Empirical Analysis of the Impact of Foreign Institutional Investment on the Indian Stock Market Volatility during World Financial Crisis 2008-09

Dr. Hojatallah Goudarzi
University of Mysore, India
E-mail: hg502003@yahoo.com

Dr. C.S.Ramanarayanan
University of Mysore, India

Received: January 15, 2010 Accepted: February 8, 2010 doi:10.5539/ijef.v3n3p214

Abstract

Capital inflows play a substantial role in developing countries. It used to increase accumulation and rate of investments to create conditions for more intensive economic growth. Capital inflows are necessary for macroeconomic stability as capital inflows affect a wide range of macro economic variables such as exchange rates, interest rates, foreign exchange reserves, domestic monetary conditions as well as saving and investments. Capital inflows, however, are not without risk. The main risk posed by large and volatile capital inflows is that they may resulted in crisis and destabilize macroeconomic management. Given the role of FII flews and its associated risks, the main purpose of this paper was to investigate the cointegration and causality between the Indian stock market and foreign institutional investment (FII) In India during world financial turmoil of 2008. The cointegration and causal relationship using Engle-Granger (1987), Johansen (1991, 1995a) and Granger (1969) methodologies were investigated .The study found that BSE500 stock index and FII series are cointegrated and causality between them is bilateral.

Keywords: Stock Markets, Foreign Institutional Investments (FII), Cointegration, Granger Causality

1. Introduction

Foreign capital has significant role for every national economy, regardless of its level of development. For the developed countries it is necessary to support sustainable development. For the developing countries, it is used to increase accumulation and rate of investments to create conditions for more intensive economic growth. For the transition countries, it is useful to carry out the reforms and cross to open economy, to cross the past long term problems and to create conditions for stable and continuous growth of GDP, as well as integration in world economy. But, to realize the potential exist in the developing countries, foreign capital plays a very crucial role. Some commonly observed effects of the capital include real exchange rate appreciation, stock market and real estate boom, reserve accumulation, monetary expansion as well as effect on production and consumption.

Since the introduction of the reform process in the early 1990s, India has witnessed a significant increase in capital inflows. The size of net capital inflows to India increased from US $ 7.1 billion in 1990-91 to US $ 108.0 billion in 2007-08. Today, India has one of the highest net capital inflows among the EMEs of Asia (Sumanjeet, 2009).

Capital inflows, however, are not without risk. The main risk posed by large and volatile capital inflows is that they may resulted in crisis and destabilize macroeconomic management. In case of India, after an impressive performance for nearly five years, foreign capital inflows lost their momentum in the second half of 2008. The most significant change was observed in the case of FIIs, which saw a strong reversal of flows. Against a net inflow of US$20.3 billion in FY2007–2008, there was a net outflow of US$15 billion from Indian markets during FY2008–2009 as foreign portfolio investors sought safety and mobilized resources to strengthen the balance sheet of their parent companies. This massive outflow of FII created panic in the stock markets. Consequently, equity markets lost more than 60% of their index value and about US$1.3 trillion of market capitalization from an index peak of about 21,000 in January 2008 to 8,867 by 20 March 2009 (Kumar and Vashisht, 2009).

The main purpose of this study is to test the relationship between FII flows and BSE500 stock index movements in the Indian stock market. The paper has been arranged as follows: Section 2 presents common methodology of cointegration and causality. Section 3 provides the review of literature. Section 4 gives the results of empirical study. In the end, paper concludes with conclusions.

2. Cointegration and Causality Theoretical Issues

According to (Watson and Teelucksingh, 2002) estimation and hypothesis testing based on OLS is justified only if the two variables involved are I(0). Since the underlying variables are I (1), a fairly reasonable expectation is that any linear combination of the variables, such as e_t, would also be I (1). This violates the basic assumptions for OLS estimation and if we insist on applying OLS, we are likely to establish nothing more than spurious correlation, i.e. a correlation that does not establish any *causal* relationship between the variables.

A tempting solution to this apparent problem is to fit the regression using the first differences of variables. But in a seminal paper, Davidson, Hendry, Srba and Yeo (1978) argue that such an approach would ignore valuable information about the "long run". They propose instead an approach that blends the variables and incorporates the short run dynamics implied by the first differences as well as the static or long run relationship between the undifferenced values which enter the relationship as an "error correction mechanism" (ECM).Engle and Granger (1987) show that the solution proposed by Davidson et al. (1978) is possible if and only if the variables involved in the relationship are cointegrated.

2.1. Cointegration

According to Brooks (Brooks, 2008) if two variables that are I (1) are linearly combined, then the combination will also be I(1). More generally, if variables with differing orders of integration are combined, the combination will have an order of integration equal to the largest. If $X_{i,t} \sim \mathrm{I}(d_i)$ for *i* = 1, 2, 3, . . . , *k* so that there are *k* variables each integrated of order d_i, and letting

$$z_t = \sum_{i=1}^{k} \alpha_i X_{i,t} \qquad 1$$

Then $z_t \sim \mathrm{I}\ (\max d_i)$.$z_t$ in this context is simply a linear combination of the *k* variables X_i. Rearranging (1)

$$x_{1t} = \sum_{i=2}^{k} \beta_i X_{i,t} + z'_t \qquad 2$$

Where $\beta_i = -\frac{\alpha_i}{\alpha_1}, z'_t = \frac{z_t}{\alpha_1}, i = 2, \ldots.., k$. All that has been done is to take one of the variables, $X1,t$, and to rearrange (2.68) to make it the subject. It could also be said that the equation has been normalized on $X_{1,t}$. But viewed another way, (2) is just a regression equation where z'_t is a disturbance term. These disturbances would have some very undesirable properties: in general z'_t, will not be stationary and is autocorrelated if all of the X_i are I (1).

As a further illustration, consider the following regression model containing variables y_t, x_{2t}, x_{3t}which are all I(1)

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + u_t \qquad 3$$

For the estimated model, the SRF would be written

$$y_t = \hat{\beta}_1 + \hat{\beta}_2 x_{2t} + \hat{\beta}_3 x_{3t} + \hat{u}_t \qquad 4$$

Taking everything except the residuals to the LHS

$$y_t - \hat{\beta}_1 - \hat{\beta}_2 x_{2t} - \hat{\beta}_3 x_{3t} = \hat{u}_t \qquad 5$$

Again, the residuals when expressed in this way can be considered a linear combination of the variables. Typically, this linear combination of I(1) variables will itself be I(1), but it would obviously be desirable to obtain residuals that are I(0). Under what circumstances will this be the case? The answer is that a linear combination of I(1) variables will be I(0), in other words stationary, if the variables are *cointegrated*.

2.1.1. Definition of Cointegration (Engle And Granger, 1987)

Let *wt* be a *k* ×1 vector of variables, then the components of *wt* are integrated of order (*d*, *b*) if:

(1) All components of *wt* are I(*d*)

(2) There is at least one vector of coefficients *α* such that

$$\alpha' \omega_t \sim \mathrm{I}(d-b)$$

In practice, many financial variables contain one unit root, and are thus I(1), so that the remainder of this section will restrict analysis to the case where *d* = *b* = 1. In this context, a set of variables is defined as cointegrated if a linear combination of them is stationary. Many time series are non-stationary but 'move together' over time - that is, there exist some influences on the series (for example, market forces), which imply that the two series are bound by some relationship in the long run. A cointegrating relationship may also be seen as a long-term or equilibrium phenomenon, since it is possible that cointegrating variables may deviate from their relationship in the short run, but their association would return in the long run (Brooks, 2008).

2.2. Methods of Parameter Estimation in Cointegrated Systems

There are (at least) three methods that could be used: Engle--Granger, Engle-Yoo and Johansen. The first and third of these will be considered in some detail in this study.

2.2.1. The Engle–Granger 2-step method

This is a single equation technique, which is conducted as follows:

Step 1

In the first step one should make sure that all the individual variables are I (1). Then the cointegrating regression using OLS should be estimated. In this step any inferences on the coefficient estimates in this regression is impossible - all that can be done is to estimate the parameter values. Then saving the residuals of the cointegrating regression, $\hat{u}_t$ is required. Finally, one should test these residuals to ensure that they are I (0). If they are I (0), proceed to Step 2; if they are I(1), estimate a model containing only first differences.

Step 2

In this step the step 1 residuals should be used as one variable in the error correction model, e.g.

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (\hat{u}_{t-1}) + \upsilon_t \quad 6$$

Where $\hat{u}_{t-1} = y_{t-1} - \hat{\tau} x_{t-1}$. The stationary, linear combination of nonstationary variables is also known as the *cointegrating vector*. In this case, the cointegrating vector would be $[1-\hat{\tau}]$. Additionally, any linear transformation of the cointegrating vector will also be a cointegrating vector. It is now valid to perform inferences in the second-stage regression, i.e. concerning the parameters $\beta 1$ and $\beta 2$ (provided that there are no other forms of misspecification, of course), since all variables in this regression are stationary.

Suppose that the following specification had been estimated as a potential cointegrating regression

$$y_t = \alpha_1 + \beta_1 x_t + u_{1t} \quad 7$$

What if instead the following equation was estimated?

$$x_t = \alpha_2 + \beta_2 y_t + u_{2t} \quad 8$$

If it is found that $u_{1t} \sim \mathrm{I}(0)$, does this imply automatically that $u_{2t} \sim \mathrm{I}(0)$? The answer in theory is 'yes', but in practice different conclusions may be reached in finite samples. Also, if there is an error in the model specification at stage 1, this will be carried through to the cointegration test at stage 2, as a consequence of the sequential nature of the computation of the cointegration test statistic.

(3) It is not possible to perform any *hypothesis tests* about the actual cointegrating relationship estimated at stage 1.

2.3. Testing For and Estimating Cointegrating Systems Using the Johansen Technique Based On Vars

Suppose that a set of *g* variables ($g \geq 2$) are under consideration that are I(1) and which are thought may be cointegrated. A VAR with *k* lags containing these variables could be set up:

$$\underset{g\times 1}{y_t} = \underset{g\times g}{\beta_1}\underset{g\times 1}{y_{t-1}} + \underset{g\times g}{\beta_2}\underset{g\times 1}{y_{t-2}} + \cdots + \underset{g\times g}{\beta_k}\underset{g\times 1}{y_{t-k}} + \underset{g}{u_t} \quad 9$$

In order to use the Johansen test, the VAR (2.76) above needs to be turned into a vector error correction model (VECM) of the form

$$\Delta y_t = \prod y_{t-k} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_{k-1} \Delta y_{t-(k-1)} + u_t$$

Where $\prod = (\sum_{i=1}^{k} \beta_i) - Ig \ and \ \Gamma_i = (\sum_{j=1}^{i} \beta_j) - Ig$ 10

This VAR contains *g* variables in first differenced form on the LHS, and $k-1$ lags of the dependent variables (differences) on the RHS, each with a Γ coefficient matrix attached to it. In fact, the Johansen test can be affected by the lag length employed in the VECM, and so it is useful to attempt to select the lag length optimally, as outlined in chapter 6. The Johansen test centers around an examination of the $\prod$ matrix. $\prod$ can be interpreted as a long-run coefficient matrix, since in equilibrium, all the Δy_{t-i} will be zero, and setting the error terms, *ut* , to their expected value of zero will leave $\prod y_{t-k} = 0$. Notice the comparability between this set of equations and the testing equation for an ADF test, which has a first differenced term as the dependent variable, together with a lagged levels term and lagged differences on the RHS.

The test for cointegration between the *y*s is calculated by looking at the rank of the $\prod$ matrix via its eigenvalues. The rank of a matrix is equal to the number of its characteristic roots (eigenvalues) that are different from zero. The eigenvalues, denoted λ_i are put in ascending order $\lambda_1 \geq \lambda_2 \geq ..... \geq \lambda_g$ If the λ_s are roots, in this context they must be less than 1 in absolute value and positive, and λ_1 will be the largest (i.e. the closest to one), while λ_g will be the smallest (i.e. the closest to zero). If the variables are not cointegrated, the rank of $\prod$ will not be significantly different from zero, so $\lambda_i \approx 0 \forall i$. The test statistics actually incorporate $\ln(1-\lambda_i) = 0$, rather than the λ_i themselves, but still, when $\lambda_i = 0$, $\ln(1-\lambda_i) = 0$.

Suppose now that rank $(\Pi) = 1$, then $\ln(1-\lambda_1)$ will be negative and $\ln(1-\lambda_i) = 0 \forall i > 1$. If the eigenvalue *i* is non-zero, then $\ln(1-\lambda_i) < 0 \forall i > 1$. That is, for Π to have a rank of 1, the largest eigenvalue must be significantly non-zero, while others will not be significantly different from zero.

There are two test statistics for cointegration under the Johansen approach, which are formulated as

$$\lambda_{trace}(r) = -T \sum_{i=r+1}^{g} \ln(1-\hat{\lambda}_i) \qquad 11$$

And

$$\lambda_{max}(r, r+1) = -T \ln(1-\hat{\lambda}_{r+1}) \qquad 12$$

where *r* is the number of cointegrating vectors under the null hypothesis and $\hat{\lambda}_i$ is the estimated value for the *i*th ordered eigenvalue from the Π matrix. Intuitively, the larger is $\hat{\lambda}_i$, the more large and negative will be $\ln(1-\hat{\lambda}_i)$ and hence the larger will be the test statistic. Each eigenvalue will have associated with it a different cointegrating vector, which will be eigenvectors. A significantly non-zero eigenvalue indicates a significant cointegrating vector.

λ_{trace} is a joint test where the null is that the number of cointegrating vectors is less than or equal to *r* against an unspecified or general alternative that there are more than *r*. It starts with *p* eigenvalues, and then successively the largest is removed. λ_{trace} = 0 when all the λ_i = 0, for *i* = 1, . . . , *g*.

λ_{max} conducts separate tests on each eigenvalue, and has as its null hypothesis that the number of cointegrating vectors is *r* against an alternative of *r* + 1.

Johansen and Juselius (1990) provide critical values for the two statistics. The distribution of the test statistics is non-standard, and the critical values depend on the value of $g - r$, the number of non-stationary components and whether constants are included in each of the equations. Intercepts can be included either in the cointegrating vectors themselves or as additional terms in the VAR. The latter is equivalent to including a trend in the data generating processes for the levels of the series.

If the test statistic is greater than the critical value from Johansen's tables, reject the null hypothesis that there are *r* cointegrating vectors in favor of the alternative that there are *r* + 1 (for λ_{trace}) or more than *r* (for λ_{max}). The testing is conducted in a sequence and under the null, *r* = 0, 1, . . . , $g - 1$ so that the hypotheses for λ_{max} are

H_0: $r = 0$ versus H_1: $0 < r \leq g$

H_0: $r = 1$ versus H_1: $1 < r \leq g$

H_0: $r = 2$ versus H_1: $2 < r \leq g$

......

H_0: $r = g - 1$ versus H_1: $r = g$

The first test involves a null hypothesis of no cointegrating vectors (corresponding to Π having zero rank). If this null is not rejected, it would be concluded that there are no cointegrating vectors and the testing would be completed. However, if H_0: $r = 0$ is rejected, the null that there is one cointegrating vector (i.e. H_0: $r = 1$) would be tested and so on. Thus the value of *r* is continually increased until the null is no longer rejected.

But how does this correspond to a test of the rank of the _ matrix? *r* is the rank of Π . Π cannot be of full rank (*g*) since this would correspond to the original y_t being stationary. If Π has zero rank, then by analogy to the univariate case, Δy_t depends only on Δy_{t-j} and not on y_{t-1}, so that there is no long-run relationship between the elements of y_{t-1} . Hence there is no cointegration. For $1 < rank(\Pi) < g$, there are *r* cointegrating vectors. Π is then defined as the product of two matrices, *α* and β' , of dimension (*g* × *r*) and (*r* × *g*), respectively, i.e.

$$\Pi = \alpha\beta' \qquad 13$$

The matrix *β* gives the cointegrating vectors, while *α* gives the amount of each cointegrating vector entering each equation of the VECM, also known as the 'adjustment parameters'.

Finally, it must be noted that the above description is not exactly how the Johansen procedure works, but is an intuitive approximation to it(Brooks,2008).

2.4. Granger Causality

2.4.1. Two-Variable Models

According to (Idris and Cheong, 2004), Correlation, does not necessarily imply causation in any meaningful sense of the word. The econometric graveyard is full of magnificent correlations that are simply spurious or meaningless (Eviews 3). As for the efficient-market hypothesis (EMH), there may not necessarily be a real association between market efficiency and co-integration. Granger (1988) argues that co-integration between two prices imply an inefficient market as the error correction model indicates that at least one of the prices is predictable. Therefore, the Granger-type causality procedure (Granger, 1969 and 1988) was applied to determine the direction of causation among the Y and X series. The causality procedure was conducted based on a bivariate system [*x, y*].

Let X_t , Y_t , be two stationary time series with zero means. The simple causal model is

$$X_t = \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \varepsilon_{t,} \qquad 14$$

$$Y_t = \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_{t,} \qquad 15$$

Where ε_t , η_t are taken to be two uncorrelated white-noise series, i.e., $E[\varepsilon_t, \varepsilon_s] = 0 = E[\eta_t, \eta_s], s \neq t$ and $E[\varepsilon_t, \varepsilon_s] = 0$ all t, s. In above equations *m* can equal infinity but in practice, of course, due to the finite length of the available data, *m* will be assumed finite and shorter than the given time series.

The definition of causality given above implies that Y_t is causing X_t provided some β_j is not zero. Similarly X_t is causing Y_t if some c_i is not zero. If both of these events occur, there is said to be a feedback relationship between X_t and Y_t . The more general model with instantaneous causality is

$$X_t + \beta_0 Y_t = \sum_{j=1}^{m} \alpha_j X_{t-j} + \sum_{j=1}^{m} \beta_j Y_{t-j} + \varepsilon_{t,} \qquad 16$$

$$Y_t + c_0 X_t = \sum_{j=1}^{m} c_j X_{t-j} + \sum_{j=1}^{m} d_j Y_{t-j} + \eta_{t,} \qquad 17$$

If the variables are such that this kind of representation is needed, then instantaneous causality is occurring and a knowledge of Y_t will improve the "prediction" or goodness of fit of the first equation for X_t (Granger,1969).

The null hypothesis is that X does not Granger cause Y.

H0: No causal relationship from X to Y

Ha: X Granger-causes Y

This hypothesis would be rejected if the coefficients of the lagged *X* were jointly significant (different from zero). The null hypothesis for equation (2) is that "*Y* does not Granger cause *X*." This hypothesis would be rejected if the coefficients of the lagged *Y* were jointly significant. If both of these null hypotheses are rejected, then a bidirectional relationship (*X* ↔*Y*) is said to exist between the two variables. The causality patterns can be unidirectional causality, *X*(Y) → Y(*X*), (Granger,1969)

3. Review of literature

Mukherjee et al(2002) **,**in their paper explored the relationship of foreign institutional investment (FII) flows to the Indian equity market with its possible covariates based on a daily data-set for the period January 1999 to May 2002. The set of possible covariates considered comprises two types of variables. The first type includes variables reflecting daily market return and its volatility in domestic and international equity markets as well as measures of co-movement of returns in these markets (viz., relevant betas). The second type of variables, on the other hand, are essentially macroeconomic ones like exchange rate, short-term interest rate and index of industrial production (IIP)—viz., variables that are likely to affect foreign investors' expectation about return in Indian equity market. Their results show that: (1) FII flows to and from the Indian market tend to be caused by return in the domestic equity market and not the other way round; (2) returns in the Indian equity market is indeed an important (and perhaps the single most important) factor that influences FII flows into the country; (3) while FII sale and FII net inflow are significantly affected by the performance of the Indian equity market, FII purchase is not responsive to this market performance; (4) FII investors do not seem to use Indian equity market for the purpose of diversification of their investment; (5) return from exchange rate variation and fundamentals of the Indian economy may have influence on FII decisions, but such influence does not seem to be strong, and; finally, (6) daily FII flows are highly auto-correlated and this auto-correlation could not be accounted for by the all or some of the covariates considered in their study. They stated that, policy implications of the findings are that a move towards a more liberalized regime in the emerging market economies should be accompanied by further improvements in the regulatory system of the financial sector. their results additionally suggested that in the case of India (and other countries having thin and shallow equity markets) the prime focus should be on regaining investors' confidence in the equity market so as to strengthen the domestic investor base of the market. Once this is achieved, a built-in cushion against possible destabilizing effects of sudden reversal of foreign inflows might develop. Only then would it be possible to reap fully the benefits of capital market integration.

Bose and Coondoo (2004), examined the impact of the FII policy reforms on FII portfolio flows to the Indian stock markets. They tried to assess the impact on FII flows of several policy revisions related to FII investment during the period January 1999 to January 2004, through a multivariate GARCH regression model. Using techniques of time series intervention analyses they incorporated the effect of each individual policy intervention in a model that includes the two most important covariates of FII flows to India, namely stock market (BSE) returns and past FII flows. The range of policies considered encompasses liberalization policies as well as restrictive ones taken to assure stability of flows. Their results strongly suggested that liberalization policies have had the desired expansionary effect and have either increased the mean level of FII inflows and/or the sensitivity of these flows to a change in BSE return and/or the inertia of these flows. On the other hand, interestingly, the restrictive measures aimed at achieving greater control over FII flows also do not show any significant negative impact on the net inflows; we find that these policies mostly render FII investments more sensitive to the domestic market returns and raise the inertia of the FII flows.

Anokye and Tweneboah (2008), Using multivariate cointegration and error correction model, examined the impact of Foreign Direct Investment (FDI) on the stock market development in Ghana. Their results indicated that there exists a long run relationship between FDI, nominal exchange rate and stock market development in Ghana. They found that a shock to FDI significantly influence the development of stock market in Ghana. Their results had several policy implications. First, it supports the policy maker's decision to slash restriction for the non-resident investors for listed companies. This would attract major investors to other sectors of the economy to bring the needed growth in the exchange market and the economy as a whole. Second, policy makers should devise strategies to increase the FDI stock (retain FDI) and offer incentives for long investing and listing on the stock market.

Sumanjeet (2009) studied Foreign Capital Flows into India .he stated that existing studies reveals that the huge surge in international capital flows since early 1990s has created unprecedented opportunities for the developing countries like India to achieve accelerated economic growth. International financial institutions routinely advise developing countries to adopt policy regimes that encourage capital inflows. Since the introduction of the reform process in the early 1990s, India has witnessed a significant increase in capital inflows. The size of net capital inflows to India increased from US $ 7.1 billion in 1990-91 to US $ 108.0 billion in 2007-08. Today, India has one of the highest net capital inflows among the EMEs of Asia. Capital inflows, however, not an unmitigated blessing. The main danger posed by large and volatile capital inflows is that they may destabilize macroeconomic management. As evident, the intensified pressures due to large and volatile capital flows in India in the recent period in an atmosphere of global uncertainties has posed new challenges for monetary and exchange rate management. He concluded that undertaking more economic reforms is not easy but has to be done: the government can either manage the process or competitive forces will bring it upon us in a lopsided manner. The ball is in the government's court. Countries that permit free capital flows must choose between the stability provided by fixed exchange rates and the flexibility afforded by an independent monetary policy.

Bansal And Pasricha (2009), studied the impact of market opening to FIIs, on Indian stock market behavior. India announced its policy regarding the opening of stock market to FIIs for investment in equity and related instruments on 14th September 1992. Using stock market data related to Bombay Stock Exchange, for both before and after the FIIs policy announcement day, they conducted an empirical examination to assess the impact of the market opening on the returns and volatility of stock return. they found that while there is no significant changes in the Indian stock market average returns, volatility is significantly reduced after India unlocked its stock market to foreign investors.

4. Cointegration and Causality Empirical Results

4.1. Data

The monthly data of BSE500 and FII series used as proxy for the Indian stock market and FII series.

4.1.2. Cointegration

In this part of study we evaluated the relationship between BSE500 stock index and FII series. One of the commonly used tools for evaluation relation between 2 variables is correlation. To test this relation between BSE500 and FII we calculated the correlation between BSE500 movements and FII series for equity, debt and their total. The results are presented in table 2.

As results show there is a high and significant correlation between BSE500 movements and FII flows in all cases.

According to (Watson and Teelucksingh, 2002) estimation and hypothesis testing based on OLS is justified only if the two variables involved are I (0). Since both BSE500 and FII series in the level form are I (1), a fairly reasonable expectation is that any linear combination of these two variables, such as e_t , would also be I (1). This violates the basic assumptions for OLS estimation and if we insist on applying OLS, we are likely to establish nothing more than spurious correlation, i.e. a correlation that does not establish any causal relationship between the two variables. Therefore, whether the above correlations are true should be evaluated comprehensively before any implication.

A tempting solution to this apparent problem is to fit the regression using the first differences of both variables. But in a seminal paper, Davidson, Hendry, Srba and Yeo (1978) argue that such an approach would ignore valuable information about the "long run". They propose instead an approach that combined the two and incorporates the short run dynamics implied by the first differences as well as the static or long run relationship between the undifferenced values which enter the relationship as an "error correction mechanism" (ECM).

Engle and Granger (1987) show that the solution proposed by Davidson et al. (1978) is possible if and only if the variables involved in the relationship are cointegrated.

To test cointegration we test whether there is any cointegration and causality between movement in the BSE500 stock Index and FII series. First we used the Engle-Granger methodology to see if there is cointegration between two series. Before testing the residuals we made sure of unit root in both series. The results are reported in the tables 3.

As tables show both series are I (1) i.e. non-stationary in the level with intercept.

To test the cointegration between two series we estimate the following equation

$$B_t = \beta F_t + e_t \tag{18}$$

Where B is BSE500 index and F is FII series.

Then we set the residuals of this equation as follows

$$\hat{e}_t = B_t - \hat{\beta} F_t \tag{19}$$

Where $\hat{\beta}$ is the OLS estimator of β

The null and alternative hypotheses are:

H_0: The variables are not cointegrated (i.e. the OLS residuals admit a unit root)

H_1: The variables are cointegrated (i.e. the OLS residuals do not admit a unit root)

If the cointegratability is established, the error correction term is established by equation 19.

Step 1 of the Engle-Granger two-step procedure requires that we fit the cointegrating regression (by OLS) and test the residuals for unit roots (we have already established that B and F are I(1)).

We estimated the cointegrating regression by OLS the results are reported in table 4.

If the residuals are stationary, then we must reject the null of no cointegration. If they are nonstationary (the null hypothesis), then they do not cointegrated.

If the null of no cointegration based on ADF test is rejected, the next step in the Engle-Granger procedure is to estimate the short-run dynamics. Using the Engle-Granger methodology we test the null, the results are presented in table 5.

Since the test statistic (−4.18) is more negative than the critical values at the 1% level, the null hypothesis of a unit root in the test regression residuals is strongly rejected. We thus conclude that the two series are cointegrated. This means that an error correction model (ECM) can be estimated, as there is a linear combination of the BSE500 and FII series that would be stationary. The ECM would be the appropriate model rather than a model in pure first difference form because it would enable us to capture the long-run relationship between the series as well as the short-run one. Therefore, we estimate an error correction model by running the following regression

$$\Delta B_t = \gamma_{11}\Delta B_{t-1} + \gamma_{12}\Delta F_{t-1} + \alpha_1(B_{t-1} - \hat{\beta}F_{t-1}) + \varepsilon_{1t} \quad 20$$

or, given that $\hat{e}_t = B_t - \hat{\beta} Y_t$

$$\Delta B_t = \gamma_{11}\Delta B_{t-1} + \gamma_{12}\Delta F_{t-1} + \alpha_1 \hat{e}_{t-1} + \varepsilon_{1t} \quad 21$$

Therefore, the next step in the Engle-Granger procedure is to estimate the short-run dynamics in an equation system such as equations above, with:

$$\hat{e}_{t-1} = B_{t-1} - 0.089F_{t-1}$$

The results are reported in the tables 6 and 7.

α_1 is estimated as -0.135103. It has correctly signed and it is significant. This is enough to confirm that BSE500 and FII are cointegrated and that the ECM form estimated here is a valid representation of the model. α_2 is estimated as 0.646659 but it is not significant. A test of the residuals of these equations verifies that they are white noise. This is further evidence of the cointegratability of BSE500 and FII. Finally, according to above results about 13.5% of disequilibrium "corrected" each month.

Since we have only two variables and Engle-Granger approach can estimate only up to one cointegrating relationship between two variables then in our study, there can be at most one cointegrating relationship since there are only two variables in the model it is enough to our purpose here. Nevertheless, we examine the issue of cointegration within the Johansen VAR framework to confirm The correlation between two variables. The results are as follows:

The first step in the Johansen is to establish whether the variables are cointegrated and, indeed, whether there may be more than one cointegrating vector (a maximum of two is possible here). We use assumption 2 (no deterministic trend in the data, intercept in the CE) and an underlying VAR with two lags. The results are displayed in table 8.

The first thing to do in the Johansen procedure is to test the null of $r = 0$ against the alternative $r \geq 1$. As results show the Johansen test statistics show rejection for the null hypothesis of no cointegrating vectors under both the trace and maximal eigenvalue forms of the test. In the case of the trace, the null of no cointegrating vectors is rejected since the test statistic of 25.68 is greater than the 1% critical value of 20.04. Therefore, there is at least one cointegrating vector.

The next step is to test the null of $r = 1$ against the alternative of $r \geq 2$. Moving on to test the null of at most 1 cointegrating vectors, the trace statistic is 0.58 while the 5% and 1% critical values are 3.76 and 6.65 respectively, so the null is not rejected at both level of significance. Thus, we conclude that there is exactly one cointegrating vector.

The normalized cointegrating vector is estimated as (including the constant term):

$$\hat{\beta}' = (1 - 8.967)$$

The corresponding cointegrating regression deduced from normalization is:

$$e_t = B_t - 8.967F_t$$

where the right-hand side of this equation is in the form $\hat{\beta}' x_t$. For purposes of comparison with OLS, it is perhaps better to write this result in the following more conventional format:

$B_t = 8.967F_t + e_t$

This is somewhat different from the OLS result (which would have been the cointegrating vector obtained by application of the EG two-step procedure) although the difference is not dramatic.

We turn now to the estimation of the ECM model. The results obtained are shown in appendix 1. .At the top of table the (normalized) cointegrating vector is displayed and, below, the ECMs involving ΔB_t and ΔF_t as "dependent" variables are shown. On the right-hand side of each equation appears the cointegrating regression (CointEq1) and the coefficient attached to it is the "adjustment parameter". Here the adjustment coefficient associated with the ΔB_t equation is negative (–0.125841) and it is also significant (t-statistic = (2.11526). This is sufficient to reject any "no cointegration" hypothesis.

4.3. Granger Causality Test

Correlation does not necessarily imply causation in any meaningful sense of that word. The Granger (1969) approach to the question of whether X causesY is to see how much of the current Y can be explained by past values of Y and then to see whether adding lagged values of X can improve the explanation. Y is said to be Granger-caused by X if X helps in the prediction of Y, or equivalently if the coefficients on the lagged X's are statistically significant. Note that two-way causation is frequently the case; X Granger causes Y and Y Granger causes X. It is important to note that the statement "X Granger causes Y" does not imply that Y is the effect or the result of X. Granger causality measures precedence and information content but does not by itself indicate causality in the more common use of the term.

Using the granger causality framework we test the null hypothesis that FII does not Granger-cause BSE and that BSE does not Granger-cause FII in the. To test these hypotheses we estimated two following bivariate regression models respectively:

$$B_t = \alpha_0 + \alpha_1 B_{t-1} + \cdots + \alpha_l B_{t-l} + \beta_1 F_{t-1} + \cdots + \beta_l F_{t-l} + \varepsilon_t \quad 22$$

$$F_t = \alpha_0 + \alpha_1 F_{t-1} + \cdots + \alpha_l F_{t-l} + \beta_1 B_{t-1} + \cdots + \beta_l B_{t-l} + u_t \quad 23$$

Where:

B_t = The dependent

F_t = The explanatory variable

u_t = A zero mean white noise error term in Eq. 5 while

F_t = The dependent

B_t = The explanatory variable in Eq. 6

For all possible pairs of(F,B) series in the group. The reported F-statistics are the Wald statistics for the joint hypothesis:

$\beta_1 = \beta_2 = \cdots = \beta_l = 0$

For each equation, the null hypothesis and alternative hypotheses are as follows:

First equation

$H_{0:}$ *B* does not Granger cause F, i. e,

$\{\beta_1, \beta_1, \cdots, \beta_l\} = 0, if\ F_C < critical\ value\ F$

$H_1 : B$ does Granger cause Y,i.e.,

$\{\beta_1, \beta_2, \cdots, \beta_l\} \neq 0, if\ F_C > critical\ value\ F$

And

Second equation

$H_0 : F$ does not Granger cause B,i.e.,

$\{\beta_1, \beta_1, \cdots, \beta_l\} = 0, if\ F_C < critical\ value\ F$

$H_1 : F$ does Granger cause B,i.e.,

$\{\beta_1, \beta_2, \cdots, \beta_l\} \neq 0, if\ F_C > critical\ value\ F$

In order to test the above hypotheses the usual Wald F-statistic test is utilized, which has the following form:

$$\mathrm{F} = \frac{(RSS_R - RSS_U)/q}{RSS_U/(T - 2q - 1)}$$

Where:

RSS_U = the sum of squared residuals from the complete (unrestricted) equation

RSS_R = the sum of squared residuals from the equation under the assumption that a set of variables is redundant, when the restrictions are imposed, (restricted equation)

T = the sample size

q = the lag length

We picked a lag length, based on our assumption about the time period over which FII could help predict the BSE500 index. Our assumption was that BSE500 affected by FII outflows from the beginning of 2008 to Dec

2008.We used the SC information criterion to select an appropriate lag length in our analysis. SC criterion indicates lag order selected by the criterion is 12.All other criterion also confirm this selection.

Given the selected lag length we tested the estimated models to test the hypothesis regard Granger causality between the variables under question. The results are reported in the table 9.

As a decision rule, if $F_{calc} > F_{crit}$ then reject $H_0 \Rightarrow$ lagged FII terms are significant, $\Rightarrow$ FII does Granger-cause BSE500.

According to the results, we do reject both null hypotheses. Econometrics results reveal that causality is bilateral. In other words, all coefficients are significant. The evidence of causal linkage between two variables implies that since each of this variables are cointegrated the predictability of each can enhanced significantly by utilizing information on other variable.

5. Conclusions and policy recommendations

Decline of about 60 percent in the index and a wiping off of about USD1.3 trillion in market capitalization since January 2008 when the Sensex had peaked at about 21,000 which was primarily due to the withdrawal of about USD 12 billion from the market by foreign portfolio investors between September and December 2008(Kumar, 2009) and effects of this on local investors responses, was a disaster for the Indian stock market.

This paper examined the cointegration and causal relationship between BSE500 and FII series in Indian economy. Since estimation and hypothesis testing based on OLS is justified only if the two variables involved are I (0) and since both BSE500 and FII series in the level form are I (1), a fairly reasonable expectation is that any linear combination of these two variables, such as e_t , would also be I (1). This violates the basic assumptions for OLS estimation and if we insist on applying OLS, we are likely to establish nothing more than spurious correlation, i.e. a correlation that does not establish any causal relationship between the two variables. Therefore, whether the above correlations are true should be evaluated comprehensively before any implications (Watson and Teelucksingh, 2002).

To overcome this problem we test whether there is any cointegration and causality between movement in the BSE500 stock Index and FII series. Using two popular econometrics methods our results confirmed the cointegration between BSE500 and FII series. Further, econometrics results reveal that causality between BSE500 and FII is bilateral. The evidence of causal linkage between two variables implies that since each of this variables are cointegrated the predictability of each can enhanced significantly by utilizing information on other variable.

Given the risk of capital inflows, it is recommended that price limits and volume quotas be used relative to the status of both the economy and Indian stock market trading cycles. When the economy is in recovery, these tools need to be loosened to attract investors and reduce the investors' pessimism about the market. However, in an extreme boom these tools need to be applied with greater intensity to reduce excess optimism that could develop into speculative behavior in the market. In edition to curb the negative impact of market volatility it is better to apply only lower level price limits and volume quotas, rather than both upper and lower limits in a time of panic or when the markets start to fragment.

Without any doubt FII play a great positive role in the Indian economy. But as world financial crisis of 2008 showed it is a potential risk as well, a risk with low frequency but high severity which classified as catastrophic risk in the literature. This type of risk is speculative risk and is not insurable. Therefore it is necessary to cover this risk to prevent future potential crisis. Given the impossibility of insuring this type of risk by direct and indirect insurance system, it is recommended that Indian financial system policy makers consider a self insurance (hedging) mechanism through establishing a captive and build up funds there based on Maximum Possible Loss(in this case capital outflows) to prevent same future probable crisis.

References

Anokye A. M. and Tweneboah. G. (2008). Foreign direct investment and stock market development: Ghana's evidence. [Online]Available: http://mpra.ub.uni-muenchen.de/11261/1/MPRA_paper_11261.pdf

Bansal. A and Pasricha. J.S. (2009). Foreign institutional investor's impact on stock prices in India, journal of academic research in economics.[Online]available: http://www.jare-sh.com/downloads/num_1_issue_2_2009/bansal.pdf

Bose.S and Coondoo .D (2004). The impact of FII regulations in India a time-series intervention analysis of equity flows, I C R A B U L L E T I N *Money AND Finance*. [Online]Available: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=755324

Brooks, C. (2008). *Introductory Econometrics For Finance*. (2nd ed.). Cambridge. Cambridge University Press.

Granger .C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods, *Econometrica*, Vol. 37, 3,424-438.

Idris. F and Cheong T.T.(2004),Inter-counter linkage in Kuala Lumpur stock exchange returns,[Online] Available :http://www.sunway.edu.my/others/vol1/faoziah.pdf.

Mukherjee. P, Bose.S And Coondoo. (2002). Foreign institutional investment in the Indian equity market an analysis of daily flows during January 1999-may 2002.[Online]Available:http://papers.ssrn.com/sol3/papers.cfm?abstract_id=430700.

Kumar .R and Pankaj Vashisht P. (2009). The global economic crisis: impact on India and policy responses, [Online]Available: http://www.adbi.org/files/2009.11.12.wp164.global.economic.crisis.india.pdf.

Sumanjeet. (2009), Foreign capital flows into India: compositions, regulations, issues and policy options[Online]Available: http://wwwarc.murdoch.edu.au/wp/wp155.pdf.

Watson P. K. Teelucksing S. S. (2002). *Econometric methods: classical and modern*. India, University of the West Indies press.

Table 1. Capital Flows into India after 1990's (US $ million)

FII FLOWS TO INDIA	
Year	**FII**
1990-91	-
1991-92	-
1992-93	1
1993-94	1665
1994-95	1503
1995-96	2009
1996-97	1926
1997-98	979
1998-99	-390
1999-00	2135
2000-01	1847
2001-02	1505
2002-03	377
2003-04	10918
2004-05	8686
2005-06	9926
2006-07	3225
2007-08	20328

Source: Hand Book of Statistics, Reserve Bank of India (RBI

Table2.Correlation Between BSE500 And FII Flows

Correlation	FII TOTAL	P-Value	FII DEBT	P-Value	FII EQUITY	P-Value
BSE500	0.87	0.0000	0.65	0.0000	0.87	0.000

Table 3.Unit root test for BSE500 and FII in the level

variables	Unit root test			
	Augmented Dickey-Fuller test statistic	Test critical values		
		1%	5%	10%
BSE	-0.721642	-3.496346	-2.890327	-2.582196
FII	-1.178706	-3.505595	-2.894332	-2.584325

BSE stands for BSE500 stock index

FII stands for foreign institutional investment

Table 4. OLS Estimation for Cointegration Analysis

Dependent Variable: INDEX				
Variables	**Coefficient**	**Std. Error**	**T-statistics**	**Probability**
FII	0.089607	0.002479	36.14222	0.0000

INDXused for BSE500 stock index

FII stand for foreign institutional investment

Table 5. Unit Root Test of Residuals of BSE500 and FII Series

variables	Unit root test for residual			
	Augmented Dickey-Fuller test statistic	Test critical values		
		1%	5%	10%
RESID01	-4.185465	-3.505595	-2.894332	-2.584325

RESID01 stand for residuals of bse500 stock index and FII series

Table 6. ECM Estimation When BSE500 Is Dependent

Dependent Variable: DINDEX				
Variables	**Coefficient**	**Std. Error**	**T-statistics**	**Probability**
DFII	-0.005603	0.007336	-0.763819	0.4468
ECM(-1)	-0.135103	0.034517	-3.914069	0.0002

DINDEX stand for first difference of BSE500 stock index

Table 7. ECM Estimation When FII Is Dependent

Dependent Variable: DFII				
Variables	**Coefficient**	**Std. Error**	**T-statistics**	**Probability**
DINDEX	-1.045588	1.368894	-0.763819	0.4468
ECM(-1)	0.646659	0.502506	1.286869	0.2011

DFII stand for first difference of FII series

Table 8. Cointegration Test Using Johansen Procedure

Johansen Cointegration Test				
Number of cointegrations	**Eigenvalue**	**Statistic**	**5 Percent Critical Value**	**1 Percent Critical Value**
None**	0.223927	25.68395	15.41	20.04
At most 1	0.005908	0.586623	3.76	6.65
*(**) denotes rejection of the hypothesis at the 5%(1%) level Trace test indicates 1 cointegrating equation(s) at both 5% and 1% levels				

Table 9. Granger Causality test

Pairwise Granger Causality Tests		
Null Hypothesis	F-Statistic	Probability
DINDEX does not Granger Cause DDFII	8.55097	2.4E-09
DDFII does not Granger Cause DINDEX	3.67542	0.00033

DINDEX is first difference of BSE500 Index

DDFII is second difference of FII series

Appendix 1: Var Model

Vector Error Correction Estimates

Standard errors in ( ) & t-statistics in [ ]

Cointegrating Eq:	CointEq1	
FII(-1)	1.000000	
INDEX(-1)	-8.967266	
	(0.73961)	
	[-12.1243]	
C	-6168.838	
Error Correction:	**D(FII)**	**D(INDEX)**
CointEq1	-0.125841	0.020557
	(0.05949)	(0.00433)
	[-2.11526]	[4.74759]