INSIDER TRADING AT THE MEXICAN STOCK EXCHANGE: EVIDENCE FROM DATA AUGMENTATION

Manuel Lobato Osorio*

Secretaría de Hacienda y Crédito Público

(Received 17 January 2003, accepted 10 March 2004)

Abstract

Prior research has suggested that insider trading is present in the Mexican Stock Exchange. Based on this evidence, the media has claimed that insider trading is common practice. The evidence to support this claim is obtained from event-study methodology and aggregate returns. Previous studies, however, have not dealt explicitly with the endemic problem of missing observations. In this paper, I study the implications of insider trading at the corporation-level and over longer horizons. My work uses Monte Carlo simulations in order to deal with the nuissance of missing data. I find statistical evidence of insider trading only in a small number of the corporations under analysis.

Resumen

Investigaciones previas han sugerido que el *insider trading* está presente en el mercado accionario mexicano. De acuerdo con este hecho, los medios han reclamado que esta práctica es común. La evidencia que apoya estos reclamos, se obtiene a partir del uso de la metodología de estudio de eventos y de los retornos agregados. Trabajos previos, sin embargo, no se han ocupado explícitamente del problema endémico de observaciones faltantes. En este trabajo, estudio las implicaciones del *insider trading* a nivel corporativo y en horizontes de largo plazo. Mi trabajo utiliza simulaciones Monte Carlo para resolver el problema de observaciones faltantes. Por último, encuentro evidencia estadística de que el *insider trading* solamente aparece en un número pequeño de las corporaciones bajo análisis.

JEL classification: G14, C15, G31

Keywords: Event Study, Monte Carlo Simulation, Aggregate Returns

* Secretaría de Hacienda y Crédito Público. Palacio Nacional, 1er. Patio Mariano, PB, Centro Histórico. México, D.F., C.P. 06066. Teléfono: 5522 8953, 9158 1207. I thank Francisco Ciocchini, Rajeev Dehejia, Jaime Díaz Tinoco and Ken Leonard for helpful comments, and specially Gur Huberman and Zhenyu Wang for suggestions and encouragement. The usual disclaimer applies. E-mail: m_lobato@lycos.com

1. Introduction

The goal of this work is to determine whether there is statistical evidence of insider trading at the Mexican Stock Exchange (La Bolsa Mexicana de Valores, BMV). This question has recently attracted the attention of the media. The Washington Post, for example, writes "Mexico's stock exchange is *rife* with insider trading and *none* of it is being punished, according to a study. The evidence lies in the movement of share prices before and after corporate news announcements..."(emphasis added).[1] For the Mexican Economy, this is a relevant issue since foreign capital is an important source of funds for the country.[2] If the claims of insider trading are true, this implies that the Mexican government has not been enforcing the law. Foreign investors would have a twofold worry: they would lose their money to the insiders and the authorities would not take action on it.

In efficient financial markets, one should expect the prices of different shares issued by the same corporation to follow the same behavior over time. In contrast, if the insider trading hypothesis holds, the time-series price behavior may be different. The fact that BMV is a market with segmented ownership allows for testing this hypothesis. Mexican corporations issue shares that can be held by Mexican citizens only (A-shares), and shares that can be held by any investor regardless of citizenship (B-shares). There is no other difference between these kinds of shares. Within this framework, the presence of insider-trading would imply an information spillover from one type of share to another. In this case, information should be flowing from the type of share where there is more insider trading to the type of share where there is less insider trading.

This is what Bhattacharya, Daouk, Jorgenson and Kehr (2000) -from now on BDJK- argue: "[the] return volatility of one series type, whose shares only citizens may hold (A-shares), unambiguously leads return volatility of another series type, whose shares can be held by foreigners (B-shares), before the public news announcement, suggesting that there is an information spillover from one series type to another... The fact that the prices of the A-shares lead the prices of the B-shares hints [to] insider trading...." This evidence is derived from Granger Causality tests on the variance of aggregate daily returns. The authors model a Vector Autoregressive (VAR) over a ninety-day period, following the standard event study methodology used to analyze the behavior of stock prices around the day some value-relevant news is made public. Their research, coupled with the fact that there has never been any prosecution of insider trading at the BMV,[3] has led the media to claim that this illegal practice is business as usual at the Mexican stock exchange.[4]

[1] The Washington Post, March 12, 1998, at page D01.

[2] The share of foreign capital in the BMV has been 29%, 31%, 35% and 43%, for 1996, 1997, 1998 and 1999, respectively. Source: BMV.

[3] At least from the day the BMV was established until the end of the period covered in this paper.

[4] Mexican and international media attention has focused on this study too. Newspapers like "El Universal" and Magazines like "Proceso" have widely publicized the issue. The Economist (UK) also has touched on this point.

This paper provides evidence against the claims of rampant insider trading. It also points out the problems that missing data can cause on Granger Causality inference. Finally, it shows how Bayesian tools can help to overcome these problems.

I find no statistical evidence to support the claim that BMV is *rife* with insider trading. I reach this conclusion in two steps. First, I apply the notion of Granger Causality in VAR models on data from daily returns of twenty-one corporations listed in that exchange. Evidence from these models suggests that insider trading would be an issue for only two corporations in the sample. However, inference from this approach could be misleading due to missing observations. Missing data are present because the market for A-shares is very thin (see Table 1). The second step is to overcome this problem. I use a Gibbs Sampler and data augmentation algorithm (DA).[5] Although results change using DA, statistical evidence of information spillover from A-shares to B-shares is found only in two corporations.

I examine daily raw returns for the period July 1994 - January 1999. Ideally, I should use data from declared insider trades as in Seyhun (1986, 1988) or on illegal trades as in Meulbroek (1992), but these data are not available. The next best alternative would be to use intra-day data. Since information about prices can be transmitted all around the world in almost real time, arbitrage of profitable opportunities would not take more than a few hours. Given that the best data are not available, the next best alternative is to use daily data. Also, I study the long run relationship between the returns on these types of shares, and not just single events. I take this approach because the superior information of insiders could have been gathered more than three months before the announcement day.

My basic specifications are at the corporation, rather than the aggregate level.[6] I do this because I cannot study the behavior of portfolios of aggregate insider trades. In addition to this, I find evidence that aggregation in the presence of missing data can lead to wrong conclusions: Granger Causality tests suggest that the behavior of aggregate returns on B-shares leads the behavior of A-shares. But this statistical relationship is present in only two out of twenty-one corporations.

In addition to potential problems with aggregation, Granger Causality tests can give a wrong conclusion in the presence of missing data. I construct an example to show this. In this same example, DA gives the right answer. I use a large number of iterations (15,000) when implementing DA, so that the Gibbs Sampler can converge to the true distribution of the parameters (Tanner and Wong, 1987). If the Gibbs Sampler converges, inference based on DA is cor-

[5] Within the framework of an autoregressive model, Sargan and Drettakis (1974) present a recursive algorithm to solve the problem of missing data. Their work is not based on Monte Carlo Methods.

[6] I also try VAR models on daily returns for periods over event windows. Table 6 shows that the number of observations is very low for several corporations. Since inference based on these models might be misleading, the basic specifications are at the corporate rather than event level.

rect.[7] Thus, DA can contribute to solve the problem that missing observations represent, within the VAR framework.

This pattern could emerge as a result of insiders trading B-shares to make their illegal trades (the law allows Mexican to hold either A- or B-shares). I offer another explanation. Because the market for B-shares is more liquid and operates more frequently, public information is being impounded into B-shares' prices more quickly than into A-shares' prices (see Table 2). Since this finding is not general at the corporation level, and there is evidence that the market for A-shares is less efficient,[8] I conclude that the leading behavior of B-shares does not constitute further statistical evidence to support the hypothesis of insider trading.[9] Therefore, throughout the paper, I only consider as evidence of insider trading the pattern in which A-shares lead B-shares.

The rest of the paper is organized as follows. Part two presents the arguments concerning insider trading at BMV and a brief review of the literature. Part three presents the VAR model. Part four describes the data. Part five shows the empirical results. Part six presents the construction of the Gibbs Sampler and the evidence from the Data Augmentation algorithm. Part seven concludes.

2. Background

In efficient financial markets information about the true value of stocks should be impounded into prices as investors get to know value-relevant news. In these markets, prices are expected either to react to news announcements, if value-relevant information has not been publicly available, or to gradually incorporate publicly available information. Questions related to how well investors do in processing and, by trading, incorporating this news into prices are an empirical matter.[10] However, there is no doubt that prices have to reflect value-relevant information.

---

[7] The implicit assumption I make is that the missing observations are independent random draws for each company over the whole period this study analyzes. This assumption is not unrealistic. From Table 2, it can be seen that there is no apparent relation between the number of observations and industry or size of the original series as I explain below. Also, there is no apparent pattern of the missing observations at the event level. The number of observations in events in general behaves as in the whole sample, see Table 6.

[8] In the VAR regressions, in those cases where there is no evidence on a lead-lagged relationship between A-shares and B-shares, in general, regressions for the A-shares have a higher $\bar{R}^2$.

[9] Another explanation consistent with the "liquidity" rationale is that of the "nonsynchronous trading effect". A simple version of this can be found in Campbell, Lo and MacKinlay (1997). Consider two stocks, and assume that their returns are governed by a common factor in a linear fashion. Due to nontrading, the stock that is always traded (have zero probabilty of nontrading) will (mistakenly) seem to Granger cause the stock with positive probability of nontrading.

[10] Huberman and Schwert (1985), for example, document that 85% the news contained in a consumer price announcement had been anticipated and being reflected in the prices of Israeli indexed bonds.

Consider two stocks that have the same underlying value. Investors should expect the prices of such stocks to follow the same behavior over time: value-relevant information should be incorporated in these prices at the same time. In contrast, if the insider trading hypothesis holds, the time-series price behavior may be different. In markets where there are no insider trading restrictions, or there is a lax enforcement of such restrictions, this activity can be widespread.[11]

Two characteristics of the BMV create an opportunity to test for insider trading. First of all, the BMV is a market with segmented ownership. Mexican corporations can issue, among other kinds,[12] shares that can be held only by Mexican citizens (A-shares) and shares that can be held by any investor regardless of her citizenship (B-shares). A-shares collectively represent at least fifty-one percent of the voting rights, while B-shares at most forty-nine percent. There is no other difference between these kinds of shares; in particular, they have full voting and cash flow privileges. Second, Mexican laws restricting insider trading are similar to those in the United States. The National Banking and Securities Commission (Comision Nacional Bancaria y de Valores) regulates the exchange and is responsible for the enforcement of insider trading laws. However, since the day it was established until the period covered in this paper, there has not been a single case filed for insider trading at the BMV. Under this circumstances, the BMV is a market in which investors trade stocks having the same underlying value and seems to have had a lax enforcement of insider trading regulations.

Within this framework, the presence of insider trading would imply an information spillover from one type of share to another. In this case, information should be flowing from the type of share where there is more insider trading to the type of share where there is less insider trading. Therefore, I test whether returns on A-shares do not lead returns on B-shares and returns on B-shares do not lead returns on A-shares. To carry out these tests, I rely on the notion of Granger Causality Tests in VAR models.

In the most closely related work to this paper, Bhattacharya *etal.* (2000) model a VAR on the variance of aggregate daily returns over a ninety-day period, following the standard event-study methodology used to analyze the behavior of stock prices around the day some value-relevant news is made public. Conducting Granger causality tests they find a lead-lagged relation between A-shares and B-shares. The authors interpret their evidence as suggesting that insider trading is the reason why prices of Mexican stocks traded at BMV seem to not react to value-relevant announcements.

BDJK present in a footnote the reason why they model a VAR on volatility rather than a VAR on returns. They argue that "... many events in our sample...

---

[11] For example, using data from securities listed in the Amsterdam Stock Exchange, Kabir and Vermaeler (1996) find that the enactment of insider trading regulations had a significant impact on the behavior (liquidity and speed of adjustment of prices) of that stock. Their results suggest that insider trading was widespread before the enactment of the insider trading restrictions.

[12] For example, shares associated with the fixed and variable parts of capital, "Series F" and "Series V", respectively. See Domowitz, Glen and Madhavan (1998) for a more detailed description of what types of shares a Mexican corporation can issue.

may be regarded as good news by one class of shareholders and bad news by the other class... To see this more clearly, suppose that an increase in prices of A-shares is sometimes followed by an increase in prices of B-shares, due to a value increase for the whole firm, and is sometimes followed by a decrease in prices of B-shares, due to a value redistribution from B to A. The VAR test in returns under these conditions will show no average linkage, an incorrect conclusion. A VAR in volatility will show a linkage."

Their argument suggests that VAR models on daily returns should not present evidence that returns on A-shares lead returns on B-shares, or *vice-versa*. As the empirical evidence in this paper shows, this is not the case for corporations in my sample. Indeed, only for one third of the corporations I find no evidence of a lead-lagged relationship at all.

The authors do not mention the problem of missing observations. From column 3 in Table 2, it is apparent that the problem of missing data is endemic. Moreover, from Table 6 one can see that estimation of VAR models using more than 30 observations is only available for ten of the events studied by BDJK. These events correspond to the five most important corporations in the sample.

As I show below, missing data can lead to a wrong Granger causality inference. In order to overcome this problem, I deal with missing observations explicitly. I use a particular version of the Data augmentation algorithm proposed by Tanner and Wong (1987). This algorithm has as objective the construction of a Markov Chain on the posterior distribution of a set of parameters, given the observed and the missing data. If the chain is stationary and its length is appropriate, the posterior distribution of the parameters is well approximated. Then, Monte Carlo integration is used to get rid of the imputed data and to make inference on the distribution of interest.

There have been many other empirical studies on insider trading. Seyhun (1986) is a classical reference of studies using declared legal insider trades. This work suggests that insider trading has a substantial impact on prices. Meulbroek (1992) uses a data set of illegal insider trades, which allows her to study this phenomenon in detail for the US case. She also finds that insider trading has an important effect on prices. There is also a growing literature on Mexican financial markets. This includes the research of Domowitz, Glen and Madhavan (1997), who are interested in the consequences of the segmentation feature of the Mexican stocks; Domowitz, Glen and Madhavan (1998) study the effects on shareholders of the international cross listing of Mexican stocks.

In addition to the literature on insider trading and Mexican financial markets, this paper also relates to the literature on Markov Chain Monte Carlo (MCMC), in particular to that concerning Gibbs sampling. The Gibbs sampler presented in this paper is based on the work by Tanner and Wong (1987). The works of Morris (1987) and Gelman *et al.* (1995) inspired me to set up the actual algorithm. A comprehensive review on MCMC, both on theoretical foundations and its application has been developed by Gilks *et al.* (1996). Tanner (1993) presents a synthesized description of the theory of Gibbs sampling and Data augmentation. Finally, a classic reference for dealing with missing data is the work of Little and Rubin (1987).

3. A Simple VAR Model

The statistical relation between returns on A-shares and returns on B-shares within the sample is modeled as a VAR. In its structural form, the simplest p-order bivariate VAR can be written as:

$$R_t^A = \beta_{11}^0 - \beta_{12}^0 R_t^B + \beta_{11}^1 R_{t-1}^A + \beta_{12}^1 R_{t-1}^B + \cdots + \beta_{11}^p R_{t-p}^A + \beta_{12}^p R_{t-p}^B + \varepsilon_t^A \quad (1)$$

$$R_t^B = \beta_{21}^0 - \beta_{22}^0 R_t^A + \beta_{21}^1 R_{t-1}^A + \beta_{22}^1 R_{t-1}^B + \cdots + \beta_{21}^p R_{t-p}^A + \beta_{22}^p R_{t-p}^B + \varepsilon_t^B$$

where $R_t^i, j = A, B$, are the returns of the j^{th} shares at time t, and β_{jk}^i, $i = 0, \cdots p$, $j = 1, 2, k = 1, 2$, are the structural parameters. It is assumed that the returns are stationary;[13] the errors are white noise with finite variances σ_A^2, σ_B^2; the series of errors $\{\varepsilon_t^A\}$, $\{\varepsilon_t^B\}$ are serially uncorrelated; and, the covariance for ε_{t+i}^A and ε_{t+j}^B is 0 for $i \neq j$.[14] From (1), we have

$$BR_t = B_0 + B_1 R_{t-1} + \cdots + B_p R_{t-p} + \varepsilon_t,$$

where

$$B = \begin{bmatrix} 1 & \beta_{12}^0 \\ \beta_{22}^0 & 1 \end{bmatrix}; R_{t-i} = \begin{bmatrix} R_{t-i}^A \\ R_{t-i}^B \end{bmatrix}, i = 0, \cdots, p; \quad B_0 = \begin{bmatrix} \beta_{11}^0 \\ \beta_{21}^0 \end{bmatrix};$$

$$B_i = \begin{bmatrix} \beta_{11}^i & \beta_{12}^i \\ \beta_{21}^i & \beta_{22}^i \end{bmatrix}, \; i = 1, \cdots, p \text{ and, } \varepsilon_t = \begin{bmatrix} \varepsilon_t^A \\ \varepsilon_t^B \end{bmatrix};$$

thus

$$R_t = A_0 + A_1 R_{t-1} + \cdots + A_p R_{t-p} + \varepsilon_t, \quad (2)$$

with

$$A_0 = B^{-1} B_0 = \begin{bmatrix} a_{11}^0 \\ a_{21}^0 \end{bmatrix}, \; A_i = B^{-1} B_i = \begin{bmatrix} a_{11}^i & a_{12}^i \\ a_{21}^i & a_{22}^i \end{bmatrix}, \; i = 1, \cdots, p$$

$$\varepsilon_t = B^{-1} \varepsilon_t = \begin{bmatrix} \varepsilon_t^A \\ \varepsilon_t^B \end{bmatrix},$$

is the reduced form to be estimated. In this case, the Variance-Covariance matrix of the errors is given by

$$\sum = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} = \begin{bmatrix} V(\varepsilon^A) & COV(\varepsilon^A \varepsilon^B) \\ COV(\varepsilon^A \varepsilon^B) & V(\varepsilon^B) \end{bmatrix}.$$

---

[13] Table 2 presents the test statistic for the Phillips-Perron tests for unit root for each stock in the sample. It is seen that, except in one corporation, the null of the presence of a unit root is rejected.

[14] I test for unit roots of the estimated residual obtained from the reduced form. In general, tests suggest that is not a problem. For autocorrelation, I present Ljung-Box tests: this does not seem to be a problem either.

The assumptions that $\{\varepsilon^A\}$ and $\{\varepsilon^B\}$ are white noise processes and the covariance for ε^A_{t+i} and ε^B_{t+j} is 0, for $i \neq j$, imply that $\{\varepsilon^A\}$ and $\{\varepsilon^B\}$ have zero mean, constant variances and are serially uncorrelated.

In order to determine the length of a VAR, models of five or less lags are considered.[15] One can choose a length by testing whether the restriction of smaller number of lags binds. Tests are carried out in pairs, $p + 1$ against p, p against $p - 1$, and so on. The test statistic is the one suggested by Sims (1980).[16] The null hypothesis is that the restricted model is not different from the unrestricted. If the null is not rejected, the restricted model with smaller lag length is chosen. Another criterion for the selection of the lag length is that of the smallest generalized AIC (Akaike Information Criterion).[17] Among the alternatives, the model with the smallest AIC is chosen. I use both criteria to select the length of VAR models to be estimated.

Once the appropriate model for a corporation is selected, the reduced-form system is estimated applying the OLS procedure to every equation. For each regression, the Ljung-Box Q test statistic for five $[LB(5)]$ and ten autocorrelations $[LB(10)]$ are computed to test for the autocorrelation of residuals (and to test for the alternative of misspecification of the model).[18]

Finally, consider the VAR model with p lags in equation (2). The sequence $\{R^B_t\}$ does not Granger cause the sequence $\{R^A_t\}$ if $a^i_{12} = 0$ for all $i = 1, \cdots, p$. For my purposes, if the lags of the returns on B-shares do not improve the forecasting performance of the returns on A-shares, it is said that the returns on B-shares do not Granger cause those on A-shares. This requires the restriction $a^i_{12} = 0$ for all $i = 1, \cdots, p$ to be tested. Likewise for A-shares and B-shares, the restriction to be tested is $a^i_{21} = 0$ for all $i = 1, \cdots, p$. In both cases the test statistic follows an $F_{p,T-2p-1}$ distribution, where T is the number of observations in the regression.

4. Data

The data were collected from the SIVA (Automated Integral System of Securi-

---

15 As I argue in the introduction, it is not unreasonable to think that all profit opportunities can be arbitraged away within a week of trading.

16 The test statistic is $(T - c)(ln \mid \sum_R \mid - ln \mid \sum_U \mid)$, where, R and U stand for the unrestricted and the restricted models, respectively; T is the number of observations; c is the number of parameters estimated in each equation of the unrestricted model, and $ln \mid \sum_i \mid$ is the logarithm of the determinant of the variance-covariance matrix of the i^{th} model. This test statistic is asymptotically distributed as a $\chi^2(l)$, where l is the number of restrictions in every system of equations.

17 For each model, considering that every equation has an intercept, the AIC test statistic is calculated according to: $AIC = T ln \mid \sum \mid + 2(n^2 p + n)$, where n is the number of equations, p is the number of lags in the system, and $ln \mid \sum \mid$ is the logarithm of the determinant of the variance-covariance matrix of the model.

18 The $Q(s)$ test statistic is computed as $Q(s) = T(T+2)\sum_{k=1}^{s} \frac{r_k^2}{T-k}$, where r_k^2 is the autocorrelation of k^{th}-order. The test statistic follows a $\chi^2(s - p - 1)$ distribution, under the null hypothesis that the error process is white noise.

ties) of the Mexican Stock Exchange.[19] The database consists of closing prices, trade quantities, and Ask and Bid prices for forty-two stocks. The forty-two stocks are twenty-one pairs of A-shares and B-shares issued by twenty-one different corporations. I follow the criteria used by BDJK to choose the corporations to be included in the sample. Two stocks issued by two of the twenty-three corporations that BDJK analyze cannot be carried over to this study because they are no longer recorded in SIVA, or they were given substitutes during the sample period.[20] ARGOS A was substituted by the series ARGOS B; PROMEX A is not recorded in SIVA.[21] This leaves twenty-one corporations in the sample, which appear listed in Table 1.

Table 1. Average Annual Return and Typical Trade Volume

Stock		Period		Return	Volume	Ratio of Volumes	Obs.
(1)		(2)		(3)	(4)	(5)=$\frac{\text{BVolume}}{\text{AVolume}}$	(6)
BAN	A	06/01/94	01/29/99	−6.66	553,306		674
	B	06/01/94	01/29/99	−6.38	2,042,385	3.69	1162
BIT	A	06/01/94	01/28/99	−7.97	40,419		302
	B	06/01/94	01/29/99	−7.97	83,867	2.07	527
CEM	A	06/01/94	01/29/99	1.82	630,986		1149
	B	06/01/94	01/29/99	1.85	1,973,518	3.13	1169
CIF	A	06/02/94	12/19/97†	21.85	1,077,727		670
	B	06/01/94	12/19/97‡	21.65	2,198,200	2.04	889
CRI	A	06/13/94	03/25/96†	141.74	293,614		140
	B	06/13/94	03/19/96†	167.07	12,391	0.04	46
DES	A	06/02/94	01/29/99	−8.74	272,748		466
	B	06/01/94	01/29/99	−8.93	374,317	1.37	1148
GFA	A	06/01/94	01/14/98†	−11.44	65,828		312
	B	06/03/94	01/13/98†	−11.25	249,902	3.80	283
GFB	A	06/01/94	01/29/99	−5.60	2,930,963		1157
	B	06/01/94	01/29/99	−5.77	9,233,015	3.15	1169
GFN	A	06/01/94	12/14/98	−5.89	61,182		372
	B	06/02/94	01/29/99	−5.62	486,364	7.95	989

[19] I thank Jaime Diaz-Tinoco at BMV for his helpful insights about SIVA.

[20] The authors say that they have twenty-four companies, but Banorte (GFNORT) and GF Norte(GFNORTE) is the same corporation.

[21] Unfortunately, no more explanations were given to us about why the series were deleted or why they are not recorded in SIVA anymore.

Stock		Period		Return	Volume	Ratio of Volumes	Obs.
(1)		(2)		(3)	(4)	(5)= $\frac{\text{BVolume}}{\text{AVolume}}$	(6)
GH	A	06/01/94	05/09/97†	−3.46	416,535		367
	B	06/01/94	05/09/97‡	−3.82	386,134	0.93	384
GIS	A	06/09/94	05/23/96	53.58	4,200		145
	B	06/02/94	05/23/96	54.47	10,976	2.61	327
INB	A	06/01/94	01/29/99	17.89	80,992		670
	B	06/02/94	01/29/99	13.31	269,176	3.32	1012
INL	A	06/01/94	11/30/95†	−36.70	127,025		187
	B	06/01/94	11/29/95†	−33.12	178,440	1.40	187
KMC	A	06/01/94	01/29/99	−8.45	774,459		1168
	B	06/03/94	01/27/99	−8.90	70,389	0.09	356
LAT	A	06/22/94	09/11/97†	18.49	10,500,000		125
	B	07/29/94	09/11/97†	16.76	16,700,000	1.59	68
PON	A	07/10/95?	10/11/95†	81.87	562,778		18
	B	07/11/95	10/11/95‡	79.69	169,000	0.30	30
SEG	A	06/03/94	01/27/99	221.78	465,089		582
	B	06/06/94	01/28/99	202.38	333,264	0.72	633
SER	A	06/01/94	01/29/99	−14.06	307,616	500	
	B	06/01/94	01/29/99	−14.08	597,457	1.94	1084
SID	A	06/01/94	01/29/99	−14.72	485,070		390
	B	06/01/94	01/20/99	−14.72	546,795	1.13	950
SIT	A	07/26/94?	01/12/98‡	−19.21	24,118		34
	B	06/01/94	01/12/98	−18.60	2,040,393	84.60	901
TIE	A	06/06/94	10/06/95†	53.90	135,198		96
	B	07/21/94	10/12/95†	61.87	115,989	0.86	92

Source: author's estimation with data from SIVA (Automated Integral System of Securities) of the Mexican Stock Exchange. Annual average return in column 3 is expresed in percentage points. This is computed as 1 minus the ratio of the price at the of the period over the price at the beginning of period, divided by the number of years in the period. Typical trade volume, column 4, is the average over the period of the number of shares traded daily. Column 5 presents the ratio of the typical trade volume of B-shares over the same figure of A-shares. The number of observations in the period is presented in column 6.

† No more observations are found in SIVA.

‡ Last observation generated on 01/29/99.

? This is the first date with non-missing data.

The first observation in the sample comes from data generated on June 1, 1994.[22] Columns 1 and 2 of Table 1 present the code of each corporation and

[22] The data used by BDJK is a subset of my sample. It is worthy to mention that while the data set BDJK used in their study was obtained from a Bloomberg Terminal, the data set I use was downloaded directly from SIVA. According to the people at BMV in charge of SIVA, that system is the supply of Mexican data for Bloomberg. In this vein, there should be no difference between both data sets.

the period where observations were available for each type of share. For sixteen corporations, observations are available until January 1999; for the rest of the corporations, the observations are available up to January 1998, November 1997, March 1996, November 1995, and October 1995.

Table 1 also presents the average annual return and the typical trade volume for each stock during the sample period, after dropping the missing observations. From column 3, one can see that for each corporation the average annual returns for the two types of stock are almost identical. This is not a surprise since both stocks represent the same future flow of dividends.

What is striking from Table 1 is the ratio of the trade volume of B-shares over the trade volume of A-shares presented in column 5. For fifteen corporations in the sample, the ratio is greater than one; in ten out of these fifteen corporations the ratio is greater than two. That is, for a half of the corporations in the sample the trade volume of B-shares at least doubles the trade volume of A-shares. For two corporations this ratio falls below 0.70, in fact below 0.10. These facts suggest that in general for the period under study the market for B-shares is more liquid than the market for A-shares.

In Table 2, the summary statistics for daily raw returns are presented. Column 6 and 7 show the average and the standard deviation, respectively, of annualized returns over the sample period. Annualized returns are daily returns scaled up by 252.

Table 2. Summary Statistics for Daily Raw Returns

Stock		Number of Observations					Stand.	
		Potential	Actual	VAR	DA	AAR	Deviation	URT
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
BAN	A	1168	492	172	1168	96.14	996.17	−14.574
	B	1168	1154			8.29	931.32	−27.259
BIT	A	1168	133	43	591	93.94	525.29	−4.975
	B	1168	361			54.53	681.11	−10.293
CEM	A	1168	1131	1055	1168	21.99	742.40	−27.784
	B	1168	1168			14.85	789.07	−31.414
CIF	A	891	548	296	855	90.95	652.71	−23.84
	B	1168	885			28.45	630.46	−30.148
CRI	A	455	63	N/A	N/A	259.13	700.10	N/A
	B	455	11			1063.13	891.97	N/A
DES	A	1168	250	18	1113	83.94	826.31	−15.688
	B	1168	1131			15.34	977.67	−30.348
GFA	A	907	195	53	854	48.00	590.14	−7.4
	B	907	147			76.31	695.75	−6.345
GFB	A	1168	1147	1120	1168	12.08	918.09	−28.938
	B	1168	1168			13.00	1103.46	−30.64
GFN	A	1136	209	108	1130	85.12	905.87	−10.138
	B	1168	912			38.68	848.43	−23.369

Stock		Number of Observations					Stand.	
		Potential	Actual	VAR	DA	AAR	Deviation	URT
(1)		(2)	(3)	(4)	(5)	(6)	(7)	(8)
GH	A	735	233	39	695	89.53	884.60	−10.56
	B	1168	245			76.03	845.92	−9.01
GIS	A	495	47	10	438	260.96	507.71	−8.468
	B	1168	229			106.14	469.81	−10.82
INB	A	1168	475	182	933	88.85	527.13	−15.66
	B	1168	940			31.49	629.69	−28.46
INL	A	378	119	32	375	−105.54	827.94	−9.022
	B	378	123			46.01	921.57	−8.563
KMC	A	1168	1166	32	1130	7.50	880.83	−33.47
	B	1168	153			176.26	824.63	−8.714
LAT	A	823	36	N/A	N/A	517.47	1065.80	N/A
	B	823	24			251.94	1257.52	N/A
PON	A	425	4	8	71	1252.33	1859.79	−1.30?
	B	1168	17			503.62	1175.28	−3.38?
SEG	A	1168	386	161	999	83.87	641.48	−12.33
	B	1168	423			68.37	622.55	−12.95
SER	A	1168	315	94	1151	68.57	1152.63	−13.85
	B	1168	1026			−23.00	917.58	−26.15
SID	A	1168	247	40	843	314.85	1355.11	−12.18
	B	1168	847			63.34	2340.21	−23.34
SIT	A	905	7	N/A	N/A	361.29	617.03	N/A
	B	1168	895			−64.41	1396.52	N/A
TIE	A	345	59	N/A	N/A	105.26	539.09	N/A
	B	345	57			42.05	393.95	N/A

Source: author's estimation with data from SIVA (Automated Integral System of Securities) of the Mexican Stock Exchange. Potential, column 2, shows the number of observations before dropping missing data. Column 3 presents the number of observations used in the computation of the statistics. In column 4, the number of observations used in the VAR is shown. Column 5 presents the number of observations recovered in Data Augmentation. Columns 6 and 7 shown the average daily annualized return (AAR) and its standard deviation respectively. The daily annualized return is computed as the daily return times 252. The Phillips-Perron Test statistic is presented in column 8 (Unit Root Test).

? The null of a unit root is not rejected at the 5% level of significance.

Only three stocks have negative average annualized returns over the period covered in the sample; two of these are B-shares. There is no case where returns on A-shares and on B-shares are both negative. For sixteen corporations, the average annualized return on A-shares is greater than the same figure for B-shares. In two of these sixteen corporations the number of days in which A-shares are traded is greater than number of days in which B-shares are traded (see column 3).

From column 3 of Table 2, the problem of the missing observations is apparent. Also, this is more of a problem for the A-shares. The number of actual observations is larger for B-shares than A-shares in sixteen corporations in the sample. The corporations with the reversed pattern are CRISOBA (CRI), ATLANTICO (GFA), KIMBERLY (KMC), LATINCASA (LAT) and ERICCSON (TIE). CRI and KMC are in the paper industry; GFA is a financial group; LAT is a copper manufacturer; and TIE is a telephone manufacturer.

Column 3 of Table 2 also show that only two corporations have an entire set of observations in the period covered in this paper. These are major firms in the Mexican Economy. BANCOMER (GFB) was, during the period of time covered in the sample, the second most important bank; CEMEX (CEM) is one of the few cement producers in Mexico. Also, note that the biggest bank in Mexico in that period, BANAMEX (BAN), has a number of actual observations which is less than a half of the potential number.

These patterns in column 3 of Table 2 suggest that there is neither an apparent correlation between missing data in A-shares and a particular industry nor between the number of missing observations and the type of industry of the corporation. What is causing the missing observations? According to the experts at BMV, this is due to the low volume of trading of some stocks. They pointed out that SIVA has recorded all prices; if one observation is missing it is because that stock was not traded on that date.

According to column 4 in Table 1, the trading volume of ten stocks, 24% of the sample, is less than 100 thousand shares per day, while the trading volume for 27 of the stocks, 64% of the sample, is less than 0.5 million shares per day. The shares of the most important corporations in the Mexican Economy trade more than 0.5 million. Bearing in mind that in a low-trading day 19 million shares are exchanged in BMV, the participation of these latter stocks is relevant for that market. See Table 3.

Table 3. Trading Volume on a Single Day for Selected Months

Month (1)	BMV High (2)	Low (3)	NYSE High (4)	Low (5)
September 1998	202,428,069	36,402,294	1,216,324,513	609,810,960
January 1999	153,094,118	18,802,208	986,639,888	727,759,630

Source: Mexican Stock Exchange and New York Stock Exchange.

Furthermore, when compared to the NYSE, the BMV is a relatively thin market. In order to illustrate this, Table 3 shows the highest and the lowest trading volume on a single day in September 1998 (this month registers one of the days with the highest volume traded in all 1998), and in January 1999 (the last month in the sample) at the BMV. It also shows the same figures for the the New York Stock Exchange (NYSE). For days with high trading volume, the number of shares exchanged in the NYSE is six times that number in the

BMV. In low-trading days, according to that table, the trading volume in NYSE exceeds the trading volume in BMV at least sixteen times.

In summary, evidence in this section suggest that in general not only are B-shares more liquid but also they are traded more frequently than A-shares. That is, the market for A-shares is thiner then the market for B-shares. This can be explained by the fact that B-shares can be traded by both Mexican and foreign investors. These findings suggest that missing data can be relevant when analyzing the results of the econometric models.

5. Empirical Evidence from VAR Models

This section provides some evidence on the statistical relationship between returns on A-shares and B-shares. In the first part, evidence of this relationship at the corporation level and event level is presented. The second part looks at a portfolio formed with each type of return. As shown below, only a small proportion of corporations in the sample are found to have a pattern consistent with information spillover from A-shares to B-shares. Also, aggregation produces different results than corporate level results.

5.1 Empirical Evidence for Daily Returns

The results for the estimation of the model in (2) on daily returns are presented in Panel A of Table 4. Data are not used for days in which there is no return on the A-shares, or on the B-shares. This can cause that for a corporation, and given a structure of missing values, estimation of a model with one additional lag, in general, reduces the number of observations in the VAR in more than one.[23] This is such a serious problem that no VAR model could be estimated for four companies: CRI, LAT, SIT, and TIE. One can see that for almost half of the corporations (seven) the VAR estimated are of length one. The length of the VAR was chosen according to the procedure presented in the previous section, but there are cases in which VAR of length greater that three could not be estimated.[24]

At the 5% of the level of significance,[25] results can be summarized as follows. For five corporations, BAN, CEM, CIF, GFB, and SEG, returns on A-shares Granger cause those on B-shares and returns on B-shares Granger cause those on A-shares. For other eight companies, BIT, DES, GFA, GFN, GH, GIS, PON, and SID, there is no lead-lagged relationship at all. Neither

[23] Consider the following example. A '.' represents a missing value and an 'x' an observation.

$$\begin{pmatrix} & t & t+1 & t+2 & t+3 & t+4 & t+5 & t+6 & t+7 \\ A & x & x & . & x & x & x & x & x \\ B & x & x & x & x & x & x & . & x \end{pmatrix}$$

The number of observations for a VAR of length 1 is 3. The number of observations for a VAR of length 2 is 1; this is the maximum length that can be obtained in this example. With this in mind, consider, say, the case of BAN: the number of observations for the estimated VAR of length 1, 2, 3, 4 and 5 are 365, 281, 217, 173 and 138, respectively.

[24] These companies are BIT, GFA, GIS, INL, KMC, and PON.

[25] All of the results considered in this paper are based on the 5% level of significance.

case is consistent with information being asymmetrically impounded in both types of shares. Note that, in general, the VAR for corporations in the first group have a larger number of observations than corporations in the second group. Excluding GFN, VAR models for corporations in this latter group have at most fifty-three observations. This may suggest that the missing data in the VAR framework could influence the inference about the lead-lagged relationship between returns on A-shares and returns on B-shares. I will get back to this point later.

For two more corporations, INB and INL, returns on B-shares lead returns on A-shares. The last two corporations, KMC and SER, presents evidence that returns on A-shares lead returns on B-shares; this is consistent with the hypothesis of insider trading.

Figure 1. Impulse Response Functions for Selected Corporations

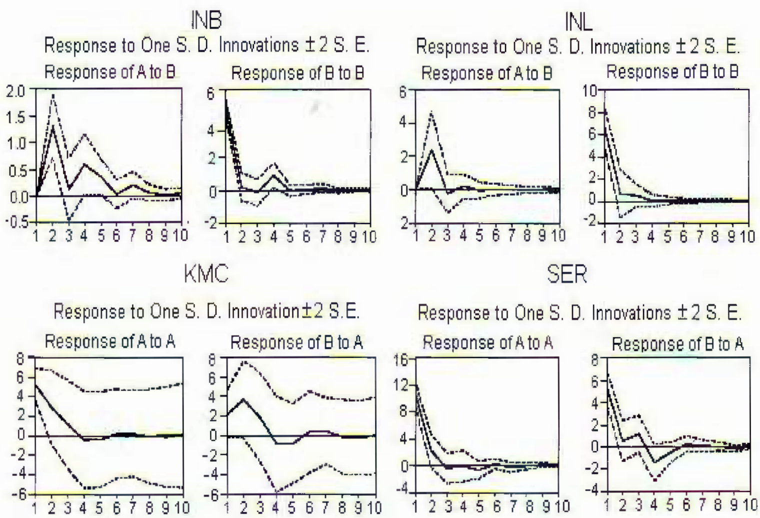

Figure 1 presents the impulse response functions (IRF) for the VAR models on KMC, INB, INL, and SER. There are four panels, one for each corporation. Every panel has two graphs: one presents the reaction of the returns on the A-shares to a one standard deviation shock on the returns on the leading share, and the other the reaction of the returns on B-shares to the same shock.[26] These impulse response functions suggest that, for INB and INL, the returns on A-shares react between one and two standard deviations to the shock on B-returns. For KMC and SER, the reaction is even greater: after the shock on returns on A-shares, B-shares jump four and six standard deviations, respectively, in the period after the shock. The lead-lagged relationships between returns on A-shares and returns on B-shares are statistically and economically significant, except in the case of KMC. In all cases, the correlation is positive. Positive

[26] More on IRF can be found in Hamilton (1994) and Enders (1995). The standard errors of the IRF are computed using 1000 Monte Carlo simulations.

returns on B-shares will induce positive returns on A-shares, for INB and INL. Also, positive returns on A-shares will bring positive returns on B-shares, for KMC and SER. The graphs show that, after the first period, the reactions become not statistically significant. All the efects of the shocks fade away almost entirely after the fifth period.

There is also evidence that the market for A-shares is less efficient. In ten out of thirteen cases, where the Granger causality tests show no lead-lagged relationships, the regressions for A-shares have a higher $\bar{R}^2$ than those for B-shares.

In order to check the robustness of the results, all VAR models are re-estimated subtracting and adding one lag to the benchmark specification presented in Table 4. This exercise could not be performed for two corporations in the sample, GIS and PON, due to the structure of the missing data. Results are presented in Table 5. This exercise alters returns for corporations where no evidence of a lead-lagged relationship exist, in two of thirteen cases. These cases are CIF and SID. For CIF, a larger number of observations presents evidence of a lead-lagged relationship where returns on B-shares lead returns on A-shares. For SID, a smaller number of observations presents evidence of a lead-lagged relationship where returns on A-shares lead returns on B-shares.

The results also change in three out of four of the corporations where a lead-lagged relationship was found. For INB and SER, where evidence supported the hypothesis of insider trading, a larger number of observations shows no evidence of information spillover form returns on A-shares to returns on B-shares. For INL, a larger number of observations shows evidence of returns on B-shares leading returns on A-shares.

In summary, the results seem to be robust to the number of observations used. Statistical evidence supporting the hypothesis of insider trading seems to fade away as the number of observations in the VAR increases. Evidence from this exercise also suggests that the market for A-shares is less efficient. Table 5 hints that missing observations can affect inferences made from Granger causality tests.

Another check I perform is to model VAR over event windows. I follow BDJK and use their definitions of events and windows to carry out the exercise. The results are presented in Table 6. The general impression is that the number of observations is very small and inference based on VAR models might not be accurate. Due to the structure of missing data, eleven out of forty possible events cannot even be analyzed.

Only ten events are available to look at VAR models estimated with more than thirty observations. These ten events correspond to five corporations: BAN, CEM, CIF, GFB, and SEG. One event in BAN and one event in SEG present evidence of the A-shares leading the B-shares. One event in CEM presents evidence of the B-shares leading the A-shares. Thus, even if this last event is considered to be evidence consistent with the hypothesis of insider trading, statistical evidence available suggests that is not a rampant activity at BMV.

In seven events there is no evidence of a lead-lagged relationship between A-shares and B-shares. In five of these, the $\bar{R}^2$ of the regressions where A-shares are the dependent variable are higher than the same figure for the B-shares. The market for A-shares seems to be less efficient than the market for B-shares at the event level.

Within events, missing observations do not seem to have a pattern. In general, their behavior in the events follow the behavior over the period that the sample covers. This suggests that missing data could be randomly generated. I use this assumption to set up the Gibbs sampler.

The evidence at the corporation and event level suggest that the hypothesis of insider trading is only consistent with a small proportion of the cases studied. When looking at events, however, bear in mind that the results could be affected by the small number of observations in the VAR models. Next, I look at the aggregate data. This allows me to have data for every day in the period the sample covers.

5.2 Results for Portfolios of Daily Returns (PDR)

For each type of share, I construct a "buy and hold" portfolio with the returns of the forty-two stocks listed in Table 2. Portfolio A is the aggregate of all daily returns on A-shares. Portfolio B is the aggregate of all daily returns on B-shares. These portfolios interpret a missing return of any stock of a given corporation as zero, i.e., as if there was no change in the price from its previous quotation.[27]

The portfolios are analyzed using the reduced form in equation (2) too. There are three different portfolios, each with a different sample period. The model for PDR0 is a VAR that includes all the observations in the sample period. PDR1 uses the same sample analyzed by BDJK, which covers the period from June 1, 1994, to June 30, 1997. PDR2 includes observations from July 1, 1997, to January 29, 1999. Table 4, Panel B shows the results for these portfolios.

Results for PDR0 suggests that overall, the behavior of the B-shares is leading the behavior of A-shares. Results for PDR0 are no different from results for PDR1. This is a completely different story than what the analysis at the corporation level reveals: only two out of twenty-one of corporations in the sample show evidence of this pattern. See Panel A of Table 4. These results suggest that aggregation can potentially matter a great deal.

Results for PRD2 suggest no lead-lagged relationship between the A-shares and the B-shares. There is evidence that both markets worked more efficiently over this period than over the period covered by PDR1 (lower $\bar{R}^2$). These results also show evidence that the market for B-shares is more efficient. This is in agreement with the results at the corporation level.

6. Data Augmentation

In this section, I use Gibbs sampling to obtain the posterior distribution of the parameters of the VAR models estimated in the previous section. The Appendix presents the derivation of the Gibbs sampler introduced here.

---

[27] I do not rebalance the portfolio to account for missing returns.

6.1 A Gibbs Sampler

The Data Augmentation algorithm includes two steps. First, the *imputation step* consists in generating a sample of m latent data from the predictive distribution of the missing data (in our case $m = 1$), given the observed data and the parameters. To generate these data, the current iteration of the conditional posterior distribution of the parameters (given the observed data and the missing data) is used. Second, the *posterior step* simulates a new set of parameters drawn from the joint distribution of the current imputation of the missing and the observed data. According to the results of Tanner and Wong (1987), iteration over these steps, from any starting point, produces a final sample of parameters that corresponds to the "true model".

In this paper, I construct a Gibbs sampler to obtain the posterior distribution of the parameters in a VAR. I derive this posterior distribution starting from a diffuse prior. I made this choice because I am interested in knowing what the sample *alone* has to say about the lead-lagged relation between A-shares and B-shares. Imposing another prior will necessary reflect my beliefs and the results of the exercise would be biased. This choice of the prior, of course, determines the posterior distribution of the parameters and the predictive distribution of the missing data.

The algorithm I use consists in the following. At iteration i, using the Variance-Covariance matrix of the equations and the set of coefficients in iteration i, Σ_i, β_i, a new set of imputed values, Z_i, is obtained. This is drawn from a Normal distribution. The mean and the variance of this distribution depends on the number of lags of the VAR, p, and is derived in the Appendix. The imputation is made with past and future values of the returns, in order to take into account all the available information.

Using the observed and the imputed data, $X = (Y, Z_i)$, the Gibbs sampler estimates a new set of coefficients, $\hat{\beta}$. With them, the estimator of the Variance-Covariance matrix of the equations, Σ, is obtained. This latter estimate is used as the parameter of an Inverse Wishart distribution. From this, the updated value of the Variance-Covariance matrix of the equations is drawn, Σ_{i+1}. This random draw, in turn, is used to compute the Variance-Covariance matrix of the coefficients, $\Sigma_{i+1} \otimes (X'X)^{-1}$. This matrix is the variance of a Normal distribution, which is used to draw the update of the set of coefficients β_{i+1}. This mean of this distribution is the vector of coefficients estimated at the beginning of the iteration, $\hat{\beta}$.

An initial set of parameters and a "complete" set of data, $\Sigma_0, \beta_0, X_0 = (X, Z_0)$, are required to start doing the iterations. I use the following procedure to estimate them. I construct the first complete set of data by imputing the missing values with the predicted returns of the lagged returns implied by the estimates presented in Table 4. Note that at this stage, I do not use future values of the returns to do the imputations. This allows the imputation of a missing value even if future returns are also missing and, thus, to maximize the number of recovered data. Nevertheless, not all data can be recovered because I need, at least, p observations in a row to impute one missing value.

I set the number of iterations equal to 15,000. The "burn-in" or "warm up" period equals 5,000 iterations.[28] The literature has discussed whether it is best to use a single long chain, or many short ones (see for example Gelman and Rubin (1992) and Geyer (1992)). To decide on what approach to use, I ran two long chains simultaneously for one corporation. Their results are not different, so I use only one large chain for the rest of the companies.

In order to assess the performance of the Gibbs sampler, I present the following exercise. I construct a VAR of order one, using errors randomly drawn from a standard normal distribution. This VAR includes two variables, A and B. Series A and series B are constructed in such a way that a Granger causality test must show that each series causes the other. A sample of 1,100 observations is generated. Drawing random numbers from a uniform distribution, I obtain the total number of observations that would be "missing". Then, from the same distribution, I obtain the actual missing observations. With these, using equal probabilities, I randomly choose if the observation has a missing variable on A, B or both returns. I do this three times. With these three random samples, I estimate a VAR for each, following the procedure used in the previous section. I also compute the Granger causality tests. The results are presented in Table7.

Results show that the approximation provided by the VAR is not satisfactory when missing observations are present. The estimated values of the parameters are not far from their true value, but the estimation is not precise. In fact, as it can be seen from the Granger causality tests, the structure of the missing observations has a great impact on the inference based on this statistic. From sections 3 and 4 in Table 7 one can see that, when the structure of missing observations is such that the VAR model is estimated with a small number of observations, the inference based on Granger causality test is misleading. In these two cases, the tests show that there is a lead-lagged relationship between series A and series B, but this is not true in reality. On the other hand, when the available information is large, *i.e.,* when the number of available observations for estimation is large, as it is the case in section 5, even though there are missing observations, the inference is right. From Table 7, one also can see that the Gibbs sampler is doing its job. Note that, except for the coefficient on the lagged values of the series B in the regression of the series A, all estimates of the slope coefficients are not very different from their true values.

In Bayesian analysis there is no analogous concept to that of Granger causality in VAR models, since the interest is centered in finding the posterior distribution of the estimated parameters. Nevertheless, for VAR models of order one, an equivalent "test of causality" can be approximated by testing that the coefficient of interest is not statistically different form zero. This is a t-test, with test statistic equal to the average of the coefficient over its standard deviation. Results also show that the t-statistics for every one of the coefficients associated to the lagged value of the i^{th} series in the regression of the j^{th} series is very large; this suggests that those coefficients are statistically different from zero; this implies that series A Granger cause series B, as well as series B Granger cause series A. This is the right inference.[29]

---

[28] For two companies, I also tried 25,000 iterations and used 10,000 as burn-in period. Results did not change significantly. All results here presented are based on 15,000 iterations.

[29] Since the series of coefficients is autocorrelated, I should use a Newey-West estimator

Finally, Table 7 also presents some statistics. In rows 'A', 'B' and 'A', 'B' the number of observations missing in series A, B or both are shown, respectively. It also present, for the random sample, the mean and the standard deviation of each series in rows "MEAN" and "S.D.", respectively. In the case of the estimation by Gibbs sampling, these rows show the average difference between the average imputed value (over the 10,000 simulations after the burn-in period) and its standard deviation. Notice that the standard deviation of the imputation is never greater than the standard deviation of the original series.

6.2 Results

All VAR models presented in Table 4 are re-estimated using Data Augmentation (DA). The number of observations recovered are presented in Table 1. These can be compared with the number of observations used in the VAR models. The number of observations used increases dramatically. As a percentage of the number of observations used in the VAR, the increase in number of observations varies from 4% (GFB) to 4280% (GIS). This suggests that the Gibbs sampler is using a much larger portion of the available information.

Results from DA are shown in Table 8. The coefficients of the VAR are the average taken from 10,000 iterations after the burn-in period. The standard deviation for every coefficient appears in parenthesis. This is computed in the standard way.

For VAR models of length one, Table 8 shows that DA changes the results for three out of seven corporations. For INL, DA presents evidence that the lead-lagged relation between B-shares and A-shares does not hold. Therefore, there is no longer evidence consistent with information spillover for this corporation. Also, while in the VAR framework the evidence on PON suggests that there is no lead-lagged relationship, evidence from DA shows a lead-lagged relationship from A-shares to B-shares. Thus, PON returns show a pattern consistent with the story of insider trading. For GFA, returns on both types of share improve the prediction of the returns on the other, the opposite of what the VAR model shows. Thus, DA changes results, but it increases the evidence supporting insider trading to only one more corporation.

For VAR models of length greater than one, after looking at the *t-statistics* of the coefficients, results do not change for most of the corporations. However, CIF and INB present patterns of returns consistent with the hypothesis of insider trading; and evidence from SER is no longer consistent with that hypothesis.

In the end, evidence of a lead-lagged relationship from A-shares to B-shares appears in only four corporations: CIF, INB, PON, and KMC, while evidence on a lead-lagged relationship from B-shares to A-shares appear only in INB. However, inference made from models with length greater than one is not equivalent to that based on Granger causality. My point here is that, even if we considered this as evidence in favor of the hypothesis of insider trading, this is not enough to claim that this activity is common at the BMV.

---

of the standard errors, as suggested by Geyer (1992). Because t-statistics are very large, my guess is that results would not change.

The combination of these results can be interpreted as showing that the hypothesis of a lead-lagged relation between B-shares and A-shares does not hold up to DA. The results presented in this section suggest that inferences made from VAR models with a small number of observations can be misleading.

7. Concluding Remarks

This paper has addressed the question of whether there is statistical evidence to support the hypothesis of insider trading at Bolsa Mexicana de Valores, the Mexican Stock Exchange. This goal has been undertaken by testing the statistical relationship between daily raw returns on A-shares (held only by Mexican investors) and daily raw returns on B-shares (held by Mexican and foreign investors) issued by twentyone corporations listed in that exchange. In the framework of VAR models, Granger causality tests show that there is statistical evidence of a lead-lagged relationship between returns on these kinds of shares only in four corporations in the sample.

I have argued that the B-shares leading A-shares is not evidence of insider trading, rather it is product of the fact that the market for B-shares is more liquid and operates more frequently than the market for A-shares (in fact, for a half of the corporations in the sample the average of daily trading volumes of B-shares at least doubles the average of the daily trading volume of A-shares). If my intuition is right, only two.

An important feature of the Mexican Stock Exchange is that it is a very thin market. As an illustration of this fact, Table 3 shows that trading volume at NSYE is at least six times the trading volume at the BMV for some selected months in the sample period. Also, less than a half-million shares on average are exchanged daily for 6.4 out of every ten secutities in my sample. In addition to this, only two corporations have a complete set of obervations for the entire sample. This nontrading behavior of stocks in this market creates the nuissance of missing observations.

Missing data can be a problem when analyzing daily returns from this stock market. Evidence gathered using a simulated random sample with randomly chosen missing data shows that missing observations can make Granger causality tests to give wrong inference.

In order to overcome this problem the Data Augmentation algorithm is used: a Gibbs Sampler is developed to obtain the distribution of the parameters of any estimated VAR, given observed and missing data. The Gibbs Sampler also imputes missing data from its conditional predictive distribution, given observed data and the parameters. In order to make an inference on the estimated parameters, missing data are integrated out. Inference based on the Data Augmentation algorithm shows that the results on the lead-lagged relationship between the returns on B-shares and on A-shares are not robust when the algorithm is used. The algorithm changes the results for some corporations, but it provides statistical evidence of insider trading in only three corporations.

The main conclusion of this paper is that there is not enough statistical evidence to claim that insider trading is a widespread activity at the Mexican Stock Exchange. This conclusion applies to the period between June 1994 and January 1999 and to the twenty-one corporations included in my sample.

Appendix

A. A Simple Gibbs Sampler for VAR

This appendix shows the derivation of the Gibbs Sampler used in this paper. For our purposes, it is convenient to present (2) again:

$$R_t = A_0 + A_1 R_{t-1} + \cdots + A_p R_{t-p} + \varepsilon_t. \tag{3}$$

We assume that

$$\varepsilon_t \sim i.i.d\ N(0, \Sigma), \tag{4}$$

with

$$\Sigma = \begin{pmatrix} \sigma_A^2 & \sigma_{AB} \\ \sigma_{AB} & \sigma_B^2 \end{pmatrix} = \begin{pmatrix} V(\varepsilon_t^A) & COV(\varepsilon_t^A \varepsilon_t^B) \\ COV(\varepsilon_t^A \varepsilon_t^B) & V(\varepsilon_t^B) \end{pmatrix},$$

where $COV(\varepsilon_{t-j}^i) = 0$ for all $i, j > 0$, $COV(\varepsilon_{t-i}^A \varepsilon_{t-j}^B) = 0$, for all $i \neq j$.

Note that the assumptions made about Σ, and the fact that the set of regressors for both equations are the same, reduces the number of elements to be drawn from $2k \times 2k$ to 2×2, as will be shown below.

Let $T + p$ be the number of potential observations in the sample, *i.e.*, let T be the number of observations net of lags, and let $k = (2p + 1)$ be the number of "explanatory" variables. Let $\mathbf{x_t}$ denote the $(k \times 1)$ vector of these explanatory variables, which includes a constant term and p lags of each element of R, $\mathbf{x_t} = [1\ R_{t-1}^A\ R_{t-1}^B \cdots R_{t-p}^A\ R_{t-p}^B]' = [1\ R_{t-1}' \cdots R_{t-p}']'$, and let β' denote the $(2 \times k)$ matrix $\beta'[A^0 A^1 \cdots A^p]$. With these, we rearrange (3) in compact form:

$$R_t = \beta' \mathbf{x_t} + \varepsilon_t. \tag{5}$$

In order to implement the Gibbs sampler, we need the posterior distribution of the parameters given the data, missing and observed, $p(\vartheta \mid Z, Y)$, and the predictive distribution of the missing data given the parameters and observed data, $p(Z \mid \theta, Y)$.

Let $\widetilde{R}^i = \{R_j^i \mid$ observation j is missing, $j6T\}, i = A, B$, be the sets of missing observations, and $\widehat{R}^i = \{R_j^i \mid j$ is observed, $j6T\}, i = A, B$, be the set of observed values of the series of the returns. Define $R^i = (\widehat{R}^i, \widetilde{R}^i), i = A, B; \widehat{R} = (\widehat{R}^A, \widehat{R}^B); \widetilde{R} = (\widetilde{R}^A, \widetilde{R}^B)$; and $R = (R^A R^B)'$, to agree with the previous notation. Also, let $\beta = \{a_{11}^0, a_{11}^1 a_{12}^1, \cdots, a_{11}^p, a_{12}^p, a_{21}^1, a_{22}^1, \cdots, a_{21}^p, a_{22}^p\}$ be the set of coefficients and $\Sigma = \{\sigma_{11}^2, \sigma_{12}, \sigma_{22}^2\}$ be the set of variances, we are interested in. Note that in agreement with our previous notation, we have $Y = \widehat{R}, Z = \widetilde{R}, R = X = (Z, Y)$, and $\theta = (\theta_1, \theta_2) = (\beta, \Sigma)$.

A.1 $p(\theta \mid Z, Y)$

It is straightforward to show that the posterior distribution of the parameters, when all data are observed, is given by:

$$p(\beta, \Sigma, \widetilde{R} \mid \widehat{R}) \propto p\ (\beta, \Sigma \mid R) p(R),$$

and, using Bayes rule,

$$p(\beta, \Sigma \mid R) \propto p(R \mid \beta, \Sigma) p(\beta, \Sigma), \tag{6}$$

where $p(R \mid \beta, \Sigma)$ is the likelihood function and $p(\beta, \Sigma)$ is the prior distribution of the parameters of interest.

Consider the following diffuse (noninformative) prior (see Gelman *et al.* (1995)):

$$p(\beta, \Sigma) \propto \mid \Sigma \mid^{-(\frac{d+1}{2})}, \tag{7}$$

where d stands for the number of equations, and in this case $d = 2$. The conditional density of the t^{th} observation, under (4), follows

$$p(R_t \mid x_t, \beta, \Sigma) \propto \mid \Sigma \mid^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}(R_t - \beta'\mathbf{x}_t)'\Sigma^{-1}(R_t - \beta'\mathbf{x}_t)]\right\},$$

and the likelihood of the sample obeys

$$\begin{aligned} p(R \mid \beta, \Sigma) \propto & \mid \Sigma \mid^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T}[(R_t - \beta'\mathbf{x}_t)'\Sigma^{-1}(R_t - \beta'\mathbf{x}_t)]\right\} \\ = & \mid \Sigma \mid^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}\sum_{t=1}^{T}[(R_t - \beta'\mathbf{x}_t)(R_t - \beta'\mathbf{x}_t)']\right)\right\} \\ = & \mid \Sigma \mid^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}S(\beta))\right\}. \end{aligned} \tag{8}$$

From (6), using (7) and (8), we get that the posterior distribution of the parameters is given by:

$$\begin{aligned} p(\beta, \Sigma \mid R) \propto & \mid \Sigma \mid^{-\frac{T}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}S(\beta))\right\} \mid \Sigma \mid^{-(\frac{2+1}{2})} \\ = & \mid \Sigma \mid^{-(\frac{T+2+1}{2})} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}S(\beta))\right\}, \end{aligned}$$

and thus,

$$p(\beta, \Sigma \mid R) \propto \begin{cases} \mid \Sigma \mid^{-\frac{T+2+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(\Sigma^{-1}S(\beta))\right\} & if \ \mid \Sigma \mid > 0 \\ 0 \text{ otherwise.} \end{cases}$$

To implement the Gibbs sampler, we need to draw from the marginal distribution of Σ and β.

A.1.1 Σ_β

Using Bayes rule, given an initial coefficients vector β_0 and initial data R_0, it is shown that

$$p(\Sigma \mid \beta, R) \propto p(R \mid \beta)p(\beta)$$

and

$$p(\Sigma \mid \beta, R) \propto \begin{cases} \mid \Sigma \mid^{-\frac{T+2+1}{2}} \exp\left\{-\frac{1}{2}\mathrm{tr}(S(\beta)\Sigma^{-1})\right\} & if \ \mid \Sigma \mid > 0 \\ 0 \text{ otherwise.} \end{cases} \tag{9}$$

Comparing (9) to the kernel of an Inverse Wishart distribution, we conclude that

$$\Sigma^{-1} \sim W_T(S(\beta)) \tag{10}$$

or that the posterior distribution of the Variance-Covariance matrix of the error term follows an Inverse Wishart distribution. In the previous expression T are the degrees of freedom, and $S(\beta)$ is the symmetric positive definite 2×2 scale matrix. Note that this distribution is the multivariate generalization of a χ^2 distribution (see Gelman *et al.* (1995)). For practical purposes, given β, the estimator of $S(\beta), S(\beta_\Sigma)$, is computed in the standard way:

$$S(\beta_\Sigma) = \frac{1}{T}\begin{bmatrix} \hat{\varepsilon}^{A'}\hat{\varepsilon}^{A} & \hat{\varepsilon}^{A'}\hat{\varepsilon}^{B} \\ \hat{\varepsilon}^{B'}\hat{\varepsilon}^{B} & \hat{\varepsilon}^{B'}\hat{\varepsilon}^{B} \end{bmatrix} = \frac{1}{T}\begin{bmatrix} \sum_{t=1}^{T}(\hat{\varepsilon}_t^{A})^2 & \sum_{t=1}^{T}\hat{\varepsilon}_t^{A}\hat{\varepsilon}_t^{B} \\ \sum_{t=1}^{T}\hat{\varepsilon}_t^{A}\hat{\varepsilon}_t^{B} & \sum_{t=1}^{T}(\hat{\varepsilon}_t^{B})^2 \end{bmatrix}$$

where $\hat{\varepsilon^i}$ is the $(T \times 1)$ vector of OLS from the i^{th} regression.

A.1.2 β_Σ

Now, given Σ and R, we compute the posterior distribution of the coefficients. Again, using Bayes rule we note that:

$$p(\beta \mid R, \Sigma) \propto p(R \mid \Sigma)p(\Sigma).$$

In order to compute the posterior distribution of the β coefficients, rearrange the system in (5) to obtain: $\begin{bmatrix} R_t^A \\ R_t^B \end{bmatrix} = \begin{pmatrix} \mathbf{x}_t' & 0' \\ 0' & \mathbf{x}_t' \end{pmatrix} \begin{bmatrix} \beta_A \\ \beta_B \end{bmatrix} + \begin{bmatrix} \varepsilon_t^A \\ \varepsilon_t^B \end{bmatrix}$, or defining the $(2 \times 2k)$ matrix $X_t' = \begin{bmatrix} x_t' \\ x_t' \end{bmatrix} = \begin{pmatrix} \mathbf{x}_t' & 0' \\ 0' & \mathbf{x}_t' \end{pmatrix} = (I_2 \otimes \mathbf{x}_t')$, and the $(2k \times 1)$ vector $\beta = [\beta_A' \beta_B']'$, in compact form

$$R_t = X_t\beta + \varepsilon_t.$$

Stacking the T equations in the sample,

$$R = \begin{bmatrix} R_1 \\ \vdots \\ R_T \end{bmatrix} = \begin{pmatrix} X_1' \\ \vdots \\ X_T' \end{pmatrix} \beta + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix} = X\beta + \varepsilon \tag{11}$$

where X is a $(2n \times 2k)$ matrix and ε is a $(2n \times 1)$ vector. Note that, given the assumptions in (4), we have that $\mathbf{V} = COV(\varepsilon\varepsilon') = (I_n \otimes \Sigma)$. Since $\mathbf{V}$ is positive difinite symmetric matrix, we can find L such that $\mathbf{V} = L'L$. Letting $\tilde{X} = L_X, \tilde{R} = LR, v = L\varepsilon$ we can respress (11) as

$$\tilde{R} = \tilde{X}\beta + v \tag{12}$$

where $\tilde{\mathbf{V}} = COV(vv') = I_{2n}$. The OLS estimator of this system is given by:

$$\begin{aligned} \beta_\Sigma &= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{R} = (X'(I_n \otimes \Sigma)^{-1}X)^{-1}X'\Sigma^{-1}R \\ &= \left(\Sigma^{-1} \otimes \left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]\right)^{-1}\left(\sum_{t=1}^{T}\left(\Sigma^{-1} \otimes R_t\right)\mathbf{x}_t\right) \\ &= \left(I_2 \otimes \left[\sum_{t=1}^{T}\mathbf{x}_t\mathbf{x}_t'\right]^{-1}\right)\sum_{t=1}^{T}(R_t \otimes \mathbf{x}_t) \end{aligned} \tag{13}$$

Using the likelihood function of (12), given Σ and R, the posterior distribution of β follows (abusing notation):

$$\begin{aligned}
p(\beta \mid R, \Sigma) \propto &\mid \Sigma \mid^{-\frac{T+2+1}{2}} \exp\left\{-\frac{1}{2}(\tilde{R}-\tilde{X}\beta)'(\tilde{R}-\tilde{X}\beta)\right\} \\
&\propto \exp\left\{-\frac{1}{2}(\tilde{R}-\tilde{X}\beta)'(\tilde{R}-\tilde{X}\beta)\right\} \\
&= \exp\left\{\tilde{R}'\tilde{R} - 2\beta'(\tilde{X}'\tilde{X})\beta_\Sigma + \beta'(\tilde{X}'\tilde{X})\beta + \beta_\Sigma'(\tilde{X}'\tilde{X})\beta_\Sigma - \beta_\Sigma'(\tilde{X}'\tilde{X})\beta_\Sigma\right\} \\
&\propto \exp\left\{-\frac{1}{2}(\beta-\beta_\Sigma)'(\tilde{X}'\tilde{X})(\beta-\beta_\Sigma)\right\}
\end{aligned} \tag{14}$$

Where the second line follows because we are taking Σ as given, and so it is only a constant, the third line uses (13) and completes the square, and the third line follows because the rest of line three is a constant that does not depend on β. Note that (14) is the kernel of a multivariate normal, $\beta \sim N(\beta_\Sigma, (X'(I_n \otimes \Sigma)^{-1}X)^{-1})$. After some algebra, one gets

$$\beta \sim N\left(\beta_\Sigma, \Sigma \otimes \left[\sum_{t=1}^{T} \mathbf{x}_t\mathbf{x}_t'\right]^{-1}\right).$$

A.2 $p(Z \mid \theta, Y)$

Given, β and Σ, we impute the missing data, starting from the likelihood in (8). First, assume the simplest case, where $p = 1$. Also assume that observation t^{th} is missing and observations $t+1$ and $t-1$ are not. In this case, for the t observation, we have

$$\begin{aligned}
p(R_{t+1}, R_t \mid R_{t-1}, \beta, \Sigma) \propto& \\
&\exp\left\{-\frac{1}{2}[(R_{t+1} - A_o + A_1R_t)'\Sigma^{-1}(R_{t+1} - A_o + A_1R_t)]\right\} \\
&\times \exp\left\{-\frac{1}{2}[(R_t - A_o + A_1R_{t-1})'\Sigma^{-1}(R_t - A_o + A_1R_{t-1})]\right\}
\end{aligned}$$

Operating we get

$$\begin{aligned}
&p(R_t \mid R_{t+1}, R_{t-1}, \beta, \Sigma) \propto \\
\exp\Big\{ &-\frac{1}{2}R_t'(A_1'\Sigma^{-1}A_1 + \Sigma^{-1})R_t - 2R_t'[(A_1'\Sigma^{-1}R_{t+1}' + \Sigma^{-1}A_1R_{t+1})] \\
&+ (A_1'\Sigma^{-1}A_0 - \Sigma^{-1}A_0)\Big\}
\end{aligned}$$

or

$$p(R_t \mid R_{t+1}, R_{t-1}, \beta, \Sigma) \propto \exp\left\{-\frac{1}{2}\left[R_t'\Omega^{-1}R_t - 2R_t'\Phi\right]\right\},$$

where

$$\Omega = \left[A_1'\Sigma^{-1}A_1 + \Sigma^{-1}\right]^{-1},$$

$$\Phi = A_1'\Sigma^{-1}R_{t+1} + \Sigma^{-1}A_1R_{t-1} - (A_1' - I)\Sigma^{-1}A_0.$$

After some manipulation, it is seen that:

$$p(R_t \mid R_{t+1}, R_{t-1}, \beta, \Sigma) \propto \exp -\frac{1}{2}[(R_t - \Omega\Phi)'\Omega^{-1}(R_t - \Omega\Phi)],$$

which is the kernel of multivariate Normal distribution with mean $\Omega\Phi$ and variance Ω. For $p \geq 1$, the most general case, the predictive distribution is given by

$$p(R_t \mid R_{t+1}, \cdots, R_{t+p}, R_{t-1}, \cdots, R_{t-1}, \cdots, R_{t-p}, \beta, \Sigma) \sim N(\Omega\Phi, \Omega), \qquad (16)$$

where

$$-\phi = \left(\begin{bmatrix} -I \\ A_1 \\ \cdots \\ A_{p-1} \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_1 \\ \\ \cdots \\ A_p \end{bmatrix}\right)' R_{t+1} +$$

$$\left(\begin{bmatrix} -I \\ A_1 \\ \cdots \\ A_{p-2} \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_2 \\ \\ \cdots \\ A_p \end{bmatrix}\right)' R_{t+2} + \cdots +$$

$$\left(\begin{bmatrix} -I \\ A_1 \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_{p-1} \\ A_p \end{bmatrix}\right)' R_{t+p-1} + \left[-I\Sigma^{-1}A_p\right]' R_{t+p} +$$

$$\left(\begin{bmatrix} -I \\ A_1 \\ \cdots \\ A_{p-1} \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_1 \\ \\ \cdots \\ A_p \end{bmatrix}\right)' R_{t-1} +$$

$$\left(\begin{bmatrix} -I \\ A_1 \\ \cdots \\ A_{p-2} \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_2 \\ \\ \cdots \\ A_p \end{bmatrix}\right)' R_{t-2} + \cdots +$$

$$\left(\begin{bmatrix} -I \\ A_1 \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_{p-1} \\ A_p \end{bmatrix}\right)' R_{t-p+1} + \left[-I\Sigma^{-1}A_p\right]' R_{t-p} +$$

$$\left(\begin{bmatrix} -I \\ A_1 \\ \cdots \\ A_p \end{bmatrix}' \begin{bmatrix} \Sigma^{-1} & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & & 0 \\ 0 & \cdots & 0 & \Sigma^{-1} \end{bmatrix} \begin{bmatrix} A_0 \\ \\ \cdots \\ A_0 \end{bmatrix}\right).$$

and

$$\Omega = \left[A_p'\Sigma^{-1}A_p + \cdots + A_1'\Sigma^{-1}A_1 + \Sigma^{-1}\right]^{-1}.$$

The Gibbs sampler iterates over (16), (10) and (15).

Table 4: VAR Models for Daily Raw Returns

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5) LB[5]	(6) LB[10]	(7) J-B	(8) B~GC A	(9) A ~GC B
				PANEL A				
BAN A	4	172	0.14645	0.52391	1.04702	17.28	4.48691	
				(0.99122)	(0.99979)	(0.00)	(0.00182)	
B			0.19637	2.85903	4.49606	19.05		8.88740
				(0.72171)	(0.92221)	(0.00)		(0.00000)
BIT A	1	43	-0.01910	1.46828	1.59370	102.33	0.21665	
				(0.91670)	(0.99861)	(0.00)	(0.64413)	
B			0.00620	1.69736	2.05243	182.29		2.02093
				(0.88923)	(0.99592)	(0.00)		(0.16289)
CEM A	5	1055	0.09377	0.23449	1.84218	1171.75	17.62091	
				(0.99870)	(0.99741)	(0.00)	(0.00000)	
B			0.04002	0.42723	4.66943	1918.65		4.68170
				(0.99455)	(0.91214)	(0.00)		(0.00031)
CIF A	5	296	0.11964	0.21796	7.49617	297.41	4.28702	
				(0.64060)	(0.67792)	(0.00)	(0.00089)	
B			0.10526	1.53966	3.55623	905.41		3.61027
				(0.90845)	(0.96515)	(0.00)		(0.00350)
DES A	5	18	0.14183	2.48495	2.48495	1.50	0.60659	
				(0.77876)	(0.99109)	(0.47)	(0.69903)	
B			0.00004	0.50854	0.50854	0.66		0.71240
				(0.99181)	(0.99997)	(0.72)		(0.63359)
GFA A	1	53	0.17824	1.05706	1.53920	302.98	0.08102	
				(0.95783)	(0.99881)	(0.00)	(0.77710)	
B			0.13039	0.78334	0.82369	208.93		2.85683
				(0.97808)	(0.99993)	(0.00)		(0.09721)

Table 4 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5) LB[5]	(6) LB[10]	(7) J-B	(8) B~GC A	(9) A ~GC B
GFB A	3	1120	0.04838	3.32407	18.16064	1521.61	10.16250	
				(0.65016)	(0.05231)	(0.00)	(0.00000)	
B			0.02242	0.64026	11.54147	2874.30		5.43441
				(0.98609)	(0.31692)	(0.00)		(0.00104)
GFN A	1	108	0.08183	2.53788	6.89467	843.68	1.25218	
				(0.77078)	(0.73535)	(0.00)	(0.26569)	
B			-0.00382	3.92364	5.36164	201.97		0.00042
				(0.56046)	(0.86575)	(0.00)		(0.98370)
GH A	2	39	0.02035	0.98710	1.02433	3.44	1.77079	
				(0.96360)	(0.99981)	(0.18)	(0.18554)	
B			0.03174	0.22297	1.40814	1.38		2.07073
				(0.99885)	(0.99919)	(0.50)		(0.14171)
GIS A	1	10	0.39515	1.25133	2.17664	1.13	3.13177	
				(0.93986)	(0.97511)	(0.57)	(0.12009)	
B			0.19723	0.39859	0.74265	0.28		0.10317
				(0.99537)	(0.99941)	(0.87)		(0.75743)
INB A	3	182	0.16287	2.03923	3.31559	24.73	7.17687	
				(0.84369)	(0.97299)	(0.00)	(0.00014)	
B			0.02211	1.21628	3.73613	63.29		1.38316
				(0.94331)	(0.95847)	(0.00)		(0.24955)
INL A	1	32	0.09484	2.73333	2.73865	20.88	5.16048	
				(0.74102)	(0.98692)	(0.00)	(0.03071)	
B			0.02498	2.63251	2.66068	7.01		1.30967
				(0.75642)	(0.98832)	(0.03)		(0.26181)

Table 4 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5) LB[5]	(6) LB[10]	(7) J-B	(8) B~GC A	(9) A ~GC B
KMC A	2	32	0.18947	0.71356	0.76768	2.88	2.70708	
				(0.98222)	(0.99995)	0.24	(0.08482)	
B			0.35164	2.46281	2.89953	4.38		4.94328
				(0.78209)	(0.98369)	0.11		(0.01481)
PON A	1	8	0.32641	7.06891	N/A	0.23	0.98926	
				(0.21557)		(0.89)	(0.36559)	
B			0.00393	3.38972	N/A	0.27		1.85125
				(0.64014)		(0.88)		(0.23176)
SEG A	1	161	0.19992	10.61563	13.68074	58.16	22.90849	
				(0.05956)	(0.18806)	(0.00)	(0.00000)	
B			0.12672	9.80620	13.90731	150.23		3.99475
				(0.08092)	(0.17726)	(0.00)		(0.04736)
SER A	3	94	0.03562	4.46147	5.16299	147.22	2.56834	
				(0.48506)	(0.88003)	(0.00)	(0.05953)	
B			0.10178	2.34060	2.52539	156.94		3.95102
				(0.80029)	(0.99050)	(0.00)		(0.01082)
SID A	5	40	0.29275	0.28710	0.44308	110.76	1.94102	
				(0.99788)	(1.00000)	(0.00)	(0.11791)	
B			0.11743	2.29817	3.44127	0.52		1.24468
				(0.80654)	(0.96906)	(0.77)		(0.31414)
PANEL B[1].								
PDR A	4	1164	0.11114	6.01418	14.41226	1032.09	5.64440	
				(0.30484)	(0.15500)	(0.00)	(0.00017)	
B			0.05681	3.37198	10.13764	4540.85		0.98695
				(0.64284)	(0.42850)	(0.00)		(0.41357)

Table 4 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5) LB[5]	(6) LB[10]	(7) J-B	(8) B~GC A	(9) A ~GC B
PDR1 A	5	745	0.15839	0.06872	6.01417	1139.50	4.22470	
				(0.99942)	(0.81407)	(0.00)	(0.00086)	
B			0.07198	0.01866	7.68608	668.95		0.90989
				(1.00000)	(0.65947)	(0.00)		(0.47393)
PDR2 A	4	392	0.15426	7.17112	14.60062	214.20	1.47762	
				(0.20822)	(0.14731)	(0.00)	0.20824	
B			0.06269	1.83052	10.35248	248.35		1.04937
			(0.87206)	(0.41013)	(0.00)		(0.38148)	

VAR models are estimated according to the procedure explained in the text. $t = 0$ is June 1, 1994. For Panel A, the dependent variable is the daily raw returns on the ith-shares listed in each row; the independent variables are the lagged daily raw returns on A-shares and the lagged daily raw returns on B-shares. For Panel B, the dependent variable is the portfolio of daily raw returns on the i^{th}-shares listed Table I; the independent variables are the lagged daily returns on the portfolios of returns on A-shares and the lagged daily returns on the portfolio of returns on B-shares. All regressions are estimated applying the OLS procedure. Regressions include a constant term. The name of the dependent variable appears in column 1. Column 2 shows the VAR length. The number of observations included in every VAR model, after dropping missing-observations and after subtracting the number of lags, appears in column 3. For each regression, the R^2 adjusted by degrees of freedom is shown in column 4. The Ljung-Box Q test statistic for five and ten autocorrelations, [LB(5)] and [LB(10)], are presented in columns 5 and 6, respectively. Their corresponding levels of significance are presented underneath in parenthesis. Column 7 presents the Jarque-Bera test statistic for normality of residuals. Its level of significance appears underneath in parenthesis. For returns on A-shares the Granger causality test statistic is presented in column 8. Its level of significance is shown underneath in parenthesis. B~GC A reads "Null: B does not Granger cause A". The analogous figures for returns on B-shares appear in column 9.

[1] The model for PDR is a VAR that includes all the observations in the sample period. In the case of PDR1, the VAR includes observations until June 30, 1997. Finally, the VAR associated to PDR2 includes observations from July 1, 1997, until the end of the sample period.

Table 5: Granger Causality Tests for Alternative Specifications

(1) Stock	(2) Lags	(3) N	(4) $Rbar^2$	(5)B~ GC A	(6)A~ GC B
BAN A	3	217	0.13071	7.61449	
				(0.00007)	
B			0.15401		7.53838
					(0.00008)
BIT A		N/A			
B					
CEM A	4	1068	0.10611	25.59522	
				(0.00000)	
B			0.03723		4.39852
					(0.00157)
CIF A	4	324	0.10763	3.80673	
				(0.00488)	
B			0.06177		0.89506
					(0.46712)

Table 5 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5)B$\sim$ GC A	(6)A$\sim$ GC B
DES A	4	28	0.15997	2.30776	
				(0.09543)	
B			0.27084		1.10337
GFA A		N/A			
B					
GFB A	2	1129	0.04403	12.93354	
				(0.00000)	
B			0.02212		7.02557
					(0.00093)
GFN A		N/A			
B					

Table 5 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5)B~ GC A	(6)A~ GC B
GH A	1	60	0.11216	1.14198	
				(0.28974)	
B			0.00836		1.35990
					(0.24841)
INB A	2	226	0.15856	19.77220	
				(0.00000)	
B			0.01736		3.92292
					(0.02118)
INL A		N/A			
B					
KMC A	1	69	0.12699	4.89757	
				(0.03036)	
B			0.05144		5.50905
					(0.02118)

Table 5 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5)B$\sim$ GC A	(6)A$\sim$ GC B
SEG A		N/A			
B					
SER A	2	134	0.05624	3.20174	
				(0.04395)	
B			0.07802		5.87675
SID A	4	53	0.18280	0.38534	
				(0.81794)	
B			0.16389		0.41145
(1) Stock	(7) Lags	(8) N	(9) Rbar^2	(10)B$\sim$ GC A	(11)A$\sim$ GC B
BAN A	5	138	0.13668	3.88734	
				(0.00260)	
B			0.18295		5.56610
					(0.00011)

Table 5 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5)B~ GC A	(6)A~ GC B
BIT A	2	21	0.15338	1.86010	
				(0.18779)	
B			-0.01544		0.42015
					(0.66399)
CEM A	6	1043	0.09104	14.23428	
				(0.00000)	
B			0.04121		4.49230
					(0.00017)
CIF A	6	274	0.10345	2.72293	
				(0.01393)	
B			0.08096		2.36819
					(0.03030)
DES A		N/A			
B					

Table 5 (continue)

(1) Stock	(2) Lags	(3) N	(4) Rbar^2	(5)B$\sim$ GC A	(6)A$\sim$ GC B
GFA A	2	29	0.09803	0.12715	
				(0.88119)	
B			0.02730		0.21627
					(0.80707)
GFB A	4	1111	0.04754	7.35402	
B			0.02137		4.27254
					(0.00195)
GFN A	2	69	0.06117	2.44474	
				(0.09481)	
B			-0.00688		0.60858
					(0.54724)
GH A	3	26	0.03519	2.27207	
				(0.11304)	
B			-0.00421		1.60288
					(0.22186)

Table 5 (continue)

(1) Stock	(7) Lags	(8) N	(9) Rbar^2	(10)B$\sim$ GC A	(11)A$\sim$ GC B
INB A	4	148	0.14556	3.66648	
				(0.00716)	
B			0.00387		0.73271
					(0.57111)
INL A	2	18	0.19882	3.09304	
				(0.07966)	
B			0.19516		3.19626
					(0.07430)
KMC A	3	14	0.54842	3.01432	
				(0.10367)	
B			0.73393		6.58968
					(0.01899)
SEG A	2	102	0.21853	6.08033	
				(0.00322)	
B			0.16673		9.15521
					(0.00022)

Table 5 (continue)

(1) Stock	(7) Lags	(8) N	(9) Rbar^2	(10)B$\sim$ GC A	(11)A$\sim$ GC B
SER A	4	68	0.03562	2.12296	
				(0.08921)	
			0.10178		5.77384
					(0.00054)
SID A	6	31	0.12025	0.18432	
				(0.97748)	
B			0.65075		7.19278
					(0.00049)

Two alternative models are presented for every VAR presented in Table II. $t = 0$ is June 1, 1994. The dependent variable is the daily raw returns on the i^{th}-shares listed in each row. The independent variables are the lagged daily raw returns on A-shares and the lagged daily raw returns on B-shares. All regressions include a constant term. For the first alternative model, results are presented in columns 2 - 6. Results for the second model are presented in columns 7-11 Columns 2 and 7 present the lenght of each VAR. The numbers of observations, after dropping missing data and subtracting the number of lags, appear in columns 3 and 8. The R^2 adjusted by degrees of freedom are shown in columns 4 and 8. For returns on A-shares the Granger causality test statistics are presented in columns 5 and 10. Their levels of significance are presented underneath in parenthesis. B $\sim$GC A reads "Null: B does not Granger cause A". Columns 6 and 11 present the analogous figures for the returns on B-shares.

Table 6 VAR Models for Single Events

(1) Event	Window	(2) Stock	(3) N	(4)Lags	(5)N VAR	(6)Rbar^2	(7)$\sim$ GC A	(8)A$\sim$GC B
27-Jun-94	31-Oct-94	BAN A	70	5	36	0.5236	5.2421	3.9370
		B	88			0.5201	(0.0020)	(0.0090)
29-Jul-94	2-Dec-94	A	63	4	34	0.1675	1.0426	5.6373
		B	85			0.5542	(0.4052)	(0.0022)
30-Sep-96	10-Feb-97	A	36	4	12	0.9648	8.1606	1.2907
		B	91			0.7456	(0.0580)	(0.4342)
27-Sep-96	7-Feb-97	BIT A	16	1	6	0.2045	0.2239	0.0126
		B	54			0.5072	(0.6684)	(0.9176)
12-Apr-95	16-Aug-95	CEM A	87	4	83	0.3865	8.9160	1.8932
		B	87			0.2686	(0.0000)	(0.1206)
18-Sep-95	22-Jan-96	A	86	4	82	0.0537	2.0601	0.9052
		B	86			-0.0170	(0.0948)	(0.4656)
25-Nov-96	9-Apr-97	A	91	1	89	0.1059	8.2899	3.2323
		B	91			0.0714	(0.0050)	(0.0757)
23-Jun-94	27-Oct-94	CIF A	52	1	43	-0.0379	0.2089	0.3689
		B	88			-0.0130	(0.6501)	(0.5471)
3-Oct-94	6-Feb-95	A	47	1	31	0.0803	2.8578	0.1630
		B	85			0.0354	(0.1020)	(0.6895)

Table 6 (continue)

(1) Event	Window	(2) Stock	(3) N	(4)Lags	(5)N VAR	(6)Rbar^2	(7)$\sim$ GC A	(8)A$\sim$GC B
5-Apr-94	9-Aug-94	CRI A	3	NA				
		B	0					
29-Nov-95	3-Apr-96	A	11	NA				
		B	7					
5-Apr-95	9-Aug-95	DES A	21	1	13	-0.1520	0.2498	0.0168
		B	85			-0.1349	(0.6281)	(0.8996)
30-Aug-95	3-Jan-96	A	7	NA				
		B	81					
15-Jan-97	29-May-97	A	27	1	14	0.4287	7.2441	11.3221
		B	91			0.4919	(0.0210)	(0.0063)
5-Oct-94	8-Feb-95	GFA A	27	NA				
		B	13					
15-May-95	18-Sep-95	A	8	NA				
		B	5					
8-Jun-94	12-Oct-94	GFB A	88	1	86	0.0391	2.0030	2.6728
		B	90			0.0571	(0.1607)	(0.1059)
1-Jul-96	5-Nov-96	A	91	1	89	0.0051	1.4901	1.7057
		B	91			-0.0033	(0.2255)	(0.1950)

Table 6 (continue)

(1) Event	Window	(2) Stock	(3) N	(4)Lags	(5)N VAR	(6)Rbar^2	(7)~ GC A	(8)A~GC B
7-Jul-95	10-Nov-95	GFN A	4	NA				
		B	42					
5-Apr-95	9-Aug-95	GH A	32	1	6	0.9639	22.4614	1.0177
		B	32			-0.1798	(0.0178)	(0.3874)
2-Nov-95	7-Mar-96	A	24	1	6	0.0208	1.8983	0.0185
		B	22			-0.6115	(0.2621)	(0.9003)
16-Nov-94	22-Mar-95	GIS A	10	1	4	-0.6456	0.3889	1.2530
		B	39			0.2500	(0.6450)	(0.4642)
27-Oct-95	1-Mar-96	A	8	NA				
		B	30					
4-Jul-94	7-Nov-94	INB A	31	NA				
		B	21					
7-Dec-95	11-Apr-96	A	49	4	15	0.1068	0.9927	1.7222
		B	82			0.1469	(0.5172)	(0.3092)
17-May-96	23-Sep-96	A	44	1	28	0.0067	2.0905	0.5022
		B	82			-0.0573	(0.1606)	(0.4851)
10-Aug-95	14-Dec-95	INL A	33	1	9	0.2077	4.0906	0.4205
		B	41			-0.0476	(0.0896)	(0.5407)

Table 6 (continue)

(1) Event	Window	(2) Stock	(3) N	(4)Lags	(5)N VAR	(6)Rbar^2	(7)∼ GC A	(8)A∼GC B
29-Dec-94	4-May-95	KMC A	87	1	11	0.2415	2.9347	2.2932
		B	22			0.0304	(0.1251)	(0.1684)
29-Nov-95	3-Apr-96	A	86	1	10	-0.0369	1.0451	0.0022
		B	18			-0.1986	(0.3407)	(0.9642)
24-May-94	28-Sep-94	LAT A	10	1	8	-0.0669	1.5452	11.0879
		B	8			0.5867	(0.2689)	(0.0208)
11-Apr-94	15-Aug-94	PON A	0	NA				
		B	2					
30-Mar-94	8-Aug-94	SEG A	15	NA				
		B	17					
8-Jul-95	8-Nov-95	A	44	1	8	0.4240	16.7242	2.4420
		B	43			0.1872	(0.0236)	(0.2347)
28-Dec-95	10-May-96	A	64	1	37	-0.0500	0.0633	4.7546
		B	64			0.0881	(0.8029)	(0.0362)
21-Nov-96	7-Apr-97	SER A	23	2	9	0.2497	0.3036	0.5979
		B	82			0.0966	(0.7538)	(0.5927)
28-Jun-96	4-Nov-96	A	28	2	10	0.2352	0.3974	1.5628
		B	87			0.2187	(0.6916)	(0.2970)

Table 6 (continue)

(1) Event	Window	(2) Stock	(3) N	(4)Lags	(5)N VAR	(6)Rbar^2	(7)~ GC A	(8)A~GC B
26-Oct-94	11-Mar-95	SID A	23	1	13	-0.1533	0.3918	0.0016
		B	90			0.0417	(0.5454)	(0.9690)
12-Dec-95	16-Apr-96	A	26	1	17	-0.0971	0.2359	2.1909
		B	81			0.0197	(0.6347)	(0.1610)
7-Jul-95	10-Nov-95	SIT A	0	NA				
		B	89					
12-Dec-95	16-Apr-96	A	1	NA				
		B	84					
18-Jul-94	21-Nov-94	TIE A	1	NA				
		B	2					

The dependent variable is the daily raw returns on the i^th-shares listed in each row; the independent variables are the lagged daily raw returns on A-shares and the lagged daily raw returns on B-shares. All regressions are estimated applying the OLS procedure. Regressions include a constant term. Column 1 presents the dates defininig the window for each event. The name of the dependent variable appears in coumn 2. The number of observations, column 3, refers to the available number of observations after dropping missing values. Column 4 shows the number of observations used in the VAR. Column 5 presents the R^2 adjusted by degrees of freedom. For returns on A-shares the Granger causality test statistic is presented in the column 7. Its the level of significance is presented underneath in parenthesis. B~ GC A reads "Null: B does not Granger cause A". Column 8 shows the analogous figures for returns on B-shares.

Table 7: VAR Models for Random Simulated Samples

	(1)True Model		(2) All data Model		(3) Random missing1			
					VAR		Gibbs Sampler	
	A	B	A	B	A	B	A	B
A(1)	0.5	-0.2	0.50811	-0.21194	0.47884	-0.19203	0.59570	-0.10292
			(0.02799)	(0.02384)	(0.06693)	(0.05709)	(0.00017)	(0.00094)
B(1)	-0.2	-0.5	-0.12528	0.52980	-0.03074	0.61990	-0.05810	0.65990
			(0.03199)	(0.02725)	(0.07327)	(0.06249)	(0.00232)	(0.00064)
C	0	0	0.01914	-0.00104	-0.05895	-0.02325	0.02214	-0.01238
			(0.03637)	(0.03098)	(0.08942)	(0.07627)	(0.00016)	(0.00016)
A								144
B								312
A,B								119
Mean			0.05016	-0.02201			-0.07844	0.07871
S. D.			(1.32475)	(1.15874)			(1.16676)	(0.95561)
			Granger Causality Tests					
H_0: B does not Granger Cause A			15.33530	0.00010	0.17599	0.67533		
H_0: A does not Granger Cause B			79.01290	0.00000	11.31393	0.00094		

Table 7 (continue)

	(4) Random missing2				(5) Random missing 3			
	VAR		Gibbs Sampler		VAR		Gibbs Sampler	
	A	B	A	B	A	B	A	B
N		241		997		714		998
A(1)	0.41158	-0.27054	0.57150	-0.13304	0.52898	-0.19061	0.51803	-0.1750
	(0.0540)	(0.0413)	(0.0001)	(0.0013)	(0.0343)	(0.0289)	(0.0002)	(0.000)
B(1)	-0.08321	0.50337	-0.12207	0.65640	-0.14399	0.53913	-0.13509	0.5507
	(0.0649)	(0.0497)	(0.0035)	(0.0011)	(0.0375)	(0.0316)	(0.0013)	(0.000)
C	-0.01116	-0.00638	-0.00261	-0.01850	0.05059	-0.00942	0.01126	-0.0046
	(0.0717)	(0.0549)	(0.0002)	(0.0001)	(0.0436)	(0.0367)	(0.0000)	(0.000)
A				131				47
B				299				99
A,B				77				11
Mean			0.15871	0.10396			0.22897	0.0072
S.D.			(1.1253)	(1.0653)			(1.0387)	(0.925)
	Granger Causality Tests							
H_0:B does not Granger cause A	1.64179	0.2013			14.6901	0.0001		
H_0:A does not Granger cause B	42.7472	0.0000			43.4595	0.0000		

Section 1 presents the coefficients used to construct simulated a random sample. The errors used follow a Standard Normal distribution. 1100 observations were modeled. The first and the last 50 observations are not used. With the simulated sample, a VAR is estimated. Results are shown in 2. Then, simulated random missing observations were added in the following fashion: the proportion of the total number of missing values was drawn from a uniform (0,1) distribution. With this number, every actual observation to be missed was drawn from a uniform $(0,1)^*$ 1000. Right after each number of missing observation is known, whether A, B or both are missing is chosen with a probability of one third. The program computes three samples. The results for every one of these samples are presented in 3, 4 and 5. The number of observations net of lags appears in the "N" row. Row A(1) presents the coefficients associated to the lagged values of returns on A-shares. Row B(1) shows the analogous figures for returns on B-shares. Standard errors appear in parenthesis. Row "A" refers to the number of missing values in the series; the same figure for series B, and for both series, are presented, respectively in rows "B" and "A, B". Row "Mean" presents, for the true model, the mean of the sample; this figure for the Gibbs estimates represent the average imputation error. The row "S. D." shows the standard deviations of the sample for the true model; for the Gibbs estimates, it presents the standard deviation of the imputation error. Granger causality tests for the regression model follows the standard test procedure.

Table 8: VAR Models for Random Simulated Samples

Dependent	A(1)	A(2)	A(3)	A(4)	A(5)	B(1)	B(2)
BAN A	0.23189	0.30844	0.21192	0.02877		0.16207	0.08282
	(0.02076)	(0.00920)	(0.04915)	(0.10040)		(0.08094)	(0.00648)
B	0.14337	0.23663	0.06451	-0.18150		0.20396	-0.21402
	(0.07385)	(0.01262)	(0.26243)	(0.35148)		(0.02318)	(0.01054)
BIT A	0.38449					-0.34345	
	(0.01849)					(0.24530)	
B	-0.44989					0.24164	
	(0.31093)					(0.02139)	
CEM A	-0.22405	-0.17858	-0.06245	-0.00522	0.03952	0.44292	0.16491
	(0.00068)	(0.00099)	(0.00074)	(0.00055)	(0.00027)	(0.00052)	(0.00091)
B	0.45213	0.22899	0.21289	0.15275	0.12611	-0.24378	-0.28130
	(0.00117)	(0.00161)	(0.00161)	(0.00186)	(0.00157)	(0.00094)	(0.00166)
CIF A	-0.03506	-0.09633	0.02929	-0.01430	-0.08712	0.12636	-0.03544
	(0.13116)	(0.09713)	(0.10594)	(0.10506)	(0.07242)	(0.09835)	(0.14269)
B	0.11053	-0.07605	0.07357	-0.04026	-0.07173	-0.11513	-0.02369
	(0.03263)	(0.09234)	(0.15371)	(0.19999)	(0.11492)	(0.02605)	(0.05575)

Table 8 (continue)

Dependent	A(1)	A(2)	A(3)	A(4)	A(5)	B(1)	B(2)
DES A	-0.01503	-0.02694	0.03874	-0.01598	-0.01237	0.06623	-0.01016
	(0.00483)	(0.00620)	(0.02335)	(0.00452)	(0.07261)	(0.02766)	(0.01948)
B	0.02688	-0.00519	0.04461	-0.00389	-0.45602	0.07204	0.00047
	(0.00740)	(0.00188)	(0.09378)	(0.03394)	(0.42584)	(0.01084)	(0.01342)
GFA A	0.92976						-0.02742
	(0.00010)						(0.00032)
B	-0.00546						0.84725
	(0.00022)						(0.00026)
GFB A	0.01174	-0.10558	0.02217			0.16396	0.07985
	(0.00111)	(0.00107)	(0.00101)			(0.00072)	(0.00076)
B	0.17262	-0.10020	0.10879			0.02006	-0.00247
	(0.00172)	(0.00121)	(0.00054)			(0.00103)	(0.00093)
GFN A	0.25771					0.28132	
	(0.17698)					(0.44009)	
B	0.35478					0.06891	
	(0.37072)					(0.16240)	

Table 8 (continue)

Dependent	B(3)	B(4)	B(5)	C	N	A	B	A,B
BAN A	-0.00252	-0.19897		-0.29826	1168	667	5	9
	(0.21577)	(0.33991)		(0.03529)				
B	-0.05123	-0.06146		0.42515				
	(0.04944)	(0.09194)		(0.04166)				
BIT A				0.42502	591	205	26	274
				(0.12676)				
B				0.66249				
				0.16978				
CEM A	0.03812	-0.00625	-0.01340	0.15620	1168	37	0	0
	(0.00080)	(0.00064)	(0.00031)	(0.00116)				
B	-0.22131	-0.16965	-0.11115	0.12599				
	(0.00162)	(0.00172)	(0.00123)	(0.00081)				
CIF A	0.04721	0.03562	0.05220	0.71209	855	313	2	0
	(0.14494)	(0.17056)	(0.09807)	(0.14672)				
B	-0.02687	0.05947	0.09661	0.29575				
	(0.08741)	(0.11431)	(0.06756)	(0.27383)				

Table 8 (continue)

Dependent	B(3)	B(4)	B(5)	C	N	A	B	A,B
DES A	0.06235	0.00106	-0.38620	0.27432	1113	848	1	17
	(0.10201)	(0.05065)	(0.39342)	(0.09007)				
B	-0.04436	0.09462	0.10634	0.12997				
	(0.00364)	(0.02549)	(0.02941)	(0.12595)				
GFA A				-0.00936	854	43	70	644
				(0.00031)				
B				0.14911				
				(0.00048)				
GFB A	0.04321			0.06255	1168	21	0	0
	(0.00088)			(0.00165)				
B	-0.06091			0.11785				
	(0.00053)			(0.00031)				
GFN A				0.38140	1130	696	21	227
				(0.08512)				
B				0.04974				
				(0.01726)				

Table 8 (continue)

Dependent	A(1)	A(2)	A(3)	A(4)	A(5)	B(1)	B(2)
GH A	0.42618	0.19292				0.11673	0.20585
	(0.03734)	(0.04921)				(0.17773)	(0.23879)
B	0.12312	0.23432				0.21224	-0.06733
	(0.14926)	(0.21167)				(0.01175)	(0.04227)
GIS A	0.97989					-0.03367	
	(0.01075)					(0.14401)	
B	-0.03073					0.41355	
	(0.03028)					(0.08572)	
INB A	0.08878	0.02370	0.02543			0.24233	-0.05979
	(0.11714)	(0.05727)	(0.03669)			(0.29007)	(0.16634)
B	0.29433	-0.10590	0.12258			-0.03770	0.01331
	(0.35719)	(0.14212)	(0.03851)			(0.08901)	(0.09799)
INL A	0.02131					0.37443	
	(0.08652)					(0.22485)	
B	0.33111					0.41468	
	0.27301)					(0.08345)	
KMC A	0.00417	0.00221				0.76918	-0.02932
	(0.02275)	(0.07367)				(0.01158)	(0.03179)
B	0.81848	0.00016				-0.01380	-1.13104
	(0.40009)	(0.02932)				(0.02229)	(0.12485)

Table 8 (continue)

Dependent	A(1)	A(2)	A(3)	A(4)	A(5)	B(1)	B(2)
PON A	0.73737					0.01946	
	(0.05660)					(0.02084)	
B	0.20607					0.08470	
	(0.08302)					(0.05134)	
SEG A	-0.02472					0.51553	
	(0.05774)					(0.08223)	
B	0.43071					0.10584	
	(0.07505)					(0.04913)	
SER A	0.03789	0.02229	-0.00991			0.14108	0.07559
	(0.01861)	(0.01058)	(0.06404)			(0.11046)	(0.08185)
B	0.13126	0.19135	0.12882			0.16855	-0.10646
	(0.10941)	(0.13517)	(0.33899)			(0.05026)	(0.00549)
SID A	0.02956	0.02173	0.00500	0.00921	-0.03158	0.13077	0.05426
	(0.02960)	(0.00993)	(0.02601)	(0.00621)	(0.08337	(0.05666)	(0.05799)
B	0.12421	0.06700	0.08529	0.16330	-0.31946	0.08663	0.10001
	(0.01480)	(0.09300)	(0.09249)	(0.08251)	(0.44115)	(0.04044)	(0.01011)

Table 8 (continue)

Dependent	B(3)	B(4)	B(5)	C	N	A	B	A,B
GH A				0.09325	695	127	114	346
				(0.01433)				
B				-0.03937				
				(0.03828)				
GIS A				0.42968				
				(0.25848)	438	155	12	237
B				(0.90477)				
				(0.14852)				
INB A	0.09767			0.23878	933	505	9	30
	(0.05327)			(0.02118)				
B	0.12379			0.21150				
	(0.03102)			(0.07470)				
INL A				-0.66613	375	63	57	196
				(0.23654)				
B				-0.36795				
				(0.06163)				
KMC A				-1.89265	1130	0	979	0
				(0.15947)				
B				5.65673				
				(0.58186)				

Table 8 (continue)

Dependent	B(3)	B(4)	B(5)	C	N	A	B	A,B
PON A				2.47461	71	8	4	48
				(0.37971)				
B				-2.97863				
				(0.62410)				
SEG A				0.21380	999	153	115	474
				(0.01219)				
B				-0.35946				
				(0.12695)				
SER A	0.13643			0.26888	1151	715	19	123
	(0.26358)			(0.01219)				
B	-0.05309			-0.35046				
	(0.04450)			(0.12695)				
SID A	0.10992	0.16648	-0.18966	2.22797	843	516	17	104
	(0.05616)	(0.06063)	(0.36832	(0.41681)				
B	0.07373	-0.01036	-0.02504	-0.27262				
	(0.02025)	(0.00802)	(0.05550)	(0.37110)				

Results are obtained using the Gibbs Sampler described in the text. $t = 0$ is June 1, 1994. The dependent variables are the daily raw returns on the i^{th}-shares listed in each row. The independent variables are the lagged daily raw returns on A-shares and the lagged daily raw returns on B-shares. The estimates are the average over 10,000 iterations. Columns "A(1)" to "A(5)" present the estimates associated with the lagged values of the returns on A-shares. The analogous figures for returns on B-shares are shown in columns "B(1)" to "B(5)". Estimates of the constant term are presented in column C. The number of observations included in every VAR model, after imputing missing observations and after subtracting the number of lags, appears in the N Column. Columns A, B and "A, B" present the number of imputed observations in the returns on A-shares, B-shares, or both. Standard errors appear in parenthesis.

References

Bhattacharya, U., H. Daouk, B. Jorgenson, and C. Kehr (2000). When an Event is Not an Event: The Curious Case of an Emerging Market. *Journal of Financial Economics*, 55, pp. 69-101.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). The Econometrics of Financial Markets, Princeton, Princeton University Press.

Domowitz, I., J. Glen, and A. Madhavan (1998). International Cross- Listing and Order FlowMigration: Evidence from an Emerging Market. *Journal of Finance*, 53, pp. 2001-27.

Domowitz, I., J. Glen, and A. Madhavan (1987). Market Segmentation and Stock Prices: Evidence from an Emerging Market. *Journal of Finance*, 52, pp. 1059-85.

Enders, W. (1995). Applied Econometric Time Series, Series in Probability and Mathematical Statistics, New York, Wiley.

Gelman, A., J. B. Carlin, H. S. Stern, and B. R. Donald (1995). Bayesian Data Analysis. Fourth edition, New York, Chapman & Hall.

Gelman, A. and B. Donald (1992). Inference from Iterative Simulation Using Multiples Sequences. *Statistical Science*, 7, pp. 457-511.

Geyer, Charles J. (1992). Practical Markov Chain Monte Carlo. Statistical Science, 7, pp. 473-511.

Gilks, W. R., S. Richardson, and D. J. Spiegelhater (1996). Markov Chain Monte Carlo in Practice. First edition, New York, Chapman & Hall.

Hamilton, J. D. (1994). Time Series Analysis, Princeton, Princeton University Press.

Huberman, G. and W. G. Schwert (1985) Information Aggregation, In- flation, and the Pricing of Indexed Bonds. *Journal of Political Economy*, 93, pp. 92-114.

Kabir R. and T. Vermaeler (1996) Insider Trading Restrictions and the Stock Market: Evidence from the Amsterdam Stock Exchange. *European Economic Review*, 40, pp. 1591-1603.

Little, J. A. and D. B. Rubin (1987). Statistical Analysis with Missing Data, first edition, New York, John Wiley & Sons.

Meulbroek, L. K. (1992) An Empirical Analysis of Illegal Insider Trading. *Journal of Finance*, 47, pp. 1661-99.

Morris, C. N. (1987) Simulation in Hierarchical Models: Comment The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82, pp. 542-43.

Sargan, J. D. and E. G. Drettakis (1974). Missing Data in an Autoregressive Model. *International Economic Review*, 15(1), pp. 39-58.

Seyhun, H. N. (1988). The Information Content of Aggregate Insider Trading. *Journal of Business*, 61, pp. 1-24.

Seyhun, H. N. (1986) Insiders Profits, Costs of Trading, and Market E.ciency. *Journal of Financial Economics*, 16, pp. 189-212.

Sims, C. (1980). Macroeconomic and Reality. *Econometrica*, 48, pp. 1-49.

Tanner, M. A. (1993) Tools for Statistical Inference. Methods for the Exploration of Posterior Distributions and Likelihood Functions, second edition, New York, Springer-Verlag.

Tanner, M. A. and W. H. Wong (1987). The Calculation of Posterior Distributions by Data Augmentation. *Journal of the American Statistical Association*, 82, pp. 528-40.