Analysts' Forecast Dispersion and Stock Returns: A Panel Threshold Regression Analysis of the Conditional Limited Market Participation Hypothesis

ABSTRACT

Prior research has investigated the association between analysts' forecast dispersion and future stock return but the evidence is not conclusive. Diether et al. (2002) address the theory of limited market participation, and provide evidence that analysts' earnings forecast dispersion is negatively associated with subsequent stock return. Johnson (2004) argues, however, that dispersion in analysts' forecasts is a proxy of risk. Grounded on Miller's (1977) market friction hypothesis, we propose a conditional limited market participation hypothesis and reexamine the relation between analysts' forecast dispersion and stock returns using the panel threshold regression approach, which allows the coefficient on the independent variable to shift when the conditioned variables exceed their respective thresholds. Our empirical results show that the degree of the negative association between analysts' dispersion and future stock return becomes considerably diminished when dispersion and size exceed the respective thresholds. Our finding does not support the view of Johnson (2004) that dispersion in analysts' forecasts serves as a proxy for risk. Although our results are consistent with the limited market participation argument in Diether et al. (2002), we modify their argument as the strength of their results is conditional on the magnitude of dispersion and the size of firm.

Keywords: Analysts' forecast dispersion, Stock returns, Short-sale constraint, Panel threshold model, Limited market participation

JEL Classification: D1; G1; C58

Analysts' Forecast Dispersion and Stock Returns: A Panel Threshold Regression Analysis of the Conditional Limited Market Participation Hypothesis

1. Introduction

How differences of analyst's opinion affect future stock returns has been of great interest to researchers. However, conflicting results and different interpretations of these results are reported in prior research. In particular, although both Diether, Malloy and Scherbina (2002) and Johnson (2004) demonstrate that investors overvalue stocks with high dispersion in analysts' forecasts, which leads to lower future stock returns, they provide different reasons to explain this relation. While Diether et al. (2002) argue that the overvaluation is due to limited market participation; Johnson (2004) attributes the finding to risk factor employing option pricing theory to explain his results.

While studies in analysts forecast dispersion abound, what does the dispersion proxy for is unsettled. Some studies consider dispersion as a proxy of divergence of opinion (e.g., Atiase and Bamber, 1994; Diether et al., 2002; Garfinkel and Sokobin, 2006 and Garfinkel, 2009), others view it as a proxy of uncertainty (e.g., Imhoff and Lobo, 1992; Barron and Stuerke, 1998; Johnson, 2004 and Barron, Stanford and Yu, 2009). Therefore, testing what analysts' forecast dispersion proxies for is valuable. To be sure, while Johnson (2004) and Diether et al. (2002) both explain the negative relation between dispersion levels and future returns, Diether et al. (2002) point out their results strongly reject the interpretation of dispersion in analysts' forecasts as a measure of risk.

Our study sheds light on this line of research and extends the debate between Diether et al. (2002) and Johnson (2004) by investigating the *degree* of the negative association between analysts' forecast dispersion and stock returns. To this end, we hypothesize that limited market participation largely diminishes when the dispersion and firm size exceed their respective threshold value, and the negative association between analysts' forecast dispersion and stock returns is conditional on the magnitude of analysts' forecast dispersion and firm size. Our results complement Diether et al. (2002), but we modify their argument as we find only firms with analysts' forecast dispersion and firm size below a threshold show lower future stock returns. This finding is consistent with Miller's (1977) overvaluation hypothesis in the presence of market frictions. However, our finding is not supportive of Johnson's (2004) contention that the analysts' forecast dispersion is a proxy of firm's idiosyncratic risk.

To examine the conditional limited market participation hypothesis, we investigate the degree of the association between analysts' forecast dispersion and stock returns conditional on the magnitude of analysts' forecast dispersion and firm size using the panel threshold regression approach. Panel threshold regression is particularly suitable for testing our hypothesis. In empirical research, the constant-coefficient regression models have been used extensively. However, these models only describe the average relation between independent and dependent variables, and the resulting coefficient estimates are not necessarily indicative of the effects of the independent variables on the tail observations.

The panel threshold regression approach was developed by Hansen (1999), and used widely in recent economics and finance research. *Threshold* describes a division of observations into intervals based on whether an observed variable passes a threshold value. The threshold model is a random-coefficient model in which the parameter of the independent variable can be expressed as regime-varying, hence capturing the non-monotonic degree of association between the dependent and independent variables. To the extent that the sample segmentation and the degree of the association between dependent (future stock returns) and independent variables (analysts' forecast dispersion) are determined jointly and endogenously, using the threshold model can address the potential problem in prior studies that assume segmentation of the sample is exogenous.

We use the standard deviation of analysts' earnings forecast scaled by the absolute value of the mean analysts' earnings forecast as a proxy for analysts' forecast dispersion, where higher/lower values imply higher/lower disagreement of professional forecasters. We measure dispersion for each month and calculate the monthly stock returns following the month in which forecast dispersion is measured. We then introduce a four-regime panel threshold regression model where the regime-switching processes are determined by two different conditioned variables: analysts' forecast dispersion (*DISPERSION*) and size of firm (*SIZE*). As an extension of Hansen's (1999) model in which a two-regime setting with one single threshold variable is considered, the observations in our study are divided into four regimes depending on both the magnitude of forecast dispersion and firm size. Our modeling thus rests upon two important considerations: The degree of dispersion and the firm size which might affect short-sale constraint. Our sample consists of 384,401 monthly observations of 4,122 U.S. non-financial firms for the period from 1983 to 2009.

Our empirical results are consistent with the following observations. First, OLS results confirm Diether et al. (2002) that analysts' forecast dispersion is negatively associated with future stock returns. However, this negative relation becomes insignificant when firm fixed effect is incorporated. Second, panel threshold regressions yield different results that are consistent with our conditional limited market participation hypothesis. Before *SIZE* exceeds the threshold (e.g., 6.8667), analysts' forecast dispersion is negatively associated with future stock return at the 5% significance level. By contrast, the degree of the association considerably weakens and becomes insignificant when *SIZE* is above the threshold. Next, we find that the regime switches when *DISPERSION* is 0.0161. In particular, the negative relation between analysts' forecast dispersion and future stock return is significant at the 1% significance level (insignificant) when *DISPERSION* is below (above) 0.0161.

Last but not least, the panel model with two conditioned variables (i.e., *DISPERSION* and *SIZE*) show that when *DISPERSION* and *SIZE* are below their respective thresholds, analysts' dispersion is negatively associated with future stock return at the 1% significance level. By contrast, this negative association is less promising when *DISPERSION* and *SIZE* are both above their respective thresholds. Our threshold results thus suggest that the negative association between analysts' forecast dispersion and future stock return is manifested in firms with lower magnitude of analysts' dispersion and firms that are small size. This result, though is consistent with Miller's (1977) market friction hypothesis, modifies the interpretations on how short-sale constraint impacts future stock returns. More importantly, our result is not consistent with the interpretation based upon risk (e.g., Johnson, 2004).

The rest of the paper proceeds as follows. Section 2 reviews related studies and develops research questions. Section 3 discusses the methodology. Section 4 describes the sample, variables, and empirical models. Section 5 discusses the empirical results. Section 6 summarizes and concludes the paper.

2. Literature review and research questions

2.1 Studies on analysts' forecast dispersion and stock returns

Prior theoretical and empirical studies investigating the relation between analysts' forecast dispersion and stock returns have mostly suggested a negative association. In particular, Miller (1977) suggests that optimism sets stock prices because pessimistic investors are reluctant to sell short. Highly dispersed analysts' earnings forecasts imply firms have large differences in future expectations, with optimistic investors believing good future prospects and pessimistic investors believing poor future prospects. Since pessimistic investors avoid such firms due to the risks or difficulties of short selling, optimistic investors dominate and drive the stock prices of these firms. This suggests a negative relation between dispersion levels and future stock returns when the overpricing is corrected. Ackert and

Athanasakkos (1997) provide support for this theory when they find that high earnings forecast dispersion is associated with lower stock returns.

Diether et al. (2002) directly addressed this issue by focusing on the dispersion of analysts' forecasts. They show that stocks with higher dispersion in analysts' earnings forecasts earn lower future returns than otherwise similar stocks. This effect is most pronounced in small stocks and stocks that have performed poorly over the past year. They argue that this result supports the view that dispersion of opinion plays a role in future stock returns because investors give the lowest valuations do not trade. In a different study, Boehme, Danielsen, and Sorescu (2006) measure differences in opinion using trading volume and find stocks with short-sale constraint and high turnover have low future returns, which is also consistent with Miller's argument.

Although also finding a negative relation between analysts' forecast dispersion and future stock returns, some studies provide alternative explanations. First, Johnson (2004) argues that dispersion levels reflect risk. He derives a pricing model in which nonsystematic risk (idiosyncratic uncertainty) increases the option value of the firm, which in turn explains the negative relation between dispersion levels and future stock returns. Therefore, while Diether et al. (2002) and Johnson (2004) both provide arguments that investors pay a premium for stocks with high dispersion in analysts' forecasts, overpricing exists for different reasons. The former claims that this overpricing is due to limited market participation, while the latter argues that it results from uncertainty that increases the option value of the firm. Barron, Stanford and Yu (2009) support Johnson's (2004) asset pricing model.

The other string of research suggests that higher dispersion serves as proxy of risk, hence lower stock price and higher future returns (e.g., Williams, 1977, Mayshar, 1983, and Varian, 1985). Other studies find positive relation between dispersion in analysts' forecasts and future stock returns include Cragg and Malkiel (1968), Friend, Westerfield, and Granito

(1978), Harris (1986), Anderson, Ghysels, and Juergens (2005), Qu, Starks, and Yan (2003) and Doukas, Kim and Pantzalis (2006). Some of the earlier studies, however, are constrained by small samples. Doukas et al. (2006) use a diversity measure and find a positive relation between diversity and future stock returns, rejecting the view that dispersion reflects divergence of opinion. In sum, the aforementioned results on the relation between stock returns and divergence opinion/analysts' forecast dispersion still represent a puzzle in the literature.

2.2Research questions and hypotheses

Our argument builds on Miller's limited market participation hypothesis; but we contend that the magnitude of analysts' forecast dispersion and the firm size will not hold a monotonic relation with future stock returns. That is, we argue that the relation between dispersion in analysts' forecasts and stock overvaluation varies over the size of firm and the magnitude of dispersion.

We provide the following arguments for the conditional limited market participation hypothesis. First, it is more difficult to short small firms' stock because these stocks are often held by individual investors. On the other hand, large firms' stocks are held by institutions and it is much easier to borrow such stock for short-selling purpose. Market is less efficient in the absence of short sellers is implied in Boehmer et al. (2008), Diether et al. (2009) and Wu (2007). Wu (2007) finds that stock price is closer to its fundamental value when it has larger short volume. Boehmer et al. (2008) report short sellers well-informed and contribute to price efficiency. Diether et al. (2002) suggest that stocks of small firms are the hardest to short and least likely to have traded options. Diether et al. (2009) also find short sellers help correct mispricing by increasing short selling. Therefore, firm size moderates the relation between dispersion and future stock returns. Accordingly, we expect to observe the negative association between analysts' forecast dispersion and subsequent stock return before the size of firm reaches its threshold. Based upon this reasoning, we formulated our first conditional limited market participation hypothesis as follows:

> *Hypothesis 1*: The degree of the negative association between analysts' forecast dispersion and subsequent stock return diminishes when the size of firm rises above a threshold.

Second, the magnitude of dispersion matters. While pessimistic investors avoid stocks with high dispersion in analysts' forecasts due to the risks or difficulties of short selling, we argue that this sidelined-investors effect weakens as the magnitude of dispersion increases. We posit that since the equilibrium prices of stocks lie somewhere between the expected prices of the pessimistic and optimistic investors, when analysts' forecast dispersion is lower than a threshold, the pessimistic investors' expected stock prices deviate from the equilibrium prices with only a small margin. This small margin may not be sufficient to compensate for the cost and risk of short-selling. As a consequence, the pessimistic investors will be sidelined and the stock price will be set by the optimistic investors. Under this scenario, the limited participation of the pessimistic investors result in a higher stock price, hence lower future returns.

Figure 1 illustrates this scenario. In Figure 1, the equilibrium price is assumed to be $85; the most pessimistic expectation of the stock price is $83. Since the expected profit margin of $2 is not attractive to cover the $4 cost and risk premium of short-selling, pessimistic investors will be sidelined and the stock price will be set by the optimistic investors at $87. The stock is thus over-valued by $2. However, when the deviation between the equilibrium prices and investors' expected prices becomes sufficiently large (or, when the dispersion exceeds a threshold) to compensate for transaction costs and risk, both pessimistic and optimistic investors enter the markets, hence the limited participation hypothesis does not hold. Under this condition, we expect to find no relation between analysts' forecast dispersion and future stock returns. Figure 2 illustrates this scenario. In Figure 2, the equilibrium price is assumed to be $85; the most pessimistic expectation of the stock price is $80. Since the expected profit margin of $5 is larger than the $4 cost and risk premium of short-selling, both pessimistic and optimistic investors will enter the market, hence the limited participation hypothesis does not hold. Therefore, we posit that limited market participation is conditioned on the degree of dispersions.

Based on the above reasoning, we formulate the second conditional limited market participation hypothesis in its alternate form as follows:

Hypothesis 2: The degree of the negative association between analysts' forecast dispersion and subsequent stock return diminishes when the magnitude of dispersion rises above a threshold.

In the above discussions, we only consider the moderating effect of one threshold variable at a time. Under a two threshold variable scenario, the negative association between analysts' forecast dispersion and subsequent stock return will be most remarkable before both the magnitude of dispersion and the size of firm reach their respective thresholds. We thus formulated our third conditional limited market participation hypothesis as follows:

Hypothesis 3: The degree of the negative association between analysts' forecast dispersion and subsequent stock return is most pronounced before dispersion and firm size exceeding their respective thresholds.

To investigate this conditional relationship between analysts' forecast dispersion and subsequent stock return, traditionally a two-step estimation procedure may be used. The first step is to apply some subjective criteria to segment sample into various subsets. In the second step, a traditional optimization procedure, such as OLS, is used to fit the data and conduct comparative analyses between the partitioned segments. However, this two-step analysis implicitly assumes that the partitioning process is exogenous. To the extent that the link between analysts' forecast dispersion and subsequent stock return is conditional on the magnitude of dispersion and the size of firm, the sample partitioning and the degree of the association between analysts' forecast dispersion and subsequent stock return should be treated endogenously and analyzed jointly.

Based on the above discussions, therefore, neither the conventional regression approach nor the two-step estimation procedure mentioned above is appropriate, while the panel threshold regression approach is suitable for our purpose. In the next section we will discuss the properties of the non-threshold and panel threshold regression models, and demonstrate that the panel threshold regression approach is suitable for our study.

3. Regression models

3.1 The non-threshold model

Let (y_{it}, x_{it}), i = 1, 2…, N and t = 1, 2…, T, be a sample population, where subscript i denotes the ith firm and t denotes the tth period. The dependent variable, y_{it}, is a firm's stock return, and x_{it} is a k×1 vector of explanatory variables for y_{it}. When the data have a panel structure, the following equation represents a fixed-effect model:

$$y_{it} = \alpha_i + \beta' x_{it} + u_{ti}, \qquad (1)$$

where α_i (i=1, 2…, N) and β' (kx1 vector) are unknown parameters to be estimated. By considering the $\acute{a}_i$ as coefficients of dummy variables, the estimator vector of β can be obtained using the conventional least squares method.[1]

As part of this study's focus is on the non-monotonic relationship between y_{it} and x_{it} (i.e. the β parameters), we redefine Equation (1) as follows:

$$y_{it}^{*} = \beta' \cdot x_{it}^{*}{}' + u_{it}^{*}, \tag{2}$$

where * denotes variables deviated from the group mean, that is, $y_{it}^{*} = y_{it} - \bar{y}_i$, $x_{it}^{*} = x_{it} - \bar{x}_i$, and $u_{it}^{*} = u_{it} - \bar{u}_i$, and $\bar{y}_i$, $\bar{x}_i$ and $\bar{u}_i$ are the means of y, x and u of firm i, respectively.

The application of the non-threshold model in Equation (2) is potentially limited due to the use of a constant loading in each identified determinant of the explanatory variable. Specifically, once the final result is derived from Equation (2), the values of all the elements in the k×1 vector, β, are fixed across all firms.

3.2 Panel data model with a single threshold variable and two regimes

As in Hansen (1999), this study first establishes the following models in which the relationship between the independent variable (e.g., analysts' forecast dispersion) and the dependent variable (subsequent stock return) is driven by one single threshold variable:

$$y_{it}^{*} = \beta^{I}{}' x_{it}^{*} I(q_{it} \le \gamma) + \beta^{II}{}' x_{it}^{*} I(q_{it} > \gamma) + u_{it}^{*}, \tag{3}$$

where q_{it} is defined as a threshold variable, γ denotes the threshold parameter, and $I(\cdot)$ is the indicator function. Two regimes are defined in the above setting: regime I is set where $q_{it} <= \gamma$, while regime II is defined where $q_{it} > \gamma$. Apparently, the non-monotonic x-to-y relationship (i.e., β^{I} vs. β^{II}) depends on where the threshold variable, q_{it} is located. Consequently, Equation (2) (i.e., the non-threshold model) represents a special case of Equation (3) (i.e., the panel threshold model) with the restriction of $\beta^{I} = \beta^{II}$.

[1] The estimation procedure for the fixed effects model is well documented in the literature, and therefore, this study omits any discussion of this.

3.3 Panel data model with two threshold variables and four regimes

As an extension of Hansen's (1999) model, we use a four-regime panel threshold regression model where the regimes are determined by two different threshold variables (i.e., q^1_{it} and q^2_{it}). The model takes the form of:

$$\begin{aligned} y^*_{it} &= \beta^{I}{}' x^*_{it} I(q^1_{it} \leq \gamma_1) I(q^2_{it} \leq \gamma_2) + \beta^{II}{}' x^*_{it} I(q^1_{it} > \gamma_1) I(q^2_{it} \leq \gamma_2) \\ &+ \beta^{III}{}' x^*_{it} I(q^1_{it} \leq \gamma_1) I(q^2_{it} > \gamma_2) + \beta^{IV}{}' x^*_{it} I(q^1_{it} > \gamma_1) I(q^2_{it} > \gamma_2) + u^*_{it}, \end{aligned} \tag{4}$$

With this model we can study the impact of the independent variable (i.e., firm size and analysts' forecast dispersion) on the dependent variable (subsequent stock return) over different phases. As before, each regime is characterized by different slopes (β^i, i= I, II, III and IV).

Our panel model with two threshold variables extends the conventional single threshold Hansen's (1999) panel model. While based on a similar framework consisting of state variables, Hansen (1999) proposes a dual-state relationship between the independent and dependent variables with one threshold variable. By contrast, a four-state system triggered by two threshold variables is addressed in our model. Hansen's (1999) model could thus be considered as a special case of our generalized model with some restrictions. Therefore, Equation (3) is a special case of Equation (4) with the restrictions of $\beta^I=\beta^{II}$ and $\beta^{III}=\beta^{IV}$.

The panel threshold approach has recently been used in many areas of applied economics and econometrics, such as the impact of inflation on output (Lopez-Villavicencio and Mignon, 2011; Omay and Kan, 2010; Bick, 2010), the effects of exchange rate volatility on exports (Hsu and Chiang, 2011), the demand of cigarettes (Huang and Yang, 2006), the tests for nonstationary dynamic panel models (Shin and Jhee, 2006), and the stochastic frontier models of dairy production (Yelou, Larue and Tran, 2010). There is also growing interest in employing panel threshold regression approach in finance and accounting research.

Applications in this field include studies on investment intensity of currencies (Chuluun, Eun and Kilic, 2011), the impact of business cycle on bank credit risk (Marcucci and Quagliariello, 2009), and cost efficiency and performances in banks (Shen, 2005). This study serves as the first attempt to investigate the non-uniform degrees of association between analysts' forecast dispersion and subsequent stock return using the panel threshold regression approach with two conditioned variables.

3.4 Estimation procedure

One of the difficulties in operating the threshold models is estimation of the threshold parameter (i.e., γ in Equation (3) or γ_1 and γ_2 in Equation (4)). This study follows Balke and Formby (1997) in designing grid-search procedures for estimating the threshold parameter. The procedures are presented as follows. (1) The threshold variable (i.e., q_{it}) is defined and obtained. (2) The series of arranged q_{it}, y^*_{it} and x^*_{it} variables in Equation (3) are established and q_{it}, y^*_{it} and x^*_{it} are ordered according to the value of q_{it}, rather than time and firm. (3) By assigning a small number to serve as the initial value of γ, e.g., 0.005, the series of arranged y^*_{it} and x^*_{it} can be split into two different regime areas: regime I against regime II. (4) The regressions of Equation (3) are estimated for each regime and the residual sum of square (RSS hereafter) is calculated and saved. (5) The value of γ is increased using one grid with very small value of 0.0001, and the above fourth step is then repeated for the new values of γ. (6) Steps 4 and 5 are then repeated and the RSS value is derived for each value of γ; γ with the minimum RSS is then chosen.

It should be noted that to avoid the problem of small samples for any regimes, it is necessary to restrict the value of the threshold parameters. This paper uses the values of 10% and 90% percentiles of the threshold variable q_{it} as the boundary values of the threshold parameter γ. That is, the percentage of observations for each regime is allowed to vary for a range from 10% to 90%.

Once the estimates of the regime-varying parameters (i.e., β^I and β^{II} in Equation (3) and β^i, i= I, II, III and IV in Equation (4)) are obtained, the next step is to examine whether the regime-varying effects are statistically significant. Since Equation (2) represents a special case of Equation (3) with the restrictions of $\beta^I=\beta^{II}$, we test the null hypothesis of no threshold effect as follows:

$$H_0: \quad \beta^I = \beta^{II} \tag{A1}$$

To test this null hypothesis, Equation (2) (i.e., the setting with the restriction of $\beta^I=\beta^{II}$) is first estimated; we then obtain the restricted residual sum of squares (i.e., RRSS). Equation (3) (i.e., the setting without the restriction of $\beta^I=\beta^{II}$) is next estimated assuming the existence of state-varying parameters, which allows for the subsequent derivation of the unrestricted residual sum of squares (i.e., URSS). Finally, the two RSS values are used to carry out a likelihood ratio test (LR), which is based on

$$\lambda = (\frac{RRSS}{URSS})^{-n/2}, \text{ n = sample size,}$$

and -2 $\log_e \lambda$ (i.e., the LR statistic) has a χ^2-distribution with k degree of freedom (number of restrictions).

4. Sample, variables, and empirical model

4.1 Sample and measures of the dependent and independent variables

We obtain analysts' earnings forecast and actual earnings per share data for the period 1983-2009 from Institutional Brokers Estimate Service (I/B/E/S) detailed database. The overall sample consists of 4,122 unique firms and 380,073 firm-month observations. The corresponding monthly stock returns are obtained from the Center for Research in Securities Prices (CRSP) US Stock and Indices database.

Following earlier studies (e.g., Payne and Thomas, 2003; Barron, et al. 2009 and others), we calculate the analysts' earnings forecast dispersion (*DISPERSION*) as the standard deviation of analysts' earnings forecast scaled by the absolute value of the mean analysts' earnings forecast. The corresponding monthly stock returns are obtained from the Center for Research in Securities Prices (CRSP) US Stock and Indices database. It should be noted that since our focus is on the relationship between analysts' earnings forecast dispersion and subsequent stock returns, the dependent variable in our regression is the monthly stock returns following the month in which forecast dispersion is measured.

4.2 Empirical model

Our empirical analyses are based on regression in which the independent variable is analysts' forecast dispersion and the dependent variable is subsequent stock return. In order to mitigate the omitted-variables problem, we include several potential determining factors for stock returns in the regression. To be sure, we include the size factor, the value factor and and the leverage[2]. The fixed-effects model is as follows:

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, SIZE_{i,t} + \beta_3\, LEV_{\,i,t} + \beta_4 B/M_{i,t} + u_{i,t}, \quad (5)$$

where *RETURN* equals the monthly stock returns of firm following the month that *DISPERSION* is measured; *DISPERSION* equals the analysts' forecast dispersion; *SIZE* is the natural log value of total assets, *LEV* is total liabilities divided by total assets and *B/M* equals book value of equity divided by market value of equity. Table 1 presents detailed variable definitions, descriptive statistics, and the correlation matrix of these variables. The average monthly stock returns is 1.14%; average analysts' forecast dispersion is 0.1805;

[2] This study obtains firm's total assets value, M/B ratio, debt ratio and total assets from Standard & Poor's Compustat database.

average debt/asset ratio is 41.45%; average *B/M* value is 2.2074 and the average firm size in logarithm is 7.5079 million. The highest correlation coefficient of 11.76% is observed between *B/M* and *SIZE*,

5. Empirical results

Three different panel models for the dependent variable (i.e., *RETURN*) are addressed by this study. We first apply the panel non-threshold technique to estimate Equation (5), and label it as '*Model 1*'. Second, we apply the panel data model with a single threshold variable on Equation (5). We first define the threshold variable as the size of firm (*SIZE*), and label this setting as '*Model 2*'. Next, we define the threshold variable as the magnitude of analysts' forecast dispersion (*DISPERSION*), and label this setting as '*Model 3*'. Finally, we apply a four-regime panel threshold regression model where the regimes are controlled by both the magnitude of analysts' forecast dispersion and the size of firm, and label this setting as '*Model 4*'.

5.1 The association between analysts' forecast dispersion and subsequent returns

In this section we employ the conventional ANOVA (Analysis of Variance) method to examine the association between analysts' forecast dispersion and subsequent returns. We first sort stocks into five quintiles based on dispersion in analyst earnings forecasts as of the month t-1 (i.e., D1, D2, D3, D4 and D5). We then calculate the monthly portfolio return as the equal-weighted average of the returns of all stocks in the portfolio for month t. Finally, we examine return of a long-short portfolio that longs D1 and shorts D5 (i.e., D1-D5). It should be noted that to control for the firm size effect, we also sort stocks based on firm size (i.e., S1, S2, S3, S4 and S5). The results are reported in Table 2.

As shown in the last column of Table 2, the annualized return on the D1-D5 strategy is 3.4008 percent (0.2834x12) and is significant at the 1% level. This result implies a negative relation between average returns and dispersion in analysts' earnings forecasts. More interestingly, the average monthly return for the long-short portfolios (i.e., D1-D5) is positive and highly significant for the two smallest stock portfolios: S1 and S2, but it largely declines and becomes insignificant for the three largest stock portfolios: S3, S4 and S5. This result implies that the negative association between dispersion in analysts' earnings forecasts and subsequent return is most pronounced in small stocks. For example, for S1, D1-D5 strategy yields an annualized return of 9.36% (0.7803 x 12); but for S5, the same statistic is negative.

It should be noted that the conclusions obtained from Table 2 are highly consistent with Diether et al. (2002). However, while the ANOVA results shown in Table 2 and Diether et al. (2002) show the existence of the association between analysts' dispersion and subsequent returns, it does not test the *degree* of association. We address this issue in the following sections.

5.2 The degree of association between stock returns and analysts' dispersion

In Table 3, we show the value of dispersion in analysts' earnings forecasts at 19 distinct percentiles: 0.05, 0.10, 0.15,…, 0.95 (the second column) and its corresponding mean stock return (the third column). For example, for the 0.05 percentile of analyst forecasts' dispersion, the corresponding mean return is the average monthly returns for stocks with dispersion under the 0.05 percentile. As for the 0.10 percentile of analysts' forecast dispersion, we calculate the average returns of stocks where their dispersion lies between the 0.05 and 0.10 percentiles. Next, we use the 0.05 percentile as the benchmark to calculate the degree of association between analysts' dispersion and subsequent return as shown in the fourth column. Degree of association is a measure of changes in mean stock return that results from changes in dispersion using 0.05 percentile as the benchmark. For example, for the 0.10 percentile of analysts' forecast dispersion, the degree of association is calculated as the ratio of (1.1283-1.2871)/ (0.0094-0.0061) = -32.0106.

Using the 0.05 percentile of dispersion as a starting point, the mean stock return with above-0.05 percentile of dispersion is lower than the mean stock return with the lowest dispersion (i.e., the 0.05 percentile). Accordingly, the values of the degree of the association between analysts' dispersion and subsequent return are negative in all cases. However, the absolute magnitude of the degree of association largely becomes small after the 0.20 percentile (see the fourth column of Table 3).

Based on Table 3, we chart a figure in which x-axis shows the dispersion value of the 0.05-to-0.95 percentile of dispersion, while the corresponding degrees of negative association between analysts' dispersion are on the y-axis (see Figure 3). Apparently, we find a strong evidence of structural change on the pattern of Figure 3: the degree of association between analysts' forecast dispersion and subsequent return (i.e., the value of the y-axis) drops off the cliff and remains at very low level when the value of the x-axis (i.e., the value of analysts' dispersion) exceeds a certain threshold.

5.3 The panel non-threshold model

Given the results shown in the previous two sections, we re-examine the relation between stock return and dispersion using panel non-threshold model and the results are presented in Table 4 (*Model 1*). We first apply OLS regression to the panel data and confirm Diether et al. (2002) that *DISPERSION* is negative and significantly (coefficient = -0.05, p-value < 0.01) associated with expected stock returns as shown in Panel A of Table 4. However, since the OLS estimation results might be biased for the data with panel structure (see Petersen, 2009), we consider the fixed effect model (see Equation (1)), and employ the "*within-group difference*" for both the explained and explanatory variables to estimate the model (see Equation (2)). The results of fixed effect model are presented in Table 4 Panel B. First, the coefficients on the two control variables, *LEV* and *SIZE*, are significant at the 1% level of significance, while B/M is not (p-value=0.1529). Second, we find that although the coefficient of *DISPERSION* is negative (-0.0161), it is not statistically significant. Therefore, the results of Diether et al. (2002) may not be robust to the inclusion of firm fixed effect.

It should be noted that the estimation results of the OLS and fixed effects model focus only on the mean behavior of the data, hence cannot capture our regime-switching argument of conditional limited market participation hypothesis. To address this issue, we examine the conditional relationships between *DISPERSION* and *RETURN* using panel threshold model in the following sections.

5.4 The panel threshold model with one threshold variable and two regimes

Table 5 reports the results of *Model 2* in which the size of firm (*SIZE*) is defined as the threshold. First, we find slightly lower *RSS* (residual sum of squares) in *Model 2*. Second, we test whether the threshold effect is statistically significant. We calculate and report the *LR* (likelihood ratio) statistics in the last line of Table 5. It should be noted that comparing *Model 1* with *Model 2*, the former specification serves as a special case for the latter under the constraint of $\beta^{I}=\beta^{II}$. The *LR* statistic for the null hypothesis of no regime switching (i.e., $\beta^{I}=\beta^{II}$) is rejected at the 1% significance level. This result implies that the *DISPERSION-to-RETURN* relationship exhibits a significant threshold effect, that is, a "regime-switching" effect exists.

Last but not least, our results show that *DISPERSION* is significantly and negatively associated with *RETURN* at the 1.3% significance level when *SIZE* is less than the threshold (6.8667). On the other hand, for firms with *SIZE* exceeding 6.8667, *DISPERSION* is insignificant. These results support our first hypothesis that limited market participation is conditioned on the size of firm, and the degree of negative association between analysts' forecast dispersion and subsequent stock returns diminishes when the size of firm rises above a threshold. The regime-switching results shown in Table 5 are also consistent with the results of the ANOVA analysis shown in Table 2.

Table 6 lists the results of *Model 3* in which the threshold variable is the magnitude of analysts' dispersion (*DISPERSION*). To test our second hypothesis, we first calculate and report the *LR* statistic (=28.3602) in Table 6. The result shows that the null of a no regime switch is rejected at the 1% significance level, indicating that the association between analysts' dispersion and future stock return exhibits a significant threshold effect.

Next, our results show that *DISPERSION* is significantly and negatively associated with *RETURN* at the 1% significance level when *DISPERSION* is less than the threshold (0.0161). Moreover, the absolute value of the estimated coefficient of *DISPERSION* in regime I (-0.9584) is seventy-two times larger than that in regime II (-0.0132). These results support our second hypothesis that limited market participation is conditioned on the degree of analysts' forecast dispersion, and the degree of negative association between analysts' forecast dispersion and subsequent stock returns diminishes when the magnitude of analysts' forecast dispersion rises above a threshold. It should be noted that the regime-switching results shown in Table 6 are consistent with the non-uniform degrees of association between return and dispersion graphed in Figure 3.

5.5 The panel threshold model with two threshold variables and four regimes

To test hypothesis 3, we estimate *Model 4* and report the results in Table 7. First, *Model 4* (the four-regime panel threshold setting) has lower *RSS* in comparison with *Model 1, Model 2 and Model 3*. Second, comparing *Model 4* with *Model 1*, the null of no regime switching (i.e., β^{I}=β^{II} =β^{III}=β^{IV}) is rejected at conventional significance level. Moreover, a dual-state setting (i.e., β^{I}=β^{II} and β^{III}=β^{IV}) is also rejected at conventional significance level according to the *LR* statistics when we compare *Model 4* against *Models 2 and 3*. According to *RSS* and *LR* statistics, the four-regime panel threshold model (*Model 4*) appears to be the most informative one. This finding is unambiguous. We thus focus our discussions on the estimates produced by this specification.

First, for regime I (i.e., *DISPERSION* <= 0.0153 *and SIZE* <= 6.8852), the estimated coefficient on *DISPERSION* (-1.3078) is negative and significant at the 1% level (p-value=0.0029), suggesting that when both the *DISPERSION* and *SIZE* are below the threshold, limited market participation is the most evident. For regime II (*DISPERSION* > 0.0153 and *SIZE* ≤ 6.8852), the coefficient is negative (-0.0566) and significant at the 5% level (p-value=0.0245). This result suggests that *DISPERSION* effect still exists in small firms even when *DISPERSSION* exceeds its threshold value. For regime III (*DISPERSION* ≤ 0.0153 and *SIZE* > 6.8852), the coefficient is negative (-0.5153), but insignificant at the conventional level (p-value=0.0925, weakly significant at the 10% level). For regime IV (i.e., *DISPERSION* > 0.0153 and *SIZE* > 6.8852), the estimated coefficient on *DISPERSION* is positive (0.0111) and statistically insignificant (p-value=0.4551).

Our conclusion is clear. The estimated coefficient on *DISPERSION* is significant in Regimes I and II. It should be noted that a general impression of the four-state under the condition of small size firm can be obtained from the integration of Regimes I and II. Moreover, the present results, as shown in Table 7, indicate that the absolute value of the estimated coefficient on *DISPERSION* (-1.3078) under Regime I is much higher than that under Regime II (-0.0566). These findings suggest that although there is a threshold effect for *SIZE*, the dispersion effect is manifested by the magnitude of dispersion. Our results are thus consistent with the argument of conditional limited market participation.

Moreover, the absolute value of the estimated coefficient of *DISPERSION* in Regime I (-1.3078) of Table 7 is higher than that in Regime I of Table 5 (-0.0616) and Table 6 (-0.9584). Obviously the degree of the negative association between analysts' forecast dispersion and subsequent stock return is most pronounced in the regime that both the magnitude of dispersion and the firm size exceed their respective thresholds.

5.6 Model specification with additional control variables

In order to assess whether the estimate parameters of our panel threshold model might be biased due to an omitted-variables problem, we now report results derived by including two more factors in our regression procedures: the momentum factor (i.e., the lagged stock returns of the firm) and the market factor (i.e., the monthly market returns measured by the CRSP value-weighted market indices), and the results are listed in Table 8[3].

As shown in Table 8, the estimated coefficient of *DISPERSION* in Regime I (-1.1626) is negative and significant at the 1% significance level, and its absolute value is higher than those in Regimes II, III, and IV of Table 8 (-0.0303, -0.3781 and 0.0023). Comparing with Table 7, we find that the pattern of regime-varying estimates for the *DISPERSION* variable is robust with respect to the inclusion of two additional control variables. In other words, while adding the two additional control variables in the model specification slightly changes the parameter estimates, it does not affect the structure of the regime-varying link between *DISPERSION* and *RETRUN* presented in Table 7.

5.7 Alternative model specifications

The panel threshold regression approaches established in this study are similar to the piecewise linear regressions which have been widely used in many areas of finance and accounting (e.g., Morck, Shleifer and Vishny, 1988). It should be noted that while both of

3 See Payne and Thomas (2003); Barron, et al. (2009) and others.

them are used to capture the non-linear associations with break-points, the former has less restrictions in that it effectively mitigates the limitations of using arbitrary and subjective criteria by jointly and endogenously analyzing the sample segmentation.

Another alternative method to investigate the nonlinear association between the dependent and independent variables is the quadratic specification which have been also extensively employed in finance and accounting fields (e.g., McConnel and Servaes, 1990). For comparative purpose, we estimate a quadratic equation and report the following results (with the p-value in the parenthesis):

$$\begin{aligned} RETURN_{i,t+1} = {} & \underset{(0.000)}{0.5475} - \underset{(0.000)}{0.0867}\, DISPERSION_{i,t} + \underset{(0.005)}{0.0005}\, DISPERSION^2_{i,t} + \underset{(0.234)}{0.0102}\, SIZE_{i,t} \\ & + \underset{(0.000)}{0.2651}\, LEV_{i,t} - \underset{(0.129)}{0.0014}\, B/M_{i,t} + \underset{(0.000)}{0.6464}\, MKT_{i,t} + \underset{(0.000)}{0.0056}\, MOM_{i,t} + u_{i,t} \qquad (6) \end{aligned}$$

In the above equation, future stock return (*RETURN*) is regressed against dispersion in analysts' forecast (*DISPERSION*) and dispersion squared ($DISPERSION^2$). Both the estimated coefficients of *DISPERSION* and $DISPERSION^2$ are significant, with the former carrying a negative sign while the latter a positive sign. This result implies a curvilinear relation between dispersion in analysts' forecast and subsequent stock return. Based on Figures 3, however, as analysts' dispersion increases, the degree of the association between dispersion in analysts' forecast and subsequent stock return declines and levels off after the 20^{th} percentile of the dispersion; hence the association does not revert. Therefore, the quadratic specification might have overestimated the reversal of the association between dispersion in analysts' forecast and subsequent stock return. By contrast, the panel threshold regression approach employed in this study essentially mitigates this limitation.

The panel threshold regression approach used in this study has several advantages over the quadratic equation used by earlier works. First, in comparison with the quadratic equation, the panel threshold regression explicitly identifies the break-points of the associations, and clearly shows the observations that belong to each regime. Second, the four-regime panel threshold regression approach (i.e., *Model 4*) extends the one-dimension framework of the quadratic specification to a multivariate system. In particular, in the quadratic specification, the u-shaped relationship between the dependent and independent variables depends only on the value of the latter. Our four-regime design is controlled by two variables (i.e., q^{1}_{it} and q^{2}_{it}), and thus effectively extends the one-dimension design of the quadratic equation.

5.8 Implications of the empirical results

We can make several observations from our findings. First, while we find strong evidence of a negative association between analysts' dispersion and subsequent return as demonstrated by Diether et al. (2002), we extend and modify their study by investigating the strength of this negative association, and hypothesize that the strength declines or diminishes as the magnitude of dispersion exceeds a threshold value. Our hypothesis is based upon the argument that sidelined investors effect weakens as the magnitude of dispersion increases because high dispersions provide discrepancy between expected price and equilibrium price that is sufficiently large to compensate for the transaction cost and risk incurred by the pessimistic short sellers. Our conditional limited market participation hypotheses are strongly supported by the empirical results shown in Tables 2-8 and Figure 3.

Second, our research designs and empirical results provide a further examination on the disagreement between studies such as Diether, et al. (2002) and Johnson (2004). To be sure, while these two studies hypothesize a negative association between dispersion in analysts' forecast and subsequent return, they attribute this negative association for two different reasons. Johnson (2004) derives an option pricing model in which nonsystematic risk (idiosyncratic uncertainty) increases the option value of the firm. To the extent that dispersion reflects idiosyncratic uncertainty of a firm, there exists a negative relationship between dispersion levels and future stock returns.

Importantly, Johnson's (2004) point is based on the well-known option value-volatility relationship: Higher volatility reflect greater fluctuations in expected underlying price levels, and this expectation results in higher option value. However, since both the first and second derivatives of the option value with respect to the volatility are normally positive[4], the degree of the association between dispersion in analysts' forecast and subsequent return should be stronger as the magnitude of dispersion increases if Johnson's (2004) point is held. Our evidence casts doubt on Johnson's (2004) argument that option value-uncertainty relation explains the negative relation between dispersion and future returns. To be sure, we show that the degree of the association becomes considerably diminished for stocks with their analysts' forecast dispersion exceeding a threshold value.

6. Summary and conclusions

Whether firms with high analysts' forecast dispersion underperform firms with low dispersion has been researched by a number of studies. These studies produce inconclusive evidence, hence interpretations differ. Our study contributes to this line of research by (1) proposing a conditional limited market participation hypothesis, and (2) using a less restrictive methodology to retest sidelined-investors effect or limited market participation hypothesis proposed by earlier studies.

Stocks of small firms are more difficult to short; hence such effect is linked to the limited market participation hypothesis. Central to our empirical design, we further argue that the sidelined-investors effect weakens as the magnitude of dispersion increases. This is

---

[4] The second derivative of the option value with respect to the volatility is also named as Vomma. Vomma is positive for deep-in-the-money and deep-out-the-money options. We show the detail of the formulation of Vomma in the Appendix.

because when analysts' forecast dispersion is lower than a threshold, the pessimistic investors' expected stock prices deviate from the equilibrium prices with only a small margin, which generates an expected return may not be sufficient to cover the cost and risk of short-selling. As a consequence, the pessimistic investors will be sidelined and the stock price will be set by the optimistic investors. The limited participation of the pessimistic investors results in a higher stock price, hence lower future returns. However, when the dispersion exceeds a threshold, the deviation between the equilibrium prices and the expected prices becomes sufficiently large to compensate for transaction costs and risk; this will induce both the pessimistic and optimistic investors to enter the markets, hence the limited participation hypothesis does not necessary hold. Under this condition, we expect to find no relation between analysts' forecast dispersion and future stock returns.

To this end, we investigate the non-monotonic degrees of association between analysts' forecast dispersion and subsequent stock return using the panel threshold regression, which allows the parameter estimates to vary over the values of the two conditioned variables: dispersion and size of firm. Unlike the traditional non-threshold regressions which rely on the mean relation between the independent and dependent variables, the panel threshold regression is capable of examining the degree of the association at the tails of observations. Moreover, since the sample segmentation and the degree of association between the independent and dependent variables are analyzed jointly and endogenously, the panel threshold regression effectively mitigates the limitations of using arbitrary and subjective criteria to partition sample firms into various subsets in the conventional two-step analysis.

We use the standard deviation of analysts' earnings forecast scaled by the absolute value of the mean analysts' earnings forecast as a proxy for analysts' dispersion. We then calculate monthly stock returns following the month in which forecast dispersion is measured. Our sample consists of 4,122 unique firms and 384,401 firm-month observations during the period from 1983 to 2009.

The panel threshold regression results show that the degree of the association between analysts' dispersion and subsequent stock return depends on whether the two conditioned variables (i.e., dispersion and size of firm) exceed their respective thresholds. In particular, our results show that analysts' dispersion is significantly and negatively associated with subsequent stock return at firms where the magnitude of dispersion and the size of firm are under their respective thresholds. The degree of this association diminishes when dispersion and/or size of firm exceed their respective threshold.

Therefore, the evidence in our study suggests that sidelined-investors effects (i.e., limited market participation) mainly inhabit in firms with characteristics of lower dispersion and smaller size. In contrast, for larger firms and/or firms with larger analysts' dispersion, our results show that the argument of limited market participation no longer holds, as the degree of the negative association between analysts' dispersion and subsequent stock return become trivial. Last, but not the least, our research designs and results shed light on the two dominant interpretations of the relation between dispersion in analysts' earnings forecasts and stock returns. Our evidence is inconsistent with the argument that dispersion in analysts' forecasts serves as a proxy for risk (e.g., Johnson, 2004). If the risk proxy argument is correct, we would see a stronger negative correlation between future stock returns and analysts' forecast dispersion as dispersion increases.

Appendix

The derivation of option Vomma

The value of a call option for a non-dividend paying underlying stock in terms of the Black–Scholes formula is:

$$C = SN(d_1) - Ke^{-r\tau}N(d_2)$$
$$d_1 = \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \tau = T - t \qquad \text{(A2)}$$
$$d_2 = d_1 - \sigma\sqrt{\tau}$$

where

S	=	The spot price of the underlying asset
K	=	The strike price
r	=	The risk free rate
σ	=	The volatility of returns of the underlying asset
τ	=	The time to maturity
N(.)	=	The cumulative distribution function of the standard normal distribution

Next, the first derivative of option value with respective to volatility (Vega hereafter) is:

$$\frac{\partial C}{\partial \sigma} = S\frac{\partial N(d_1)}{\partial d_1}\frac{\partial d_1}{\partial \sigma} - Ke^{-r\tau}\frac{\partial N(d_2)}{\partial d_2} * \frac{\partial d_2}{\partial \sigma} = S\sqrt{\tau}n(d_1) > 0\,. \qquad \text{(A3)}$$

The positive value of Vega implies that higher volatility reflects greater expected fluctuations (in either direction) in underlying price levels, and this expectation results in higher option value.

The second derivative of option value with respect to volatility (Vomma hereafter) is:

$$
\begin{aligned}
\frac{\partial^2 C}{\partial \sigma^2} &= \frac{\partial\left(\partial C / \partial \sigma\right)}{\partial \sigma} = \frac{\partial\left(S\sqrt{\tau}\, n(d_1)\right)}{\partial \sigma} \\
&= S\sqrt{\tau}\frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}\left(-d_1\right)\frac{\partial d_1}{\partial \sigma} \\
&= S\sqrt{\tau}\frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}\left(-d_1\right)\frac{-1}{\sigma} d_2 \\
&= S\sqrt{\tau}\frac{1}{\sqrt{2\pi}} e^{-\frac{d_1^2}{2}}\frac{d_1 d_2}{\sigma}
\end{aligned} \tag{A4}
$$

As shown in Equation (A4), Vomma is positive in most cases. In particular:

(1) As $S >> K$ (i.e., deep in the money), we have $d_1 > 0$ and $d_2 > 0$, thus $\frac{\partial^2 C}{\partial \sigma^2} > 0$.

(2) As $S << K$ (i.e., deep out of money), we have $d_1 < 0$ and $d_2 < 0$, thus $\frac{\partial^2 C}{\partial \sigma^2} > 0$

(3) As $S = K$ (i.e., at the money):

(a) If $r > \frac{\sigma^2}{2}$, we have $d_1 > 0$ and $d_2 > 0$, thus $\frac{\partial^2 C}{\partial \sigma^2} > 0$

(b) If $r < \frac{\sigma^2}{2}$, we have $d_1 > 0$ and $d_2 < 0$, thus $\frac{\partial^2 C}{\partial \sigma^2} < 0$

Therefore, (3b) is the only scenario that Vomma will be negative. However, during our sampling period from 1983 to 2009, the annualized standard deviation of the S&P500 index is 15.58%, implying a σ^2 of approximately 0.024. Therefore, $\sigma^2/2 = 0.012$ is far less than the average T-bill rate of approximately 4.73% during the period.

References

Ackert, L., and G. Athanassakos, 1997, Prior uncertainty, analyst bias, and subsequent abnormal returns, *Journal of Financial Research* 20, 263-273.

Anderson, E.W., E. Ghysels and J.L. Juergens, 2009, The impact of risk and uncertainty on expected returns, *Journal of Financial Economics* 94, 233-263.

Atiase, R.K. and L.M. Bamber, 1994, Trading volume reactions to annual accounting earnings announcements: The incremental role of predisclosure information asymmetry, *Journal of Accounting and Economics* 17, 309-329.

Balke, N.S. and T.B. Fomby, 1997, Threshold cointegration, *International Economic Review* 38, 627-645.

Barron, O.E. and P.S. Stuerke, 1998, Dispersion in analysts' earnings forecasts as a measure of uncertainty, *Journal of Accounting, Auditing and Finance* 13, 245-270.

Barron, O.E., M.H. Stanford and Y. Yu, 2009, Further Evidence on the relation between analysts' forecast dispersion and stock returns, *Contemporary Accounting Research* 26, 329-357.

Basu, S., 1977, The investment performance of common stocks in relation to their price-earnings ratios: a test of the efficient market hypothesis, *Journal of Finance* 32, 663-682.

Bick, A. 2010, Threshold effects of inflation on economic growth in developing countries, *Economics Letter*s 108, 126-129.

Boehme, R.D., B.R. Danielsen, and S.M. Sorescu, 2006, Short-sale constraints and overvaluation, *Journal of Financial and Quantitative Analysis* 41, 455-488.

Boehmer, E., Jones, C., Zhang, X., 2008. Which shorts are informed? *Journal of Finance* 63, 491-527.

Chan, L. K., Y. Hamao and J. Lakonishok, 1991, Fundamentals and stock returns in Japan, *Journal of Finance* 46, 1739-1765.

Chuluun, T., C.S. Eun and R. Kilic, 2011, Investment intensity of currencies and the random walk hypothesis: Cross-currency evidence, *Journal of Banking and Finance* 35, 372-387.

Cragg, L., and B. Malkiel, 1968, The consensus and accuracy of some predictions of the growth of corporate earnings, *Journal of Finance* 23, 67-84.

Diether, K.B., C.J. Malloy and A. Scherbina, 2002, Differences of opinion and the cross section of stock returns, *Journal of Finance* 57, 2113-2141.

Diether, K., Lee, K., Werner, I., 2009a. Short-sale strategies and return predictability, *Review of Financial Studies* 22, 575-607.

Doukas, J.A., C. Kim and C. Pantzalis, 2006, Divergence of opinion and equity returns, *Journal of Financial and Quantitative Anaylsis* 41,573-606.

Fama, E.F. and K.R. French, 1992, The cross-section of expected stock returns, *Journal of Finance* 47, 427-465.

Fama, E.F. and K.R. French, 1993, Common risk factors in the returns on stocks and bonds, J*ournal of Financial Economics* 33, 3-57

Friend, E., R. Westerfield and M. Granito, 1978, New evidence on the capital asset pricing model, *Journal of Finance* 33, 903-917.

Garfinkel, J.A., 2009, Measuring investors' opinion divergence, *Journal of Accounting Research* 47, 1317-1348.

Garfinkel, J.A. and J. Sokobin, 2006, Volume, opinion divergence, and returns: A study of post-earnings announcement drift, *Journal of Accounting Research* 44, 85-112.

Johnson, T.C., 2004, Forecast dispersion and the cross section of expected returns, *Journal of Finance* 59, 1957-1978.

Hansen, B.E., 1999, Threshold effects in non-dynamic panels: estimation, testing, and inference, *Journal of Econometrics* 93, 334-368.

Harris, R., 1986, Using analysts' growth forecasts to estimate shareholder required rates of return, *Financial Management* 15, 58-67.

Huang, B.N. and C.W. Yang, 2006, Demand for cigarettes revisited: an application of the threshold regression model, *Agricultural Economics* 34, 81-86.

Hsu, K.C. and H.C. Chiang, 2011, The threshold effects of exchange rate volatility on exports: Evidence from US bilateral exports, *Journal of International Trade & Economic Development* 20, 113-128.

Imhoff, E.A. and G.J. Lobo, 1992, The effect of ex ante earnings uncertainty on earnings response coefficients, *The Accounting Review* 67, 427-439.

Lopez-Villavicencio, A. and V. Mignon, 2011, On the impact of inflation on output growth: Does the level of inflation matter? *Journal of Macroeconomics* 33, 455-464.

Marcucci, J. and M. Quagliariello, 2009, Asymmetric effects of the business cycle on bank credit risk, *Journal of Banking and Finance* 33, 1624-1635.

Mayshar, J., 1983, On divergence of opinion and imperfections in capital markets, *American Economic Review* 73, 114-128.

McConnell, J.J. and H. Servaes, 1990, Additional evidence on equity ownership and corporate value, *Journal of Financial Economics* 27, 595-612.

Miller, E., 1977, Risk, Uncertainty, and Divergence of Opinion, *Journal of Finance* 32, 1151-1168.

Morck, R., A. Shleifer, and R.W. Vishny, 1988, Management ownership and market valuation : An empirical analysis, *Journal of Financial Economics* 20, 293-315.

Omay, T. and E.O. Kan, 2010, Re-examining the threshold effects in the inflation-growth nexus with cross-sectionally dependent non-linear panel: Evidence from six industrialized economies, *Economic Modelling* 27, 996-1005.

Payne, J., and W. Thomas, 2003, The implications of using stock-split adjusted I/B/E/S data in empirical research, *Accounting Review* 78, 1049-1067.

Petersen, M., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435-480.

Qu, S., L. Starks and H. Yan, 2003, Risk, dispersion of analyst forecasts and stock returns, *Working paper*, University of Texas, Austin.

Shen C.H., 2005, Cost efficiency and banking performances in a partial universal banking system: application of the panel smooth threshold model, *Applied Economics* 37, 993-1009.

Shin D.W. and W.C. Jhee, 2006, Tests for asymmetry in possibly nonstationary dynamic panel models, *Economics Letters* 91, 15-20.

Yelou, C., B. Larue and K.C. Tran, 2010, Threshold effects in panel data stochastic frontier models of dairy production in Canada, *Economic Modelling* 27, 641-647.

Varian, H.R., 1985, Divergence of opinion in complete markets: A note, *Journal of Finance* 40, 309-317.

Williams, J.T., 1977, Capital asset prices with heterogeneous beliefs, *Journal of Financial Economics* 5, 219-239.

Wu, J., 2007, Short selling and the informational efficiency of price. Working Paper, Texas A&M University.

Table 1
Descriptive statistics and correlation coefficients of variables

Panel A: Descriptive statistics of variables

Variable	Mean	Standard Dev.	Q1	Median	Q3
RETURN	1.1411	10.2215	-3.4161	0.0842	5.5814
DISPERSION	0.1805	1.3323	0.0184	0.0407	0.1011
LEV	0.4145	0.2288	0.2786	0.4260	0.5431
SIZE	7.5079	1.8564	6.2430	7.3749	8.6408
B/M	2.0274	16.6415	0.3066	0.5138	0.8304

Panel B: Correlation coefficients of variables

Variable	*RETURN*	*DISP.*	*LEV*	*SIZE*	*B/M*
RETURN	1.0000				
DISPERSION	-0.0064	1.0000			
LEV	0.0036	0.0251	1.0000		
SIZE	-0.0117	-0.0223	-0.0883	1.0000	
B/M	-0.0019	0.0220	-0.0197	0.1176	1.0000

Variable definitions:

RETURN	=	The monthly stock returns (%) on the individual firm stock following the month that *DISPERSION* is calculated
DISPERSION	=	The standard deviation of earnings forecasts scaled by the absolute value of the mean earnings forecast
SIZE	=	The natural logarithm of the total assets of the firm in $million
LEV	=	Total liabilities/total assets
B/M	=	Book value of equity divided by market capitalization

We obtain analysts' earnings forecast and actual earnings per share data for the period of 1983-2009 from Institutional Brokers Estimate Service (I/B/E/S) detailed database. The overall sample consists of 4,122 firms and 384,401 monthly observations.

Table 2
The association between analysts' dispersion and subsequent returns

Mean returns						
Dispersion	Size quintiles					All
quintiles	S1(Small)	S2	S3	S4	S5(Large)	stocks
D1(Low)	1.6112	1.3398	1.0552	1.1285	0.8650	1.1741
D2	1.4870	1.3362	1.1963	1.0581	1.0595	1.2256
D3	1.2720	1.4215	1.1138	1.0388	1.0444	1.1930
D4	1.1252	1.1947	1.2713	1.1729	1.0012	1.1657
D5(High)	0.8309	0.9534	0.9374	0.8579	0.8709	0.8907
D1–D5	0.7803**	0.3863**	0.1177	0.2706	-0.0059	0.2834**
t-statistics	(5.2397)	(2.7851)	(0.9996)	(2.3862)	(-0.0604)	(5.0735)

We sort each firms in five groups based on their size and dispersion in analyst earnings forecasts. D1 represents the lowest quintile dispersion, while D5 the highest. S1 represents the smallest quintile firms, while S5 the largest. Dispersion and firm size are defined as in Table 1. The table reports average monthly portfolio returns; t-statistics are in parentheses. The ** denotes significance at the 1% level.

Table 3
The degree of the negative association between stock return and analysts' dispersion

Percentile	Dispersion	Mean return	Degree of association
0.05	0.0061	1.2871	
0.10	0.0094	1.1283	-32.0106
0.15	0.0123	1.0916	-24.3356
0.20	0.0153	1.2166	-6.4437
0.25	0.0184	1.2327	-3.9040
0.30	0.0218	1.1978	-5.1941
0.35	0.0256	1.2293	-2.7826
0.40	0.0299	1.2595	-1.1149
0.45	0.0349	1.2061	-2.7552
0.50	0.0407	1.2562	-0.8875
0.55	0.0479	1.2245	-1.5175
0.60	0.0567	1.1360	-3.0666
0.65	0.0678	1.2462	-0.6909
0.70	0.0821	1.2292	-0.8082
0.75	0.1011	1.1155	-1.9471
0.80	0.1290	1.1045	-1.6400
0.85	0.1739	1.1173	-1.1600
0.90	0.2593	0.9550	-1.5932
0.95	0.5000	0.8493	-1.2435

This table shows the value of the dispersion in analyst earnings forecasts at 19 distinct percentiles: 0.05, 0.10, 0.15…, 0.95 (the second column) and its corresponding mean stock return (the third column). More specifically, for the 0.05 percentile of dispersion, the corresponding mean return is the average monthly returns of stocks with their dispersion under the 0.05 percentile. Moreover, for the 0.10 percentile dispersion, the mean return is the average returns of stocks with dispersion lying between the 0.05 and 0.10 percentiles. The 0.05 percentile is used as the benchmark to calculate the degree of the association between analysts' dispersion and subsequent return (the fourth column). For example, for the 0.10 percentile of analysts' forecast dispersion, the degree of association is calculated as the ratio of (1.1283-1.2871)/ (0.0094-0.0061) = -32.0106.

Table 4
Regression results of the panel non-threshold model

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, LEV_{i,t} + \beta_3\, SIZE_{\,i,t} + \beta_4\, B/M_{i,t} + u_{i,t}$$

Panel A: OLS Model

Variables	*Coefficient*	*Standard error*	*P-value*
Intercept	0.5387	0.0749	0.0000**
DISPERSION	-0.0500	0.0119	0.0000**
LEV	0.2612	0.0694	0.0000**
SIZE	0.0109	0.0086	0.2050
B/M	-0.0015	0.0010	0.1140
RSS (Residual sum of squares)		3.5927X 10^7	

Panel B: Fixed Effect Model

Variables	*Coefficient*	*Standard error*	*P-value*
Intercept	0.0000	0.0164	0.9995
DISPERSION	-0.0161	0.0129	0.2121
LEV	1.1483	0.1484	0.0000**
SIZE	-0.7234	0.0253	0.0000**
B/M	-0.002	0.0014	0.1529
RSS (Residual sum of squares)		3.9566 X10^7	

The ** denotes significance at the 1% level. *RETURN* is the monthly stock returns following the month that *DISPERSION* is calculated; *DISPERSION* is the analysts' forecast dispersion; *SIZE* is the firm size; *LEV* is the leverage; and *B/M* is the book-to-market of equity.

Table 5
Regression results of the panel threshold model with one threshold variable (*SIZE*) and two regimes

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, LEV_{i,t} + \beta_3\, SIZE_{i,t} + \beta_4\, B/M_{i,t} + u_{i,t}$$

Variables	*Coefficient*	*Standard error*	*P-value*
Panel A: Regime I where *SIZE* <= 6.8667 (Obs% =38.52%)			
Intercept	0.0534	0.0322	0.0964
DISPERSION	-0.0616	0.0248	0.0131*
LEV	1.6419	0.2353	0.0000**
SIZE	-0.5971	0.0436	0.0000**
B/M	0.0099	0.0107	0.3539
Panel B: Regime II where *SIZE* >6.8667 (Obs% =61.48%)			
Intercept	0.004	0.0191	0.8347
DISPERSION	0.0099	0.0144	0.4925
LEV	0.516	0.1932	0.0076**
SIZE	-0.8461	0.0332	0.0000*
B/M	-0.0025	0.0013	0.0509
RSS (Residual sum of squares)		3.9561×10^7	
LR statistic of $\beta^I = \beta^{II}$		49.2017**	

The ** and * denote significance at the 1% and 5% level, respectively. *RETURN* is the monthly stock returns following the month that *DISPERSION* is calculated; *DISPERSION* is the analysts' forecast dispersion; *SIZE* is the firm size; *LEV* is the leverage; and *B/M* is the book-to-market of equity.

Table 6
Regression results of the panel threshold model with one threshold variable (*DISPERSION*) and two regimes

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, LEV_{i,t} + \beta_3\, SIZE_{\,i,t} + \beta_4\, B/M_{i,t} + u_{i,t}$$

Variables	*Coefficient*	*Standard error*	*P-value*
Panel A: Regime I where *DISPERSION* <= 0.0161 (Obs% =21.25%)			
Intercept	0.0321	0.0349	0.3574
DISPERSION	-0.9584	0.2368	0.0001**
LEV	1.0612	0.2865	0.0002**
SIZE	-0.8706	0.0493	0.0000**
B/M	0.0057	0.0063	0.3644
Panel B: Regime II where *DISPERSION* >0.0161 (Obs% =78.75%)			
Intercept	-0.0196	0.019	0.3019
DISPERSION	-0.0132	0.0133	0.3203
LEV	1.1812	0.1719	0.0000**
SIZE	-0.6899	0.0294	0.0000**
B/M	-0.0023	0.0015	0.1183
RSS (Residual sum of squares)		$3.9563 \text{ X}10^7$	
LR statistic of $\beta^I=\beta^{II}$		28.3602**	

The ** denote significance at the 1% level. *RETURN* is the monthly stock returns following the month that *DISPERSION* is calculated; *DISPERSION* is the analysts' forecast dispersion; *SIZE* is the firm size; *LEV* is the leverage; and *B/M* is the book-to-market of equity.

Table 7
Regression results of the panel threshold model with two threshold variables and four regimes

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, LEV_{i,t} + \beta_3\, SIZE_{\ i,t} + \beta_4\, B/M_{i,t} + u_{i,t}$$

Variables	*Coefficient*	*Standard error*	*P-value*
Panel A: Regime I where *DISPERSION* <= 0.0153 and *SIZE* <= 6.8852 (Obs% = 6.59%)			
Intercept	0.0254	0.0801	0.7514
DISPERSION	-1.3078	0.4394	0.0029**
LEV	1.3049	0.4835	0.0070**
SIZE	-0.7684	0.0986	0.0000**
B/M	-0.0581	0.0607	0.3384
Panel B: Regime II where *DISPERSION* > 0.0153 and *SIZE* <= 6.8852 (Obs% = 32.33%)			
Intercept	0.0302	0.0356	0.3969
DISPERSION	-0.0566	0.0252	0.0245*
LEV	1.706	0.266	0.0000**
SIZE	-0.5674	0.0483	0.0000**
B/M	0.0112	0.011	0.3072
Panel C: Regime III where *DISPERSION* <= 0.0153 and *SIZE* >6.8852 (Obs% = 13.44%)			
Intercept	0.0489	0.0408	0.2306
DISPERSION	-0.5153	0.3063	0.0925
LEV	0.6148	0.3803	0.1060
SIZE	-0.8906	0.0632	0.0000**
B/M	0.006	0.0056	0.2817
Panel D: Regime IV where *DISPERSION* > 0.0153 and *SIZE* > 6.8852 (Obs% = 47.64%)			
Intercept	-0.0084	0.0222	0.7036
DISPERSION	0.0111	0.0148	0.4551
LEV	0.5134	0.2237	0.0217*
SIZE	-0.8441	0.0389	0.0000**
B/M	-0.0028	0.0013	0.036
RSS (Residual sum of squares)		$3.9559 \text{ X}10^{7}$	
LR statistic of $\beta^{I}=\beta^{II}=\beta^{III}=\beta^{IV}$ *(Models 4 vs. 1)*		74.0186**	
LR statistic of $\beta^{I}=\beta^{II}$ *and* $\beta^{III}=\beta^{IV}$			
Models 4 vs. 2		24.8170**	
Models 4 vs. 3		45.6585**	

The ** and * denote significance at the 1% and 5% level, respectively. *RETURN* is the monthly stock returns following the month that *DISPERSION* is calculated; *DISPERSION* is the analysts' forecast dispersion; *SIZE* is the firm size; *LEV* is the leverage; and *B/M* is the book-to-market of equity.

Table 8
Regression results for the panel threshold model with two additional control variables

$$RETURN_{i,t+1} = \alpha_i + \beta_1 DISPERSION_{i,t} + \beta_2\, LEV_{i,t} + \beta_3\, SIZE_{\,i,t} + \beta_4 B/M_{i,t} + \beta_5 MKT_{i,t} + \beta_6 MOM_{\,i,t} + u_{i,t}$$

Variables	*Coefficient*	*Standard error*	*P-value*
Panel A: Regime I where *DISPERSION* <= 0.0174 and *SIZE* <= 7.0980 (Obs% = 8.81%)			
Intercept	0.0090	0.0640	0.8877
DISPERSION	-1.1626	0.3660	0.0015**
LEV	1.4396	0.4215	0.0006**
SIZE	-0.4872	0.0858	0.0000**
B/M	0.0805	0.0527	0.1265
MKT	0.7085	0.0124	0.0000**
MOM	-0.0308	0.0053	0.0000**
Panel B: Regime II where *DISPERSION* > 0.0174 and *SIZE* <= 7.0980 (Obs% = 34.70%)			
Intercept	-0.0413	0.0322	0.1994
DISPERSION	-0.0303	0.0221	0.1711
LEV	1.5684	0.2543	0.0000**
SIZE	-0.3229	0.0461	0.0000**
B/M	0.0077	0.0077	0.3166
MKT	0.8081	0.0068	0.0000**
MOM	-0.0073	0.0026	0.0055**
Panel C: Regime III where *DISPERSION* <= 0.0174and *SIZE* >7.0980 (Obs% = 14.74%)			
Intercept	0.0438	0.0363	0.2280
DISPERSION	-0.3781	0.2698	0.1611
LEV	1.1627	0.3568	0.0011**
SIZE	-0.4577	0.0577	0.0000**
B/M	0.0025	0.0059	0.6734
MKT	0.4719	0.0067	0.0000**
MOM	-0.0317	0.0040	0.0000**
Panel D: Regime IV where *DISPERSION* > 0.0174 and *SIZE* > 7.0980 (Obs% = 41.75%)			
Intercept	-0.0102	0.0226	0.6506
DISPERSION	0.0023	0.0149	0.8800
LEV	0.6934	0.2334	0.0030**
SIZE	-0.4733	0.0409	0.0000**
B/M	-0.0026	0.0013	0.0455*
MKT	0.5800	0.0046	0.0000**
MOM	0.0066	0.0024	0.0058**

RSS (Residual sum of squares)	3.5260 X10^7

The ** and * denote significance at the 1% and 5% level, respectively. *RETURN* is the monthly stock returns following the month that *DISPERSION* is calculated; *DISPERSION* is the analysts' forecast dispersion; *SIZE* is the firm size; *LEV* is the leverage; *B/M* is the book-to-market of equity; *MKT* is the market return; and *MOM* is the lagged stock returns. The overall sample consists of 380,073 monthly observations.

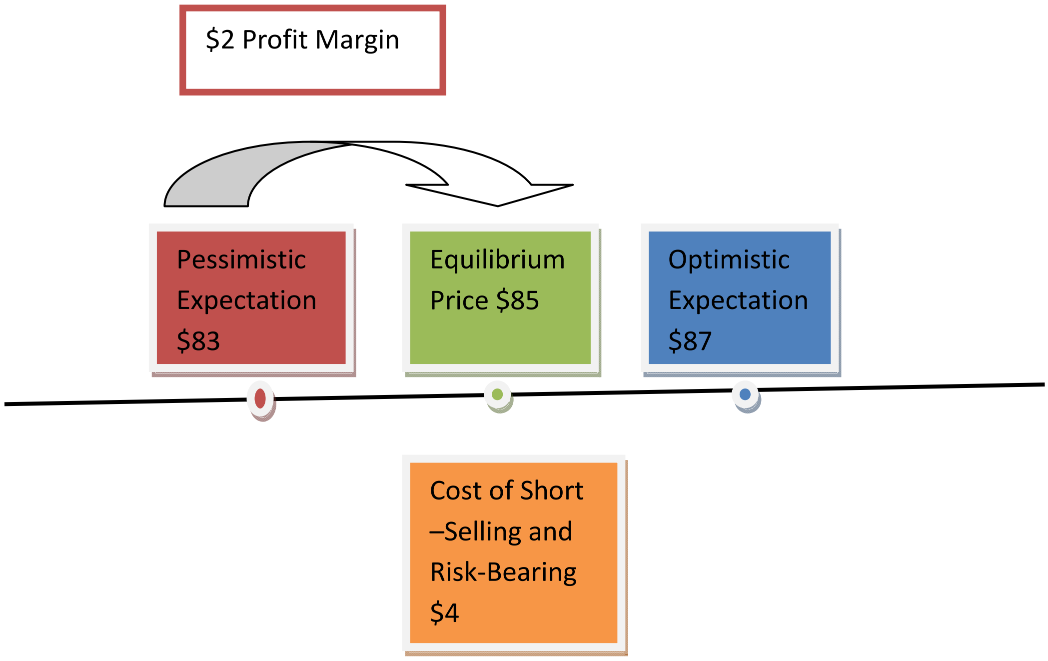

Figure 1. Low analysts' forecast dispersion

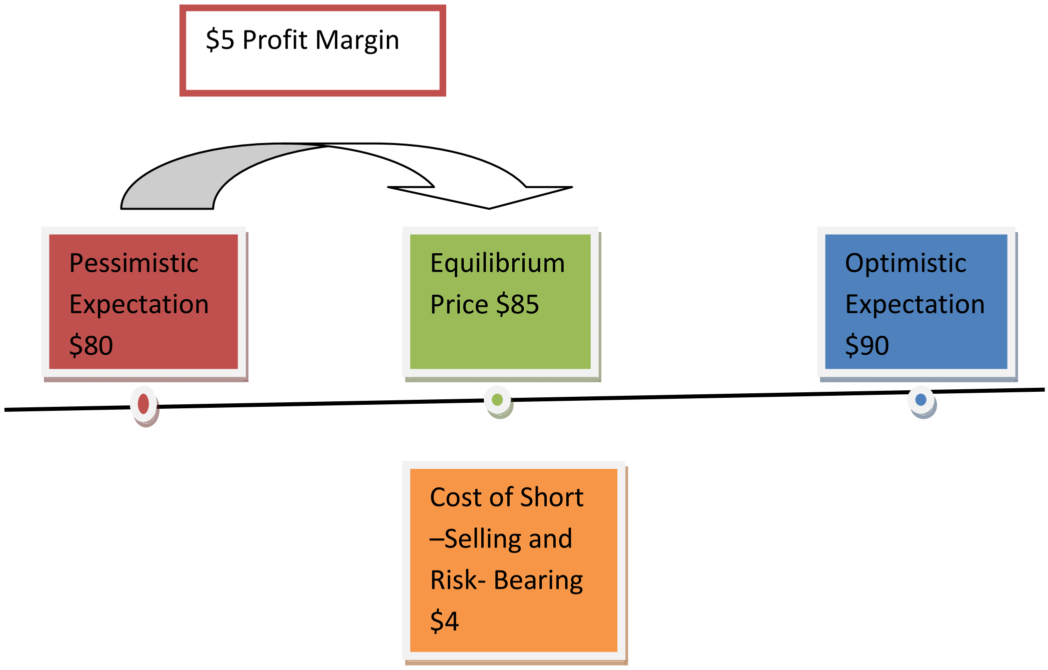

Figure 2. High analysts' forecast dispersion

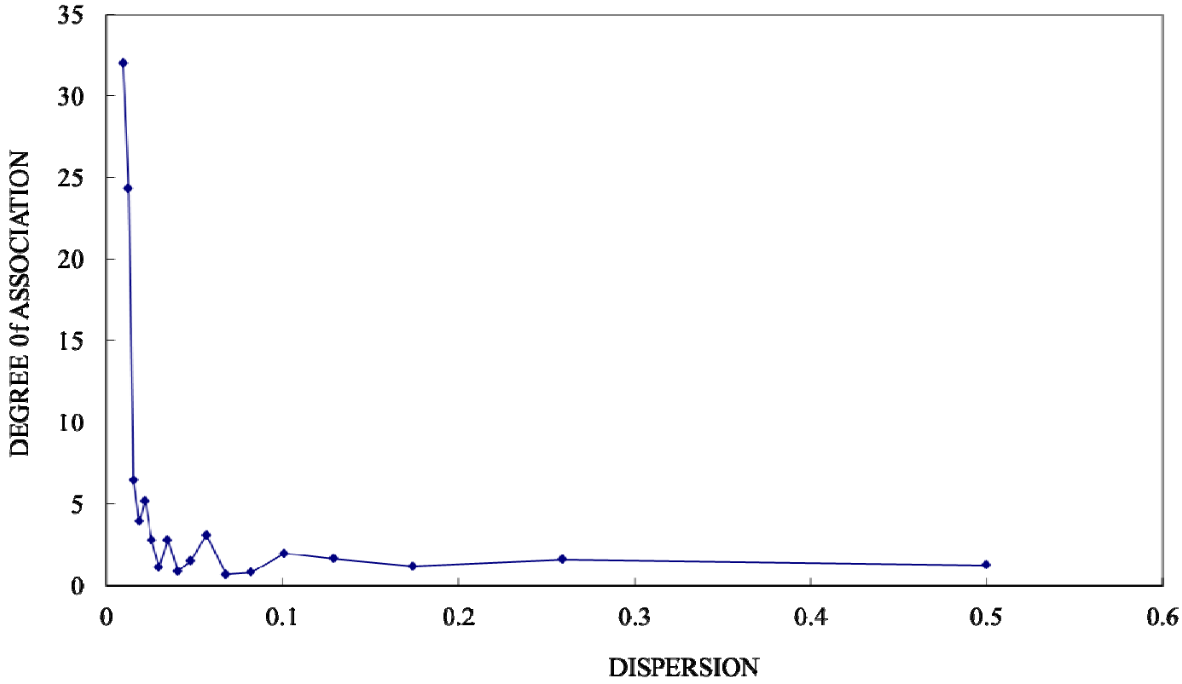

Figure 3: The non-uniform degrees of association between return and dispersion