11	14	17	18	19	19	23	28	30
A- & B-shares firms	54	58	69	76	80	82	86	88	87	87
B-shares only firms	4	12	16	25	26	26	28	24	24	24
Total A-share firms	287	311	514	720	825	923	1060	1136	1200	1263
Total B-share Firms	58	70	85	101	106	108	114	112	111	111
Total firms	291	324	530	745	851	949	1088	1160	1224	1287
Panel B: No. of shares (100 million)										
State-owned shares	296.47	328.67	432.01	612.28	865.51	1116.07	1475.13			
	(43.31%)	(38.74%)	(35.42%)	(31.52%)	(34.25%)	(36.13%)	(38.90%)			
Sponsors legal persons shares	73.87	135.18	224.63	439.91	528.06	590.51	642.54			
	(10.79%)	(15.95%)	(18.42%)	(22.64%)	(20.90%)	(19.12%)	(16.95%)			
Foreign legal persons shares	7.52	11.84	14.99	26.07	35.77	40.51	46.20			
	(1.10%)	(1.40%)	(1.23%)	(1.34%)	(1.42%)	(1.31%)	(1.22%)			
Private placement of legal persons shares	72.82	61.93	91.82	130.48	152.34	190.10	214.20			
	(10.64%)	(7.30%)	(7.53%)	(6.72%)	(6.03%)	(6.15%)	(5.65%)			
Staff shares	6.72	3.07	14.64	39.62	51.70	36.71	24.29			
	(0.98%)	(0.36%)	(1.20%)	(2.04%)	(2.05%)	(1.19%)	(0.64%)			
Others	1.10	6.27	11.60	22.87	31.47	33.20	35.07			
	(0.16%)	(0.74%)	(0.95%)	(1.18%)	(1.25%)	(1.07%)	(0.92%)			
A-shares	143.76	179.94	267.32	442.68	608.03	813.18	1078.16			
	(21.00%)	(21.21%)	(21.92%)	(22.79%)	(24.06%)	(26.33%)	(28.43%)			
B-shares	41.46	56.52	78.65	117.31	133.96	141.92	151.56			
	(6.06%)	(6.66%)	(6.45%)	(6.04%)	(5.30%)	(4.59%)	(4.00%)			
H-shares	40.82	65.00	83.88	111.45	119.95	124.54	124.54			
	(5.96%)	(7.66%)	(6.88%)	(5.74%)	(4.75%)	(4.03%)	(3.28%)			
Total number of shares	684.54	848.42	1219.54	1942.67	2526.79	3088.95	3791.71			
	(100.0%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)	(100.0%)			

Table 3: Basic listing details on each of the three stock exchanges

As at	30th Dec 2005	30th Dec 2004	30th Dec 2003	30th Dec 2002	30th Dec 2005	30th Dec 2004	30th Dec 2003	30th Dec 2002
Panel A: HKEx	**Main Board**				**GEM**			
No. of listed companies	934	892	852	812	201	204	185	166
No. of listed H shares	80	72	64	54	40	37	28	20
No. of listed red-chips stocks	86	81	72	71	3	3	0	1
Total no. of listed securities	2448	1971	1598	1416	201	205	187	170
Total market capitalisation (bil)	HKD 8,113	HKD 6,629	HKD 5,478	HKD 3,559	HKD 67	HKD 67	HKD 70	HKD 52
Total negotiable capitalisation (bil)	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Average P/E ratio (Times)	15.57	18.73	18.96	14.89	22.94	28.65	38.79	21.75
Total turnover (m shares)	21104	20122	6504	2943	106	52	116	68
Total turnover (m)	HKD 18,500	HKD 9,724	HKD 10,921	HKD 1,693	HKD 85	HKD 33	HKD 98	HKD 37
Panel B: SHSE	**A-Share**				**B-Share**			
No. of listed companies	824	827	770	705	54	54	54	54
Total market capitalisation (bil)	RMB 2,286	RMB 2,571	RMB 2,940	RMB 2,492	RMB 24	RMB 30	RMB 40	RMB 44
Total negotiable capitalisation (bil)	RMB 651	RMB 705	RMB 780	RMB 702	RMB 24	RMB 30	RMB 40	RMB 44
Average P/E ratio (Times)	16.38	24.29	36.64	34.5	12.4	20.15	30.32	30.61
Total turnover (m shares)	1980	810	1428	633	23	21	28	18
Total turnover (m shares)	RMB 10,128	RMB 4,624	RMB 11,004	RMB 5,752	RMB 68	RMB 74	RMB 145	RMB 81
Panel B: SZSE	**A-Share**				**B-Share**			
No. of listed companies	531	522	491	496	55	56	57	56
Total market capitalisation (bil)	RMB 895	RMB 1,060	RMB 1,212	RMB 1,247	RMB 36	RMB 42	RMB 50	RMB 34
Total negotiable capitalisation (bil)	RMB 351	RMB 395	RMB 451	RMB 463	RMB 35	RMB 37	RMB 44	RMB 30
Average P/E ratio (Times)	16.96	0	0	0	9.11	0	0	0
Total turnover (m shares)	1137	517	635	374	25	42	74	13
Total turnover (m shares)	RMB 5,629	RMB 2,850	RMB 4,551	RMB 3,079	RMB 88	RMB 185	RMB 402	RMB 48

Table 4: Proportion of trading by investor type

HKEx (Source from HKEx website)	**2001/2002**	**2002/2003**	**2003/2004**
• Principal trading	7%	3%	7%
• Local (retail + institutional)	32 + 24 = 56%	30 + 28 = 58%	34 + 22 = 57%
• Foreign (retail + institutional)*	2 + 35 = 37%	4 + 35 = 39%	3 + 33 = 36%
Total:	100%	100%	100%
SHSE/SZSE (Source from CSRC report)			
• A-board turnover	87%	97%	97%
• B-board turnover	13%	3%	3%
Total:	100%	100%	100%
• Institutional accounts	0.52%	0.5%	0.48%
• Retail accounts	99.48%	99.5%	99.52%
Total:	100%	100%	100%

*Foreign investors on the HKEx are mainly institutional investors, with US and Europe constituting around 75% of HKEx foreign institutional investor trading.

Table 5: Selected firms, number (millions) and proportion of issued capital shares

Industry	Matched AB-AH firms	RIC code	Total	A	B/H	Non-tradable
Heavy machinery	China Shipping Development	600026.SS 1138.HK	3326 (100%)	451 (14%)	1296 (39%)	1579 (47%)
	Shanghai Zhenhua Port Machinery	600320.SS 900947.SS	1541 (100%)	435 (28%)	330 (21%)	776 (51%)
Liquor	Tsingtao Brewery	600600.SS 168.HK	1308 (100%)	200 (15%)	655 (50%)	453 (35%)
	Yantai Changyu Pioneer Wine	000869.SZ 200869.SZ	406 (100%)	50 (12%)	137 (34%)	219 (54%)
Transport	China Southern Airline	600029.SS 1055.HK	4374 (100%)	1000 (23%)	1174 (27%)	2200 (50%)
	Hainan Airline	600221.SS 900945.SS	730 (100%)	372 (51%)	185 (25%)	173 (24%)
Pharmacy	Shandong Xinhua Pharmaceutical	000756.SZ 719.HK	457 (100%)	76 (17%)	150 (33%)	231 (51%)
	Livzon Pharmaceutical	200513.SZ 000513.SZ	306 (100%)	116 (38%)	122 (40%)	68 (22%)
Textile	Jingwei Textile Machinery	000666.SZ 350.HK	604 (100%)	203 (34%)	181 (30%)	330 (36%)
	Inner Mongolia Cashmere	600295.SS 900936.SS	1032 (100%)	192 (18%)	420 (41%)	420 (41%)
Telecom	ZTE Corporation	000063.SZ 763.HK	960 (100%)	376 (39%)	160 (17%)	424 (44%)
	Eastern Communications	600776.SS 900941.SS	628 (100%)	118 (19%)	150 (24%)	360 (57%)
Utility	Huadian Power	600027.SS 1071.HK	5590 (100%)	569 (10%)	1431 (26%)	3590 (64%)
	Heilongjiang Energy	600726.SS 900937.SS	1126 (100%)	195 (17%)	432 (38%)	499 (45%)
Shipyard	Guangzhou Shipyard	600685.SS 317.HK	495 (100%)	126 (25%)	157 (32%)	212 (43%)
	Shenzhen Chiwan Wharf	000725.SZ 200725.SZ	645 (100%)	85 (13%)	180 (28%)	380 (59%)
Electrical	Guangdong Kelon Electrical	000921.SZ 921.HK	992 (100%)	195 (20%)	460 (46%)	337 (34%)
	BOE Tech Group	000725.SZ 200725.SZ	2196 (100%)	262 (12%)	1116 (51%)	818 37%)
R&D	NE Electric Development	000585.SZ 42.HK	873 (100%)	144 (16%)	258 (30%)	471 (54%)
	Guangdong Electric Power Development	000539.SZ 200539.SZ	2659 (100%)	513 (19%)	665 (25%)	1481 (56%)

Table 6: Basic trading features

Matched AB-AH firms	Total		Buyer-initiated trades		No trade		Seller-initiated trades	
	A	B/H	A	B/H	A	B/H	A	B/H
China Shipping Development	20,733 (100%)	20,733 (100%)	9,888 (48%)	5,356 (26%)	1,253 (6%)	10,741 (52%)	9,592 (46%)	4,636 (22%)
Shanghai Zhenhua Port Machinery	21,480 (100%)	21,480 (100%)	10,424 (48%)	7,670 (36%)	844 (4%)	7810 (36%)	10,212 (48%)	6,000 (28%)
Tsingtao Brewery	21,000 (100%)	21,000 (100%)	6,257 (30%)	1,700 (8%)	9,442 (45%)	18,035 (86%)	5,301 (25%)	1,265 (6%)
Yantai Changyu Pioneer Wine	21,360 (100%)	21,360 (100%)	5,128 (24%)	5,266 (25%)	11,795 (55%)	12,832 (60%)	4,437 (21%)	3262 (15%)
China Southern Airline	20,880 (100%)	20,880 (100%)	10,121 (48%)	4274 (20%)	656 (4%)	13,533 (65%)	10,103 (48%)	3,073 (15%)
Hainan Airline	21,360 (100%)	21,360 (100%)	8,412 (39%)	2,571 (12%)	4,104 (19%)	17,194 (80%)	8,844 (42%)	1,595 (8%)
Shandong Xinhua Pharmaceutical	21,000 (100%)	21,000 (100%)	6,533 (31%)	457 (2%)	8,727 (42%)	20,333 (97%)	5,740 (27%)	210 (1%)
Livzon Pharmaceutical	21,840 (100%)	21,840 (100%)	6,763 (31%)	5,839 (27%)	8,917 (42%)	11,151 (51%)	6,160 (28%)	4,850 (22%)
Jingwei Textile Machinery	20,760 (100%)	20,760 (100%)	7,709 (37%)	567 (3%)	5,438 (26%)	19,991 (96%)	7,613 (37%)	202 (1%)
Inner Mongolia Cashmere	21,720 (100%)	21,720 (100%)	7,448 (34%)	4,435 (20%)	7,418 (34%)	13,876 (64%)	6,854 (32%)	3,409 (16%)
ZTE Corporation	20,760 (100%)	20,760 (100%)	9,185 (44%)	5,354 (26%)	3,283 (16%)	12,192 (59%)	8,292 (40%)	3,214 (15%)
Eastern Communications	21,480 (100%)	21,480 (100%)	8,474 (39%)	3,471 (16%)	4,936 (23%)	15,854 (74%)	8,070 (38%)	2,155 (10%)
Huadian Power	18,120 (100%)	18,120 (100%)	8,564 (47%)	2,239 (12%)	350 (2%)	14,133 (78%)	9,206 (51%)	1,748 (10%)
Heilongjiang Energy	21,480 (100%)	21,480 (100%)	7,420 (35%)	4,984 (23%)	6,787 (32%)	12,820 (60%)	7,273 (34%)	3,676 (17%)
Guangzhou Shipyard	20,880 (100%)	20,880 (100%)	8,195 (39%)	2,628 (13%)	4,851 (23%)	16,650 (79%)	7,834 (38%)	1,602 (8%)
Shenzhen Chiwan Wharf	21,600 (100%)	21,600 (100%)	6,798 (31%)	6,510 (30%)	9,884 (46%)	11,222 (52%)	4,918 (23%)	3,868 (18%)
Guangdong Kelon Electrical	11,400 (100%)	11,400 (100%)	4,722 (41%)	2201 (19%)	2,157 (19%)	7,909 (69%)	4,521 (40%)	1,290 (12%)
BOE Tech Group	21,000 (100%)	21,000 (100%)	9,751 (46%)	9,578 (46%)	1,422 (7%)	4,434 (21%)	9,827 (47%)	6,988 (33%)
NE Electric Development	6,360 (100%)	6,360 (100%)	2,284 (36%)	339 (5%)	2,038 (32%)	5,776 (91%)	2,038 (32%)	245 (4%)
Guangdong Electric Power Development	21,480 (100%)	21,480 (100%)	8,960 (42%)	6,484 (30%)	3,583 (17%)	10,203 (48%)	8,937 (41%)	4,793 (22%)
	Average (A-H firms)		40.1%	13.4%	21.5%	77.2%	38.4%	9.4%
	Average (A-B firms)		36.9%	26.5%	27.9%	54.6%	35.4%	18.9%

Table 7 Panel A: Descriptive statistics on heavy machinery group

Matched AB-AH firms	p_t		r_t		X_t or Y_t		ΔX_t or ΔY_t	
	X	Y	X	Y	X	Y	X	Y
China Shipping Development (AH)								
Mean	8.396	6.443	6.443	-0.00012	0.00823	0.03218	-0.00081	0.00064
Median	8.19	6.35	6.35	0	0	0	0	0
Max	12.04	7.7	7.7	0.3	1	1	2	2
Min	5.69	5.15	5.15	-0.47	-1	-1	-2	-2
Std Dev	1.476	0.5549	0.5549	0.0142	0.9687	0.68197	1.30339	0.84559
Skewness	0.5336	0.2667	0.2667	-1.1036	-0.0164	-0.0402	-0.02595	-0.0592
Kurtosis	2.522	2.2957	2.2957	107.223	1.0659	2.1493	2.2821	3.3427
Shandong Zhenhua Port Machinery (AB)								
Mean	9.477	0.925	0.000	0.000	0.0035	0.0753	-0.0005	0.0014
Median	9.500	0.946	0.000	0.000	0.0000	0.000	0.000	0.000
Max	14.010	1.293	0.130	0.025	1.000	1.000	2.000	2.000
Min	5.340	0.583	0.400	-0.022	-1.000	-1.000	-2.000	-2.00
Std Dev	2.216	0.177	0.016	0.0014	0.979	0.791	1.323	1.067
Skewness	0.0592	0.022	0.843	0.3676	-0.007	-0.135	-0.0153	-0.117
Kurtosis	1.963	1.973	33.907	40.476	1.0418	1.6102	2.2389	2.749

Panel B: Correlation Matrix

China Shipping Development	p_t^X	p_t^Y	r_t^X	r_t^Y	X_t	Y_t	ΔX_t	ΔY_t
p_t^X	1							
p_t^Y	0.858307	1						
r_t^X	-0.007229	-0.008689	1					
r_t^Y	-0.000624	0.015783	-0.012836	1				
X_t	0.004533	-0.001206	0.553062	-0.001885	1			
Y_t	0.047601	0.015056	-0.007908	0.444665	0.008185	1		
ΔX_t	-0.000902	-0.000380	0.448942	-0.008421	0.672929	-0.000882	1	
ΔY_t	-0.001157	0.003321	-0.017484	0.522293	-0.010745	0.618276	-0.007614	1
Shandong Zhenhua Machinery								
p_t^X	1							
p_t^Y	0.844229	1						
r_t^X	-0.001983	-0.009850	1					
r_t^Y	0.002721	0.001662	0.048429	1				
X_t	-0.004521	-0.000277	0.592019	0.030745	1			
Y_t	-0.000965	-0.004582	0.046139	0.413852	0.034051	1		
ΔX_t	0.001524	0.001523	0.469936	-0.005841	0.674973	-0.007013	1	
ΔY_t	-0.000376	-0.000218	0.008468	0.366508	0.010626	0.675092	0.003562	1

Table 8: GMM estimation results

Heavy Machinery	$\hat{\phi}^{X}$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^{Y}$
China Shipping Development (A-H)					
Coefficient	0.000838	0.000288	0.088411	0.231042	0.006205
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Shandong Zhenhua Port Machinery (A-B)					
Coefficient	0.000796	0.000311	0.088458	0.092499	0.000166
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Liquor	$\hat{\phi}^{X}$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^{Y}$
Tsingtao Brewery (A-H)					
Coefficient	0.001861	6.12E-05	0.079854	0.087344	0.001671
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Yantai Changyu Winery (A-B)					
Coefficient	4.06E-03	0.001819	0.082055	0.066193	0.001668
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Transport	$\hat{\phi}^{X}$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^{Y}$
China Southern Airline (A-H)					
Coefficient	0.002576	0.001081	0.127089	1.77E-01	0.002391
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Hainan Airline (A-B)					
Coefficient	0.002056	6.40E-03	1.38E-01	0.036267	9.69E-05
p-value	(0.000)**	(0.000)**	(0.000)**	(0.001)**	(0.000)**
Pharmacy	$\hat{\phi}^{X}$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^{Y}$
Shandong Xinhua Pharmaceutical (A-H)					
Coefficient	0.001317	-0.00035	0.076999	2.33E-01	-1.72E-04
p-value	(0.000)**	(0.000)**	(0.000)**	(0.113)	(0.413)
Livzon Pharmaceutical (A-B)					
Coefficient	1.48E-03	6.86E-04	8.49E-02	0.105711	0.001378
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Textile	$\hat{\phi}^{X}$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^{Y}$
Jingwei Textile (A-H)					
Coefficient	0.00153	0.000658	0.095061	0.06626	-0.00039
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.095)
Mongolia Cashmere (A-B)					
Coefficient	0.001461	-0.00018	0.070886	0.060823	9.79E-05
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**

p-values in parentheses
*: 5% significant level
**: 1% significant level

Table 8 (cont): GMM estimation results

Telecom	$\hat{\phi}^X$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^Y$
ZTE Corporation (A-H)					
Coefficient	0.004099	0.000411	0.065591	0.094522	0.002104
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Eastern Communications (A-B)					
Coefficient	0.001488	0.00259	0.074699	0.062622	6.81E-05
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**

Utility	$\hat{\phi}^X$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^Y$
Huadian Power (A-H)					
Coefficient	0.002432	0.001128	0.17897	0.18228	0.001709
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Heilongjiang Energy (A-B)					
Coefficient	0.001529	-0.00027	0.083236	0.064494	9.33E-05
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**

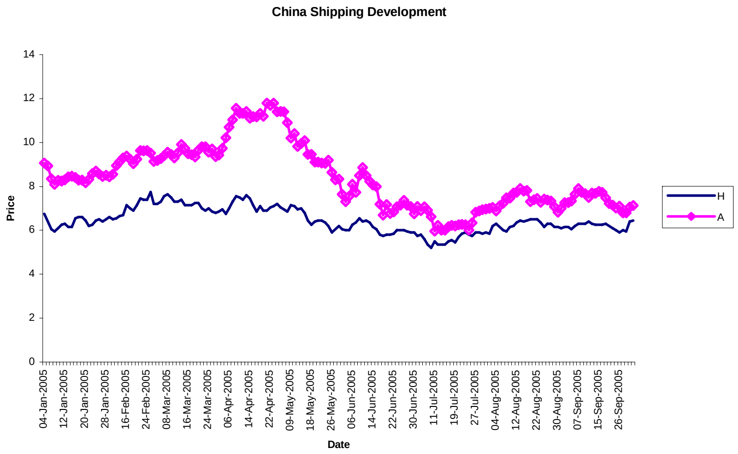
Shipyard	$\hat{\phi}^X$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^Y$
Guangzhou Shipyard (A-H)					
Coefficient	0.001456	0.000158	0.098398	0.072669	0.000795
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
Shenzhen Chiwan Wharf (A-B)					
Coefficient	0.003955	0.000197	0.082724	0.075364	0.001971
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.041)*

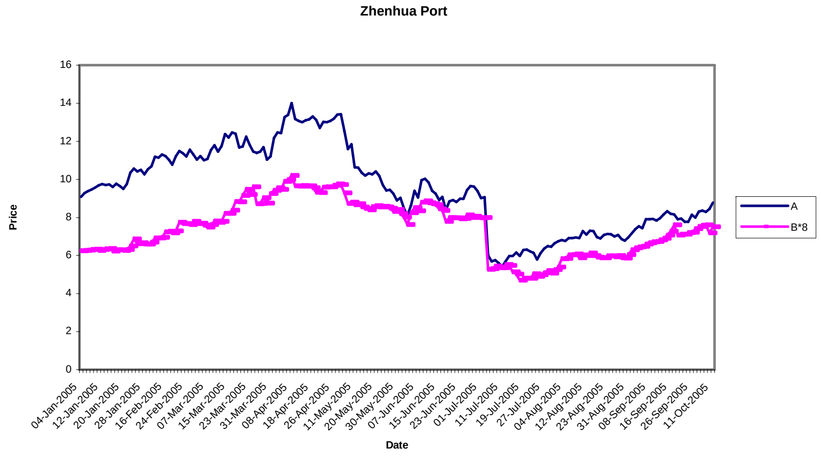
Electrical	$\hat{\phi}^X$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^Y$
Guangdong Kelon (A-H)					
Coefficient	0.001561	0.000724	0.068708	0.075617	0.00111
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
BOE Tech Group (A-B)					
Coefficient	0.000949	0.00442	0.116296	0.12537	0.003182
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**

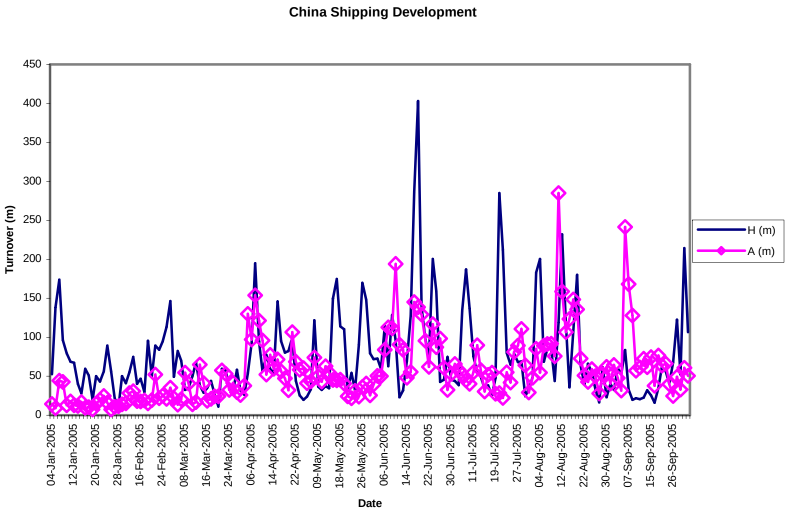
R&D	$\hat{\phi}^X$	$\hat{\theta}$	$\hat{\rho}$	$\hat{\delta}$	$\hat{\phi}^Y$
NE Electric Development (A-H)					
Coefficient	0.001842	0.000173	0.120581	0.051888	0.00026
p-value	(0.000)**	(0.000)**	(0.000)**	(0.030)*	(0.049)*
Guangdong Electric Power (A-B)					
Coefficient	0.002139	0.000513	0.066172	0.073904	0.00141
p-value	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**

p-values in parentheses
*: 5% significant level
**: 1% significant level

Figure 3a: Heavy Machinery







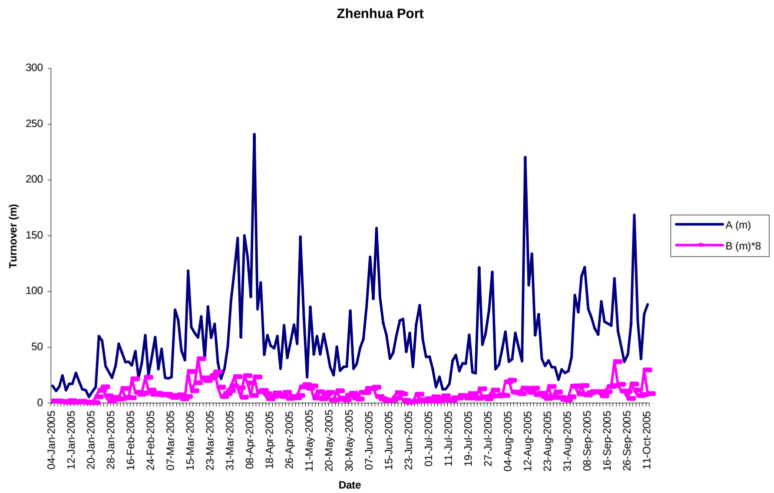
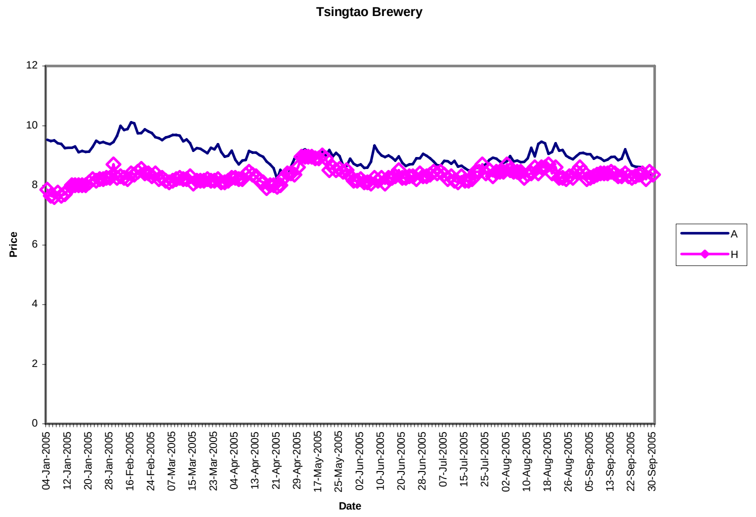
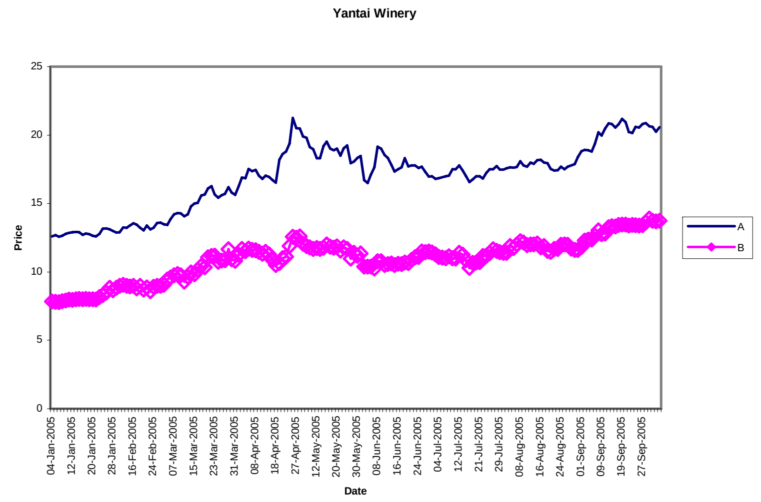
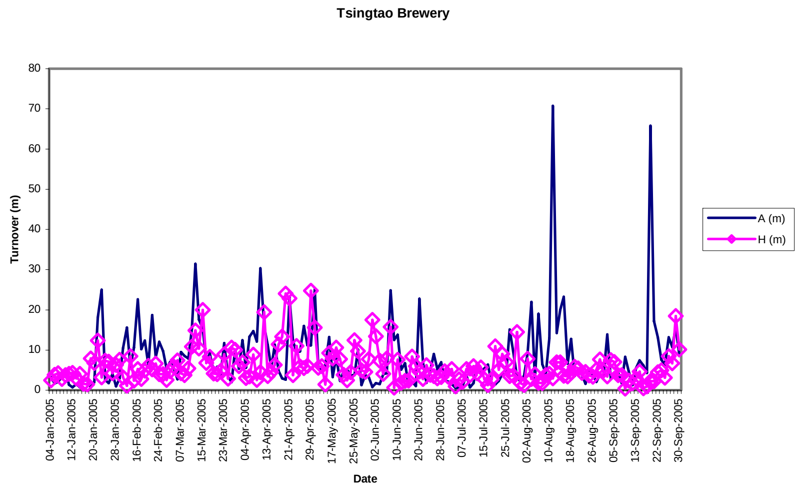


Figure 3b: Liquor







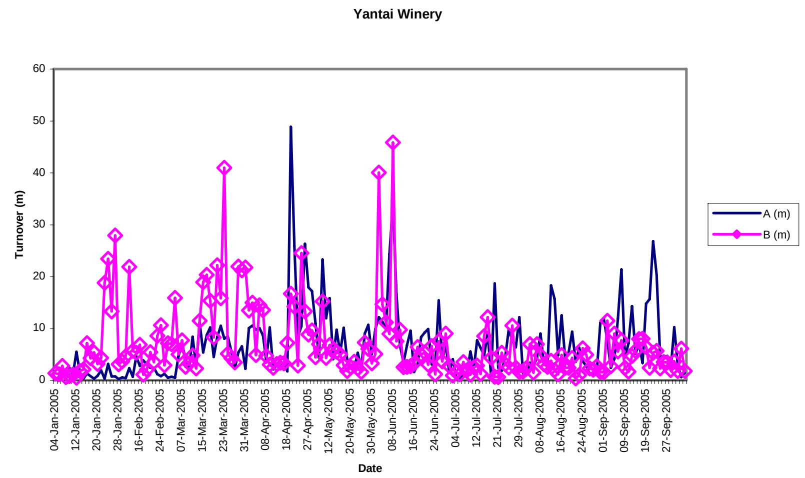
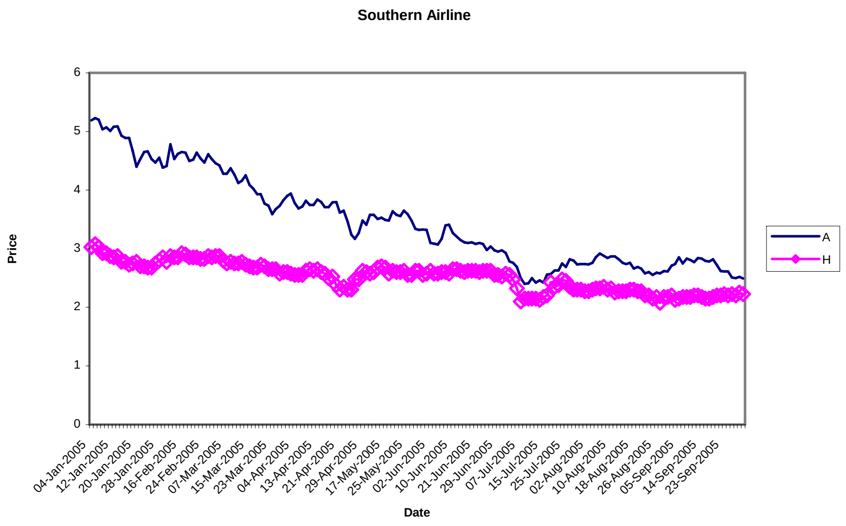
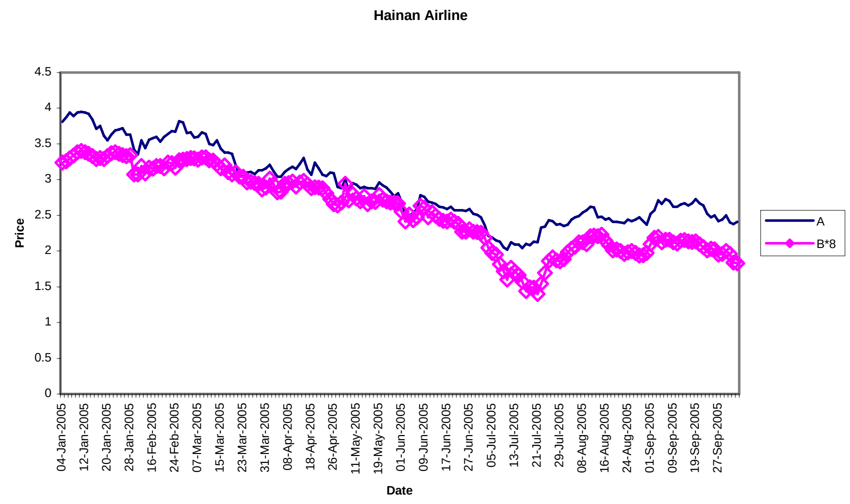
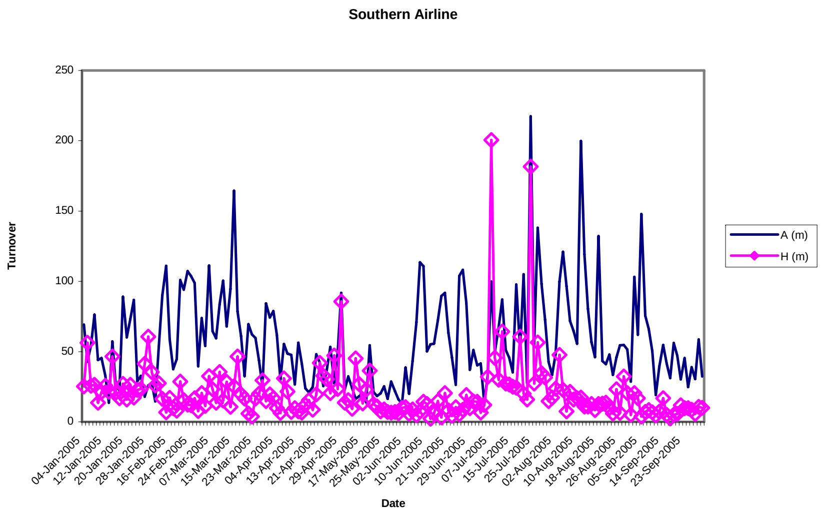


Figure 3c: Transportation







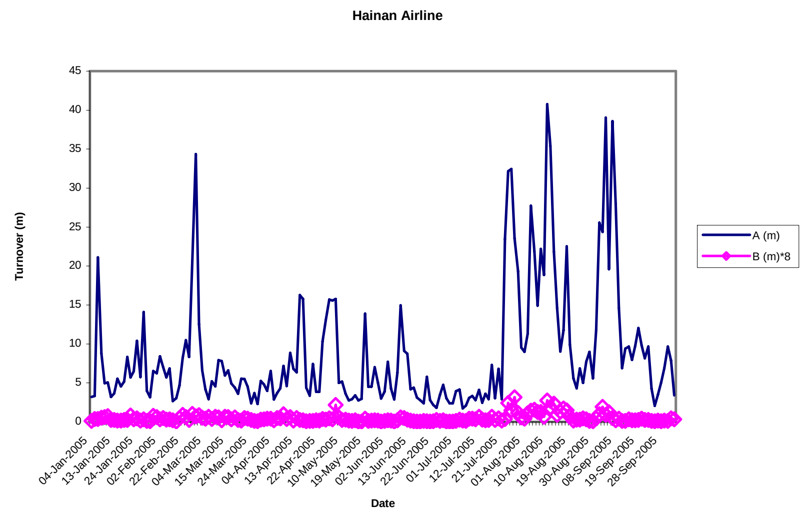
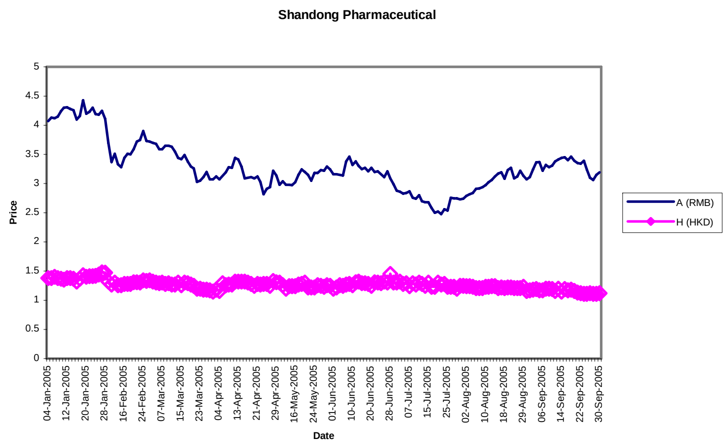
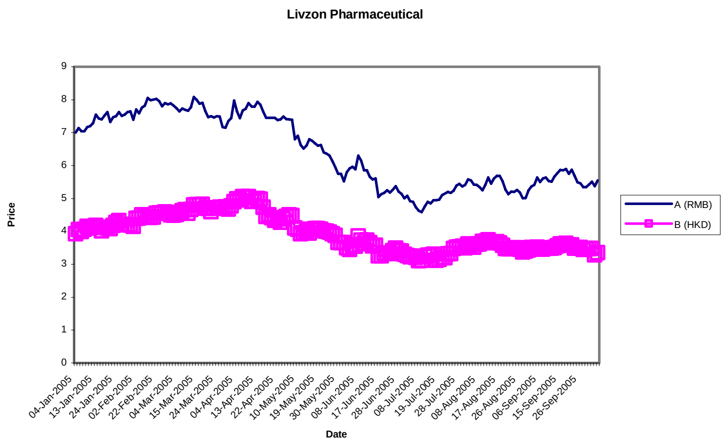
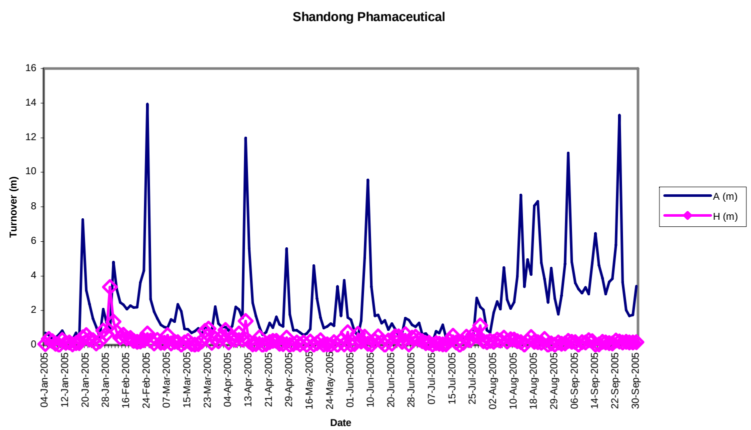


Figure 3d: Pharmaceutical







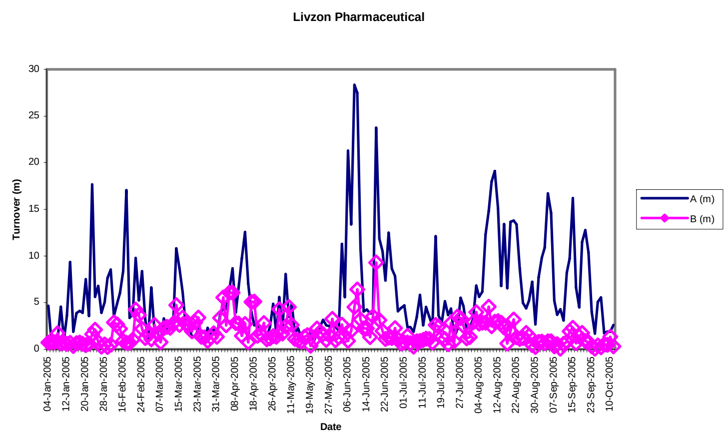
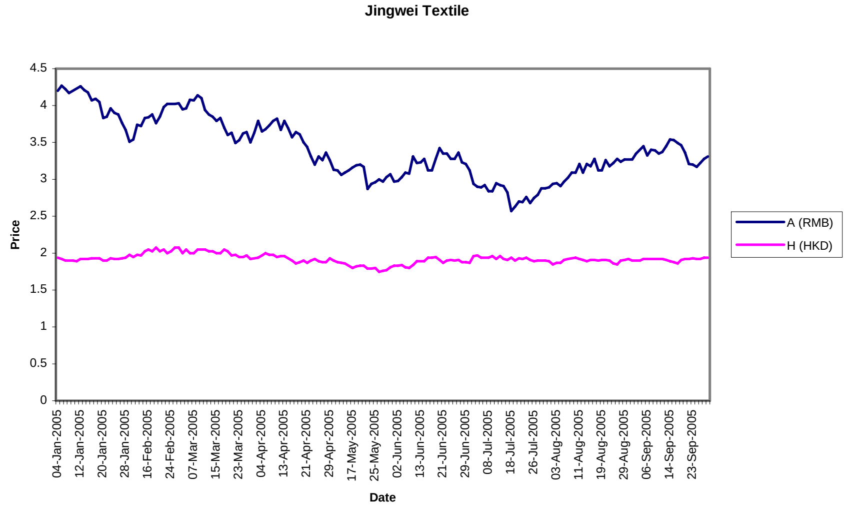
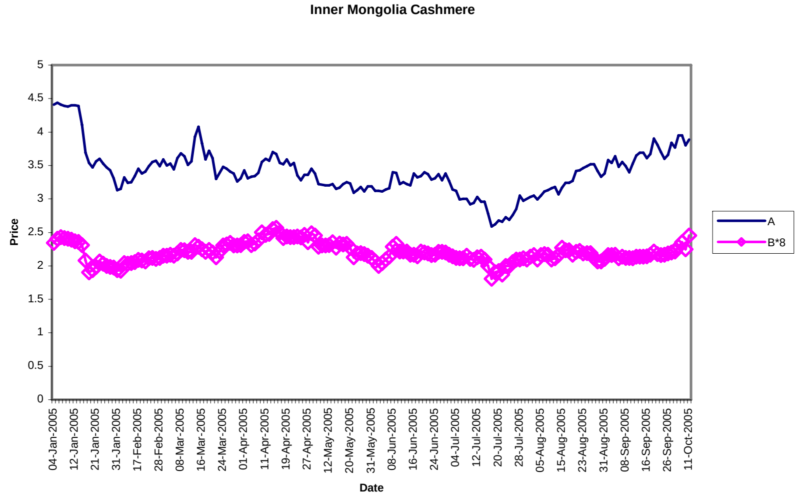
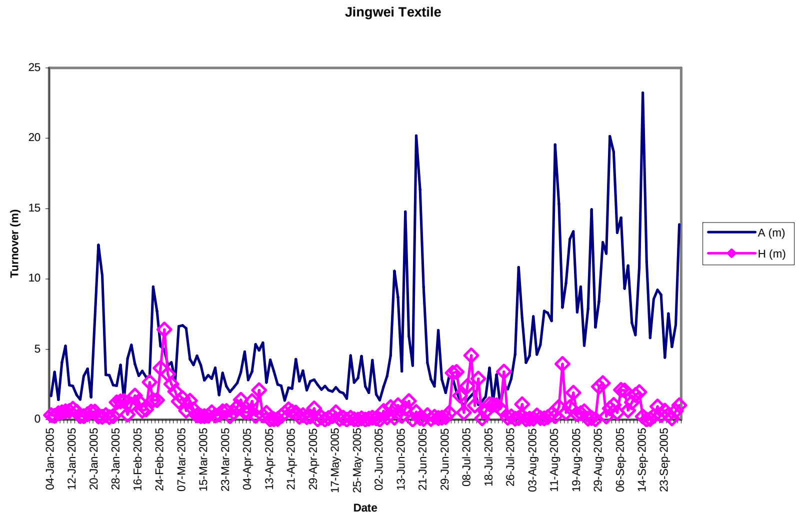


Figure 3e: Textile







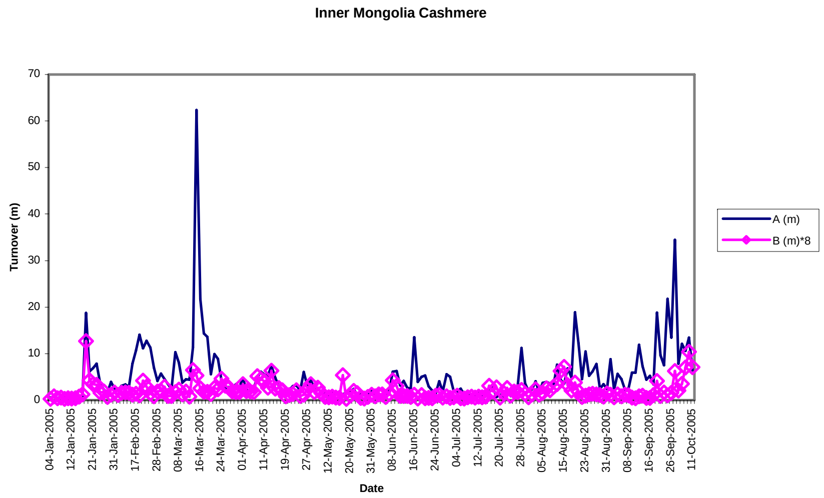
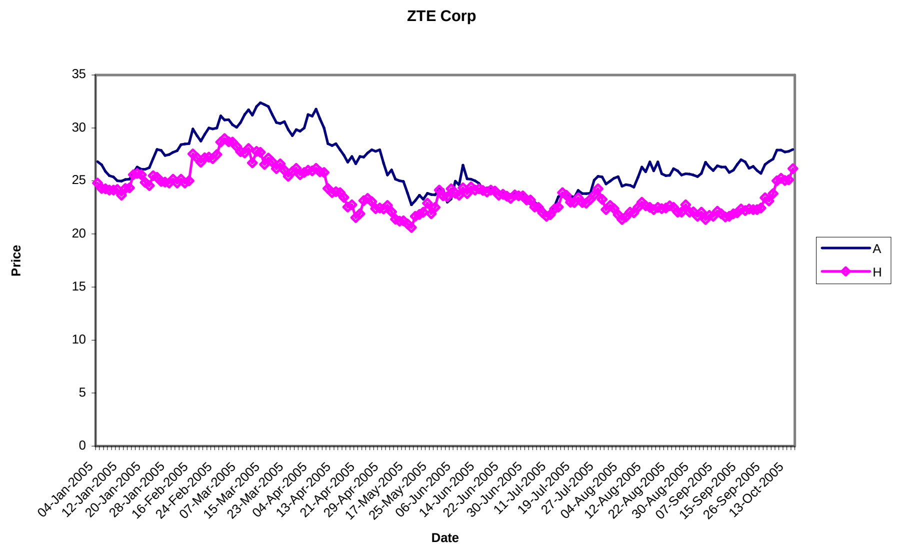
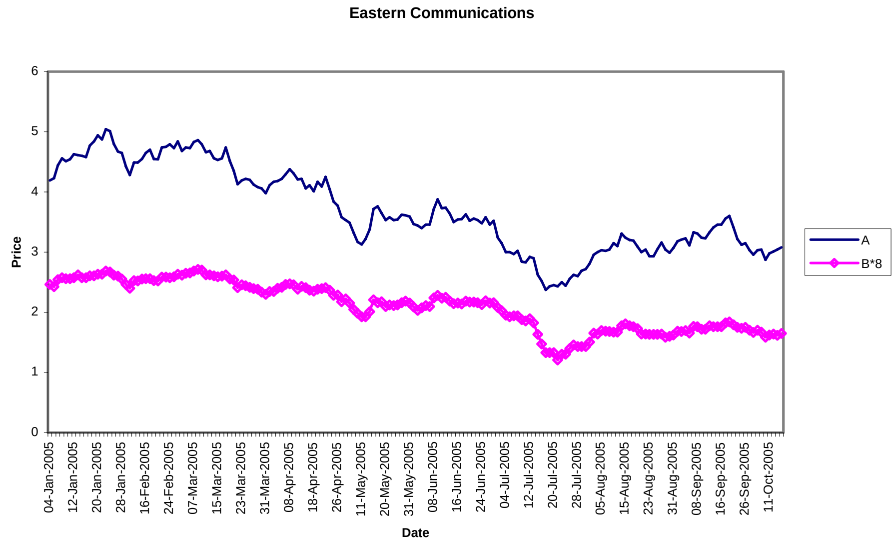
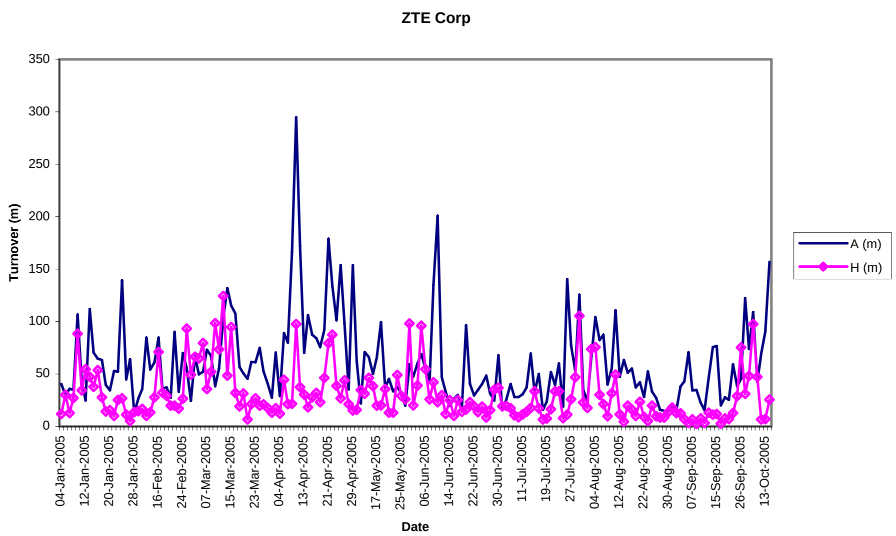


Figure 3f: Telecommunication







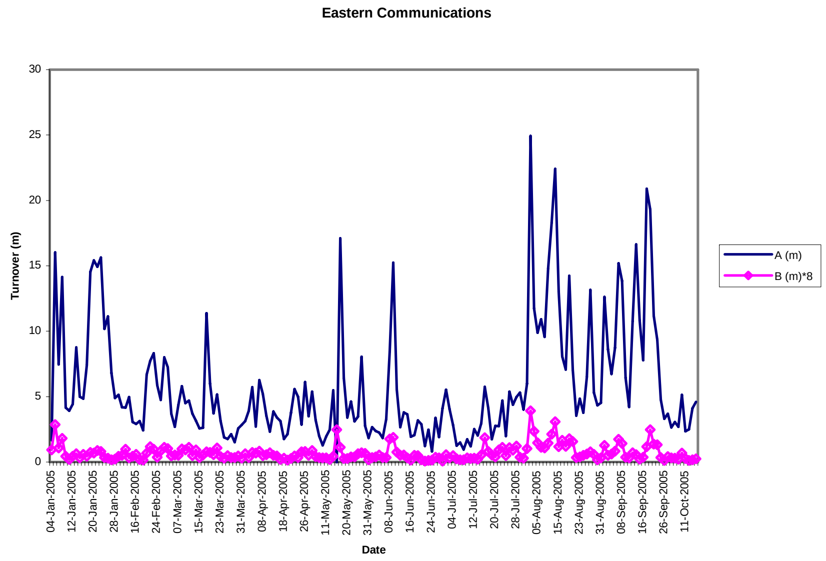
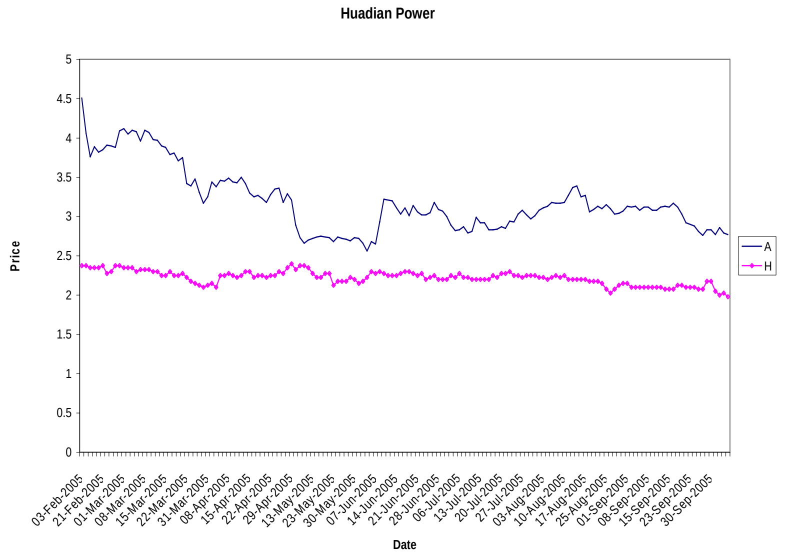
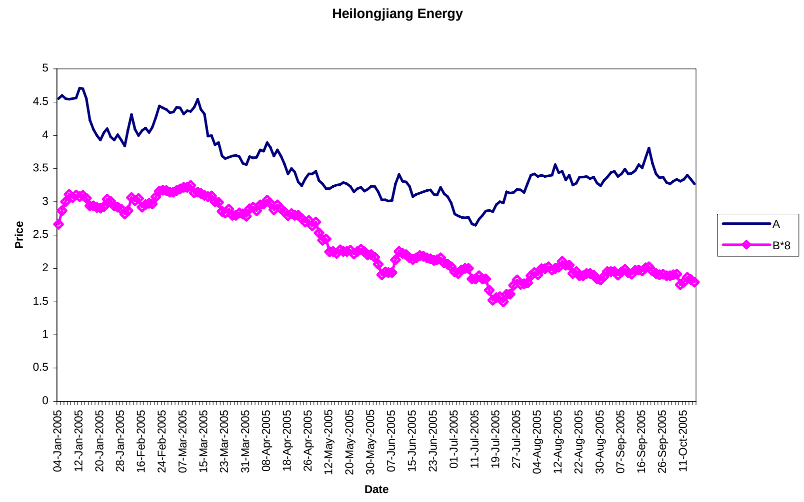
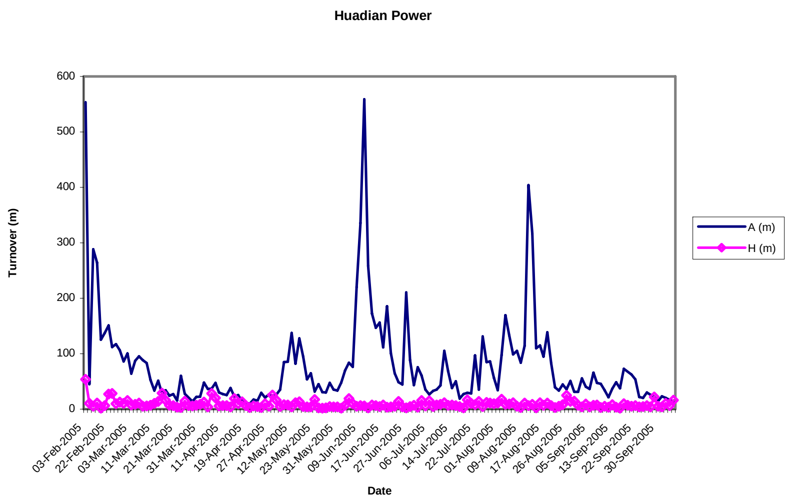


Figure 3g: Utility







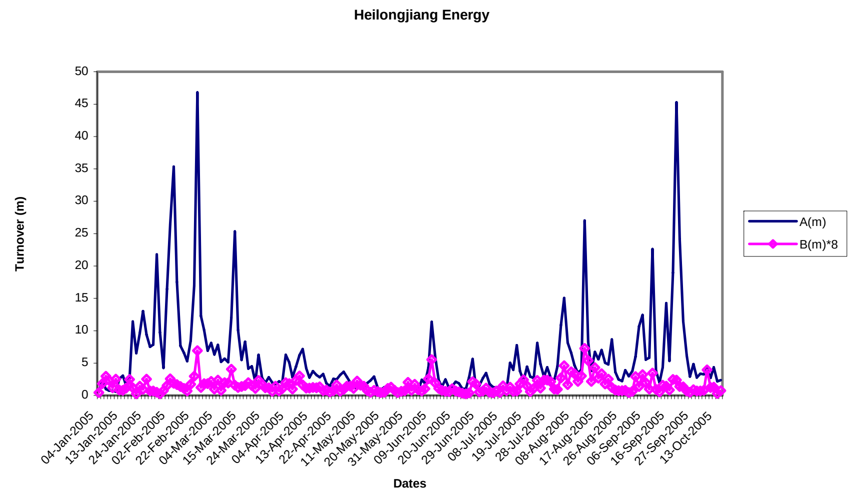
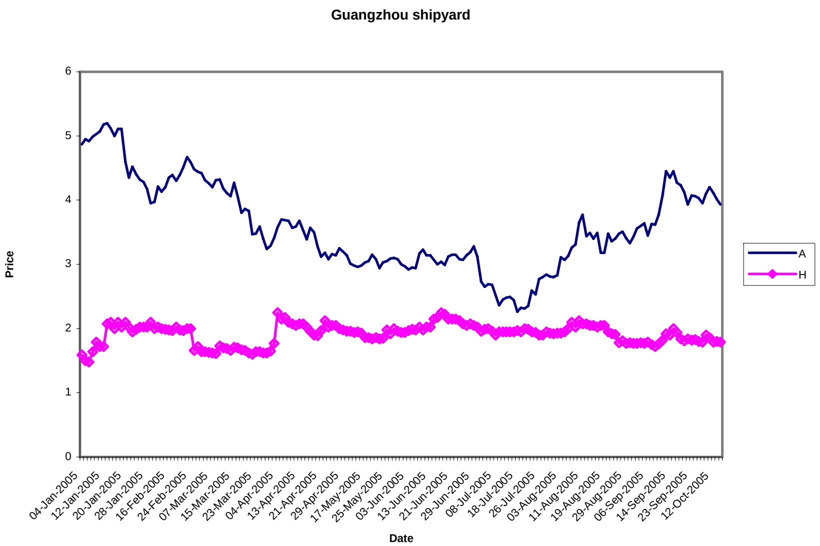
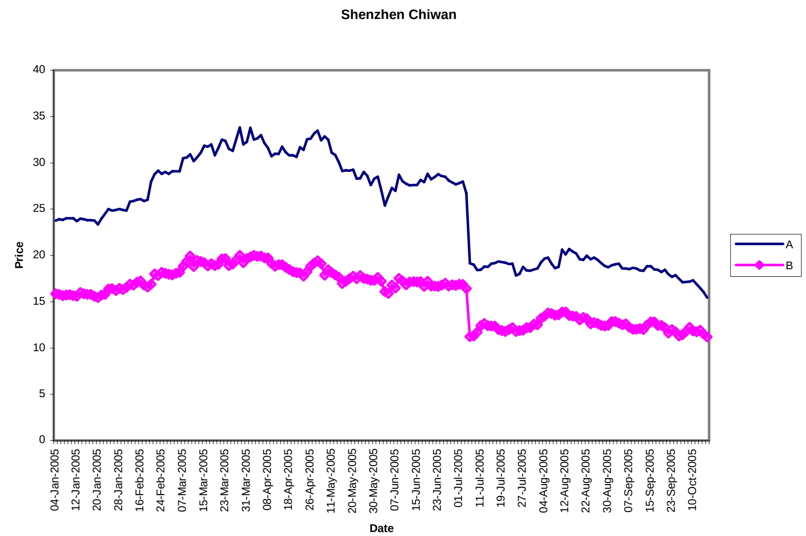
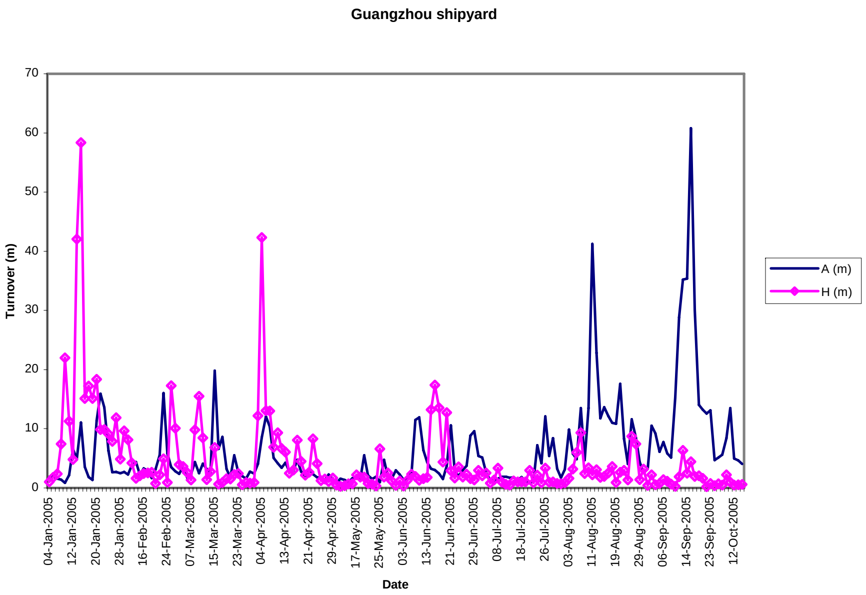


Figure 3h: Shipyard







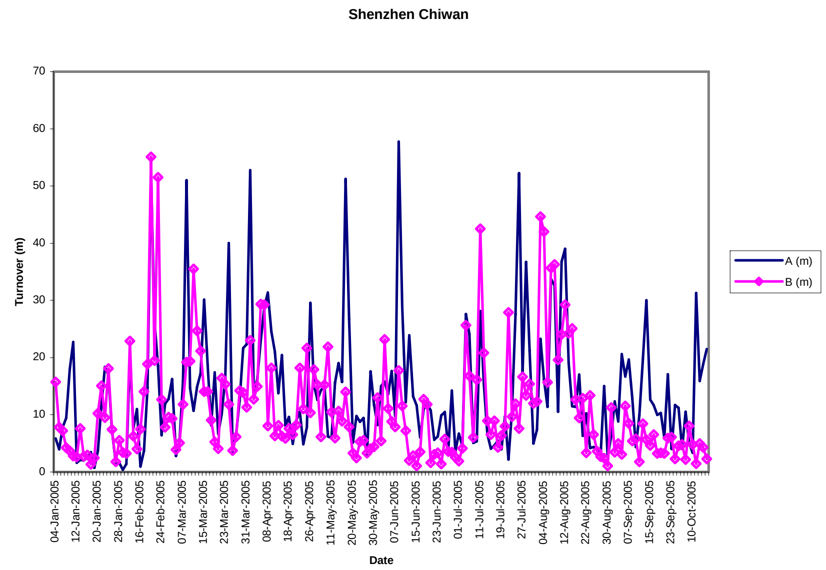


Figure 3i: Electrical

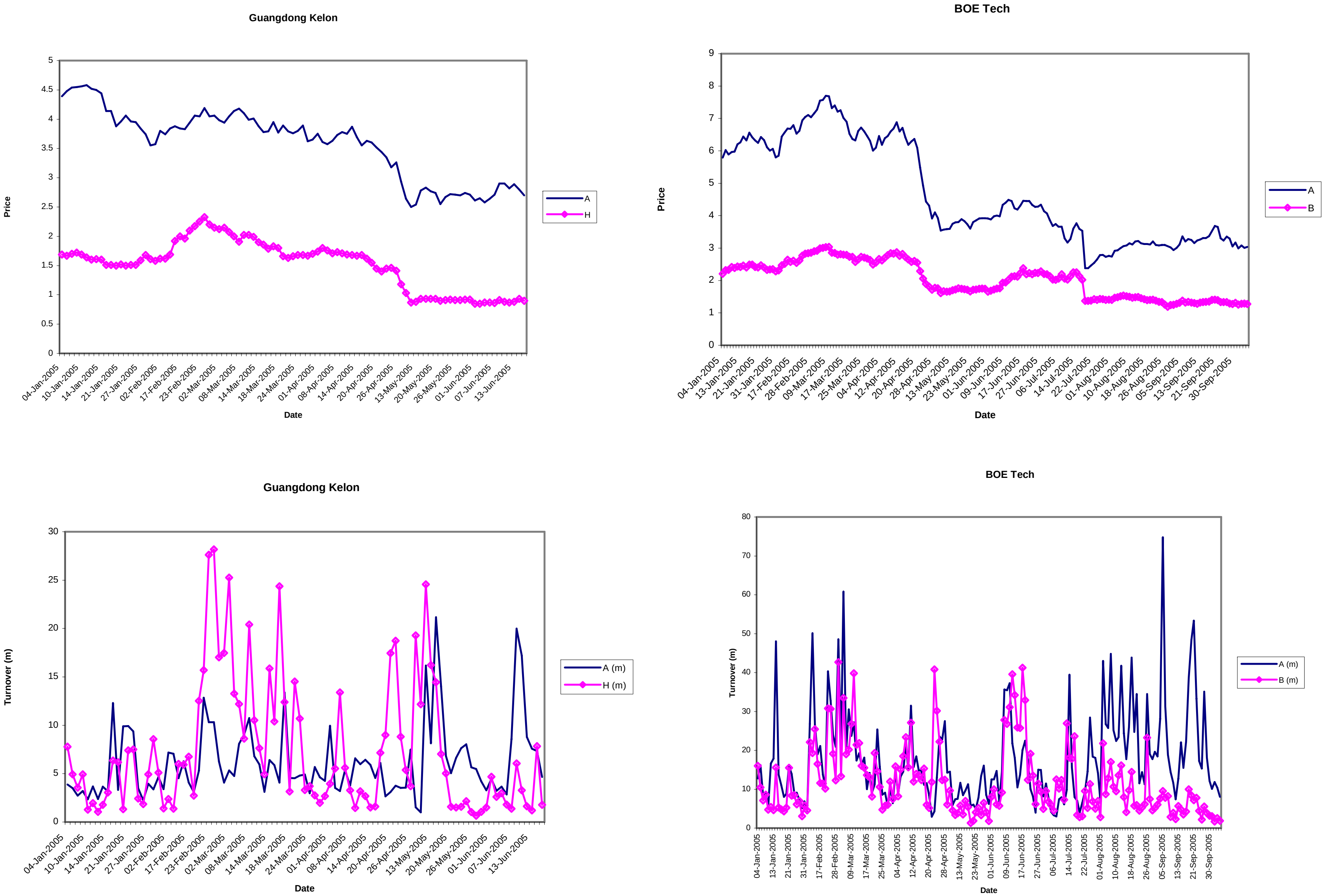
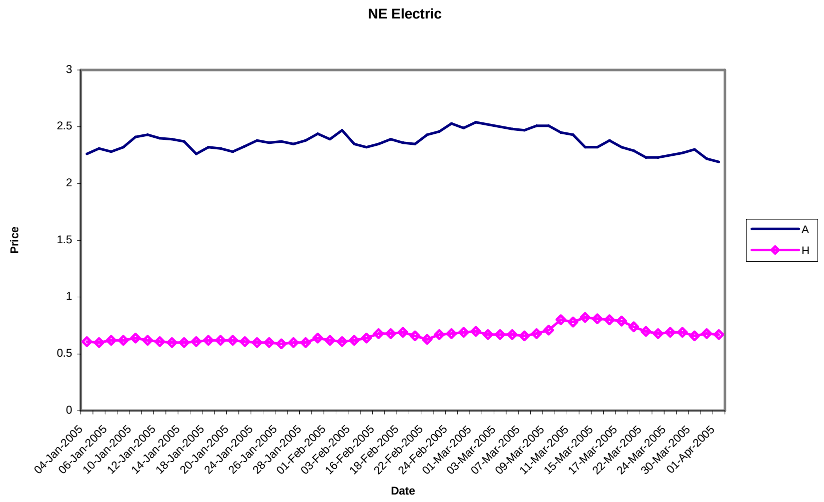
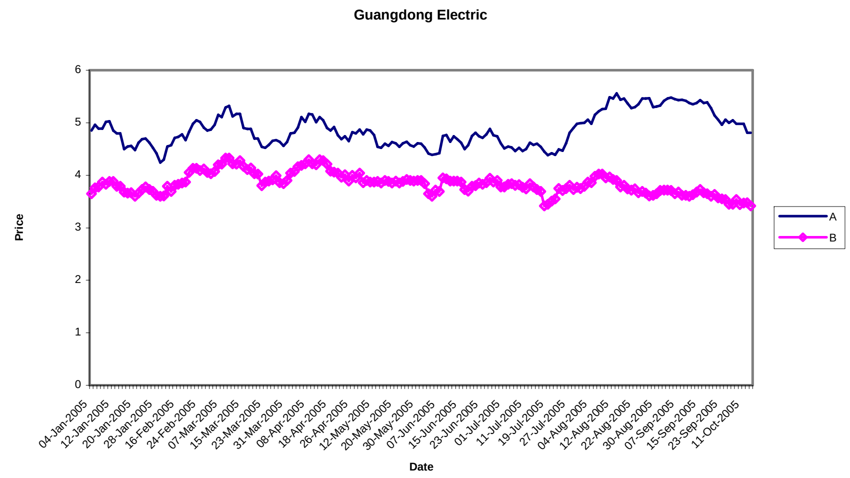
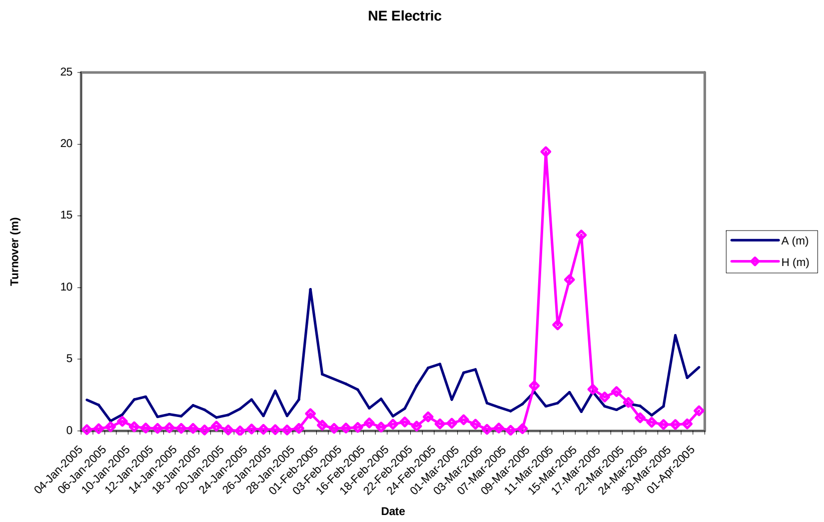


Figure 3j: R&D


NE Electric
Price
Date
A
H


Guangdong Electric
Price
Date
A
B


NE Electric
Turnover (m)
Date
A (m)
H (m)

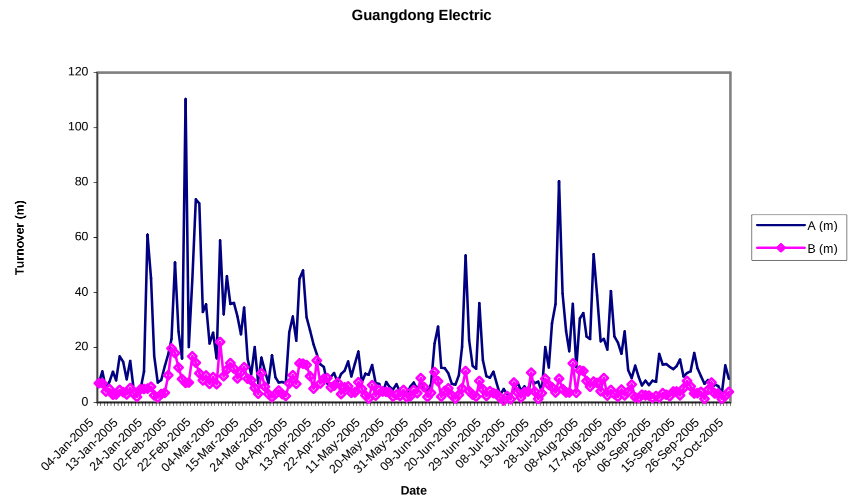

Guangdong Electric
Turnover (m)
Date
A (m)
B (m)

Figure 4: Proportion of trade directions on the various boards