Herd Behavior towards the Market Index: Evidence from 21 Financial Markets

Daxue Wang

(*corresponding author*)
IESE Business School
Tel.: +34-651 99 05 87
dwang@iese.edu

Miguel Canela

University of Barcelona
Tel: +34-93 402 1636
miguelcanela@ub.edu

ABSTRACT

This paper uses the cross-sectional variance of the betas to study herd behavior towards market index in major developed and emerging financial markets (categorized as developed group, Asian group, and Latin American group). We propose one of the robust regression techniques to calculate the betas of the CAPM and those of the Fama-French three factor model, with an intention to diminish the impact of multivariate outliers in return data. Through the estimated values obtained from a state space model, we examine the evolution of herding measures, especially their pattern around sudden events such as the 1997-1998 financial crises. This 1997-1998 turmoil turns out to have formed a turning point for most of the financial markets. We document a higher level of herding in emerging markets than in developed markets. We also find that the correlation of herding between two markets from the same group is higher than that between two markets from different groups. This paper will shed light on the calculation of beta and on the financial policy to understand the dynamics of herding in financial markets.

JEL Code: C1, C60, G12, G14, G15

Keywords: Beta, Herding, Kalman Filter, Outlier, Robust Regression, Cycle.

Introduction

In financial markets, herding is usually termed as the behavior of an investor to imitate the observed actions of others or the movements of market instead of following his own beliefs and information. Possibly herding is among the most mentioned but least understood terms in the financial lexicon. Difficulties to measure and quantify the existence of the behavior form obstacles for extensive research. Even so, there are at least two points people tend to unanimously agree upon. First, as one of the founding pillars in newly developed behavioral asset pricing area, herding helps to explain market wide anomalies. Since individual biases are not influential enough to move market prices and returns, they have real anomalous effect only if there is a social contamination with a strong emotional content, leading to more widespread phenomena such as herding. Second, it is generally accepted that the flood of herding may lead to a situation in which the market price fails to reflect all relevant information and thus the market becomes unstable and moves towards inefficiency. Policy makers often express concerns that herding by financial market participants destabilizes markets and increases the fragility of the financial system. As a result, it is in their interest to curtail herding (Bikhchandani and Sharma, 2001).

Theoretical and empirical research on herding has been conducted in an isolated manner. Theoretical study focuses on the causes and implications of herding. The main consensus is that herding can be construed as being either a rational or irrational form of investor behavior. According to Devenow and Welch (1996), the irrational view focuses on investor psychology where an investor follows others blindly. On the other hand, imperfect information, concerns for reputation, and compensation structures foster rational herd behavior.[1]

The empirical studies thus far do not test a particular model of herding behavior described in the theoretical literature; instead, they gauge whether clustering of decisions, in a purely statistical sense, is taking place in financial markets or within certain investor groups. Two streams of empirical literature have been developed to investigate the existence of herding

[1] See Bihkchandani and Sharma (2000) for an overview of the theoretical research on rational herd behavior in financial markets.

in financial markets. The first stream analyzes the tendency of individuals or certain groups of investors, such as mutual fund managers and financial analysts, to follow each other and trade an asset at the same time. In this case, detailed records of investors' trading activities are required. For instance, Lakonishok *et al.* (1992) measure herding as the average tendency of a group of money managers to buy or sell particular stocks at the same time, relative to what could be expected if the managers make their decision independently; Grinblatt *et al.* (1995) use data on portfolio changes of mutual funds between 1974 and 1984 to examine herding among fund managers and the relation of such behavior to momentum investment strategy. Wermers (1995) proposes a portfolio-change measure, by which herding is measured by the extent to which portfolio weights assigned to the various stocks by different money managers move in the same direction;

The second stream focuses on market-wide herding, that is, the collective behavior of all participants towards the market views and therefore buying or selling particular asset at the same time. Christie and Huang (1995) regress the cross-sectional (market wide) standard deviation of individual security returns on a constant and two dummy variables designed to capture extreme positive and negative market returns. They argue that during periods of market stress rational asset pricing would imply positive coefficients on these dummy variables, while herding would suggest negative coefficients. However, the introduction of dummy variables is entirely arbitrary since the choice of what is meant by "extreme" is subjective. And they do not control for movements in fundamentals, so it is hard to tell whether the negative coefficient, if there is any, is herding or just a sign of independent adjustment to fundamentals that is taking place.

Based on the cross-sectional dispersion concept of Christie-Huang, Hwang and Salmon (2004) use the cross-sectional dispersion of beta to detect herding towards the market index. The authors apply their model to the US and Korean stock markets, and they find this herd behavior shows significant movements and persistence over the sample period. One merit of their paper is that they separate the herding from "spurious herding", common movements in asset returns being induced by movements in fundamentals. Herding potentially leads to market inefficiency whereas "spurious herding", or, fundamental adjustment, reflects just an efficient reallocation of assets on the basis of common information on fundamentals. However, they derive the monthly beta of an asset with daily return data over monthly intervals, a period that is too short to diminish the influence of unusual bad or good events of the company on the beta.[2]

The main purpose of our paper is to improve the Hwang-Salmon model and to investigate the herding towards the market in major financial markets form.[3] We do this in two dimensions. First, it is not realistic to us that Hwang-Salmon assume that the log of cross-sectional standard deviation of betas is normally distributed with a static mean. Starting from the assumptions of the stock returns, we explore the distribution of the cross-sectional variance of betas. By doing this, we get a time-varying cross-sectional dispersion of betas, which we believe is more realistic than the static one.

We then apply the model to various financial markets, thereby obtaining the monthly herding measures in each market, through which we can work on our concept of relative herd behavior. Pairwise correlations of herding measures are calculated among these markets. We also identify the pattern of herd behavior during sudden events such as the 1997-1998 global financial crises. It is hypothesized that the break out of financial crisis has a direct connection with the herding in the market. With the US and Korea markets, Hwang-Salmon conclude that financial crises stimulate a return towards efficiency and investors turn to fundamentals rather than the overall market movement during market stress. We test this argument in a global setting with constituents of more financial markets.

To obtain the cross-sectional dispersion of the betas, we propose the use of the "right" beta of the CAPM or the linear factor model, a highly debated area in empirical finance. In this paper, we adopt a rolling robust regression approach to calculate the betas. The purpose of using robust regression is to diminish the influence of outliers on the point estimate of beta. To our knowledge, this paper is the first one to calculate outlier robust beta on emerging markets. [4]

---

[2] Regarding the time period for which beta should be estimated, five years is a widely accepted alternative. With very short periods of time, there is the risk of capturing an unusually good or bad period for the company; with very long periods of time, the data could be less representative for the company.

[3] These markets are Australia, France, Germany, Hong Kong, Japan, United Kingdom, United States, China, India, Indonesia, Korea, Malaysia, Philippines, Thailand, Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Venezuela.

[4] It is worth noting that outliers could be bad noise or the most important information revealing aspect of the data. Hence further analysis to the identified outliers is necessary to provide more complete information. In other words, when influential outlier returns (containing important information) exist, neither the OLS betas

The original framework of the CAPM is developed in a single-period setting. In most of the empirical studies the beta is assumed to be constant over a defined period of time. However, this treatment of constant beta is contradicted by increasingly more evidence that beta is time-varying (Blume, 1971; Fabozzi and Francis, 1977; Fernandez, 2004). Several alternative models have been developed to capture the time-varying character of the beta: Fama and Macbeth (1973) propose a rolling regression approach to estimate the beta; Braun *et al.* (1995) use a bivariate EGARCH model to estimate a beta influenced asymmetrically by the market returns; Fama and French (1993) and Ferson and Harvey (1999) bring in macroeconomic variables to account for the beta; Faff and Brooks (1998) apply the Kalman filter approach to explain the stochastic evolution of the beta.

As for the accuracy of these alternative estimation methods, Groenewold and Fraser (2000) conclude that the rolling regression, although simple, is no less accurate than those more complicated models. Under the rolling regression, each month only one observation is new and therefore this overlapping problem leads to a high degree of autocorrelation in the beta time series. Regarding this, Greonewold and Fraser (2000) use non-overlapping sub-periods and conclude that this alteration does not change the results significantly. This evidence justifies the fact that rolling regression remains most popular among practitioners and in academic research. For instance, commercial resources such as Bloomberg Professional, Baseline, Value Line and Datastream provide betas of certain securities. Although each resource gives a different result for the beta of a security due to the several differences involved in the calculations, they typically use ordinary least square (OLS) to regress the return of a security on a market index over a certain period, typically 2 to 5 years.

However, these conventional calculations of the CAPM beta fail to consider the existence of bivariate or multivariate outliers, which may be quite large in real data. These outliers have substantial influence on the OLS point estimate of beta. The differences between the OLS estimate and the robust estimate might be viewed as financially significant by investors. Therefore, in calculating the rolling regression, we propose the robust estimation rather than the traditional OLS estimation.

---

nor the robust betas provide an adequate picture of the risk-return characteristics, and they may be combined to achieve more convincing results.

We also calculate, in addition to the CAPM beta, the beta under the three factor framework of Fama and French (1993). For comparison purpose, both the OLS and robust estimators are calculated. The establishment and analysis of Fama-French *HML* and *SMB* factor for stock markets other than the USA help broaden our understanding of global portfolio management.

The estimated time-varying herding measures allow us to examine the concept of relative herding which we propose in the paper. In brief, we find that herding towards the market is stronger in emerging markets than in developed markets. Additionally, we find that the herding measure, like some macroeconomics aggregate variables, follows a pattern of cycles, and some sudden events can sometimes be identified as turning points of the cycles. Furthermore, we do not observe any trend in the magnitude and the volatility of the herding measure over time. Finally, we see a higher correlation of herding between two markets from the same group than between markets from different groups.[5]

The remainder of the paper is organized as follows. Section I proposes the concept of relative herding and form the hypotheses. Section II develops the model and introduces the Kalman filter, together with the robust regression technique used in this paper. Section III describes the data. Empirical results on the distribution of betas and herding patterns are discussed in Section IV. Section V closes the paper with concluding remarks and directions of future research.

I. Herd Behavior in Financial Markets

There are various types of herding in both theoretical and empirical literature. In this paper, we focus on the evolution of herding towards the market index, a particular type of herding within the second stream of literature, as mentioned in the introduction. We will discuss how this type of herding affects the market and traditional asset pricing model in *Section II*.

1.1 A Concept of "Relative Herding"

[5] As it will be seen later, we group the sample markets by their development stage and geographical location.

In our opinion, no market is free of herd behavior. The notion of relative herding, i.e., the herding of one market measured against another or one period against another, may be more useful than the all-or-nothing view taken by the conventional literature.

1.1.1 Relative Herding: The Cross-Sectional Perspective

Jirasakuldech *et al.* (2004) point out that high informational efficiency countries are associated with a lower level of equity market volatility, which, according to Christie and Huang (1995), is an indication of less serious herd behavior in the market. Contrary to investors in developed markets, investors in newly established or emerging markets find it difficult or expensive to gather and collect information in order to conduct fundamental analysis. Instead, observing and imitating other investors' decision or the market index is relative easy and cheap, and as a result, herding can and often ensues in emerging markets (Komulainen, 2001). Empirically, Choe *et al.* (1999) show that herd behavior existed in South Korean stock market eventually led to economic instability with the financial crisis in 1997; Komulainen (2001) and Chang *et al.* (2000) report the existence of herding in stock markets of South Korea and Taiwan.

Based on these studies, we build our first hypothesis as follows:

H1: *Herding towards the market index is stronger in emerging than in developed markets.*

1.1.2 Relative Herding: The Time-Varying Perspective

In this section we propose that herding measures fluctuate over time, and some turning points can be associated with the occurrence of sudden events. The study of the time-varying character of herding is loosely inspired by the business cycle theory in economics. [6]

Business cycle study has a long tradition in economics, referring to the periodic fluctuations of economic activity along with its long term growth trend. There are many explanations to the existence of cycles. For instance, the psychological cycle explanation (by Arthur C. Pigou, among others) attributes it to the change of entrepreneurs' expectation of profits and confidence. It is assumed that the sense of optimism and pessimism motivates businesses to enlarge or contract investment. When the market is rife with optimistic expectation on

[6] In recent years economic theory has a trend of moving from the study of "business cycle" to "economic fluctuation", even though some economists still use the phrase "business cycle" as a convenient shorthand.

consumptions, firms will attempt to increase production, leading to a general overshooting of output. This overshooting of output then results in an oversupply of goods, causing the bankruptcy of some firms, a general collapse in output, and a wave of pessimism. In turn, this will lead to an undersupply of goods, thereby paving the way for the next wave of optimism.

Any further discussion of business cycle is beyond the interest of this paper, since our purpose is to draw an analogy between the investors in stock markets and the entrepreneurs. Imagine that when speculative prices in stocks or the whole market increase, the success story of some investors or the market index may attract public attention and promote word-of-mouth enthusiasm. Herding investors may thus become more herding-oriented, heightening expectations for further price increases. This process in turn increases investor demand and thus generates another wave of price increase. If the feedback is not interrupted, it may produce a speculative "bubble", in which high expectations, instead of fundamental values, support high current prices, thus making the bubble vulnerable and easy to burst, whereby people start changing their herd behavior. Since the feedback that propelled the bubble carries the seeds of its own destruction, the end of the bubble may be unrelated to new stories about fundamentals as well. A similar feedback process applies to negative bubbles. [7] In combination, we can expect fluctuations in the herding towards the market, an interwoven of upward and downward trends.

Like the observations in the business cycles, we predict that the lengths of the cycles in herding toward the market (from peak to peak, or from trough to trough) vary. Here, loosely following the definition to the macroeconomic business cycle by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER)[8], we identify a peak of herding as the beginning of a decrease which lasts more than a few months, and a trough as the beginning of an increase which lasts more than a few months.

We build the second hypothesis as follows:

---

[7] Consistent with some combination of feedback effects and other demand factors driving the stock market independently of fundamentals, De Bondt and Thaler (1985) and Jegadeesh and Titman (1993) report that stock prices tend to continue in the same direction over short intervals of six months to a year, but to reverse themselves over longer intervals.

[8] For details of the NBER definition on business cycle, refer to "The NBERs Business-Cycle Dating Procedure", Business Cycle Dating Committee, National Bureau of Economic Research, October 2003.

H2: *Herding, like some macroeconomics aggregate variables, follows a pattern of cycles. Some sudden events can be identified as turning points of the cycles.*

To our knowledge, herding is an inherently coded human behavior, and it could be changed through the learning and experience of people. The increasingly easy procurement of information and the advancement of technology, among other factors, would attract more investors back to fundamental value of firms, thereby decreasing the magnitude of the herding. In other words, we expect to see a downward trend in the magnitude of the herding measure, accompanied by a decreasing volatility.

H3: *There is a downward trend in the magnitude of herding measure, and the volatility of the herding decreases over time.*

1.2 Correlation of Herd Behavior in Different Markets

Hwang and Salmon (2004) find a low correlation between the herding of financial markets of United States and Korea. They conclude that market sentiment may not always be transferred internationally. This observation is interesting given the fact that it is rather counterintuitive. Relevant questions arise naturally. For instance, is this low correlation between the US and Korea the mainstream or just a special case in global markets? And if it is a special case, what is the main pattern we can expect on herding correlation? To answer these questions, a broader study, covering more sample markets, is necessary.

Aiken (2005) states that the more open an economy is, the greater influence global equity markets have on changes in investor sentiment. Since the past two decades has witnessed the global trend of capital market liberalization and increasing cross-boarder investment, we have good reasons to expect that there is comovement in herding between different markets. Due to the differences in market development stage, listed securities, market participants, investor philosophy, etc., the transfer of the herding sentiment may exhibit different patterns across different markets. We conjecture that herding is more correlated in financial markets with similar development stage, or economic characters. For instance, the positive correlation of herding in developed markets such as the USA and the UK is higher than that between the USA and Argentina. As it will be seen in the data section, we divide our sample markets into three groups, covering major developed markets, emerging Asian markets, and emerging Latin American markets, respectively. The division is made, somewhat arbitrarily, by their development stage and geographical location.

H4: *Herding sentiment towards the market travels across international markets. There is a higher correlation of herding between markets within the same group than that between markets from different groups.*

II. The Methodology

2.1 Risk-Return Equilibrium with the Existence of Herding towards the Market

The CAPM (Sharpe, 1964; Lintner, 1965) is widely used in defining the risk-return equilibrium relationship of equities. In the model, the market beta defines the return of equities. Not surprisingly, accumulating evidence shows that beta cannot be the sole variable to explain the equity returns. Fama and French (1993) develop their three factor model (F-F model, hereafter) to incorporate factors associated with size and BE/ME. Fama and French (1996) further show that F-F model can explain most of the departures from the CAPM predictions.

Basically, both the CAPM and the F-F model are within the risk-return framework: risk determines the asset return. However, Hirshleifer (2001) argues that expected return of an asset is not only compensated by its fundamental risk, but also related to the investor mis-valuation caused by cognitive imperfection of investors and social dynamics such as herding. In this section, we explore how the CAPM deviate from their original form with the existence of herding towards the market index. The conclusion also applies to the F-F model.

Our model of measuring herding is based on Hwang and Salmon (2004). Hwang-Salmon assume that the log of cross-sectional standard deviation of betas is normally distributed with a static mean. We argue against this assumption on two grounds. First, imposing assumptions on the cross-sectional standard deviation of betas is not appropriate. Stock beta is actually derived from stock return, on which we usually make assumptions. Second, it is not obvious to us that the mean of the cross-sectional standard deviation of the betas should be static. Given the assumptions in this paper, we explore the distribution of the cross-sectional dispersion of betas by making assumptions on stock returns; by doing this, we get a time-varying series of the cross-sectional variance of beta, which we believe is more realistic than the static one.

2.1.1 The model

In essence, Hwang and Salmon (2004) measure herding on observed deviations from the equilibrium beliefs expressed in the CAPM. In a market with rational investors, the CAPM in equilibrium can be expressed as:

$$\mathrm{E}_{\mathrm{t}}(r_{it}) = \beta_{imt}\,\mathrm{E}_{\mathrm{t}}(r_{mt}) \tag{2}$$

where r_{it} and r_{mt} are the excess returns on asset *i* and the market at time *t* respectively; β_{imt} is the systematic risk measure; and $\mathrm{E}_{\mathrm{t}}(\cdot)$ is the conditional expectation at time *t*.

We follow the assumption of Hwang-Salmon that investors form firstly the common market-wide view, $\mathrm{E}_{\mathrm{t}}(r_{mt})$, and their behavior is then conditional on it. When herding towards the market occurs, the investors shift their beliefs to follow the performance of the overall market more than they should in the CAPM. In other words, they ignored the equilibrium relationship in the CAPM and move towards matching the return on individual assets with that of the market. For instance, it is a common strategy that investors buy "underperforming" assets and sell "overperforming" assets. When herding towards the market occurs, if the market goes up significantly, then an asset with an intrinsic beta of 1.5 will become the target of selling, since its price increases more than the market index and looks more expensive. This selling of the asset leads to the decrease of the asset price. On the other hand, an asset with an intrinsic beta of 0.5 will become the target of buying, since its price increases less than the market index and looks cheaper. The buying of the asset leads to the increase of the asset price. Similar behavior happens when market goes down significantly.

Thus, when there exists herding towards the market portfolio, the conventional CAPM no longer holds, and the expected returns on the asset and the observed beta will be biased, denoted as $\mathrm{E}^{\mathrm{b}}_{\mathrm{t}}(r_{it})$ and β^{b}_{imt}, respectively. The mis-valuation mechanism can be described as follows.

When the market increases significantly, for an equity whose β_{imt} is larger than one, $\mathrm{E}_{\mathrm{t}}(r_{it}) > \mathrm{E}_{\mathrm{t}}(r_{mt})$ as the CAPM says. However, the selling-herding of the investors will push the equity's price downward, making $0 < \mathrm{E}_{\mathrm{t}}^{\mathrm{b}}(r_{it}) < \mathrm{E}_{\mathrm{t}}(r_{it})$ and, therefore, $1 < \beta_{imt}^{b} < \beta_{imt}$. For the same asset, when the market decreases significantly and thus $\mathrm{E}_{\mathrm{t}}(r_{it}) < \mathrm{E}_{\mathrm{t}}(r_{mt})$ as the CAPM predicts, buying-herding of the investors will push the asset's price upwards, making $\mathrm{E}_{\mathrm{t}}(r_{it}) < \mathrm{E}_{\mathrm{t}}^{\mathrm{b}}(r_{it}) < 0$ and, therefore, $1 < \beta_{imt}^{b} < \beta_{imt}$. The inverse process applies to the situation when $\beta_{imt} < 1$, and in this case, the biased beta will become larger when market changes, i.e., $1 > \beta_{imt}^{b} > \beta_{imt}$. For an equity whose $\beta_{imt} = 1$, it is neutral to herding.

The above process can be expressed with the following mean reverting process:

$$\beta_{imt}^{b} = \mathrm{E}_{\mathrm{t}}^{\mathrm{b}}(r_{it}) \Big/ \mathrm{E}_{\mathrm{t}}(r_{mt}) = \beta_{imt} - h_{mt}(\beta_{imt} - 1) \qquad (3)$$

where h_{mt} is a latent herding parameter that changes over time. When $h_{mt} = 0$, there is no herding and the equilibrium CAPM holds; when $h_{mt} = 1$, there is perfect herding towards the market portfolio and all the individual assets move in the direction and with same magnitude as the market portfolio. In general, $0 < h_{mt} < 1$, and some degree of herding exists. When $h_{mt} < 0$, there is reversed herding.[9]

The form of herding under discussion represents market-wide behavior. So it is preferable to use all assets in the market than a single asset to eliminate the effects of idiosyncratic movements in any individual β_{imt}^{b}. Then,

$$\begin{aligned} \mathrm{var}_c(\beta_{imt}^{b}) &= \mathrm{var}_c(\beta_{imt} - h_{mt}(\beta_{imt} - 1)) \\ &= (1 - h_{mt})^2 \, \mathrm{var}_c(\beta_{imt}) \qquad (4) \end{aligned}$$

where $\mathrm{var}_c(\cdot)$ is the cross-sectional variance.

[9] The opposite form of the behavior, called "adverse herding", could also happen when individual returns become more sensitive for large beta securities but less sensitive for low beta securities. In this case, high betas become higher and low betas become lower.

In plain words, the existence of herding makes the cross-sectional dispersion of the individual betas smaller than it would be in equilibrium.

Taking logarithms of *Eq.4* on both sides,

$$\log[\mathrm{var}_c(\beta^b_{imt})] = \log[\mathrm{var}_c(\beta_{imt})] + 2\log(1 - h_{mt}) \tag{5}$$

Define $H_{mt} = 2\log(1 - h_{mt})$ and $\mu_t = \mathrm{E}(\log[\mathrm{var}_c(\beta_{imt})])$, and write

$$\log[\mathrm{var}_c(\beta_{imt})] = \mu_t + v_{mt}$$

where we assume $v_{mt} \sim \mathrm{iid}(0, \boldsymbol{\sigma}^2_{\mathbf{mv}})$.

In *Appendix I*, we prove that $\log[\mathrm{var}_c(\beta_{imt})]$ can be better approximated by a normal distribution with time-varying mean,

$$\mu_t = \mathrm{E}(\log[\mathrm{var}_c(\beta_{imt})]) = \log\left(\frac{\sum_{i=1}^{\mathrm{N_t}} [\sigma^2_{i,t}(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))]}{\mathrm{N_t}}\right) \tag{6}$$

where $\mathbf{F_t}$ is the matrix of linear factors, $\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1})$ is the diagonal term of the matrix $(\mathbf{F_t}'\mathbf{F_t})^{-1}$ for the market beta, $\sigma^2_{i,t}$ is the disturbance variance, and $\mathrm{N_t}$ is the number of the stocks at time *t*. $\sigma^2_{i,t}$ can be estimated with $\sum_{j=1}^{\mathrm{J}} e^2_{i,j} / (\mathrm{J} - \mathrm{K} - 1)$, and $e_{i,j}$, *J*, and *K* represent the residual term of the regression, number of observations, and number of factors, respectively.

We estimate $\log[\mathrm{var}_c(\beta^b_{imt})]$ with $\log[\mathrm{var}_c(\hat{\beta}^b_{imt})]$, then *Eq.5* can be rewritten as

$$\log[\mathrm{var}_c(\hat{\beta}^b_{imt})] = \mu_t + H_{mt} + v_{mt} \tag{7}$$

Allow H_{mt} to evolve over time and we assume it follows an AR(1) process:

$$H_{mt} = \phi H_{mt-1} + \eta_{mt} \tag{8}$$

where $\eta_{mt} \sim iid(0, \boldsymbol{\sigma}^2_{\mathbf{m\eta}})$.

The model is a standard state space model with *Eq.7* as the measurement equation and *Eq.8* as the transition equation. It can be estimated with the Kalman filter, which is briefly introduced in Appendix II. [10]

2.2 The Calculation of Beta: A Robust Technique

2.2.1 Robust Estimate of the Beta

Ordinary Least Squares (OLS) estimation is the most commonly used technique in estimating beta. However, the OLS estimate has an obvious drawback, i.e., it can behave badly when the errors are not from a normal i.i.d. distribution, particularly when they are heavy-tailed, as revealed by the return data in real financial world. It turns out that a few outliers can have a very strong influence on the OLS beta, thus leading to a distorted perspective on the relationship between equity returns and index returns. For instance, the fact that a small number of exceptionally large outlier returns giving rise to a beta of 3.0 does not justify an expectation of future return movement twice bigger than those of the market. Under situations like this, robust estimation of beta can provide a very good fit to the bulk of the equity returns versus index data (Martin and Simin, 1997). [11]

The most applied method of robust regression is M-estimate, a generalization of maximum-likelihood estimation. Consider the linear model:

$$y_i = X_i\beta + \varepsilon_i \tag{9}$$

where i =1,…n. The fitted model is:

$$y_i = X_i b + e_i \tag{10}$$

The M-estimate principle is to minimize the objective function:

[10] For further introduction of Kalman filter, please refer to "Time series analysis by state space methods" , 2001, by J. Durbin and S. Koopman, Oxford University Press.

[11] Despite their superior performance over least squares estimation in many situations, robust methods for regression are still very seldom used. One possible reason is that computation of robust estimates is much more demanding than least squares estimation, although this is no more a problem given today's standard. Another reason for their lack of popularity may be that some popular statistical software packages failed to implement the methods (Stromberg, 2004). The belief of many statisticians that classical methods are robust (Hampel *et al.*,1986) also leads to the slow uptake of robust methods.

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho(y_i - X_i b) \tag{11}$$

where the function $\rho(\cdot)$ gives the contribution of each residual to the objective function.

Define $\psi = \rho'$ as the first-order derivative of ρ. By differentiating the objective function with respect to b and setting the partial derivatives to *0*, we obtain a system of estimating equations:

$$\sum_{i=1}^{n} \psi(y_i - X_i b) X_i' = 0 \tag{12}$$

Define the weight function $w(e) = \psi(e)/e$ and let $w_i = w(e_i)$, then the estimating equations become:

$$\sum_{i=1}^{n} w_i e_i X_i' = 0 \tag{13}$$

Solving the estimating equations is a weighted least-squares problem, with the objective of minimizing $\sum_{i=1}^{n} w_i^2 e_i^2$. The weights depend on the residuals, the residuals depend on the estimated coefficients, and the estimated coefficients depend on the weights. An iteration solution is required to solve the problem.

In this paper, we apply Huber estimation, one of the most applied techniques in robust regression practice. See *Appendix III* for the comparison between the OLS estimation and the Huber estimation.

2.2.2 Monte Carlo Experiments

As an example, we perform Monte Carlo simulations for the following bivariate model to test whether or not the least square and Huber beta estimates are significantly different from each other:

$$y_i = \alpha + \beta x_i + \varepsilon_i \qquad i=1,\ldots,N \tag{14}$$

where the true α and β are set to zero.

We try 10,000 Monte Carlo replicates of 100 observations of (x_i, ε_i) for each of the following two situations. In the first situation, we use for every replicate a fixed set of 100 x_i's, where $x_i \sim N(0,1)$, and 10,000 sets of 100 ε_i's, where $\varepsilon_i \sim N(0,1)$. In the second situation, we use the same 10,000 replicate samples of (x_i, ε_i), but we replace one (among 100) set of (x_i, ε_i) with an independent pair of (x_i, ε_i), where $x_i \sim N(5,2), \varepsilon_i \sim N(10,0.5)$. The results of the simulation are presented in *Table 1* and *Figure 1*.

Conditional on the value of the fixed independent variables x_i, $\hat{\beta} \sim N(0, \sigma^2 / \sum x_i^2) = N(0,0.0102)$ when $\varepsilon_i \sim N(0,1)$. This normal density is overlaid as a reference in all the panels in *Figure 1*.

When both the independent variable and the error terms are normally distributed, as shown in the top panels in *Figure 1*, the histogram of the OLS estimate is very close to the theoretical normal distribution. The robust estimates also behave quite well, being reasonably normal in shape and well centered on zero, with a standard deviation slightly higher the OLS standard deviation. This increased standard deviation represents the lowered efficiency of robust regression when the errors are normal.

When the independent variable and the error terms are normally distributed but with outliers (bottom panels in *Figure 1*), the distribution of the OLS estimate is radically shifted in location and in shape from the former situation when no outliers exist: The mean increases from 0.0015 to 0.3785, the standard error increases from 0.102 to 0.142, and the skewness decreases from -0.023 to -0.692. On the contrary, the distribution of the robust estimates is very close to that obtained when there are no outliers.

From this simple simulation, we can tell that robust regressions achieve almost the efficiency of OLS with ideal data while substantially better-than-OLS efficiency in non-ideal situations.[12]

[12] The existence of outlier in our example is very obvious even to eyes. For comparison purpose, we also run simulations when the outlier pair (x_i, ε_i) is not obvious, for instance, $x_i \sim N(0,1.5)$ and $\varepsilon_i \sim N(1,0.5)$. In this case, we do not find a significant difference among the three techniques. Even so, we still give credentials to the robust techniques since, for large samples such as this paper, it is not possible to identify outlier with eyes, although their existence might be very obvious.

III. The Data

3.1 Securities, Market Index and Proxy for Risk-Free Interest Rate

Monthly total returns (in local currencies) of equities listed in the 21 stock markets are obtained from Datastream, covering the period from January of 1985 (or the earliest available date) to December of 2005. The 1997 Asian crisis, the 1998 Russian crisis, and the bull market until early 2001 and subsequent bear market are covered in the sample period.

We divide the 21 markets into three groups: the developed markets, the emerging Latin American markets, and the emerging Asian markets. Included in the developed markets are France, Germany, Hong Kong, Japan, United Kingdom, and the United States; included in the Latin American group are Argentina, Brazil, Chile, Colombia, Peru, Mexico, and Venezuela; and included in the Asian group are China, India, Indonesia, Korea, Malaysia, Philippines, and Thailand. The selected emerging markets are those marked as major emerging markets in EMDB of S&P/COMPUSTAT.

In each market, the asset beta is calculated against commonly quoted market index in the market. We have examined the correlation between these market indices and Morgan Stanley Country Index (MSCI), which covers above 85 per cent of total market capitalization of each country. It turns out that the correlations are larger than 0.90 for all the markets except Brazil (0.761) and Peru (0.824). We can therefore conjecture that the results with these common market indices are not significantly different from those obtained with the MSCI.[13]

To calculate the excess market return, we approximate the risk-free rate of return with the short-term treasury bill rate and alternative short-term interest rate (if treasury bill rate is not available), which are obtained from Datastream and Global Financial Market Database. *Table 2* lists the descriptive statistics of the above mentioned variables. All the markets (except China) have an average positive monthly return on market index, ranging between

[13] According to Shanken (1987), if the correlation between the proxy and the true market exceeds about 0.7, then the rejection of the CAPM with a proxy would also lead to the rejection of the CAPM with the true market portfolio.

8.14% (Brazil) and 0.32% (Japan). The general rule of "high return vs. high return" pattern is observed in the market index change. Similar situation happens to the interest rate.

3.2 Fama-French Three Factors

Since Fama-French factor values are only readily available for the United States at the website of Kenneth French, we calculate these factor values by our own, and form the 6 size-BE/ME stock portfolios based on all the equities from Datastream. In forming the factors, we follow the method described in Fama and French (1993), with a minor modification: Since there are different financial reporting periods in these markets, we set the date of forming the portfolio in January of each year, instead of July in Fama and French. We conjecture this modification will not change the results significantly.

The descriptive statistics of the *SMB* and *HML* for the sample markets are reported in *Table 3*. The third column in the table indicates the number of stocks used in forming the 6 size-BE/ME stock portfolios, which covers the majority of the stocks in each market. According to the table, the *SMB* factor earns significantly positive returns for all developed markets (except the United States) and less than half of the total emerging markets. *HML* factor earns significantly positive returns for France, Germany and seven emerging markets. The excess market returns are not significantly different from zero for most of the markets. On exceptional cases, we observe significantly negative returns on the market factor (for Colombia) and on the *HML* factor (for Hong Kong). In other words, the strategy of investing on these factor portfolios over monthly horizons earns significant positive returns for some of the markets, insignificant positive returns for others, but negative returns for the very few others.

Figure 2 depicts the return vs. risk relationship for the three factors in all markets. A regression line plot is placed on top of the scatterplot. Basically, the return and the risk (represented by standard deviation) follow a positive linear pattern for each of the factors, with emerging markets on the *high risk-high return* end and developed markets on the *low risk-low return* end. It is interesting to point out that, for factors of excess market return and *SMB*, Brazil has a much higher average return than other countries, followed by its much higher standard deviation. Although we will not dwell on this, it is worth studying the potentials of expanding the efficient frontier in portfolio management by taking advantage of global financial markets.

IV. Empirical Results and Analysis on the Herd Behavior

This section first reports the results on the beta estimation under CAPM and F-F model, respectively. Then we use the estimated betas to calculate the herding in each market. With the herding measure, we study the hypotheses developed in *Section I*, along with some discussions on other aspects of the herding.

4.1 Estimated Cross-sectional Mean and Cross-sectional Variance of Betas

Several filters are used in estimating betas of equities in these markets: First, first two months' data for each security are deleted to eliminate the effect of IPO (Initial Public Offerings) underpricing. Second, securities with a history of less than 1.5 years are deleted. Third, observations with number of equities less than 10 percent of the number of equities at the end-period are eliminated.

A set of monthly estimates of beta for all the equities in the sample are thus obtained for each market. Statistics of the betas of the CAPM and the F-F model, estimated with the robust Huber technique, are reported in *Table 4*. The large number of calculated betas (ranging from 2,118 in Venezuela to 953,058 in the United States) makes it possible for us to examine the distribution of the beta coefficients of the CAPM and the F-F model. We find that they are significantly different from zero in all cases, justifying the roles played by the markets, size and BM/MV in explaining the cross-sectional equity returns.

Figure 3 depicts the evolution of the estimated cross-sectional mean of betas, $E_c(\hat{\beta}_{imt}^b)$, and the estimated cross-sectional variance of betas, $Var_c(\hat{\beta}_{imt}^b)$, which are calculated as

$$E_c(\hat{\beta}_{imt}^b) = \sum_{i=1}^{N_t} \hat{\beta}_{imt}^b \Big/ N_t \tag{15}$$

$$Var_c(\hat{\beta}_{imt}^b) = (\hat{\beta}_{imt}^b - E_c(\hat{\beta}_{imt}^b))^2 \big/ (N_t - 1) \tag{16}$$

where N_t is the number of the stock at time *t*.

Statistics about the two series are reported in *Table 5*. Under the CAPM, $E_c(\hat{\beta}^b_{imt})$ for these market ranges between 0.243 (China) and 1.365 (Colombia), with a mean value of 0.776; $Var_c(\hat{\beta}^b_{imt})$ ranges between 0.058 (China) and 1.045 (Colombia), with a mean value of 0.301. Under the F-F model, $E_c(\hat{\beta}^b_{imt})$ ranges between 0.354 (China) and 1.558 (Colombia), with a mean value of 0.838; $Var_c(\hat{\beta}^b_{imt})$ ranges between 0.044 (China) and 0.863 (Colombia), with a mean value of 0.320.

The correlation between the cross-sectional variances of betas ($Var_c(\hat{\beta}^b_{imt})$) obtained by CAPM and by F-F model ranges between 0.485 (USA) and 0.983 (Philippines), with a mean value of 0.887. As the difference between the results with the CAPM and the F-F model does not seem to be large enough to change our interpretation of the herding measure, in the remaining of the paper, we only calculate the herding measure obtained with the F-F model.

4.2 The Properties of the Estimated Herding Measures

We use the Kalman filter to estimate the herding indicator ($h_{mt} = 1 - e^{H_{mt}/2}$) with *Eq.7* and *Eq.8*. The main results are reported in Panel *A* of Table 6. The average herding value ranges from 0.004 (USA) to 0.055 (Colombia), with an average of 0.031.

As we have mentioned, the Kalman filter algorithm provides two series, a filtered one and a smoothed one. Here we only report the filtered series, since the smoothed one resembles the filtered one in all markets and does not alter our conclusions (the average correlation between the two series is as high as 0.97).

Figure 4 depicts the evolution of the herding measure in each market. A visual observation tells us there might be high correlation between two markets from the same group. For most of the markets, we see the trend from peak to trough over the period of early 1997 to early 1999. We will check these points in next section.

4.3 An Examination on the Herding Behavior

4.3.1. Cross-Sectional Comparison

In order to check if there is difference in the magnitude of herding towards the market between emerging and developed markets, we run the two sample *t-test* on the mean of herding measures, as shown in Panel *B* of *Table 6*.[14] With a *t-value* of 11.41, we reject the null hypothesis that there is no difference in the mean of herding measure. In order words, the evidence support the hypothesis *H1*, and the emerging markets have a higher level of herding towards the market.

4.3.2. Turning Points of the Herding Measures

The description of a cycle always starts with the identification of turning points in the series. There are various methods to accomplish this. In this paper, we follow the classical Bry and Boschan (1971) procedure.

In essence, the Bry-Boschan procedure is to isolate "true" turning points from some "false" turning points which are either short lived or of insufficient amplitude. It starts with a highly smoothed series to find initial estimates of local peak (trough) at time t, which is defined as the local maximum (minimum) over an interval from $t-k$ to $t+k$, where k is generally set to five. These peaks and troughs must alternate. With these initial estimates, a less smoothed curve is investigated to refine the dates of the turning points. This process is then repeated with a short-term (3 to 5 months) moving average. Final turning points are determined using the unsmoothed series, with a set of predefined restrictions, for instance, the cycle must be no less than 15 months in length and all phases must be over 5 months in duration. Interested readers are referred to King and Plosser (1989) for detailed description of the procedure.

We apply the Bry-Boschan procedure to the herding measures, with a minor modification of removing the minimum length requirement on the cycle and the phase. Panel *A* of *Table 7* lists the months of peak and trough for each market. These turning points scatter without easily identifiable rules. Even though, we can tell that the majority of the markets have turning points between early 1997 and late 1998 when the 1997-1998 financial crisis broke out starting from Southeast Asia.

---

[14] Here we ignore the existence of autocorrelation in each series of herding measures, since we believe this will not affect our test result significantly.

According to the table, developed markets have more turning points than emerging markets. We think it is due to the fact that the former has a longer sample history. We report the average length of cycle, calculated as twice of the arithmetic mean of monthly intervals from peak to trough, or trough to peak. Results show that the length ranges from 19 months (China) to 69 months (Thailand), with an average of 42 months.

Panel *B* of *Table 7* tests the null hypothesis that there are no differences among the average length of the cycles of the three groups. With an *F-value* of 2.21, we do not reject the null hypothesis at conventional levels.

In a nutshell, the evidence supports the hypothesis *H2* that herding follows a pattern of cycles, and sudden events such as the financial turmoil can be identified as turning points of the cycles.

4.3.3 The Volatility of the Herding Measures

Figure 5 depicts the volatility of the herding measures (σ_h^2) in the developed group, Asian group, and Latin American group. Most of the time, σ_h^2 fluctuates within a small range between 0 and 0.01. There are several huge leptokurtic with peak value over 0.02 in the following markets: Australia (early 1990), Hong Kong (late 1999), Japan (early 1990), United States (early 1990), Korea (early 1999), Thailand (early 2000), Malaysia (early 2001), Chile (early 1996), Colombia (mid-1998), and Venezuela (mid-1998). No general trend can be observed within each group from the figure. Thus, we do not find evidence to support hypothesis *H3* that the volatility of the herding decreases over time.

Panel *A* of *Table 8* reports the first two moments of the volatility of the herding measures. In panel *B* of the table, we test the null hypothesis that the means of the volatility of the herding of the three groups are equal ($\mu_{\sigma_h^2}^{developed} = \mu_{\sigma_h^2}^{Asian} = \mu_{\sigma_h^2}^{LatinAmerican}$). With an F-value of 2.69, we do not reject the null hypothesis at 5% level of significance, although it is weakly rejected at 10% level.

4.3.4 The Comovement of the Herding Measures

The correlation coefficients of herding measures between various markets are shown in *Table 9*. Among the 210 pair wise correlation coefficients, 94 (46%) are significantly positive, 61 (29%) are significantly negative, while the remaining 55 (26%) are insignificant. A closer look at the table reveals that most of the correlations between two countries from the same groups (e.g., Australia and France, Indonesia and India) are significantly positive, while the signs of the correlations between two countries from different groups (for instance, between China and Australia) are mixed. It is worth noting that, in the developed group, Japan, as the only case, has significant negative correlations with all the other markets. Similar situation happens to Brazil in the Latin American group.

Panel *A* of *Table 10* tests the equality between herding correlation within the groups (i.e., between countries in the same group) and those across the groups (i.e., between countries from different groups). The null hypothesis is that the mean of the correlation coefficients within the group is equal to those across the groups. With a *t-value* of 3.60, we reject the null hypothesis.

Panel *B* of *Table 10* tests the equality of the means of the herding correlation of the groups, with a null hypothesis of equal mean. With an *F*-value of 1.33, we do not reject the null hypothesis at any conventional level of significance.

To check whether the above correlation is spurious, we calculate the correlation of the first difference of the herding measures, $\Delta h_{mt} (= h_{mt} - h_{mt-1})$. The result is reported in *Table 11*. Now, among the 210 pair wise correlation coefficients, 43 are significantly positive, 16 are significantly negative, while the remaining 151 are insignificant. The highlighted values in the table (all of them are insignificant) indicate that their counterpart in *Table 9* is significant at 5% level. From the table, a majority of the correlations in the developed group keep its original level of significance at *Table 9*, so it is highly unlikely that the correlations of the original series within this group are spurious. On the other hand, since most of the correlations between markets from different group are no longer significant, we cast doubt on the significance of correlations between two markets from different groups.

In sum, we find support for the hypothesis *H4,* i.e., the correlation of herding between markets from the same group is higher that that between markets from different groups.

VI. Conclusions

In this paper, we use the cross-sectional variance of the betas to study herd behavior towards market index in major developed and emerging financial markets. We propose the robust regression technique to calculate the betas of the CAPM and those of the Fama-French three factor model, with an intention to diminish the impact of multivariate outliers in return data. Through the estimated values obtained from a state space model, we examine the evolution and cross-sectional relationship of the herding measures, especially their pattern around sudden events such as the 1997-1998 financial crises.

As a result, we find a higher level of herding in emerging markets than in developed markets. Additionally, the herding measure, like some macroeconomics aggregate variables, follows a pattern of cycles. And some sudden events, such as the 1997-1998 financial turmoil, can be identified as turning points of the cycles. Furthermore, we do not observe any trend in the magnitude and the volatility of the herding measure over time. Finally, we witness a higher correlation of herding between two markets from the same group than those from different groups.

One direct question related to this paper is: What are the possible factors influencing the herd behavior towards the market? To answer this question, future research is suggested on the robustness check of this paper's conclusion, in the presence of variables reflecting either the state of the market, for instance, the market volatility and macroeconomic fundamentals, or the history and cultural ingredients.

Given the fragility of emerging financial markets, it is imperative to study the effect of policy change in capital markets on herd behavior in these markets. For instance, what is the impact of financial market liberalization on the herd behavior of investors? Had the herd behavior been weakened or strengthened upon the liberalization?

Other interesting questions include: Why in the developed group, Japan has negative herding correlations with all the other developed markets? Do professional investors like mutual funds show different herding pattern from individual investors?

Appendix I. The Distribution of Cross-Sectional Dispersion of Betas

According to the APT, the excess return of asset *i* follows the linear factor model:

$$\mathbf{R_{i,t}} = \sum_{k=1}^{K} \beta_{ik,t}\mathbf{F_{k,t}} + \varepsilon_{i,t}\,, \qquad i=1,...,\mathrm{N_t} \text{ and } t=1,\ldots,\mathrm{T}. \qquad \text{(A-1)}$$

where factor $\mathbf{F_k}$ is assumed to be uncorrelated (k=1,…,K), and ε_i is uncorrelated across assets.

One factor used actually in all models is $\mathbf{R_{m,t}}$, the excess market portfolio return, and *Eq.A-1* can be written as:

$$\mathbf{R_{i,t}} = \beta_{im,t}\mathbf{R_{m,t}} + \sum_{k=1}^{K-1} \beta_{ik,t}\mathbf{F_{k,t}} + \varepsilon_{i,t}\,. \qquad \text{(A-2)}$$

Taking cross-sectional expectation on both sides:

$$\mathrm{E_C}(\mathbf{R_{i,t}}) = \mathrm{E_C}(\beta_{im,t})\mathbf{R_{m,t}} + \sum_{k=1}^{K-1} \mathrm{E_C}(\beta_{ik,t})\mathbf{F_{k,t}} + \mathrm{E}_C(\varepsilon_{i,t})\,. \qquad \text{(A-3)}$$

Since $\mathrm{E_C}(\mathbf{R_{i,t}}) = \mathbf{R_{m,t}}$, and $\mathbf{R_{m,t}}$ and $\mathbf{F_{k,t}}$ are uncorrelated,

$$\mathrm{E_C}(\beta_{im,t}) = 1\,. \qquad \text{(A-4)}$$

This is consistent with the intuition that beta of the whole market is always one.

In estimating the above model, we assume that market betas of each stock are constant over a fixed interval, e.g. 60 months, but are variable under a longer time period. In addition, we assume that the first moment of time-varying beta of each stock is one.

Imagine we have *J* observations (over the above mentioned fixed interval) to estimate $\beta_{im,t}$ under OLS framework. Then the OLS estimator $b_{im,t} \sim \mathrm{N}(\beta_{im,t}, \sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1})))$, and

$$(\mathrm{b_{im,t}} - 1) \sim \mathrm{N}((\beta_{im,t} - 1), \sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))) \qquad \text{(A-5)}$$

where $\text{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1})$ is the diagonal term of the $(\mathbf{F_t}'\mathbf{F_t})^{-1}$ for the market beta, $\sigma_{i,t}^2$ is the disturbance variance. $\sigma_{i,t}^2$ can be estimated with $(\sum_{j=1}^{\mathrm{J}} e_{i,j}^2/(\mathrm{J}-\mathrm{K}-1))$, where $e_{i,j}$ is the residual term of the regression. [15]

Assume the expectation of the beta of each stock is one, then

$$(\mathrm{b}_{\mathrm{im,t}}-1) \sim \mathrm{N}(0, \sigma_{i,t}^2(\text{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))). \qquad \text{(A-6)}$$

At time t, let $X = \begin{bmatrix} (\mathrm{b}_{\mathrm{1m,t}}-1) \\ \ldots \\ (\mathrm{b}_{\mathrm{Nm,t}}-1) \end{bmatrix} \sim \mathrm{N}(0, \mathbf{\Phi})$, then $\mathbf{\Phi}$, the covariance matrix, is positive-definite, with diagonal terms $\sigma_{i,t}^2(\text{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))$ (i=1,…,$\mathrm{N_t}$).

Define Y as a quadratic polynomial of X:

$$Y = X'(\mathbf{I}/\mathrm{N}_t)X, \qquad \text{(A-7)}$$

where $\mathbf{I}$ is identity matrix.

Let $\mathbf{z}$ be the Cholesky matrix of $\mathbf{\Phi}$, and define $\mathbf{u}$ as a square matrix whose rows comprise orthonormal eigenvectors of ($\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}$). By construction, $\mathbf{u}$ is orthogonal.

Define the change of variables:

$$Z = \mathbf{u}\mathbf{z}^{-1}X. \qquad \text{(A-8)}$$

Then Z is multivariate normal with mean vector

$$\mathrm{E}(Z) = \mathrm{E}(\mathbf{u}\mathbf{z}^{-1}X) = \mathbf{u}\mathbf{z}^{-1}\,\mathrm{E}(X) = 0 \qquad \text{(A-9)}$$

And covariance matrix

$$\mathrm{Cov}(Z) = (\mathbf{u}\mathbf{z}^{-1})\mathbf{\Phi}(\mathbf{u}\mathbf{z}^{-1})' = (\mathbf{u}\mathbf{z}^{-1})\mathbf{\Phi}(\mathbf{z}^{-1})'\mathbf{u}' = \mathbf{u}\mathbf{I}\mathbf{u}' = \mathbf{I} \qquad \text{(A-10)}$$

So $Z \sim \mathrm{N}_{N_t}(0, \mathbf{I})$.

Then $Y = X'(\mathbf{I}/\mathrm{N}_t)X = (\mathbf{z}\mathbf{u}^{-1}Z)'(\mathbf{I}/\mathrm{N_t})(\mathbf{z}\mathbf{u}^{-1}Z) = Z'(\mathbf{u}\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}\mathbf{u}')Z = (Z'\mathbf{c}Z)/\mathrm{N_t}$ (A-11)

where $\mathbf{c} = \mathbf{z}'\mathbf{z}$.

[15] For the robust M-estimation, the estimated coefficients are asymptotically normal, i.e., $b_{im,t} \overset{a}{-} \mathrm{N}(\beta_{im,t}, \sigma_{i,t}^2(\text{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1})))$, where $\sigma_{i,t}$ can be estimated with the larger of robust estimate of sigma and a weighted average of OLS estimate of sigma and robust estimate of sigma. Here we only show the case of OLS estimation, and the conclusion applies to the robust M-estimation asymptotically.

Since we define $\mathbf{u}$ as a matrix whose rows comprise orthonormal eigenvectors of $\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}$, by the Spectral Theorem of linear algebra, the matrix $\mathbf{uz}'(\mathbf{I}/\mathrm{N_t})\mathbf{zu}'$ is diagonal with diagonal elements equal to the eigenvalues of $\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}$. Then, Y depends on only diagonal terms of the form $c_{i,i}Z^2$, and can be written as:

$$Y = (\sum_{i=1}^{\mathrm{N_t}} c_{i,i}Z^2)/\mathrm{N_t} \tag{A-12}$$

When $\mathrm{N_t}$ is large, we have simulated the distribution of variable Y with different combinations of eigenvalues of $\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}$, finding that only in one case, when one of the eigenvalues is extremely larger than the others, Y deviates from normal distribution with large skewness and kurtosis; otherwise, Y is very close to normal distribution.

We then go through the real data in various markets for the eigenvalues of matrix $\mathbf{z}'(\mathbf{I}/\mathrm{N_t})\mathbf{z}$, and observe that the exceptional case of one large eigenvalue rarely happens in reality. So *we suggest approximating Y with normal distribution*.

Since $\mathbf{zz}' = \Phi$, the sum of eigenvalues of $\mathbf{z}'\mathbf{z}$ equal to the sum of diagonal terms of Φ, i.e., $\sum_{i=1}^{\mathrm{N_t}}[\sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))]$.

So the expectation of Y is $$\mathrm{E}(Y) = \mathrm{E}(\frac{\sum_{i=1}^{\mathrm{N_t}} c_{i,i}Z^2}{\mathrm{N_t}}) = \frac{\sum_{i=1}^{\mathrm{N_t}}[\sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))]}{\mathrm{N_t}} \tag{A-13}$$

Define the variance of Y as Var(Y),

$$\mathrm{Var}(Y) = \mathrm{Var}(\frac{\sum_{i=1}^{\mathrm{N_t}} c_{i,i}Z^2}{\mathrm{N_t}}) < \frac{2*\{\sum_{i=1}^{\mathrm{N_t}}[\sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))]\}^2}{\mathrm{N_t^2}} = 2[\mathrm{E}(Y)]^2 \tag{A-14}$$

We check the relative magnitude of $\mathrm{E}(Y)$ and $\mathrm{Var}(Y)$, and observe that, in all the markets, $\mathrm{Var}(Y) << \mathrm{E}(Y)$. Thus, we suggest approximating the distribution of log*(Y)* by a normal distribution with mean *m* and variance s^2, where

$$\mathrm{E}(\log(Y)) \cong \log(\mathrm{E}(Y) = \log(\frac{\sum_{i=1}^{\mathrm{N_t}}[\sigma_{i,t}^2(\mathrm{Diag}^m((\mathbf{F_t}'\mathbf{F_t})^{-1}))]}{\mathrm{N_t}}) \tag{A-15}$$

QED

Appendix II. The Kalman Filter

In general, a state space model can be defined with two equations:

$$\mathbf{Y_t} = \mathbf{c} + S\mathbf{X_t} + e_t \tag{A-16}$$

$$\mathbf{X_t} = \mathbf{d} + H\mathbf{X_{t-1}} + z_t \tag{A-17}$$

where $\mathbf{X_t}$ is the hidden vector at time *t*, $\mathbf{Y_t}$ is the observation vector at time *t*, *c* and *d* are vectors with constants, e is the measurement error, and *z* is the state error. e and *z* are both multivariate normally distributed, with mean zero and covariance matrices of **R** and **Q**, respectively.

The Kalman filter is an algorithm to perform filtering on this state-space model. The goal is to minimize the difference between the observation Y_t and the prediction based on the previous observations $Y_{t|t} = P[Y_t | Y_1,...,Y_{t-1}]$. This can be accomplished by recursive maximum likelihood estimation. The estimation of the state equation by the Kalman filter algorithm also offers a smoothed time series, by performing fixed-interval smoothing, i.e. computing $Y_{t|t} = P[Y_t | Y_1,...Y_{T-1}]$, for $t \le T$.

The Kalman filter can be regarded as an online estimation procedure, which is used to estimate the parameters online when new observations are coming in only after they have been estimated. In contrast, the Kalman smoother can be thought of as an offline procedure, which is only used when the total series have been observed. The Kalman filter results in approximations of the maximum likelihood estimates, while the smoother results in exact maximum likelihood estimates.

Appendix III. Comparison of Ordinary Least Square and Robust Regression

Table A1 shows the objective functions and weight functions for the ordinary least squares estimator and the Huber estimator. Both of them increase without bound as the residual departs from 0, but the Huber objective function increases more slowly. Least squares assigns equal weight to each observation; the weights of the Huber estimator decline for $|e| > k$, where e is the residual term, and k is called a tuning constant for the Huber estimator.

A smaller k provides more resistance to outliers, but at the expense of lower efficiency when the errors are normally distributed. In general, $k = 1.345\sigma$ for the Huber (where σ is the conventional standard deviation), producing 95% efficiency when the errors are normal, and still offering protection against outliers.

Table A1. Objective function and weight function of least squares and Huber estimations.

Method	Objective Function (ρ)	Weight Function (w_i)
Ordinary Least Squares	e^2	1
Huber Robust	$\begin{cases} e^2/2 & (\text{when } \lvert e\rvert \le k) \\ k\lvert e\rvert - k^2/2 & (\text{when } \lvert e\rvert > k) \end{cases}$	$\begin{cases} 1 & (\text{when } \lvert e\rvert \le k) \\ k/\lvert e\rvert & (\text{when } \lvert e\rvert > k) \end{cases}$

REFERENCES

Aiken, A, 2005, The existence and causes of investor sentiment in emerging markets, *Working Paper*.

Bikhchandani, S. and S. Sharma, 2001, Herd behavior in financial markets, *IMF Staff Papers* 47, 279–310.

Blume, 1971, On the assessment of risk, *Journal of Finance* 26, 1-10.

Braun P., D. Nelson, and A. Sunier, 1995, Good news, bad news, volatility and betas, *Journal of Finance* 50, 1575-1603.

Bry, G. and C. Boschan, 1971, Cyclical analysis of time series: Selected procedures and computer programs, *NBER*.

Chang, E. C., J. W. Cheng, and A. Khorana, 2000, An examination of herd behavior in equity markets: An international perspective", *Journal of Banking and Finance* 24, 1651–1679.

Choe, H., B. C. Kho, and R. Stulz, 1999, Do foreign investors destabilize financial markets? The Korean experience in 1997, Journal of Financial Economics 54, 227-264.

Christie W. and R. Huang 1995, Following the pied piper: Do individual returns herd around the market? *Financial Analysts Journal*, July-August, 31-37.

DeBondt, W and R. Thaler, 1985, Does the stock market overreact, *Journal of Finance* 40, 793-805.

Devenow, A. and I. Welch, 1996, Rational herding in financial economics, *European Economic Review* 40, 603–615.

Fabozzi F. and J. Francis, 1977, Stability tests for alphas and betas over bull and bear market conditions, *Journal of Finance* 32, 1093-1099.

Faff R. and R. Brooks, 1998, Time varying beta risk for Australian industry portfolios: An exploratory analysis, *Journal of Business Finance and Accounting* 25, 721-745.

Fama E. and K. French, 1993, Common risk factors in the returns of stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, E, and K. French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama E. and J. Macbeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607-636.

Fernandez, P., 2004, Are calculated beta worth for anything? *IESE Business School Working Paper*.

Ferson W. and C. Harvey, 1999, Conditioning variables and the cross section of stock returns, *Journal of Finance 54*, 1325-1360.

Grinblatt, M., S. Titman, and R. Wermers, 1995, Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior, *American Economic Review* 85, 1088-1105.

Groenewold N. and P. Fraser, 2000, Forecasting beta: How well does the 'five-year rule of thumb' do? *Journal of Business Finance and Accounting* 27, 953-982.

Hampel, F.R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, Robust statistics: The approach based on influence functions, *Wiley*, 1986.

Hirshliefer, D., 2001, Investor psychology and asset pricing, *Journal of Finance* 56, 1533-1597.

Huang, S. and M. Salmon, 2004, Market stress and herding, *Journal of Empirical Finance* 11, 585-616.

Jegadeesh, N. and S. Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

Jirasakuldech, B., D. DUDNEY, and T. Zorn, 2004, The P/E multiple, market efficiency and bubbles, *University of Nebraska-Lincoln Working Paper*.

King R., C. Plosser, 1989, Real business cycles and the test of the Adelmans, *NBER Working Paper* 3160.

Komulainen, T., 2001, Currency crisis in emerging markets: Capital flows and herding behavior, *Bank of Finland, Institute for Economies in Transition*, Discussion Paper 10.

Lakonishok, J, A. Shleifer, and R.W. Vishny, 1992, The impact of institutional trading on stock prices, *Journal of Financial Economics* 32, 23-43.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in security portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Martin, R. and T. Simin, 1997, Robust beta mining, *IEEE Proceedings of KDD'97* (Workship on Issues in the Integration of Data Mining and Data Visualization).

Shanken, J., 1987, Multivariate proxies and asset pricing relations, *Journal of Financial Economics* 18, 91-110.

Sharpe, W. F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 19, 425-442.

Stromberg, A. J., 2004, Why write statistical software? The case of robust statistical methods, *Journal of Statistical Software* 10, 1-8.

Wermers, R., 1995, Herding, trade reversals, and cascading by institutional investors, *University of Colorado Working Paper*.

Table 1. Monte Carlo Simulated Betas with OLS and Robust Regression

Sample	Method	Mean	Std. Dev.	Skewness	Kurtosis
without outliers	OLS	0.0015	0.102	-0.023	3.09
	Huber	-0.0013	0.104	-0.024	3.07
with outliers	OLS	0.3785	0.142	-0.692*	3.86*
	Huber	0.0723	0.108	0.002	3.10

* represents significance at 5% level.

Table 2. Descriptive Statistics of Equities on Sample Markets

Market	Time Period	Monthly Market Index Change				Risk-Free Interest Rate		
		Name	Mean	Std. Dev.	Correlation with MSCI Index	Type	Mean	Std.Dev.
Australia (AUS)	Jan-85 to Dec-05	MSCI	0.0092	0.0491	1.000	3M-T-Bill	0.0066	0.0033
France (FRA)	Jan-85 to Dec-05	MSCI	0.0101	0.0592	1.000	3M-T- Bill	0.0048	0.0024
Germany (GER)	Jan-85 to Dec-05	FAZ General	0.0075	0.0561	0.986	3M-Benchmark Bond	0.0036	0.0016
Hong Kong (HK)	Jan-85 to Dec-05	Hang Seng Index	0.0134	0.0803	0.978	1M-Deposit	0.0040	0.0021
Japan (JAP)	Jan-85 to Dec-05	NIKKEI 225	0.0032	0.0613	0.907	3M-T-Bill	0.0019	0.0020
United Kingdom (UK)	Jan-85 to Dec-05	FTSE All Share	0.0073	0.0461	0.992	1M-T-Bill	0.0060	0.0024
United States(USA)	Jan-85 to Dec-05	S&P 500 Composite	0.0111	0.0446	0.999	3M-T-Bill	0.0039	0.0016
China (CHI)	Jan-94 to Dec-05	MSCI	-0.0046	0.1113	1.000	3M- Deposit	0.0025	0.0015
India (IND)	Jan-91 to Dec-05	BSE National	0.0169	0.0964	0.986	3M-T-Bill	0.0071	0.002
Indonesia (INDO)	Jan-91 to Dec-05	Jakarta SE Composite	0.0095	0.0868	0.956	1M-Deposit	0.0120	0.0069
Korea (KOR)	Jan-86 to Dec-05	Korea SE Composite	0.0131	0.0932	0.974	1M-Deposit	0.0067	0.0021
Malaysia (MAL)	Jan-86 to Dec-05	KLCI Composite	0.0094	0.0852	0.992	1M-Deposit	0.0038	0.0015
Philippines (PHI)	Jan-90 to Dec-05	SE Composite	0.0078	0.0919	0.949	3M-T-Bill	0.0094	0.0039
Thailand (THA)	Jan-89 to Dec-05	Bangkok S.E.T.	0.0080	0.1007	0.972	3M-Deposit	0.0057	0.0034
Argentina (ARG)	Jan-92 to Dec-05	Merval	0.0122	0.1181	0.940	1M-Deposit	0.0087	0.007
Brazil (BRA)	Jan-92to Dec-05	Bovespa	0.0814	0.186	0.761	3M-Deposit	0.0093	0.0084
Chile (CHL)	Jan-90 to Dec-05	IGPA	0.0148	0.0584	0.961	1M-CD	0.0045	0.0017
Colombia (COL)	Jan-92 to Dec-05	CSE	0.0076	0.0385	0.913	3M -Deposit	0.0150	0.0071
Mexico (MEX)	Jan-91 to Dec-05	IPC (BOLSA)	0.0223	0.0836	0.982	1M-CetesYield	0.0133	0.0078
Peru (PER)	Jan-92 to Dec-05	Lima SE General (IGBL)	0.0288	0.1110	0.824	1M-Deposit	0.0074	0.0043
Venezuela (VEN)	Jan-93 to Dec-05	S&P/IFCG	0.0244	0.1187	0.952	1M-Deposit	0.0179	0.0092
		Average	**0.0149**	**0.0846**	**0.954**		**0.0073**	**0.0039**

Table 3. Descriptive Statistics of Fama-French Three Factors

Market	Time Period	No. of Stocks	R_m-R_b		R_{SMB}		R_{HML}	
			Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
AUS	Jan-85 to Dec-05	1442	0.0026	0.0490	0.0217**	0.0620	0.0001	0.0544
FRA	Jan-85 to Dec-05	770	0.0052	0.0593	0.0091**	0.0416	0.0081**	0.0486
GER	Jan-85 to Dec-05	1041	0.0040	0.0621	0.0103**	0.0499	0.0060†	0.0561
HK	Jan-85 to Dec-05	961	0.0096†	0.0805	0.0196**	0.0828	-0.0104*	0.0702
JAP	Jan-85 to Dec-05	3659	0.0013	0.0614	0.0077**	0.0463	0.0011	0.0378
UK	Jan-85 to Dec-05	1723	0.0013	0.0461	0.0172**	0.0537	0.0032	0.0306
USA	Jan-85 to Dec-05	N/A	0.0072*	0.0445	0.0011	0.0315	0.0020	0.0359
CHI	Jan-94 to Dec-05	1557	-0.0071	0.1114	-0.0008	0.0999	0.0083	0.1377
IND	Jan-91 to Dec-05	490	0.0098	0.0967	0.0100*	0.0677	0.0109	0.0886
INDO	Jan-91 to Dec-05	319	-0.0025	0.0876	0.0108	0.1179	0.0387**	0.1447
KOR	Jan-86 to Dec-05	851	0.0064	0.0934	0.0108*	0.0740	0.0084†	0.0748
MAL	Jan-86 to Dec-05	984	0.0055	0.0853	0.0063	0.0628	0.0060	0.0696
PHI	Jan-90 to Dec-05	241	-0.0016	0.0919	0.0228**	0.1004	0.0298**	0.1158
THA	Jan-89 to Dec-05	498	0.0023	0.1010	0.0113	0.0988	0.0106	0.1153
ARG	Jan-92 to Dec-05	70	0.0036	0.1189	0.0056	0.0915	0.0169*	0.1011
BRA	Jan-92to Dec-05	478	0.0722**	0.1810	0.0424**	0.1512	0.0162	0.173
CHL	Jan-90 to Dec-05	196	0.0103*	0.0585	0.0053	0.055	0.0174**	0.0724
COL	Jan-92 to Dec-05	35	-0.008**	0.0392	-0.0019	0.074	0.0198*	0.1118
MEX	Jan-91 to Dec-05	129	0.0089	0.0837	0.0063	0.0812	0.0017	0.1176
PER	Jan-92 to Dec-05	94	0.0214*	0.1102	0.0165	0.1313	0.0384**	0.1827
VEN	Jan-93 to Dec-05	39	0.0065	0.1191	-0.0023	0.1147	0.0054	0.1313
		Average	**0.0076**	**0.0848**	**0.0105**	**0.0804**	**0.0118**	**0.0938**

The factor values for the U.S. are from French's website. In other markets, we use the market index and risk-free rate described in Table 2 to obtain the excess market returns (**R_m-R_b**). For the SMB and HML factors, we form the 6 size-BE/ME portfolios based on the equities from Datastream (the number of equities used are indicated in the third column). **, *and † represent significance at 1%, 5% and 10% level, respectively.

Table 4. Properties of the Robust Regression Betas

	Sample Period	No. of betas	The CAPM β_m mean	The CAPM β_m std.dev.	The F-F model β_m mean	The F-F model β_m std.dev.	β_{smb} mean	β_{smb} std.dev.	β_{hml} mean	β_{hml} std.dev.
AUS	Jan-90 to Dec-05	109475	0.816	0.844	0.815	0.851	0.620	0.832	0.178	0.706
FRA	Jan-90 to Dec-05	76829	0.614	0.664	0.818	0.664	0.535	0.685	0.105	0.541
GER	Jan-90 to Dec-05	112701	0.591	0.586	0.783	0.712	0.492	0.721	-0.027	0.560
HK	Jan-90 to Dec-05	92655	0.785	0.459	0.844	0.471	0.586	0.537	-0.193	0.564
JAP	Jan-90 to Dec-05	452902	0.764	0.449	0.734	0.437	0.692	0.618	0.426	0.733
UK	Jan-90 to Dec-05	148710	0.857	0.659	0.895	0.651	0.537	0.624	0.106	0.743
USA	Jan-90 to Dec-05	953058	0.835	0.840	0.781	0.769	0.680	0.985	0.244	0.957
CHI	Jan-00 to Dec-05	85922	0.243	0.242	0.340	0.235	0.637	0.655	-0.142	0.492
IND	Jan-96 to Dec-05	89584	0.862	0.464	0.916	0.461	0.560	0.630	-0.210	0.430
INDO	Jan-96 to Dec-05	27397	0.926	0.555	0.920	0.54	0.331	0.456	0.203	0.350
KOR	Jan-93 to Dec-05	137852	0.816	0.412	0.973	0.433	0.639	0.539	0.062	0.538
MAL	Jan-93 to Dec-05	87680	1.164	0.425	0.983	0.355	0.734	0.540	0.394	0.504
PHI	Jan-95 to Dec-05	23206	0.831	0.584	0.900	0.587	0.371	0.485	0.152	0.393
THA	Jan-94 to Dec-05	46136	0.728	0.545	0.878	0.591	0.340	0.540	0.234	0.504
ARG	Jan-97 to Dec-05	6484	0.665	0.340	0.673	0.347	0.347	0.536	0.117	0.397
BRA	Jan-97 to Dec-05	31776	0.546	0.444	0.614	0.468	0.203	0.487	0.049	0.350
CHL	Jan-96 to Dec-05	18650	0.730	0.564	0.828	0.594	0.264	0.515	0.104	0.416
COL	Jan-97 to Dec-05	3389	1.352	0.957	1.530	0.959	0.130	0.415	0.055	0.310
MEX	Jan-96 to Dec-05	11508	0.612	0.438	0.669	0.455	0.269	0.466	0.124	0.334
PER	Oct-97 to Dec-05	8970	0.695	0.574	0.676	0.580	0.173	0.495	0.046	0.391
VEN	Dec-97 to Dec-05	2118	0.780	0.409	0.800	0.426	0.335	0.409	0.086	0.402
		Average	**0.772**	**0.545**	**0.827**	**0.552**	**0.451**	**0.580**	**0.101**	**0.505**

Table 5. Properties of the Cross-sectional Mean and Variance of the Betas

	$E_c(\hat{\beta}^b_{imt})\vert_{CAPM}$ (a)		$E_c(\hat{\beta}^b_{imt})\vert_{F-F}$ (b)		**Correlation between (a) and (b)**	$Var_c(\hat{\beta}^b_{imt})\vert_{CAPM}$ (c)		$Var_c(\hat{\beta}^b_{imt})\vert_{F-F}$ (d)		**Correlation between (c) and (d)**
	Mean	Std. Dev.	Mean	Std. Dev.		Mean	Std. Dev.	Mean	Std. Dev.	
AUS	0.816	0.058	0.869	0.119	0.345	0.626	0.180	0.653	0.134	0.824
FRA	0.621	0.135	0.818	0.070	0.563	0.283	0.150	0.350	0.188	0.916
GER	0.569	0.095	0.756	0.124	0.895	0.278	0.165	0.393	0.241	0.910
HK	0.823	0.155	0.890	0.114	0.609	0.173	0.061	0.186	0.060	0.947
JAP	0.786	0.105	0.764	0.115	0.897	0.179	0.053	0.168	0.048	0.969
UK	0.840	0.119	0.901	0.062	0.405	0.367	0.160	0.377	0.148	0.954
USA	0.859	0.081	0.867	0.101	0.829	0.626	0.210	0.793	0.077	0.485
CHI	0.243	0.039	0.354	0.117	0.456	0.058	0.012	0.044	0.012	0.672
IND	0.863	0.115	0.913	0.088	0.903	0.195	0.062	0.198	0.053	0.985
INDO	0.907	0.133	0.911	0.056	0.956	0.285	0.057	0.286	0.058	0.886
KOR	0.808	0.055	0.958	0.076	0.514	0.153	0.051	0.159	0.072	0.948
MAL	1.177	0.092	0.979	0.038	0.428	0.173	0.028	0.124	0.030	0.829
PHI	0.814	0.108	0.885	0.085	0.867	0.320	0.156	0.328	0.138	***0.983***
THA	0.727	0.044	0.872	0.074	0.786	0.285	0.092	0.330	0.128	0.962
ARG	0.665	0.045	0.673	0.035	0.753	0.116	0.023	0.122	0.029	0.925
BRA	0.558	0.126	0.627	0.103	***0.991***	0.179	0.040	0.204	0.053	0.966
CHL	0.730	0.038	0.828	0.027	0.517	0.319	0.076	0.350	0.100	0.946
COL	***1.365***	***0.184***	***1.558***	***0.242***	0.926	***1.045***	***0.268***	***0.863***	***0.288***	*0.817*
MEX	0.624	0.077	0.675	0.045	0.882	0.185	0.046	0.203	0.061	0.953
PER	0.696	0.027	0.677	0.048	0.265	0.332	0.068	0.338	0.080	0.933
VEN	0.797	0.113	0.813	0.094	0.895	0.145	0.099	0.159	0.121	0.815
Average	**0.776**	**0.093**	**0.838**	**0.087**	**0.699**	**0.301**	**0.098**	**0.320**	**0.101**	**0.887**

This table reports the first two moments of $E_c(\hat{\beta}^b_{imt})$, the cross-sectional mean of the betas, and $Var_c(\hat{\beta}^b_{imt})$, cross-sectional variance of the betas in each market, both under the CAPM and the F-F model. Column 6 and 11 show the correlation between the series calculated under the CAPM and the F-F model. The underlined values represent the minimum one in the series, and the italicized values represent the maximum one in the series.

Table 6. Properties of the Herding Measure Estimated under the F-F model

Panel A. Properties of Herding Measure in the Markets

	$h_{mt} = 1 - e^{H_{mt}/2}$ Mean	Std. Dev.	ϕ	**Maximum likelihood values**	**Correlation between filtered series and smoothed series**
AUS	0.010	0.011	0.997	5.29	0.976
FRA	0.032	0.015	0.992	5.30	0.973
GER	0.030	0.021	0.993	4.52	0.981
HK	0.029	0.018	0.995	5.29	0.985
JAP	0.051	0.019	0.998	4.30	0.991
UK	0.036	0.011	0.993	5.79	0.983
USA	0.004	0.025	0.999	5.85	0.997
CHI	0.015	0.007	0.983	4.33	0.964
IND	0.035	0.010	0.997	4.88	0.983
INDO	0.033	0.011	0.986	4.83	0.965
KOR	0.032	0.017	0.998	5.30	0.962
MAL	0.041	0.027	0.994	5.10	0.984
PHI	0.032	0.015	0.990	4.93	0.976
THA	0.038	0.015	0.994	5.01	0.983
ARG	0.020	0.025	0.994	4.72	0.978
BRA	0.013	0.005	0.954	4.73	0.885
CHL	0.040	0.019	0.995	4.84	0.981
COL	0.055	0.028	0.995	4.72	0.987
MEX	0.053	0.009	0.977	4.83	0.927
PER	0.021	0.011	0.977	4.83	0.944
VEN	0.041	0.021	0.988	4.61	0.968
Average	**0.031**	**0.016**	**0.990**	**4.95**	**0.970**

Panel B. *t* test on the difference of the mean of herding between developed and emerging markets

Source	**Observations**	**Mean**	**Std. Dev.**
$h_{mt}^{developed}$	1344	0.027	0.015
$h_{mt}^{emerging}$	1659	0.034	0.018

H_0: $mean(h_{mt}^{developed}) = mean(h_{mt}^{emerging})$

H_a: $mean(h_{mt}^{developed}) < mean(h_{mt}^{emerging})$

Result: t-value=11.41, p=0.00

Panel *A* reports the Kalman filtered state space model of *Eq.7* & *Eq.8*. Column 6 lists the correlation between the Kalman filtered series and Kalman smoothed series. Panel *B* tests the difference of the mean of herding between developed markets ($h_{mt}^{developed}$) and emerging markets ($h_{mt}^{emerging}$).

Table 7. Turning Points of the Herding Measure

Panel A. Turning Points

	Peak*	Trough*	Average length of the cycles (μ_L ,month)
AUS	Feb90,Apr95,Oct97,Jun00,Apr04	Mar91, Apr97, Sep98,Mar02	41
FRA	Jan92,Aug97,Apr00,Feb02,Apr04	Sep92,Jul98,Aug01,Jun03,Nov04	37
GER	Oct91,Jul98,Dec04	Jan90,Jun92,Feb02	68
HK	Dec92,Jan98,Jun01,Jun04	Apr96, Apr99,Jun03,Mar05	42
JAP	Aug00,Jun04	Aug95,Jul03,Apr05	58
UK	Mar96,Oct99,May04	Aug91,Jul97,Dec02	61
USA	Feb94,Jun96,May98	Oct90,Mar95,Jul97,Dec98	33
CHI	Apr03,Nov04	May04	19
IND	Jun98,May00,Apr04	Dec96,Jul99,May03,Nov04	32
INDO	Apr97,Jul03	Oct00,Jun04	57
KOR	Dec96,Jul00,May04	Dec98,Nov02	45
MAL	Apr94,May97,Dec02	Nov95,Jul01,May03	44
PHI	Sep95,Jul97,Sep00,Aug04	Dec96,Oct98,Jun03	36
THA	Feb96,Jun97	Oct99,Sep04	69
ARG	Sep98,Dec00	Nov99,Sep02	32
BRA	Jul99,Jul03	Jul98,Nov01,Sep04	37
CHL	Feb00,Jun02,Apr05	Jul99,Feb01,Jul03	28
COL	Apr98,Jun00,Aug03	Jan00,Sep02,Oct04	31
MEX	Aug96, Feb01,Jul03	Feb99,Apr02,Aug04	38
PER	Jul00,Nov03	Jan00,Oct01	31
VEN	Nov04	Jan03	44
		Average	**42**

Panel B. One-way ANOVA F test for equality of the average length of herding cycle of each group

Source	Observations	Mean	Std. Dev.
$\mu_L^{Developed}$	7	48	13.5
μ_L^{Asian}	7	43	16.4
$\mu_L^{LatinAmerican}$	7	34	5.5

H_0: Mean($\mu_L^{Developed}$)= Mean(μ_L^{Asian})=Mean($\mu_L^{LatinAmerican}$)

H_a: Not H_0

Result: F-value=2.21, p-value=0.14

Panel *A* lists the Bry-Boschan turning points of herding measures. μ_L represents the average length of the cycles, calculated as twice of the arithmetic mean of monthly intervals from peak to trough, or trough to peak. Panel *B* tests the null hypothesis that there are no differences among the mean length of the cycles of the three groups. $\mu_L^{Developed}$, μ_L^{Asian} , $\mu_L^{LatinAmerican}$ represent the average length of the cycles of the developed group, the Asian group, and the Latin American group, respectively.

* The first three digits represent the month and the last two digits represent the year. For instance, Feb90 means February, 1990.

Table 8. Volatility of Herding Measures

Panel A. Descriptive Statistics of Herding Measure Volatility $(\sigma^2_{h,t})$

	Mean	Std. Dev.		Mean	Std. Dev.		Mean	Std. Dev.
AUS	0.0018	0.0061	CHI	0.0011	0.0014	ARG	0.0117	0.0146
FRA	0.0040	0.0056	IND	0.0027	0.0033	BRA	0.0005	0.0008
GER	0.0064	0.0031	INDO	0.0019	0.0021	CHL	0.0058	0.0060
HK	0.0061	0.0097	KOR	0.0045	0.0078	COL	0.0127	0.0122
JAP	0.0049	0.0084	MAL	0.0146	0.0150	MEX	0.0013	0.0015
UK	0.0020	0.0022	PHI	0.0039	0.0052	PER	0.0028	0.0037
USA	0.0104	0.0141	THA	0.0047	0.0075	VEN	0.0078	0.0077
Average	**0.0051**	**0.0070**	**Average**	**0.0048**	**0.0060**	**Average**	**0.0061**	**0.0066**

Panel B. One-way ANOVA F test for equality of the mean of herding volatility of each group

Source	Observations	Mean	Std. Dev.
$\sigma^2_{h,Developed}$	1344	0.0051	0.0084
$\sigma^2_{h,Asian}$	900	0.0053	0.0092
$\sigma^2_{h,LatinAmerican}$	759	0.0060	0.0093

H_0: $\mu^{developed}_{\sigma^2_h} = \mu^{Asian}_{\sigma^2_h} = \mu^{LatinAmerican}_{\sigma^2_h}$

H_a: Not H_0

Result: F-value=2.69, p-value=0.07

Panel *A* lists the mean and standard deviation for the variance of herding measure. Panel *B* tests the null hypothesis that there are no differences among the means of herding volatility for the three groups.

Table 9. Correlation Coefficient of Herding Measures

	AUS	FRA	GER	HK	JAP	UK	USA	CHI	IND	INDO	KOR	MAL	PHI	THA	ARG	BRA	CHL	COL	MEX	PER	VEN
AUS	1.00																				
FRA	0.71*	1.00																			
GER	0.37*	0.73*	1.00																		
HK	0.36*	-0.04	-0.11	1.00																	
JAP	-0.38*	-0.29*	-0.23*	-0.06	1.00																
UK	0.66*	0.80*	0.74*	-0.16*	-0.03	1.00															
USA	0.90*	0.53*	0.19*	0.46*	-0.15*	0.49*	1.00														
CHI	-0.35*	-0.85*	-0.53*	-0.02	-0.81*	-0.60*	0.51*	1.00													
IND	-0.60*	0.84*	0.88*	-0.31*	-0.22	0.56*	-0.70*	-0.84*	1.00												
INDO	-0.13	-0.17	-0.13	0.15	-0.73*	-0.67*	0.21*	0.79*	-0.13	1.00											
KOR	0.54*	-0.09*	-0.28*	0.33*	-0.14	-0.50*	0.41*	0.19	-0.16	0.58*	1.00										
MAL	-0.45*	0.23*	0.49*	0.05	-0.67*	-0.11	-0.51*	0.09	0.58*	0.58*	0.09	1.00									
PHI	-0.48*	0.50*	0.53*	0.01	-0.69*	-0.03	-0.55*	-0.31*	0.40*	0.36*	0.23*	0.75*	1.00								
THA	0.24*	-0.50*	-0.40*	0.46*	-0.64*	-0.77*	0.10	0.81*	0.80*	0.80*	0.56*	0.45*	0.33*	1.00							
ARG	0.20*	-0.53*	0.07	0.51*	0.08	-0.26*	0.59*	0.32*	-0.10	0.33*	0.30*	0.36*	0.01	0.46*	1.00						
BRA	0.12	0.19	0.05	-0.48*	0.08	0.22*	-0.15	-0.38*	0.03	-0.09	-0.00	-0.14	-0.35*	-0.35*	-0.21*	1.00					
CHL	-0.22*	0.04	0.22*	0.24*	-0.56*	-0.37*	0.16	0.35*	0.32*	0.67*	0.49*	0.74*	0.62*	0.61*	0.63*	-0.47*	1.00				
COL	-0.79*	0.74*	0.84*	-0.16	-0.30*	0.49*	-0.78*	-0.85*	0.84*	-0.07	-0.42*	0.62*	0.76*	-0.37*	-0.17	-0.05	0.08	1.00			
MEX	0.49*	-0.33*	-0.31*	0.37*	-0.13	-0.57*	0.47*	-0.16	-0.20*	0.38*	0.67*	-0.07	0.01	0.50*	0.37*	0.15	0.20*	-0.30*	1.00		
PER	-0.21*	-0.42*	0.06	0.32*	-0.63*	-0.31*	0.13	0.66*	-0.16	0.49*	-0.01	0.26*	0.43*	0.47*	0.44*	-0.28*	0.31*	0.22*	0.26*	1.00	
VEN	-0.63*	0.19	0.65*	0.11	-0.06	0.42*	-0.25*	0.29*	0.52*	-0.11	-0.41*	0.74*	0.64*	-0.06	0.39*	-0.50*	0.48*	0.57*	0.47*	0.22*	1.00

This table shows the correlation coefficients of herding measures from the Fama-French three factor model. * represents significance at 5% level. Among them, values in bold indicate significantly positive, and italicized values indicate significantly negative.

Table 10. Test of Correlation Coefficients of Herding Measures among Various Groups

Panel A. Two-sample *t* test for correlation coefficients within single group and between groups

Source	Observations	Mean	Std. Dev.	95% Confidence Interval
Corr_within-group	63	0.24	0.340	[0.14, 0.34]
Corr_between-group	147	0.02	0.447	[-0.06, 0.09]

H_0: Mean(Corr_within-group)= Mean(Corr_between-groups)
H_a: Mean(Corr_within-group) > Mean(Corr_between-groups)

Result: t-value = 3.60, p-value= 0.0002

Panel B. One-way ANOVA F test for equality of the mean of herding correlation in each group

Source	Observations	Mean	Std. Dev.
Corr_Developed	21	0.26	0.413
Corr_Asian	21	0.33	0.430
Corr_Latin American	21	0.13	0.341

H_0: Mean(Corr_Developed)= Mean(Corr_Asian)=Mean(Corr_Latin American)
H_a: Not H_0

Result: F-value=1.33, p-value=0.27

In Panel *A*, we tests the null hypothesis that the mean of pair wise herding correlation between countries from the same groups (i.e., for instance, Australia and France, China and India) is the same as the mean of pair wise correlation between countries from different groups (for instance, China and Australia). Panel *B* is to test the null hypothesis that there are no differences among the means of the herding correlation between countries from the Developed group, from the Asian group, and from the Latin American group.

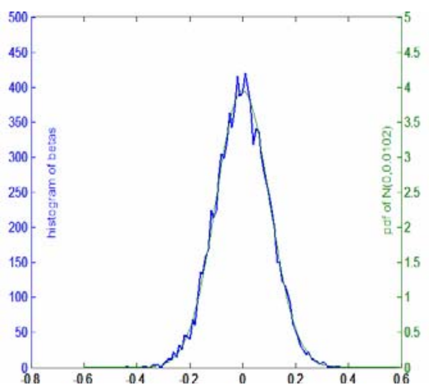
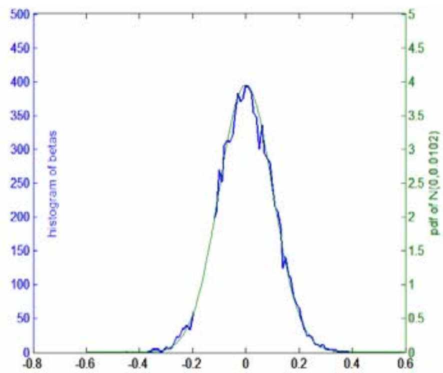
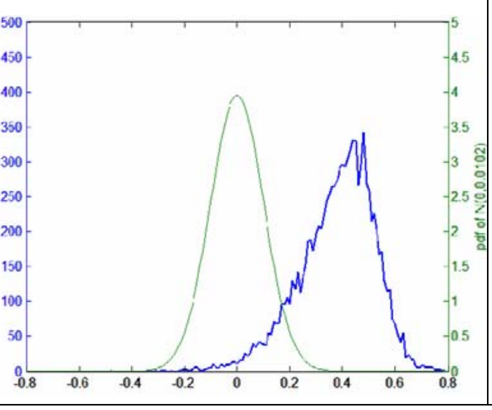
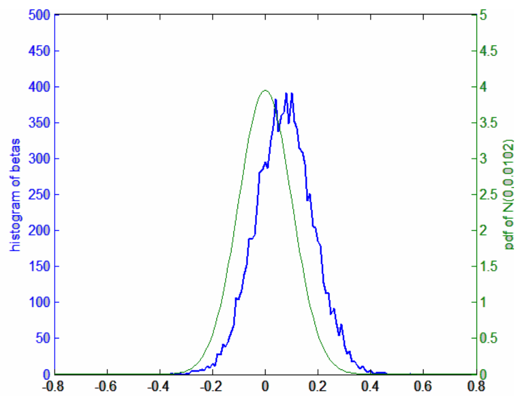
Table 11. Correlation Coefficient of the First Difference of Herding Measures

	AUS	FRA	GER	HK	JAP	UK	USA	CHI	IND	INDO	KOR	MAL	PHI	THA	ARG	BRA	CHL	COL	MEX	PER	VEN
AUS	1.00																				
FRA	**0.33***	1.00																			
GER	**0.19***	0.11	1.00																		
HK	**0.27***	0.06	**0.19***	1.00																	
JAP	0.12	**0.20***	**0.16***	**0.23***	1.00																
UK	**0.32***	**0.20***	**0.15***	*-0.16**	**0.32***	1.00															
USA	**0.53***	0.11	**0.28***	**0.26***	0.10	0.06	1.00														
CHI	*-0.32**	-0.17	-0.15	*-0.36**	-0.23	**0.34**	*-0.39**	1.00													
IND	0.01	0.03	0.15	0.03	0.13	-0.08	0.11	*-0.38**	1.00												
INDO	-0.17	0.05	-0.17	-0.13	*-0.39**	*-0.32**	-0.05	0.15	0.03	1.00											
KOR	**0.35***	**0.28***	0.05	0.13	0.11	-0.01	0.13	-0.15	0.05	**0.22***	1.00										
MAL	**0.24***	**0.24***	0.07	**0.17***	-0.10	-0.05	-0.13	0.03	0.06	**0.38***	**0.18***	1.00									
PHI	0.07	0.05	0.13	**0.34***	0.05	-0.04	0.15	-0.09	0.09	-0.05	**0.21***	0.01	1.00								
THA	0.04	0.02	-0.13	**0.27***	*-0.19**	-0.12	0.03	0.08	-0.10	**0.23***	0.04	**0.27***	**0.19***	1.00							
ARG	0.03	*-0.27**	**0.41***	**0.22***	0.07	0.05	0.03	-0.09	0.15	-0.04	-0.02	0.04	0.04	0.07	1.00						
BRA	0.17	0.15	-0.09	-0.09	-0.14	0.16	-0.07	-0.12	0.02	0.15	0.06	0.01	*-0.27**	-0.04	0.02	1.00					
CHL	0.06	0.16	0.01	**0.19***	-0.04	-0.15	0.08	0.12	**0.37***	-0.01	0.19	0.16	**0.35***	0.03	0.00	*-0.28**	1.00				
COL	-0.13	-0.13	0.10	**0.20***	0.07	-0.04	-0.01	-0.05	0.21	-0.02	-0.03	0.09	0.15	0.01	-0.09	-0.10	0.03	1.00			
MEX	0.15	-0.16	-0.01	**0.23***	-0.03	-0.17	**0.23***	*-0.33**	0.13	-0.04	0.15	-0.10	-0.11	0.15	**0.31***	0.19	*-0.23**	0.18	1.00		
PER	0.15	-0.18	0.17	0.16	0.17	0.01	-0.04	0.04	0.07	-0.23	-0.06	-0.06	**0.20***	-0.03	0.19	*-0.23**	0.14	**0.29***	0.14	1.00	
VEN	*-0.31**	-0.03	0.08	0.09	0.08	**0.22***	0.13	**0.33***	0.09	-0.10	*-0.35**	-0.09	0.11	**0.21***	0.15	-0.10	0.09	0.12	-0.15	-0.12	1.00

This table shows the correlation coefficients of the first difference of herding measures from the Fama-French three factor model.

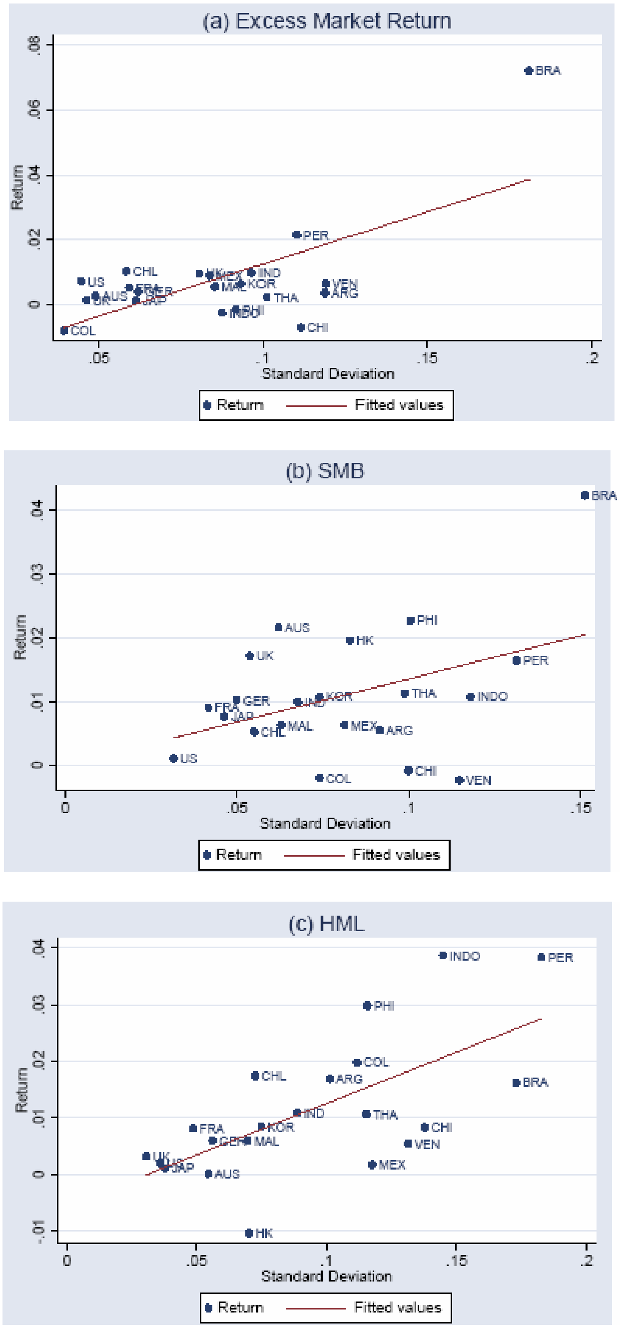
* represents significant at 5% level. Among them, values in bold indicate significantly positive, and italicized values indicate significantly negative. The highlighted values (all of them are insignificant) indicate that their counterpart in Table 9 is significant at 5% level.

Figure 1. Monte Carlo Simulation for OLS and Robust Beta Estimates

	OLS estimate	Huber robust estimate
without outliers		
with outliers		

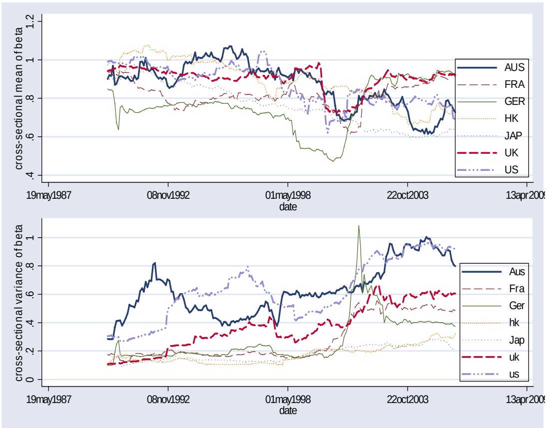
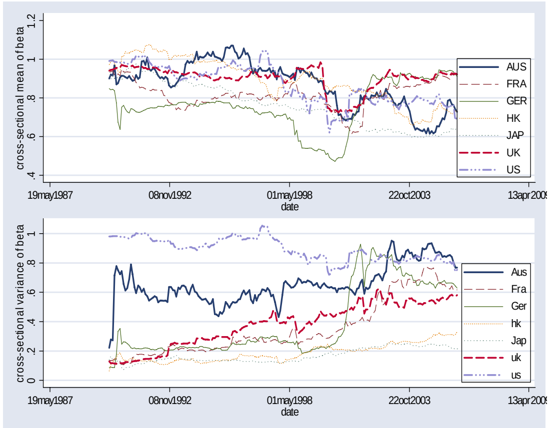
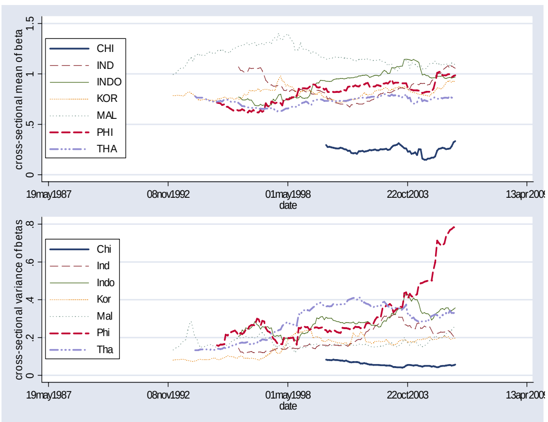
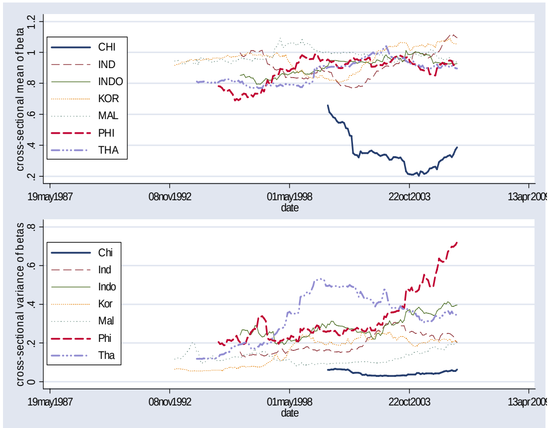
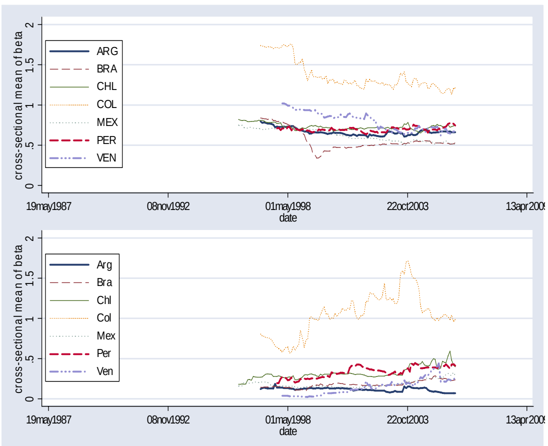
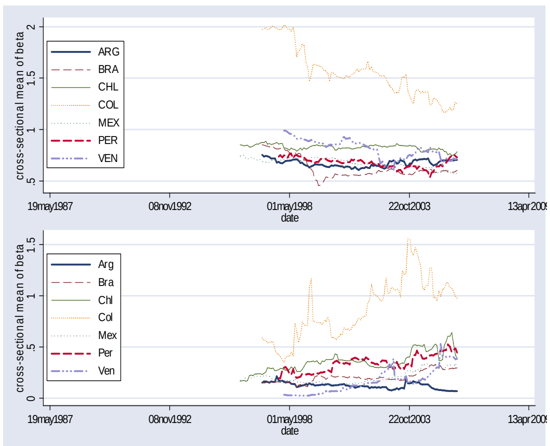
The thick lines in the top two panels represent the histogram profile (10,000 replicates) of the distribution of the OLS, and Huber robust estimated beta when both the independent variable and the error terms are standard normally distributed. The bottom two panels show histogram profiles of the distribution when both the same data is used with the exception that there is one percent probability of an (x_o, ε_o) outlier, where $x_o \sim$ N(5, 2), and $\varepsilon_o \sim$ N(10, 0.5). The overlaid density (thin line) is the "true" distribution of *N(0,0.0102)*.

Figure2. Fama-French Three Factor in Global Markets



The figures depict the scatterplot of the return vs. risk relationship for three factors in Fama-French model (*a* for excess market return, *b* for *SMB*, and *c* for *HML*), with each point representing one of the 21 markets. A regression line plot is placed on top of each scatterplot.

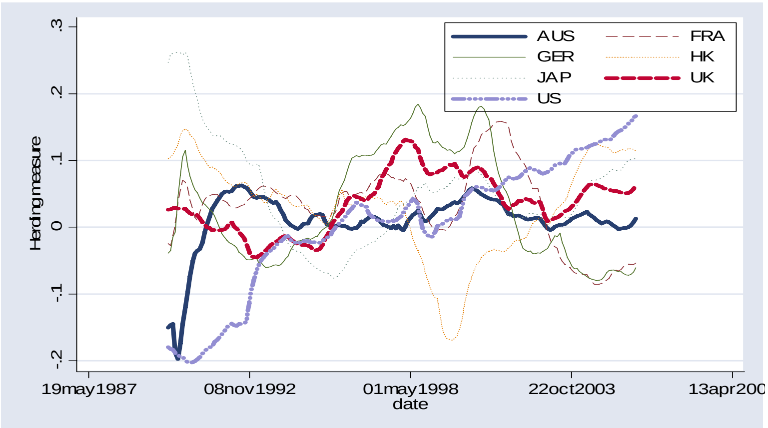
Figure 3. Evolution of Cross-Sectional Mean and Variance of Betas

<table>
<tr><th></th><th>The CAPM</th><th>F-F model</th></tr>
<tr><td>Developed Market</td><td>
</td><td>
</td></tr>
<tr><td>Asian Markets</td><td>
</td><td>
</td></tr>
<tr><td>Latin American Markets</td><td>
</td><td>
</td></tr>
</table>

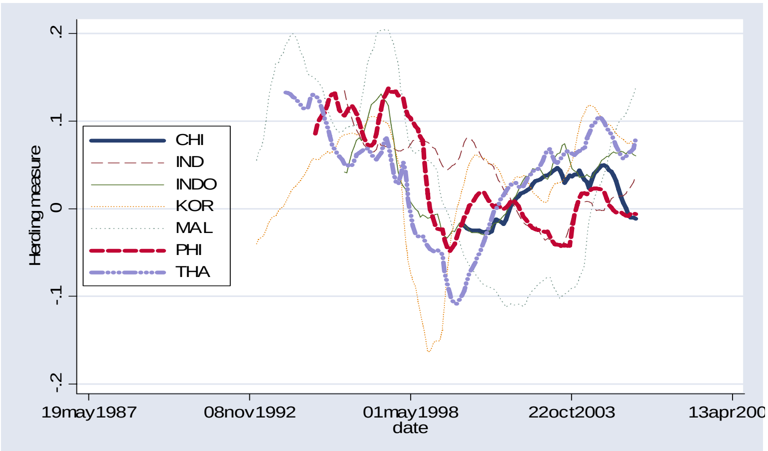
This figure shows the evolution of the cross-sectional mean and variance of the betas for the developed groups, Asian group and Latin American group, under the CAPM and the F-F model respectively. In each combinations of the graph, the top one is for the cross-sectional mean of the betas, and the bottom one for the cross-sectional variance of the betas.

Figure 4. Evolution of Herding Measures

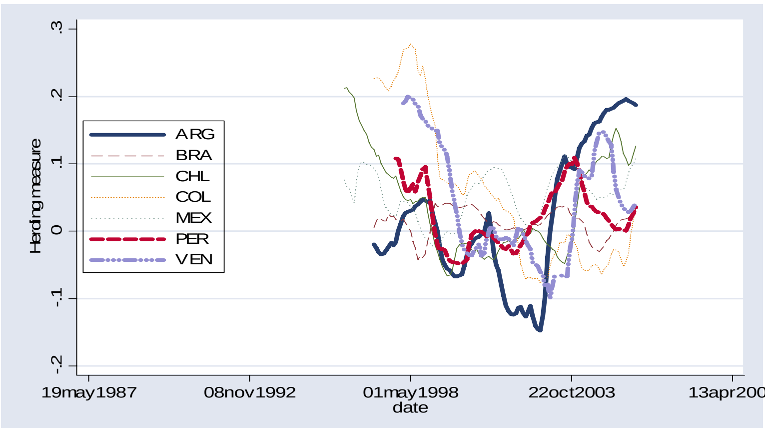
(a) Developed Markets



(b) Asian Markets



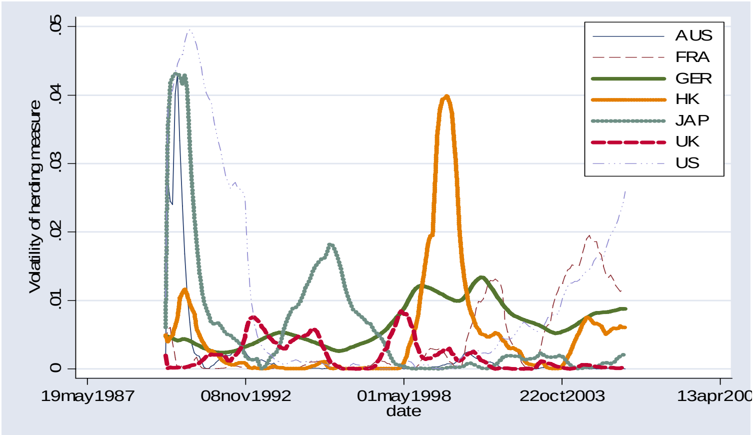
(c) Latin American Markets



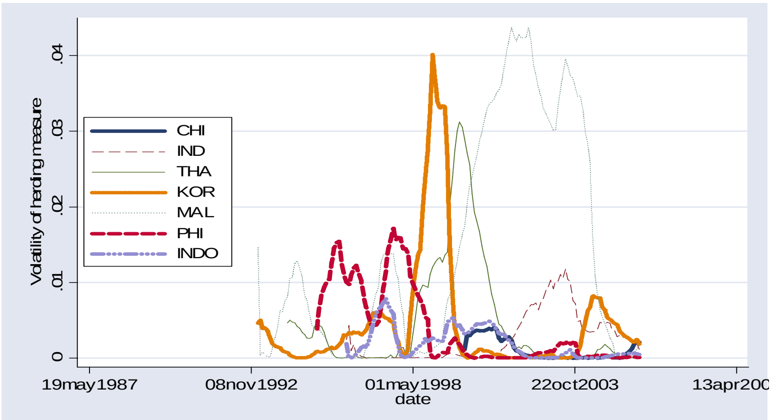
The figures depict the evolution of herding measure, obtained through the market betas of the F-F model, for the developed group, Asian group, and Latin American group.

Figure 5. Evolution of Volatility of Herding Measures

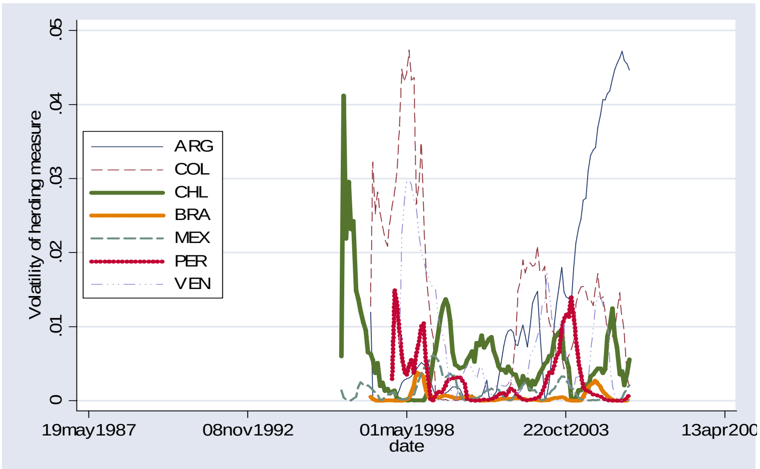
(a) Developed Markets



(b) Asian Markets



(c) Latin American Markets



The figures depict the evolution of the volatility of the herding measures (σ_h^2) for the developed group, Asian group, and Latin American group.