EFA Eastern Finance Association

The Financial Review 38 (2003) 591–609

The Financial Review FR

Market Segmentation and Information Asymmetry in Chinese Stock Markets: A VAR Analysis

Jian Yang*

Prairie View A&M University

Abstract

This study examines the market segmentation and information asymmetry patterns in Chinese stock markets. The recursive cointegration analysis confirms that each of six markets is not linked with other markets in the long run. Further, the result from data-determined forecast error variance decomposition clearly shows that foreign investors in the Shanghai B-share market are better informed than Chinese domestic investors in two A-share markets and foreign investors in Shenzhen and Hong Kong markets over time. The finding challenges a widespread assumption of less informed foreign investors in the literature, but suggests that foreign investors could be more informed in emerging markets.

Keywords: market segmentation, information asymmetry, Chinese stock markets, directed acyclic graphs, forecast error variance decomposition

JEL Classifications: G15/G32

* *Corresponding author*: Department of Accounting, Finance and Information Systems, Prairie View A&M University, P.O. Box 638, Prairie View, TX 77446; Phone: (936) 857-4011; Fax: (936) 857-2797; E-mail: jian_yang@pvamu.edu

An earlier version of this paper was presented at Texas A&M University and the 2002 Financial Management Association annual meeting. The author thanks David A. Bessler, FMA session participants, seminar participants at Texas A&M University, and two anonymous referees for their many helpful comments.

1. Introduction

Chinese stock markets have grown rapidly and received much attention since their inception in the early 1990s. In particular, Chinese stock markets are the only equity markets covered by the International Finance Corporation that have completely segmented trading between domestic and foreign investors (Chui and Kwok, 1998). An A-share market is open only to Chinese domestic investors and a B-share market only to foreign investors. Baily (1994) first documented the big price discounts of foreign class B shares relative to comparable domestic A shares, which could be evidence for effective segmentation between A-share and B-share markets. Although segmented trading in two markets does not always result in segmented markets because cross-market informational links or arbitrage can work, Fung, Lee, and Leung (2000) provided supportive evidence for the argument of segmented A-share and B-share markets. However, Fung, Lee, and Leung's evidence is based on the use of stock return data, which does not allow for nonstationary Chinese stock prices (Long, Payne, and Feng, 1999) and the possibility of market integration in the long run (i.e., cointegration).

Furthermore, big discounts of B shares have been particularly puzzling in light of the fact that foreign shares typically trade at a price premium relative to domestic shares in other cases. In a comprehensive study of 11 countries with similar stock market segmentation structures, Bailey, Chung, and Kang (1999, p. 508) conclude that China is a "strange" case and "difficult to explain." More recently, Sun and Tong (2000) offer a new explanation to this puzzle. In particular, Sun and Tong propose a differential demand argument: that the existence of H-share and "red chip" stock markets provides good substitutes for B-share markets and thus makes demand for B shares quite elastic. In such cases, the number of shares outstanding of Chinese issuing firms for foreign investors can be too large to maintain reasonable prices. This argument, however, should be subject to further empirical tests.

Another interesting issue, likely to be more interesting and closely related to Chinese stock market segmentation, is the information asymmetry pattern in Chinese stock markets. The existence of information asymmetry (i.e., different information sets) on values of domestic assets between foreign investors and domestic investors has been well recognized in finance literature (Brennan and Cao, 1997; Kang and Stulz, 1997; Stulz and Wasserfallen, 1995). Some equilibrium pricing models of Chinese market segmentation (e.g., Chakravarty, Sarkar, and Wu, 1998) are based on the assumption of the information asymmetry pattern in Chinese stock markets. Many other works concerning Chinese stock markets, such as return volatility (Su and Fleisher, 1999) and initial public offerings (Mok and Hui, 1998), have also produced important implications for the information asymmetry issue, regarding whether foreign or domestic investors are better informed in these markets.

It is typically assumed in the finance literature that domestic investors are better informed than foreign investors about the value of local assets (Brennan and Cao, 1997; Kang and Stulz, 1997; Stulz and Wasserfallen, 1995). However, the issue of whether domestic or foreign investors have better information is controversial in the case of Chinese stock markets. Some researchers (Chui and Kwok, 1998; Mok and Hui, 1998) argue that foreign investors can be better informed than Chinese domestic investors about the value of Chinese stocks. The reasons include the Chinese government's control of mass media, the more advanced technology for processing and analyzing information by foreign investors, and higher corporate information disclosure requirements applicable to B shareholders. In contrast, others (Chakravarty, Sarkar, and Wu, 1998; Su and Fleisher, 1999) argue that foreign investors have less information on Chinese stocks than domestic investors. This could be due to language barriers, different accounting standards, and a lack of reliable information about the local economy and firms available to foreign investors. Lastly, Chen, Lee, and Rui (2001, p. 144–45) argue that no informational linkage exists, and thus no information asymmetry exists between A-share and B-share markets in Chinese stock markets.

This study employs a vector autoregression (VAR) analysis to address the issues of the A-share and B-share market segmentation/integration and the information asymmetry patterns in Chinese stock markets. The analysis also sheds light on the validity of recent explanations of price discounts of B shares relative to A shares by Sun and Tong (2000) and Chen, Lee, and Rui (2001). The study contributes to the literature in two aspects. First, the study addresses the possible segmentation or integration pattern in the long run among various stock markets as well as the stability of such a pattern. It is well known now that the long-run segmentation or integration pattern would be robust against short-run deviations, which differentiate this work from Fung, Lee, and Leung (2000). Furthermore, examining the temporal stability of the (long-run) market segmentation or integration pattern is equally important, particularly considering that Chinese stock markets were established only in the early 1990s and that the Chinese economy as a whole is in transition to a market economy. A recursive cointegration technique (Diamandis, Georgoutsos, and Kouretas, 2000; Hansen and Johansen, 1993), in addition to the standard cointegration technique, is employed in the study.

Second, an innovative directed graphs technique (Pearl, 1995; Spirtes, Glymour, and Scheines, 1993; Swanson and Granger, 1997) is used to discover the important contemporaneous causal pattern across various Chinese stock markets.[1] The introduction of such a method is essential to make causal inferences without the requirement of time lags (as is required by Granger-type causality). In this study, following the recent advance in VAR analysis (Bessler and Yang, 2003; Haigh and Bessler, in press; Swanson and Granger, 1997), the contemporaneous causal pattern as discovered by directed graphs suggests a specific and data-determined pattern for the structural decomposition of VAR residuals. This enables us to go beyond the serious limitation

[1] The contemporaneous causal patterns may be inherent in the strong contemporaneous correlations of VAR residuals (Swanson and Granger, 1997). The strong contemporaneous correlations of market innovations have been well documented in the United States and international stock markets (Bessler and Yang, 2003; Eun and Shim, 1989; Hasbrouck, 1995; Yang, Min, and Li, in press).

of a recursive causal structure embedded in the widely used Choleski decomposition and substantially improves the VAR analysis (i.e., forecast error variance decomposition) in this study (Bernanke, 1986; Cooley and LeRoy, 1985; Swanson and Granger, 1997). The rest of this paper is organized as follows. Section 2 describes the data used in this study. Section 3 outlines the empirical framework. Section 4 presents empirical results, and section 5 discusses conclusions.

2. Data

China has two stock exchanges, the Shanghai Securities Exchange and the Shenzhen Securities Exchange. Both exchanges trade A shares limited to domestic investors and B shares limited to foreign investors. A Chinese company can also issue H and N shares traded in Hong Kong and New York, respectively. N shares, listed in the New York Stock Exchange (in the form of ADRs), have a very thin market (Sun and Tong, 2000) and thus are not included in this study. There are also some state-backed Chinese companies incorporated and listed in Hong Kong that have acquired a nickname: “red chips.” Sun and Tong (2000) argued that red chip stocks can be a substitute for B shares to foreign investors.

In this study, a comprehensive data set including all relevant Chinese stock market price indexes is used. The six Chinese stock market price indexes under study are Shanghai A-share (SHA), Shanghai B-share (SHB), Shenzeng A-share (SZA), Shenzeng B-share (SZB), Hong Kong H-share (HKH), and red chip stock (HKR) market price indexes. As discussed in Poon and Fung (2000) and Sun and Tong (2000), although both Hong Kong H-share and red chip stocks are China-backed securities, the companies included in the H-share stock index are typically different from those included in the red chip stock index, and market reaction to these two types of stocks also appears to be different (Poon and Fung, 2000, p. 317). The data for this study cover a six-year period, from January 2, 1995 to December 29, 2000, totaling 1,565 daily observations for each stock index price series. The selected period follows the unification of the dual exchange rates of the Chinese currency renminbi (RMB) and implementation of a managed-float system on January 1, 1994. This timeframe effectively avoids the early period of the Chinese stock markets, when erratic events occurred in the markets’ infancy stage (Sun and Tong, 2000). The data are provided by Datastream. All data series are also converted into U.S. dollars using appropriate exchange rates.

Two standard procedures are applied to establish the nonstationarity of each individual series. The null hypothesis for both procedures is that a unit root exists. One is the augmented Dickey-Fuller test (Dickey and Fuller, 1981), and the second procedure is the Phillips-Perron test (Phillips and Perron, 1988). Table 1 reports results for I(1) versus I(0) (level prices), applying the two tests. Similar to Long, Payne, and Feng (1999), we find that there is one unit root in each of the stock price indexes under study, but no unit root in their first differences (not reported here) at the 5% significance level.

Table 1

Results of unit root tests

This table reports the results on the augmented Dickey-Fuller (Dickey and Fuller, 1981) and Phillips-Perron (Phillips and Perron, 1988) unit root tests. The numbers of lags are selected according to the Akaike Information Criterion. The critical values of the augmented Dickey-Fuller unit root tests without trend and with trend are –2.86 and –3.41 at the 5% level, respectively. The critical values of the Phillips-Perron unit root tests without trend and with trend are –14.1 and –21.7 at the 5% level, respectively. ADF = augmented Dickey-Fuller test. PP = Phillips-Perron test. SHA = Shanghai A-share market price index. SHB = Shangai B-share market price index. SZA = Shenzeng A-share market price index. SZB = Shenzeng B-share market price index. HKH = Hong Kong H-share market price index. HKR = red chip stock market price index.

	Without trend		With trend	
Market	ADF	PP	ADF	PP
SHA	−0.36	−0.56	−2.59	−16.60
SHB	−1.01	−4.51	−0.72	−3.37
SZA	−0.59	−1.16	−2.21	−11.59
SZB	−1.94	−7.33	−2.01	−7.72
HKH	−1.68	−6.12	−2.79	−13.27
HKR	−1.78	−5.07	−1.76	−4.95

3. Empirical framework

Let X_t denote a vector which includes the price indexes (p) for the six markets under consideration ($p = 6$), and it can be modeled in an error correction model (ECM):

$$\mathrm{H}_0\colon \Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu + e_t (t = 1, \ldots, T) \qquad (1)$$

Equation (1) resembles a VAR model in first differences, except for the presence of the lagged level of X_{t-1}. Its associated parameter matrix, Π, contains information about the long-run (cointegrating) relations among p variables. There are three cases of interest: (a) if Π is of full rank, then X_t is stationary in levels, and a VAR in levels is an appropriate model; (b) if Π has zero rank, then it contains no long-run information, and the appropriate model is a VAR in differences; and (c) if the rank of Π is a positive number, r, and is less than p, there exist matrices α and β, with dimensions $p \times r$, such that $\Pi = \alpha\beta'$. In such a case, $\beta' X_t$ is stationary, even though X_t is not.

The cointegration analysis in this study employs the procedure developed by Johansen (1991). We jointly test the hypotheses of A-share and B-share market segmentation/integration and the differential demand argument (Sun and Tong, 2000) based on cointegration analysis. This can be done by examining the number of cointegrating vectors in a six-market system, r, as follows:

$$H_1(r)\colon \Pi = \alpha\beta' \qquad (2)$$

A trace test is conducted to determine r. The null hypothesis for the trace test (Johansen, 1991) is that there are at most $r(0 \leq r \leq p)$ cointegrating vectors or long-run relations. Whether segmented trading between A- and B-share markets results in informationally segmented markets in the long run is to be tested, which is equivalent to testing no cointegration between A-share and B-share markets on both Shanghai and Shenzhen exchanges.[2] Also, the differential demand argument suggests that there is at least one long-run relation among two B-share markets, H-share, and red chip stock markets. These predictions should be captured by the number of cointegrating vectors in the six-market system. Of course, if we find one or more long-run relations in the system, we should proceed to examine whether the evidence is simply for the differential demand argument or against the A-share and B-share market segmentation argument. In other words, we need to further test whether the identified long-run relation is between A-share and B-share markets or among B-share, H-share and red chip stocks. Incidentally, this last step is not needed in this study as shown below.

To examine the stability of the identified (non)cointegration relation, we apply a recursive cointegration technique to test constancy of cointegration rank as described in Hansen and Johansen (1993). This rank constancy test has been recently used in other studies such as Diamandis, Georgoutsos, and Kouretas (2000) to serve a similar purpose. Hansen and Johansen (1993) suggest that the rank constancy test can be done under the two VAR representations of equation (1). In the "Z-representation" all the parameters of the ECM are reestimated during the recursive estimations. Whereas under the "R-representation" the short-run parameters Γ_i are fixed to their full sample values, only the long-run parameters in equation (1) are reestimated. Hansen and Johansen also remarked that the results from the R-representation are more relevant in recursive cointegration analysis.

In the case of no cointegration (which is applicable in this study), the first difference VAR with the appropriate lags can be estimated and used to summarize dynamic influence each market has on other markets in the six-market system. However, the individual dynamics coefficients of the VAR are hard to interpret, and the strength of these dynamic relations remains unclear. Thus, forecast error variance decomposition is conducted to summarize dynamic interactions among the six market prices. The forecast error variance decomposition provides a direct test on the information asymmetry pattern in the short-run dynamic sense. If a market informationally leads others, this market's returns should most significantly be explained by its own market innovations and not as much by other markets' innovations. Instead, innovations from this market should be able to significantly explain other market returns.

[2] As pointed out by a referee, Chinese stock markets could be segmented in trade, but still connected informationally. However, the extent to which these Chinese stock markets segmented in trade are linked informationally in the long run ultimately should be a matter of empirical testing.

The method for treating contemporaneous innovation correlation is critical to such a forecast error variance decomposition (Bernanke, 1986; Cooley and LeRoy, 1985; Swanson and Granger, 1997). The VAR analysis in the literature most often relies on a Choleski decomposition to achieve a just-identified system in contemporaneous time. A big problem of implementing the Choleski decomposition is a restrictive and often unrealistic assumption of the existence of a recursive contemporaneous causal structure (Bernanke, 1986; Cooley and LeRoy, 1985; Swanson and Granger, 1997). In this study, we follow a very flexible structural decomposition commonly referred to as the "Bernanke ordering," which allows for nonrecursive contemporaneous causal structure (Bernanke, 1986). The structural decomposition requires writing the innovation vector (u_t) from the estimated error correction model as: $\mathbf{A}u_t = v_t$, where $\mathbf{A}$ is a 6×6 matrix and v_t is a 6×1 vector of orthogonal shocks in this study. We apply a directed graph algorithm (i.e., PC algorithm) as given in Spirtes, Glymour, and Scheines (1993) to place zeros on the $\mathbf{A}$ matrix. A similar procedure was also recently employed by Swanson and Granger (1997), Bessler and Yang (2003), and Haigh and Bessler (in press).[3]

As is well known, Granger causality exploits the time sequence asymmetry that a cause precedes its associated effect (and thus an effect does not precede its cause). By contrast, the literature of *directed graphs* (Pearl, 1995; Spirtes, Glymour, and Scheines, 1993; Swanson and Granger, 1997) builds on the insight of a nontime sequence asymmetry in causal relations. Consider a causally sufficient set of three variables X, Y, and Z. A causal fork that X causes Y and Z can be illustrated as: Y ← X → Z. Here the unconditional association between Y and Z is nonzero (as both Y and Z have a common cause in X), but the conditional association between Y and Z, given knowledge of the common cause X, is zero: common causes screen off associations between their joint effects. Illustrate the inverted causal fork, X and Z cause Y, as: X → Y ← Z. Here the unconditional association between X and Z is zero, but the conditional association between X and Z, given the common effect Y, is not zero: common effects do not screen off association between their joint causes.

Essentially, a directed graph is an assignment of contemporaneous causal flow (or lack thereof) among a set of variables (vertices) based on observed correlations and partial correlations. The edge relation characterizing each pair of variables represents the causal relation (or lack thereof) between them. In the context of this study, no edge (XY) indicates (conditional) independence between two variables, whereas an undirected edge (X—Y) signifies a covariance that is given no particular causal interpretation. A directed edge (Y → X) suggests that a variation in Y, with all other

[3] The procedure of Swanson and Granger (1997) is essentially a special case of more general directed graph algorithms (e.g., the PC algorithm) as discussed in Spirtes, Glymour, and Scheines (1993). One of the major differences between these two algorithms is that the PC algorithm calculates and tests up to k-2 order conditional correlations for k variables whereas Swanson and Granger only tests up to first order conditional correlations. Readers are referred to Spirtes, Glymour, and Scheines (1993); Pearl (1995); and Bessler and Yang (2003) for a more complete account of directed graphs.

variables held constant, causes a (linear) variation in X that is not mediated by any other variable in the system. A bidirected edge ($X \leftrightarrow Y$) indicates the bidirection of causal interpretation between the two variables.

Spirtes, Glymour, and Scheines (1993) provided a directed graph algorithm (i.e., PC algorithm) for removing edges between markets and directing the causal flow of information between markets. The algorithm begins with a complete undirected graph, where innovations from every market are connected with innovations with every other market. The algorithm removes edges from the complete undirected graph by first checking for (unconditional) correlations between pairs of variables. Edges connecting variables having zero correlation are removed. Remaining edges are then checked for first order partial correlation (correlation between two variables conditional on a third variable) equal to zero. Similarly, edges connecting variables having zero first order conditional correlation are removed. Edges that survive this check of first order conditional correlation are then checked against zero second order conditional correlation, and so on. For N variables, the algorithm continues to check up to $N - 2$ order conditional correlation.

In applications, Fisher's z statistic is used to test whether conditional correlations are significantly different from zero. Fisher's z statistic can be applied to test for significance from zero; where $z([i, j \mid k]n) = 1/2(n - |k| - 3)^{1/2} \times \ln\{(|1 + [i, j \mid k]|) \times (|1 - [i, j \mid k]|)^{-1}\}$ and n is the number of observations used to estimate the correlations, $(i, j \mid k)$ is the population correlation between series i and j conditional on series k (removing the influence of series k on each i and j), and $|k|$ is the number of variables in k (that we condition on). If i, j, and k are normally distributed and $r(i, j \mid k)$ is the sample conditional correlation of i and j given k, then the distribution of $z([i, j \mid k]n) - z(r[i, j \mid k]n)$ is standard normal.

The notion of *sepset* is further used to assign the direction of contemporaneous causal flow between variables that remain connected after all possible conditional correlations have been passed as nonzero. The sepset of a pair of variables whose edge has been removed is the conditioning variable(s) on the removed edge between two variables. For vanishing zero order conditioning (unconditional correlation), the sepset is the empty set. Edges remaining connected are directed by considering triples X ---- Y ---- Z, such that X and Y are adjacent as are Y and Z, but X and Z are not adjacent. Direct the edges between triples X ---- Y ---- Z as X ---- > Y <---- Z if Y is not in the sepset of X and Z. If X ----> Y, Y and Z are adjacent, X and Z are not adjacent, and there is no arrowhead at Y, then Y ---- Z should be positioned as Y ----> Z. If there is a directed path from X to Y, and an edge between X and Y, then X ---- Y should be positioned as X ----> Y. More illustration of the notion of the sepset concept and its use will be provided in the following section on the empirical results.

4. Empirical results

We first select the optimal lag in equation (1) by minimizing the Akaike Information Criterion. The maximum lag is set at 15 (three trading weeks). The Akaike

Table 2

Johansen trace test statistics for Chinese stock markets

We compare trace test statistics (T) to the critical values (C) from Osterwald-Lenum (1992; Tables 1.1∗ and 1.1). We determine the lags in the underlying VARs by the consideration of the minimization of the Akaike Information Criterion and by the absence of significant serial corrections in the residuals. The use of 10% critical values yields qualitatively the same conclusion.

	Without linear trend		With linear trend	
H_0:	T	C (5%)	T	C (5%)
$r = 0$	90.80	102.14	82.83	94.15
$r = 1$	49.10	76.07	43.40	68.52
$r = 2$	29.63	53.12	23.99	47.21
$r = 3$	13.67	34.91	8.25	29.68
$r = 4$	6.26	19.96	4.32	15.41
$r = 5$	2.34	9.24	1.51	3.76

Information Criterion suggests the optimal lag of $k = 12$ for the level VAR and thus $k = 11$ for the ECM and the first difference VAR. This lag is also the only lag length that does not produce the autocorrelation problem of VAR residuals. The larger lag can be attributed to segmented trading of four A and B share markets and the relative newness of Chinese domestic investors to information processing.

The Johansen (1991) trace test results are reported in Table 2. We fail to reject that there is a zero cointegrating vector either with a constant included in the cointegrating space or with a linear trend at the 5% significance level (as well as at the 10% significance level, as confirmed in Figure 1). Also, residuals on the ECM estimation are reasonably well behaved. Lagragian multiplier-type tests on fourth order autocorrelation on residuals (chi-squared tests) can only reject the null of white noise residuals at a p-value of 0.15. There are some autoregressive conditional heteroskedasticity (ARCH) effects on each of the six market innovations, which, however, do not seriously affect our conclusion on the cointegration rank (Gonzalo, 1994).

Applying Hansen and Johansen's (1993) recursive cointegration analysis, we show in Figure 1 normalized trace tests calculated at each data point over the period from January 1, 1996 through December 29, 2000. The one-year period in 1995 is used as the base period. Statistics in Figure 1 are normalized by the 10% critical values given in Osterwald-Lenum (1992; figure entries greater than 1.0 indicate that the null hypothesis is rejected at that data point). It is clear that no cointegration exists over most of the sample period, because our plots of the trace test statistics under both representations are all below the line 1.0 after the second half of the year 1997. This is most obvious in the R-representation, in which only a few very short peaks go beyond the line 1.0 over the sample period. Nevertheless, cointegration might exist among some or all of the six markets for some short periods prior to the first half of year 1997. In sum, the results of recursive cointegration analysis confirm that the first difference VAR with 11 lags would be an appropriate specification (i.e., the case b of equation (1)), particularly in the light of the R-representation results.

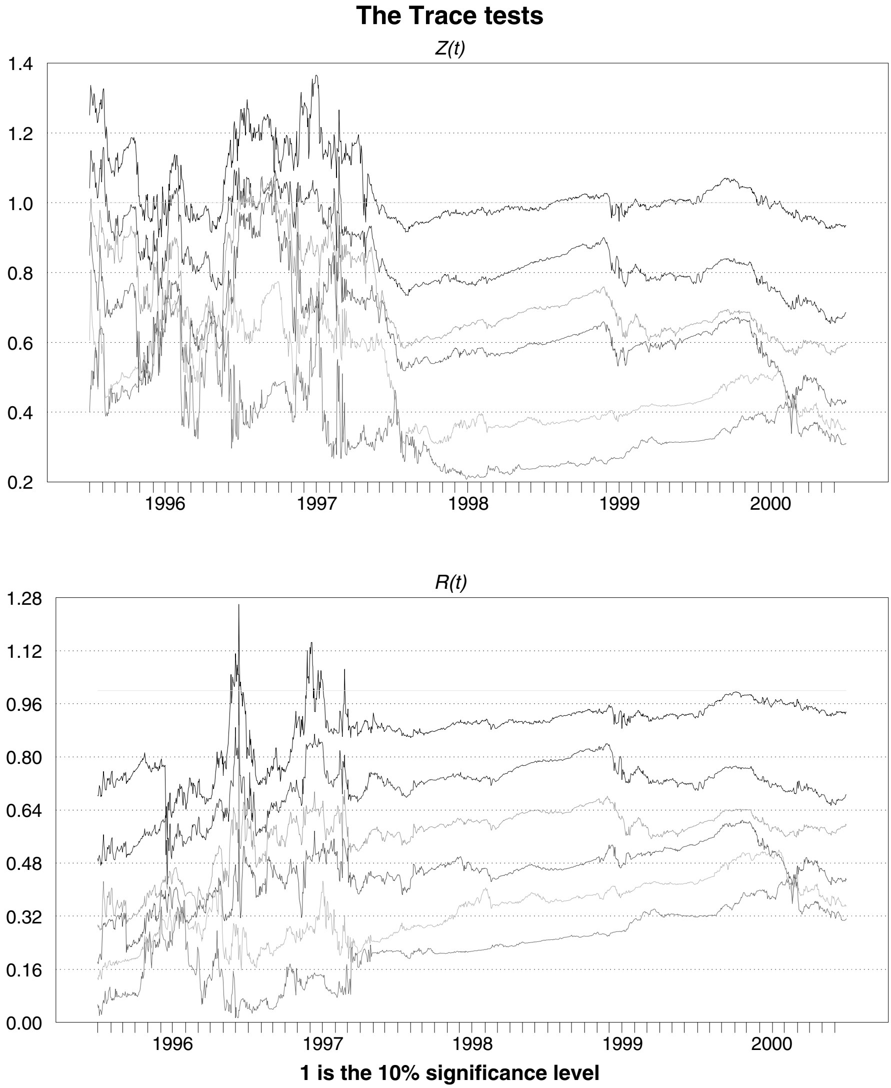


Figure 1

Recursive cointegration analysis: Plots of trace test statistics

Two important implications of these findings immediately follow. First, extending the work of Fung, Lee, and Leung (2000), the results suggest that A-share and B-share markets on both Shanghai and Shenzhen exchanges are segmented in the long run. Thus, long run diversification benefits for foreign investors through investing in B shares (as well as H shares and red chip stocks) do not simulate diversification benefits that they truly desire through investing in much larger Chinese A-share stocks. Second, the differential demand argument (Sun and Tong, 2000) does not hold well in the long run; big discounts of B shares relative to A shares cannot be attributable to close substitutability between B shares and H shares and red chip stocks to foreign investors. This also implies that foreign investors do not perceive that they can gain the same long-run diversification benefits through investing in H shares and red chip stocks as through investing in B shares. In addition, A-share markets in Shanghai and Shenzhen do not share comovements in the long run and neither do B-share markets on the two exchanges. This can be an indication of two different stock portfolios respectively represented by A-share (or B-share) markets on Shanghai and Shenzhen exchanges. For example, Chakravarty, Sarkar, and Wu (1998, p. 327) observed that B-share stocks listed on the Shanghai Securities Exchange are dominated by former state-owned enterprises and B-share stocks listed on the Shenzhen Securities Exchange are dominated by joint-venture companies with foreign investors.

The six-variable (first difference) VAR model results in the following innovation correlation matrix (lower triangular entries only are printed in the following order: Δx_1, Δx_2, Δx_3, Δx_4, Δx_5, Δx_6 where 1-SHA, 2-SHB, 3-SZA, 4-SZB, 5-HKH, 6-HKR)

$$V = \begin{bmatrix} 1 & & & & & \\ 0.46 & 1 & & & & \\ 0.89 & 0.47 & 1 & & & \\ 0.48 & 0.70 & 0.50 & 1 & & \\ 0.08 & 0.26 & 0.10 & 0.18 & 1 & \\ 0.10 & 0.28 & 0.12 & 0.22 & 0.76 & 1 \end{bmatrix} \tag{3}$$

Directed graph theory explicitly points out that the off-diagonal elements of the scaled inverse of this matrix (V or any correlation matrix) are the negatives of the partial correlation coefficients between the corresponding pair of variables, given the remaining variables in the matrix. For example, if we want to compute the conditional correlation between innovations in Δx_{1t} (SHA) and innovations in Δx_{5t} (HKH) given innovations in Δx_{3t} (SZA), we can calculate the inverse of the following matrix (taking the corresponding elements from the matrix V above):

$$V_1 = \begin{bmatrix} 1 & & \\ 0.89 & 1 & \\ 0.08 & 0.10 & 1 \end{bmatrix} \tag{4}$$

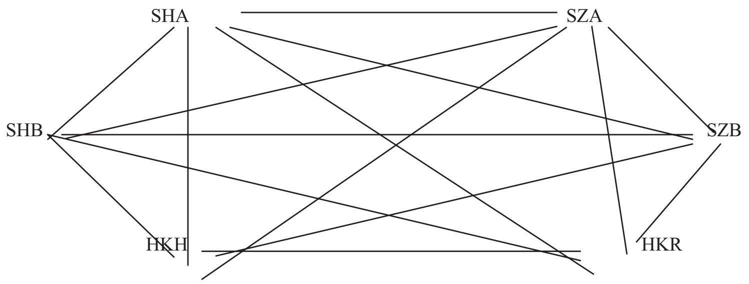
The matrix V_1 is the 3×3 matrix with lower triangular elements $V_{1,1}$, $V_{3,1}$, $V_{3,3}$, $V_{5,1}$, $V_{5,3}$, and $V_{5,5}$ of the V matrix given previously. The off-diagonal elements of the scaled inverse of the V_1 matrix are the negatives of the partial correlation coefficients between the corresponding pair of variables given the remaining variables. So, for example, the partial correlation between innovations from the market SHA and innovations from the market HKH given innovations in the market SZA is -0.02. Under the assumption of multivariate normality, Fisher's z can be applied to test for significance from zero. The marginal significance level is 0.41. Thus, at any conventional levels of significance we can conclude that innovations in the Shanghai A-share market (SHA) and the Hong Kong H-share market (HKH) are conditionally independent in contemporaneous time. Also, the SZA is in the sepset of (SHA, HKH). However, the SHB is not in the sepset of (SHA, HKH), and we can direct the edges as SHA ----> SHB <---- HKH. Similarly, the partial correlation between innovations from the market SHA and innovations from the market HKR, given innovations in the market SZA, is also –0.02 with the p-value $= 0.42$. The edge between the SHA and the HKR is also removed. Again, the SZA is in the sepset of (SHA, HKH), and the SHB is not in the sepset of (SHA, HKR). Thus, we can direct the edges as SHA ----> SHB <---- HKR.

Figure 2 gives both the complete undirected graph and the final directed graph on innovations from our six-market VAR model. Panel A is the starting point from which edges are removed and directed, as explained previously. At the 10% significance level, we find the directed edges as given in Panel B.[4] Specifically, we see edges running from the Shanghai A-share market to the Shanghai B-share and the Shenzhen A-share markets, from the Shanghai B-share market to the Shenzhen B-share market, from the Hong Kong H-share and red chip stock markets to the Shanghai B-share market, and from the Shenzhen B-share market to the Shenzhen A-share market. There is one undirected edge between the Hong Kong H-share and red chip stock markets.[5] Except for the above edges that connect different markets, there is no other instantaneous linkage between these six markets.

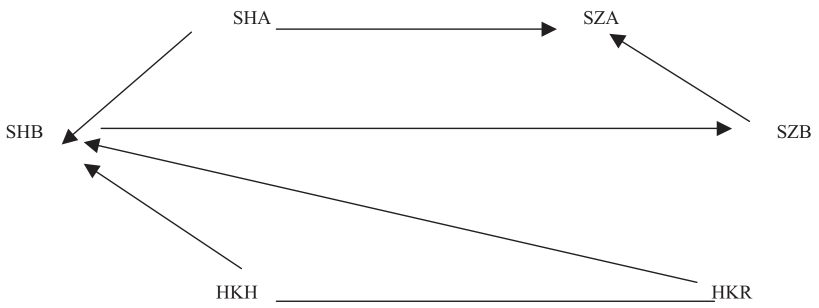
It is noteworthy, however, that the variance decomposition result in contemporaneous time (Day 1 in Table 3) clearly shows that evidence for some contemporaneous

[4] The PC algorithm is programmed into the TETRAD II software (Scheines, Spirtes, Glymour, and Meek, 1994), which is actually used to conduct the directed graph analysis in this study. The directed graph results at the 1% and 5% significance levels are somewhat different as they contain fewer (significant) edges than those at the 10% significance level. However, as pointed out by Scheines, et al. (p. 105), a slightly higher significance level may be appropriate here, as the PC algorithm is known to underfit, that is, to include too few (significant) edges. In the following, we use the forecast error variance decomposition (and impulse response analysis) result in contemporaneous time to double-check the robustness of the directed graph result.

[5] In the forecast error variance decomposition in the following, we direct this edge as the one running from the Hong Kong H-share market to red chip stocks based on some (albeit weak) evidence found at the 15% significance level. To examine how sensitive the forecast error variance decomposition results may be in response to the alternative ordering of the two variables, we also conduct the forecast error variance decomposition based on the alternative directed edge from red chip stocks to the Hong Kong H-share market. The main inference from variance decomposition results is qualitatively unchanged.



Part A
Complete undirected graph on innovations from equation (1)



Part B
Final directed graph on the model

Figure 2

Contemporaneous causal flow patterns using directed graphs

flows in Figure 2 can be fragile due to their economic insignificance.[6] Specifically, the contemporaneous causal flows from SHA to SHB (0.5%) and from SZB to SZA (0.3%) are statistically significant (at the 10% significance level) but not economically significant, whereas contemporaneous causal flows between all other pairs of

[6] We thank both referees for pointing this out. We also conduct impulse response analysis based on the estimated VAR and the directed graphs (results available on request). The basic inference from the impulse response functions is generally consistent with that from forecast error variance decomposition. In particular, with regard to strength of the contemporaneous causal flows from SHA to SHB and from SZB to SZA, the impulse responses of SHB to SHA and SZA to SZB in contemporaneous time appear to be moderate. It is noteworthy, however, that allowing for the low economic significance of contemporaneous flows from SHA to SHB and from SZB to SZA does not change the basic inference of this study, which is based on forecast error variance decomposition results at such longer horizons as ten to 30 days.

Table 3

Forecast error variance decomposition results (percentage)

We conduct forecast error variance decomposition based on the directed graph on innovations given in Figure 2 with additional directed edge (found at the 15% level): HKH → HKR. Day 1 is the contemporaneous period. The following table shows in each subsection how the variance of a particular market is explained by price shocks to the six markets in the first row (in percentage). SHA = Shanghai A-share market price index. SHB = Shangai B-share market price index. SZA = Shenzeng A-share market price index. SZB = Shenzeng B-share market price index. HKH = Hong Kong H-share market price index. HKR = red chip stock market price index.

Day	SHA	SHB	SZA	SZB	HKH	HKR
Variance of SHA explained by price shocks to the six markets						
1	100.0	0.0	0.0	0.0	0.0	0.0
2	57.8	17.9	3.8	0.6	13.2	6.6
10	14.5	40.1	1.7	1.4	28.4	14.0
30	14.3	39.5	2.9	1.5	28.3	13.6
Variance of SHB explained by price shocks to the six markets						
1	0.5	48.2	0.0	0.0	34.4	16.9
2	0.4	47.8	0.1	0.3	34.7	16.7
10	0.8	47.4	0.6	0.4	34.3	16.5
30	0.9	47.2	0.7	0.5	34.3	16.3
Variance of SZA explained by price shocks to the six markets						
1	53.3	12.0	21.6	0.3	8.6	4.2
2	48.0	15.5	19.7	0.4	10.9	5.5
10	31.6	26.5	13.2	0.8	18.5	9.4
30	29.1	27.8	12.7	0.9	19.7	9.8
Variance of SZB explained by price shocks to the six markets						
1	0.5	47.7	0.0	1.1	34.0	16.7
2	0.5	47.6	0.0	1.1	34.1	16.7
10	0.6	47.5	0.1	1.1	34.0	16.7
30	0.6	47.4	0.3	1.1	34.0	16.6
Variance of HKH explained by price shocks to the six markets						
1	0.0	0.0	0.0	0.0	100.0	0.0
2	2.3	0.8	1.9	0.4	94.6	0.1
10	4.0	31.7	4.7	2.0	46.8	10.8
30	3.9	36.3	4.5	1.5	41.3	12.6
Variance of HKR explained by price shocks to the six markets						
1	0.0	0.0	0.0	0.0	57.6	42.4
2	1.4	10.6	0.6	1.1	52.5	33.8
10	3.6	33.1	1.8	1.7	38.4	21.5
30	3.0	38.0	1.9	1.5	35.6	19.9

markets in Figure 2 (SHA to SZA [53.3%], from SHB to SZB [47.7%], from HKH to SHB [34.4%], from HKR to SHB [16.9%], and from HKH to HKR [57.6%]) are both statistically and economically significant. Hence, allowing for little economic significance of contemporaneous flows from SHA to SHB and from SZB to SZA clearly suggests informationally segmented A-share and B-share markets in both Shanghai and Shenzhen in contemporaneous time.

More importantly, as vigorously argued in Swanson and Granger (1997), the contemporaneous causal pattern as identified through the directed graphs analysis of the correlation (covariance) matrix of $\hat{e}_t$ (equation (1)) provides a data-determined solution to the basic problem of orthogonalization of VAR residuals and thus is critical to forecast error variance decompositions of a VAR. The advantages of employing the forecast error variance decomposition in this study include its allowance for time-lagged information transmission in addition to contemporaneous information transmission and description of economic significance of dynamic causal linkages between the markets.

Under the ordering of innovations as suggested by the directed graph (Figure 2), 30-day forecast error variance decomposition results are given in Table 3. Our multivariate framework reveals a number of new insights, which can not be discovered in a bivariate framework (e.g., Chui and Kwok, 1998). Though the SZA's price variations (29–53%) can also be significantly explained by the SHA, the Shanghai B-share appears to be the best informed market among four A- and B-share markets. A large amount (about 40% for SHA, 47% for SZB, and about 27–28% for SZA) of the price variations for other A-share and B-share markets can be explained by the SHB market shock in 10- to 30-day horizons. The Shenzhen B-share market contains little new information beyond the information originated from three other markets for foreign investors (i.e., SHB, HKH, and HKR). The relatively unimportant informational role of the Shenzhen B-share market could be due to the fact that the Shenzhen market is much smaller and less liquid than the Shanghai market. It could also be a reflection of the fact that many market players in the Shenzhen B-share market are Hong Kong investors or local investors who can get access to the Hong Kong market, due to geographical as well as economic affinity (Sun and Tong, 2000, p. 1900).

Thus, the evidence contradicts the argument of no information asymmetry (Chen, Lee, and Rui, 2001) or the argument of better informed domestic investors (Chakravarty, Sarkar, and Wu, 1998; Su and Fleisher, 1999). Rather, it extends the argument of Chui and Kwok (1998) and Mok and Hui (1998) in that only the Shanghai B-share market, but not the Shenzhen B-share market, leads both A-share markets in Shanghai and Shenzhen. The finding is also at odds with the argument that the Shenzhen market informationally leads the Shanghai market, as reported in Fung, Lee, and Leung (2000) and Poon and Fung (2000).

Another new finding is that the Hong Kong H-share market, and, to a much lesser extent, red chip stocks can significantly explain the price variations of all four A- and B- share markets. This is generally consistent with the argument that much of Chinese news is collected by or funneled through Hong Kong (Chui and Kwok, 1998, p. 352), but contradicts the previous finding that the prices of B shares have little relation with those of China-related stocks traded in Hong Kong (Baily, 1994; Chakravarty, Sarkar, and Wu, 1998). Neither does the finding agree with the argument of Poon and Fung (2000) that the red chip stocks process information faster than the H share. In addition, these two markets (HKH and HKR) can explain the price variations of B-share markets somewhat better than they explain those of A-share markets. This provides mild evidence for Sun and Tong's (2000) argument that H shares and red chip stocks are close substitutes for B shares. However, little evidence is found to support their projections that red chip stocks are closer substitutes for Shenzhen B shares and that H shares are closer substitutes for Shanghai B shares (Sun and Tong, 2000, p. 1900).

Finally, we also find that the Shanghai B-share market has a significant explanatory power for the price movements of H shares and red chip stocks traded in Hong Kong. This further supports the key finding of this study that foreign investors in the Shanghai B-share market are best informed among the four A- and B-share markets. Moreover, the information observed from this market is also useful for foreign investors' trading in Hong Kong China-backed stocks (i.e., HKH and HKR). Specifically, about 32% to 36% and 33% to 38% of price variations in the Hong Kong H-share market (HKH) and red chip stocks (HKR), respectively, can be explained by innovations from the Shanghai B-share market (SHB) in the 10- to 30-day horizons. By contrast, 34% to 35% and 16% to 17% of the Shanghai B-share market price variations can be explained by the Hong Kong H-share market and red chip stocks, respectively. Furthermore, the proportions of the price variations in other A-share and B-share markets explained by the Hong Kong H-share market innovations in the 10- to 30-day horizons (about 28% for SHA, 34% for SZB, and about 19–20% for SZA) are much greater than those by red chip stock market innovations (about 14% for SHA, 17% for SZB, and about 9–10% for SZA), but much smaller than those by the Shanghai B-share market innovations (about 40% for SHA, 47% for SZB, and about 27–28% for SZA). Hence, we can conclude that the Shanghai B-share market is somewhat more informed than the Hong Kong H-share market and much more informed than red chip stocks, a fact that has not yet been documented in the literature.

5. Conclusion

This paper examines the market segmentation and information asymmetry issues in Chinese stock markets, using Shanghai A-share, Shanghai B-share, Shenzhen A-share, Shenzhen B-share, Hong Kong H-share, and red chip stock market price indexes. A number of empirical regularities have been uncovered and compared with previous studies. The multivariate cointegration analysis suggests, and recursive cointegration analysis confirms, that A-share markets are segmented with B-share markets in the long run in both Shanghai and Shenzhen, verifying the finding of Fung, Lee, and Leung (2000) in a more robust sense. Moreover, the finding also implies that the differential demand argument recently proposed by Sun and Tong (2000) does not hold well in the long run. In other words, big discounts of B shares relative to A shares cannot be attributable to close substitutability between B shares and H shares and red chip stocks to foreign investors.

The data-determined forecast error variance decomposition based on an innovative directed graphs technique is further employed to shed light on the controversial issue of information asymmetry between Chinese domestic and foreign investors. The forecast error variance decomposition result clearly shows that foreign investors in the Shanghai B-share market are better informed than Chinese domestic investors in both A-share markets over time. Further evidence on the leading informational role of the Shanghai B-share market shows that foreign investors in this market are also better informed than foreign investors in the Shenzhen B-share, the Hong Kong H-share, and the red chip stock markets. Contradictory to Poon and Fung (2000), it is also found that another important source of information is the Hong Kong H-share market rather than red chip stocks.

The information asymmetry pattern identified in this study contradicts the arguments of no informational asymmetry (Chen, Lee, and Rui, 2001), better informed domestic investors (Chakravarty, Sarkar, and Wu, 1998; Su and Fleisher, 1999), or the informational leading role of the Shenzhen market over the Shanghai market (Fung, Lee, and Leung, 2000; Poon and Fung, 2000). The information asymmetry pattern extends the argument of Chui and Kwok (1998) and Mok and Hui (1998) in that only the Shanghai B-share market but not the Shenzhen B-share market leads both A-share markets in Shanghai and Shenzhen. In a broad sense, these findings challenge a widespread assumption in the finance literature that foreign investors are less informed than domestic investors about the value of local assets (Brennan and Cao, 1997; Kang and Stulz, 1997; Stulz and Wasserfallen, 1995). However, the findings of this study are consistent with more recent research (Froot, O'Connell, and Seasholes, 2001; Pan, Chan, and Wright, 2001). Froot, O'Connell, and Seasholes (2001) document that foreign investors' portfolio inflows have a noticeable ability to predict positive future returns in emerging markets but not in developed markets. Pan, Chan, and Wright (2001) also find that foreign investors are better informed than domestic investors in six East Asian emerging markets. Overall, our finding in Chinese stock markets suggests that foreign investors in China's emerging markets are better informed than domestic investors whereas, by contrast, foreign investors in developed markets can be less informed, as widely assumed in the literature. Further work along this line could be a potentially fruitful area of research.

References

Bailey, W., 1994. Risk and return on China's new stock markets: Some preliminary evidence, *Pacific-Basin Finance Journal* 2, 243–260.

Bailey, W., P. Chung, and J. Kang, 1999. Foreign ownership restrictions and equity price premiums: What drives the demand for cross-border investments? *Journal of Financial and Quantitative Analysis* 34, 489–511.

Bernanke, B.S., 1986. Alternative explanations of the money-income correlation, *Carnegie-Rochester Conference Series on Public Policy* 25, 49–99.

Bessler, D.A. and J. Yang, 2003. The structure of interdependence in international stock markets, *Journal of International Money and Finance* 22, 261–287.

Brennan, M. and H. Cao, 1997. International portfolio investment flows, *Journal of Finance* 52, 1851–1880.

Chakravarty, S., A. Sarkar, and L. Wu, 1998. Information asymmetry, market segmentation and the pricing of cross-listed shares: Theory and evidence from Chinese A and B shares, *Journal of International Financial Markets, Institutions and Money* 8, 325–355.

Chen, G.M., B. Lee, and O. Rui, 2001. Foreign ownership restrictions and market segmentation in China's stock markets, *Journal of Financial Research* 24, 133–155.

Chui, A. and C. Kwok, 1998. Cross-autocorrelation between A shares and B shares in the Chinese stock market, *Journal of Financial Research* 21, 333–353.

Cooley, T. and S. LeRoy, 1985. Atheoretical macroeconometrics: A critique, *Journal of Monetary Economics* 16, 283–308.

Diamandis, P., D. Georgoutsos, and G. Kouretas, 2000. The monetary model in the presence of I(2) components: Long-run relationships, short-run dynamics and forecasting of the Greek Drachma. *Journal of International Money and Finance* 19, 917–941.

Dickey, D.A. and W.A. Fuller, 1981. The likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica* 49, 1057–1072.

Eun, C. and S. Shim, 1989. International transmission of stock market movements, *Journal Financial and Quantitative Analysis* 24, 241–256.

Froot, K., P. O'Connell, and M. Seasholes, 2001. The portfolio flows of international investors, *Journal of Financial Economics* 59, 151–193.

Fung, H., W. Lee, and W. Leung, 2000. Segmentation of the A- and B-share Chinese equity markets, *Journal of Financial Research* 23, 179–195.

Gonzalo, J., 1994. Five alternative methods of estimating long-run equilibrium relationships, *Journal of Econometrics* 60, 203–233.

Haigh, M.S. and D.A. Bessler, in press. Causality and price discovery: An application of directed acyclic graphs, *Journal of Business*.

Hansen, H. and S. Johansen, 1993. Recursive estimation in cointegrated VAR models. Unpublished manuscript, University of Copenhagen.

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

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* 59, 1551–80.

Kang, J. and R. M. Stulz, 1997. Why is there a home bias? An analysis of foreign portfolio equity ownership in Japan, *Journal of Financial Economics* 46, 3–28.

Long, D., J. Payne, and C. Feng, 1999. Information transmission in the Shanghai equity market. *Journal of Financial Research* 22, 29–45.

Mok, H. and Y. Hui, 1998. Underpricing and aftermarket performance of IPOs in Shanghai, China, *Pacific-Basin Finance Journal* 6, 453–474.

Osterwald-Lenum, M., 1992. A note with quantiles of the asymptotic distribution of the maximum likelihood cointegration rank test statistics, *Oxford Bulletin of Economics and Statistics* 54, 461–472.

Pan, M., K.C. Chan, and D. J. Wright, 2001. Divergent expectations and the Asian financial crisis of 1997, *Journal of Financial Research* 24, 219–238.

Pearl, J., 1995. Causal diagrams for empirical research, *Biometrika* 82, 669–710.

Phillips, P. and P. Perron, 1988. Testing for a unit root in time series regression, *Biometrica* 75, 335–46.

Poon, W. and H. Fung, 2000. Red chips or H shares: Which China-backed securities process information the fastest? *Journal of Multinational Financial Management* 10, 315–343.

Scheines, R., P. Spirtes, C. Glymour, and C. Meek, 1994. *TETRAD II: User's Manual and Software* (Lawrence Erlbaum Associates, NJ).

Spirtes, P., C. Glymour, and R. Scheines, 1993. *Causation, Prediction, and Search* (Springer-Verlag, NY).

Stulz, R. and W. Wasserfallen, 1995. Foreign equity investment restrictions, capital flight, and shareholder wealth maximization: Theory and evidence, *Review of Financial Studies* 8, 1019–1057.

Su, D. and B. Fleisher, 1999. Why does return volatility differ in Chinese stock markets? *Pacific-Basin Finance Journal* 7, 557–586.

Sun, Q. and W. Tong, 2000. The effect of market segmentation on stock prices: The China syndrome, *Journal of Banking and Finance* 24, 1875–1902.

Swanson, N.R. and C.W.J. Granger, 1997. Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions, *Journal of the American Statistical Association* 92, 357–367.

Yang, J., I. Min, and Q. Li, in press. European stock market integration: Does EMU matter? *Journal of Business Finance and Accounting*.