

Managerial Efforts and the Nature of Skewness in Stock Returns

Daruo Xie

daruo.xie@anu.edu.au

The Australian National University

First version: November 2019

This version: December 2019

Abstract

Firm-level stock returns are positively skewed. This paper presents new evidences to challenge existing explanations of this well-known stylized fact. This paper then provides a simple and entirely new explanation. I revert to the textbook assumption of a frictionless stock market with complete information, and my simple model suggests that stock returns can exhibit positive skewness even in a perfectly efficient market. My model proposes a direct connection between managerial value-creation efforts and stock return skewness. The model predictions are consistent with the empirical evidences.

1. Introduction

How stock price movements reflect the flow of new information is one of most important questions about financial market efficiency. It is a well-known stylized fact that the distribution of firm-level stock returns is positively skewed (Figure 2). Prior literatures attempting to explain the positive skewness largely fall into two categories. Literatures in the first category, e.g. Duffee (1995, 2002), and Albuquerque (2012), suggest that, the distribution of stock returns is conditionally normal, and the unconditional distribution, however, is a mixture of normal distributions, thereby displaying positive skewness due to a positive correlation between the first and second moments. Literatures in the second category, e.g. Damodaran (1985), Dye (1990), and Acharya, DeMarzo and Kremer (2011), suggest that, self-interested firm managers strategically time the releasing of internal information. Managers may choose to promptly release good news, but delay the release of bad news. This asymmetry in disclosure policy leads to the positive skewness.

This paper challenges prior explanations by providing new evidences. First, I measure firm-level stock return skewness by using daily returns with different measurement time windows. The results show that the mean skewness measured with the shortest 5-day window is almost identical to that measured with the longest 240-day window. In fact, skewness across different measurement windows varies very little (Figure 3). This observation is a direct challenge to conditional normality. Secondly, I measure stock return skewness on the dates of the Federal Open Market Committee (FOMC) announcements. FOMC policy announcements are external information shocks that firm managers are unable to manipulate. The results show that the mean skewness on FOMC-announcement dates is significantly more positive than that on non-FOMC-announcement dates. The increase of skewness on FOMC announcement dates will be a puzzle for the theories that attribute skewness to disclosure policies.

In this paper, I propose a simple and entirely new explanation for the positive skewness of stock returns. The economic intuition is straightforward. New information's ultimate impact on firm value will be a sum of two components: the impact of the information itself, and the impact of the firm managerial efforts induced by the information. When

a firm manager learns good news about the firm, the manager’s reactive efforts will make the good news marginally better, an effect of “icing on the cake”. In contrast, when the firm manager learns bad news, the manager’s reactive efforts will make the bad news marginally less bad, an effect of “damage control”. There is an intrinsic asymmetry in managerial efforts as reacting to good news and bad news. Consequently, when the information-induced managerial efforts create firm value, the efforts will lengthen the right tail of positive price movements, while shorten the left tail of negative movements. This asymmetry will ultimately lead to the positive skewness of returns (illustrated in Figure1).

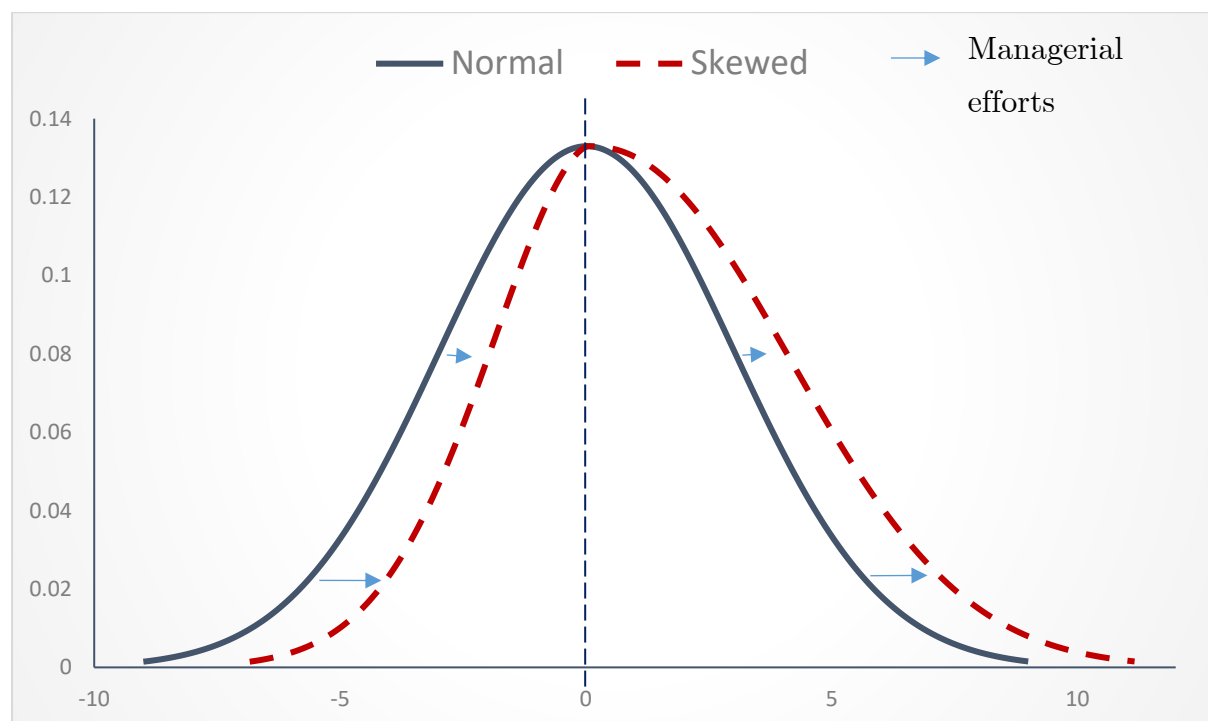


Figure 1 The intuition: An effect of “icing on the cake” on the right tail, along with an effect of “damage control” on the left tail, skew the normal distribution.

My new explanation is not in conflict with the central limit theorem. According to the central limit theorem, when a large number of independent and identically distributed small price changes are added, their sum tends toward a normal distribution, even if the original small changes are not normally distributed. In real world, corporate decision-making, which represents managerial efforts, is at a speed (in hours or days) much slower than information generating speed (in milliseconds). Hence, firm

managers do not react to each atomic information. Instead, managers only react to the aggregation of a large number of atomic information. This means that the central limit theorem already takes effect before firm managers can possibly take any actions. The central limit theorem does not eliminate the asymmetry caused by managerial efforts.

I construct a simple model to frame the reasonings above. I revert to the textbook assumption that stock market is efficient and frictionless. The point of departure is that three model assumptions are added: First, new information is an indispensable input in a manager's value production function. Second, the magnitude of new information's shock on firm value approximates the amount of new information.¹ Third, corporate decision-making is at a speed much slower than information generating speed. The simplicity of the model serves my aim to demonstrate that stock returns can exhibit positive skewness even in a perfectly efficient market. In the latter part of the paper, I extend the basic model to incorporate the information asymmetry between firm managers and the market.

My model is consistent with the empirical evidences presented in this paper. The model attributes skewness to managerial efforts, rather than a positive correlation between the first and the second moments. The model implies that stock returns exhibits positive skewness conditionally as well as unconditionally, which is consistent with the empirical observations in Figure 3. My model also provides a simple explanation for the otherwise puzzling observation that return skewness is more positive on FOMC-announcement dates. According to the model, a large bulk of external information on FOMC-announcement dates induces greater managerial efforts, which leads to higher skewness. Moreover, my model also explains the previous findings in McNichols (1988) that stock return skewness is less positive in earnings announcement periods.

¹ As shown in section 3.2 in this paper, the second assumption can be relaxed as: the magnitude of new information's shock on firm value is significantly related to the amount of new information.

In focusing on the role of information in value production, my paper relates to the literatures in the economics of information, set off from Hayek (1945). In his 1945 article “*The Use of Knowledge in a Society*”, Hayek wrote: “*It is, perhaps, worth stressing that economic problems arise always and only in consequence of change. So long as things continue as before, or at least as they are expected to, there arise no new problems requiring a decision, no need to form a new plan*”. This paper contributes to the literatures by providing a simple theory that connects the role of information to firm-level skewness.

The speed limit for firm managers reacting to information takes a critical role in my theory. However, this paper does not consider this speed limit as a type of information frictions.² The limit of reacting speed essentially comes from the nature of the conscious mind. The speed of a conscious human mind cannot match the high speed of information generating (in milliseconds), even if the economy is entirely free from information acquisition and dissemination costs. Moreover, the reacting speed of firm managers is further constrained by organizational processes of corporate decision-making.

The negative skewness of aggregate market returns is another well-known stylized fact. Many literatures have attempted to explain the negative skewness.³ This paper treats the positive skewness of individual stock returns and the negative skewness of market returns as two separate phenomena.⁴ My model does not fit to explain the market return skewness. The market is not a conscious entity, which may not be constrained

² A model class where investors have incomplete information, for example, Levy (1978), Mayshar (1979, 1981), and Merton (1987).

³ For example, Black (1976), Christie (1982), French, Schwert, and Stambaugh (1987), Hong and Stein (2003), and Veronesi (2004).

⁴ Albuquerque (2012) observes a lack of time-series co-movement between market return skewness and firm-level return skewness. For example, market return skewness was large negative near 2008 financial crisis, while the firm-level skewness did not decrease near the crisis.

by any speed limit. Moreover, the skewness of market portfolio is far from a sum of firm-level skewness. As the market portfolio contains thousands of stocks, the co-skewness terms dominate the skewness of the market portfolio. My model is silent about the cause of co-skewness and treats it as a separate issue.⁵ Many finance literatures have focused on the relationship of skewness and the cross-sectional stock returns.⁶ My paper focuses on the nature of skewness rather than the pricing of skewness.

The rest of the paper is organized as follows: Section 2 presents historical facts and new empirical evidences. Section 3 presents a simple model. Section 4 concludes.

2. Historical Facts and Evidences

The average positive skewness in firm-level stock returns has been documented by finance literatures since Simkowitz and Beedles (1978). Figure 2 plots the average of firm-level stock return skewness from 1952 to 2018. The skewness is computed using rolling 12-month daily returns. As Figure 2 illustrates, the average firm-level skewness is always positive, even in the periods over the 1987 stock market crash.

⁵ As an analogy, variance (idiosyncratic risk as main component) and co-variance (systematic risk) are often treated as two separate issues in academia.

⁶ For example: Chen et al (2001), Ang and Chen (2002), Boyer et al (2010), and Daniel and Moskowitz (2015).

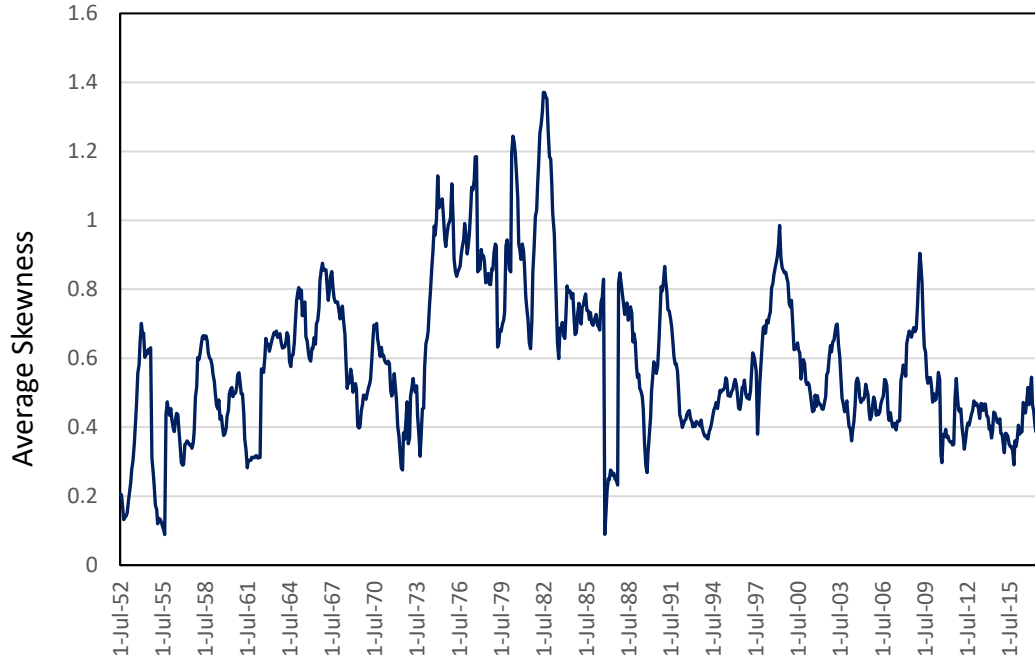


Figure 2 I compute the average skewness for all stocks. Stock skewness is measured using daily stock returns in a 12-month rolling window, and then calculate a simple average across all stocks. The sample covers the US market data from July 1, 1952 to December 31, 2018.

2.1 Is stock return distribution conditionally normal?

Finance literatures have attempted to explain the well-known stylized fact of positive skewness. One hypothesis is that stock return distribution is conditionally normal but unconditionally skewed due to a correlation between the first and second moments. I do a simple test of this hypothesis using various measurement time windows. The rationale of this test is that the skewness measured with a short time window will approximate the conditional skewness, while that measured with a long time window will be the unconditional skewness. Figure 3 plots the mean non-standardized skewness in daily stock returns with various measurement windows: 5 days, 10 days, 20 days, 40 days, 60 days, 120 days, and 240 days. This test adopts the non-standardized skewness because it is the unbiased estimator of skewness, which enables me to compare the skewness across different measurement windows.⁷ For a stock i , the non-standardized sample skewness in a window of T days from day t_0 to day $t_0 + T$ is:

⁷ An unbiased estimator is necessary for a small number of observations such as 5 days or 10 days. In other parts of this paper, I use the standardized skewness. When the number of observations are large enough, estimation bias will not be a serious concern.

$$\gamma_{T,i,t_0} = \frac{T}{(T-1)(T-2)} \sum_{t=t_0}^{t_0+T} (r_{i,t} - \bar{r}_i)^3 \quad (1)$$

Then the time-series mean of average skewness in a window of T consecutive days is:

$$\bar{\gamma}_T = \frac{1}{M_T} \sum_{t_0} \left(\frac{1}{N_{t_0}} \sum_{i=1}^{N_{t_0}} \gamma_{T,i,t_0} \right) \quad (2)$$

In the parentheses in (2), I first compute a cross-sectional average across N_{t_0} stocks. Then I compute a time-series mean across all possible T -day windows rolling from July 1, 1952 to December 31, 2018. The M_T in (2) is the total number of T -day windows in the entire sample period.

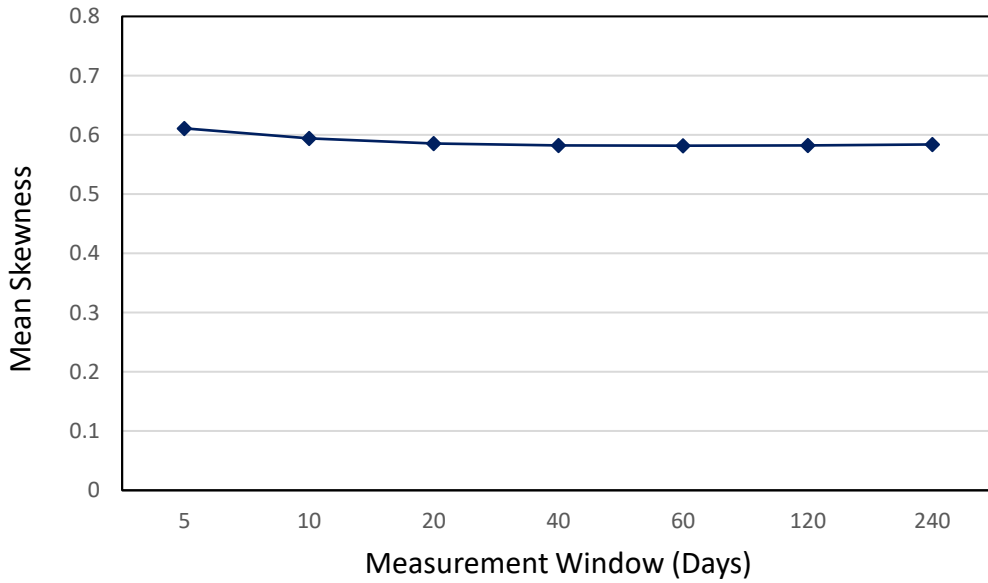


Figure 3 I use various time windows to compute the corresponding time-series mean of average skewness for each window. First, for each stock, I measure its sample non-standardized skewness using daily stock returns in a T -day window. Then, I compute the cross-sectional average skewness across all stocks in the T -day window. Finally, I compute the time-series mean across all possible T -day windows in the entire sample period. The data sample covers the US market from July 1, 1952 to December 31, 2018. The time windows are 5 days, 10 days, 20 days, 40 days, 60 days, 120 days, and 240 days.

If stock return distribution were indeed conditionally normal but unconditionally skewed, the skewness measured with a short window would be close to zero, and the skewness would increase as measurement window becomes longer. I would expect to

see an upward-sloping curve in Figure 3. However, Figure 3 shows that skewness across different measurement windows varies very little. The mean skewness measured with the shortest 5-day window is almost identical to that measured with the longest 240-day window. This is a direct challenge to skewness theories basing on conditional normality, such as Albuquerque (2012).

2.2 A tale of two announcements

Some literatures, e.g. Acharya, DeMarzo, and Kremer (2011), suggest that self-interested firm managers choose to promptly release good news, but delay the release of bad news. The timing of disclosure of internal information gives rise to a positive skewness in stock returns. If disclosure policy were the cause of skewness, firm-level skewness would be unrelated to the external information shocks that firm managers cannot preempt. As Acharya et al. wrote in their 2011 paper: “*absent preemption there is no relation between the (external) news announcement and the timing of disclosure.*”⁸ The Federal Open Market Committee (FOMC) monetary policy announcements are events of external information shocks, about which firm managers are no better informed than the market.

I empirically study whether stock return skewness in FOMC announcements periods is different from that in subsequent non-announcements periods. Table 1 reports the average skewness computed in FOMC announcement periods and non-FOMC-announcement periods. The FOMC announcement period includes the FOMC announcement day and the day after the announcement day. For each stock, the subsequent non-FOMC-announcement period is also a two-day period that begins a random number of days (R) after each announcement day. The R values are uniformly chosen from $[4, 8]$, a 5-day interval. The motivation for using a random number R is to avoid arbitrarily fix a non-announcement period. The 5-day interval is chosen so that a non-announcement period will not cross beyond the next FOMC announcement

⁸ According to their theory in Acharya et al. (2011), alternatively, if the firm manager can preempt the external news announcement, disclosures will be clustered after the announcement of bad news, and firm stock returns will be negatively skewed (rather than positively skewed) in the announcement period.

day.^{9 10} For each firm, the daily returns at the FOMC announcement periods and non-FOMC-announcement periods are grouped respectively. Firms are included in the analysis if at least 10-year trading data are available. FOMC held 757 meetings from 1936 to 2016, a minimum of 43 meetings in any 10-year interval. Accordingly, 10-year trading data means at least 43 announcement periods and 43 non-announcement periods for computing skewness for each firm.

Table 1 This table reports the average of mean return, variance and skewness computed for each firm in FOMC-announcement periods and subsequent non-FOMC-announcement periods respectively. The FOMC announcement periods includes the announcement day and the day after the announcement day. The non-FOMC-announcement period is also a two-day period that begins a random number of days (R) after each announcement day. The R values are uniformly chosen from the interval $[4, 8]$. Firms are included if at least 10-year trading data are available. The historical FOMC meeting dates come from the Federal Reserve website: www.federalreserve.gov, where they are available starting in 1936. The stock return data from CRSP covers the US market from 1936 to 2016.

Number of firms	Average value in FOMC announcement period	Average value in non-FOMC- announcement period	Difference	t -statistic
<i>Mean of returns</i>				
9577	0.132%	0.049%	0.083%	20.89
<i>Variance of returns</i>				
9577	0.00218	0.00184	0.00034	1.24
<i>Skewness of returns</i>				
9577	0.833	0.764	0.069	2.73

The results in Table 1 show that the average of skewness on FOMC-announcement dates is significantly more positive than that on non-FOMC-announcement dates. The

⁹ Still around 1% of FOMC announcements are less than 10 days from each other. I replicated the study here by removing the 1% announcements, and found qualitatively identical results to those reported in table 1.

¹⁰ As a robustness check, I replicated the study using a 10-day interval or a 20-day interval, and found similar results.

difference is 0.069 with a t -statistic 2.73.¹¹ The external information shocks increase the firm-level skewness. This increase will be a puzzle for the theories that attribute skewness to firm manager's timing of disclosure of internal information. To be clear, I do not deny the possibility that disclosure policy may contribute to skewness. The empirical evidences in table 1 present difficulties for the skewness theories basing on disclosure policies, therefore open the door to new explanations.

Table 2 This table reports the average of mean return, variance and skewness computed for each firm in earnings-announcement periods and subsequent non-earnings-announcement periods respectively. The earnings announcement periods includes the announcement day and the day after the announcement day. The non-earnings-announcement period is also a two-day period that begins a random number of days (R) after each announcement day. The R values are uniformly chosen from the interval [4, 8]. Firms are included if at least 10-year trading data are available. Earnings announcement dates come from the Compustat. Stock return data from CRSP. The sample covers the US market from 1962 to 2016.

Number of firms	Average value in earnings announcement period	Average value in non-earnings- announcement period	Difference	t -statistic
<i>Mean of returns</i>				
6851	0.122%	0.087%	0.035%	3.75
<i>Variance of returns</i>				
6851	0.00377	0.00179	0.00197	21.81
<i>Skewness of returns</i>				
6851	0.444	0.594	-0.150	-5.86

It is noteworthy that the increase of skewness is far from being universal for any types of announcement events. McNichols (1988) documented that stock return skewness is lower around a firm's quarterly earnings announcements.¹² Similar to McNichols (1988), I empirically measure the skewness in earnings-announcement periods using

¹¹ Due to the generation of random number R , the outcomes will be slightly different each time I repeat the computing. I run the computing 5 times, and report the one with the median t -statistic.

¹² McNichols (1988) also pointed out that the explanation for her finding is not very clear.

most recent data. Table 2 reports the average of skewness computed at the firm level in earnings announcement periods and subsequent non-earnings-announcement periods. For the aim of being comparable to the results in table 1, I adopt the same methods: each earnings-announcement period includes the earnings announcement day and the day after the announcement day. The subsequent non-earnings-announcement period is also a two-day period that begins a random number of days (R) after each announcement, where R values are uniformly chosen from [4, 8].

The results in Table 2 show that stock return skewness is significantly lower in earnings-announcement periods, compared to subsequent non-earnings-announcement periods. The difference is -0.150 with a t -statistic -5.86.¹³ This decrease of skewness in earnings-announcement periods (table 2) is in strikingly contrast to the increase of skewness in FOMC-announcement periods (table 1). The two types of announcements have exact opposite effects on skewness, which present a challenge for any skewness theory to explain both.

In the next section, I will provide a simple and entirely new explanation for skewness, and my theory will be consistent with the empirical evidences.

3 A Simple Model

3.1 A Basic Model

In this section I construct a simple model on the distribution of stock price changes. I revert to the textbook assumption that stock market is frictionless and informationally efficient. The economy has N firms, and a total of S state variables. Let firm i be a representative firm. Stock prices of firm i already reflect all information relevant to this firm. The unexpected changes in state variables will generate new information.

¹³ For same reason as before in table 1, here I also run the computing 5 times and report the one with the median t -statistic.

New information causes the firm's stock price to change from transaction to transaction. Transactions in stock i happens at discrete time τ . Assume that transactions are uniformly spread across time with a constant time interval $\Delta\tau$ between two sequential transactions. The unexpected change in state variables is random during each transaction interval. The direct impact of new information on the firm value is an independent, identically distributed random variable ε_τ .

The firm i has its corporate decision-making body. Without loss of generality, this paper treats the firm's top manager as a representative for the firm's entire decision-making body, which may include CEO, COO, CFO, the board of directors, etc.¹⁴ The firm manager makes an effort to maximize the firm's market value. The manager learns new information and reacts to new information by making corporate decisions at discrete time t . The manager's effort is represented by making a set of corporate decisions $\{D_j\}_t$, where $j=1,\dots,J$, which represents the space of corporate decisions. The decision makings are uniformly spread across time with a constant time interval Δt between two sequential corporate decisions. The manager's decision at each time t exploits all information relevant to the firm up to the time t . The managerial efforts create positive value for the firm. A value production function maps managerial efforts onto firm value:

$$V_{i,t+1} = V_{i,t} + Y_{i,t+1} + \delta_{i,t+1} \quad (3)$$

where:

$V_{i,t+1}$ is the value of firm i at time $t + 1$.

$V_{i,t}$ is the value of firm i at time t .

$\delta_{i,t+1}$ is a random component of value change from t to $t + 1$, caused by changes of state variables.

$Y_{i,t+1}$ is the contribution of managerial efforts to firm value growth. $Y_{i,t+1}$ follows a classic value production function of Cobb–Douglas form:

¹⁴ The representative top manager(s) in this paper can trace back to the term entrepreneur in Ronald Coase (1937) "The Nature of Firm", which refers to a person or persons who take the place of the price mechanism in the direction of resources.

$$Y_{i,t+1} = A_i K_{i,t+1}^\beta L_{i,t+1}^{1-\beta} \quad (4)$$

where:

$K_{i,t+1}$ is the capital of firm i at $t + 1$.

$L_{i,t+1}$ is the managerial labor efforts at $t + 1$.

A_i is total factor productivity.

β and $1 - \beta$ are the output elasticities of capital and manager's labor respectively. The production function (4) displays constant returns to scales.

Now I impose my first model assumption.

Model Assumption 1: New information is an indispensable factor input in the manager's value production function.

Few researchers doubt that information is a valuable resource in production. And yet in (4), the role of information appears implicitly embodied in the total factor productivity A_i . Now I explicitly highlight the role of information by rewriting the value production function (4) as:

$$Y_{i,t+1} = C_i I_{i,t+1}^\alpha K_{i,t+1}^\beta L_{i,t+1}^{1-\beta} \quad (5)$$

where:

$C_i I_{i,t+1}^\alpha$ is equivalent to the total factor productivity A_i in (4).

$I_{i,t+1}$ is the amount (quantity) of new information generated from t to $t + 1$.

C_i is the new total factor productivity in (5).

α is the output elasticities of information.

Note that if $\alpha = 0$, information will be irrelevant in value production, and equation (5) will return to the conventional form of production function (4). My first model

assumption above is equivalent to assume $\alpha > 0$, so new information becomes an indispensable factor input in the value production function (5).

The factor input I_{it+1} is the amount (quantity) of new information, regardless of good or bad information. $I_{i,t+1}$ is measured in unit of bits. I_{it+1} reflects the scale of state variable changes from t to $t + 1$. By its definition, I_{it+1} is non-negative. The production function (5) adopts new information, rather than all available information. The rationale is as follows: Given a fully efficient market, firm value $V_{i,t}$ already reflects all information and the managerial efforts up to time t . The manager would not need to make new efforts at $t + 1$, if all state variables have remained unchanged from t to $t + 1$ and hence no new information has been generated. Hence, only new information demands new efforts from the manager. According to (5), given the same level of managerial efforts, a greater amount of new information would better facilitate the manager in value creation. The value production function (5) that explicitly highlights the role of new information also resonates with Hayek (1945): *“It is, perhaps, worth stressing that economic problems arise always and only in consequence of change. So long as things continue as before, or at least as they are expected to, there arise no new problems requiring a decision, no need to form a new plan”*.

Note that the new information’s direct impact on firm value is not through firm manager’s production function (5), but through the random component δ_{it+1} in equation (3) instead. In the value production function (5), the other factor input $L_{i,t+1}$ is the labor effort of the manager in making the corporate decision at $t + 1$. Decision making is costly. The firm manager has disutility for effort:^{15 16}

$$C(L_{i,t+1}) = \frac{L_{i,t+1}^2}{2} \quad (6)$$

¹⁵ In the equation (6), the choice of quadratic function for disutility is to be consistent to those in CEO incentive literatures, such as Baker and Hall (2002). The model’s main conclusion doesn’t depend on this particular choice.

¹⁶ A more capable manager could be thought of as having lower cost of effort. However, the equation (6) assumes that all managers share the same dislike for their effort. This is because I subsume cross-sectional differences in managerial effort cost into the difference in total factor productivity C_{it} in equation (5). Thus managers with low effort cost are considered to have high productivity.

Assume the manager is risk-neutral and the manager's utility for decision-making can be captured with the sum of utility for monetary reward from the efforts and disutility for the effort:

$$U(W_{i,t+1}, L_{i,t+1}) = W_{i,t+1} - C(L_{i,t+1}) \quad (7)$$

where $W_{i,t+1}$ is the monetary reward from the manager's effort. Assume a simple linear relationship between CEO monetary reward and firm value growth:

$$W_{i,t+1} = S_{i,t+1} + b_i(V_{i,t+1} - V_{i,t}) \quad (8)$$

where $S_{i,t+1}$ is a fixed component of the manager's rewards, independent of firm performance, b_i is the manager's ownership percentage.¹⁷ From equations (3-8), the manager's optimal effort is the solution to the following problem:

$$\max_{\{L_{i,t+1}\}} \left\{ S_{i,t+1} + b_i C_i K_{i,t+1}^\beta I_{i,t+1}^\alpha L_{i,t+1}^{1-\beta} + b_i \delta_{i,t+1} - \frac{L_{i,t+1}^2}{2} \right\} \quad (9)$$

The first-order condition for optimal managerial effort choice is:

$$L_{i,t+1}^* = [(1 - \beta)b_i C_i]^{\frac{1}{1+\beta}} K_{i,t+1}^{\frac{\beta}{1+\beta}} I_{i,t+1}^{\frac{\alpha}{1+\beta}} \quad (10)$$

From (5) and (10), The value creation from managerial efforts is

$$Y_{i,t+1}^* = [(1 - \beta)b_i]^{\frac{1-\beta}{1+\beta}} C_i^{\frac{2}{1+\beta}} K_{i,t+1}^{\frac{2\beta}{1+\beta}} I_{i,t+1}^{\frac{2\alpha}{1+\beta}} \quad (11)$$

From (3) and (11), the firm value at $t + 1$ is

$$V_{i,t+1} = V_{i,t} + D_{i,t+1} I_{i,t+1}^{\frac{2\alpha}{1+\beta}} + \delta_{i,t+1} \quad (12)$$

¹⁷ Finance literatures, e.g. ... , offer explanations for an optimal CEO ownership percentage, but for my purpose, it is exogenous and outside my model.

where $D_{i,t+1} = [(1 - \beta)b_i]^{\frac{1-\beta}{1+\beta}} C_i^{\frac{2}{1+\beta}} K_{i,t+1}^{\frac{2\beta}{1+\beta}}$ represents the productivity of information. From (12), the firm's stock return $r_{i,t+1}$ from t to $t + 1$ is

$$r_{i,t+1} = \frac{V_{i,t+1} - V_{i,t}}{V_{i,t}} = \frac{1}{V_{i,t}} \delta_{i,t+1} + \frac{D_{i,t+1}}{V_{i,t}} I_{i,t+1}^{\frac{2\alpha}{1+\beta}} \quad (13)$$

Equation (13) demonstrates that the stock return is the sum of two components: The first term $\frac{1}{V_{i,t}} \delta_{i,t+1}$ is the direct impact of new information (state variables changes) on the firm value, and the second term is $\frac{D_i}{V_{i,t}} I_{i,t+1}^{\frac{2\alpha}{1+\beta}}$ is the indirect impact of new information, which represents information-induced managerial efforts.

Now I impose two more model assumptions:

Model Assumption 2: The magnitude of new information's shock on firm value approximates the amount of new information: $I_{i,t+1} \approx g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m$.

Model Assumption 3: Corporate decision-making is at a speed much slower than information generating speed: $\Delta t \gg \Delta \tau$.

The model assumption 2 is based on the rationale as follows: The amount of new information, I_t , in the unit of bits, can hardly be directly observed. Given the market is fully efficient, stock prices reflect all available information. It is reasonable to assume that a large/small quantity of new information will cause a large/small shock on the firm value. Namely, the magnitude of new information's shock on the firm value approximates the amount of the new information:¹⁸

$$I_{i,t+1} \approx g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m \quad (14)$$

¹⁸ The assumption of equation (14) can be relaxed. In section 3.2 in this paper, I will relax the assumption as: $I_{i,t+1} = I_0 + g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m + e_{i,t+1}$, which contains a baseline amount of information I_0 and an orthogonal term e_i . The assumption (14) is equivalent to assume that both I_0 and e_i are small. In section 3.2, the relaxed version of the model assumption 2 will be: New information's shock on firm value is significantly related to the amount of new information. This relaxed version would deliver similar skewness results, with an additional parallel shift caused by I_0 and an additional symmetric noise caused by e_i .

where in (14) $I_{i,t+1}$ is the amount of new information, $\delta_{i,t+1}$ is the direct impact of the new information on firm value, and g_i , m are positive constants. Absolute value in (14) is taken because $|\delta_{i,t+1} - E_t(\delta_{i,t+1})|$ measures the magnitude of a shock, rather than positive or negative shocks. If the power $m = 1$, equation (14) will simply become a proportional relationship. If there is a diminishing marginal effect of information on price change, then $m > 1$.

The rationale for the model assumption 3 is as follows: Information generation is at a speed of $\Delta\tau$ between two sequential transactions, which is likely in the order of magnitude of milliseconds or even faster. In contrast, practically, Δt is likely in the order of magnitude of hours or days.¹⁹ Firm manager reacts to information at a much slower speed than the information generating speed, i.e. $\Delta t \gg \Delta\tau$. There are two main reasons. First, a firm manager is a conscious entity who ultimately relies on human mind to learn information and make decisions. There are biological limits on the manager's information-learning and decision-making speed. Moreover, in this paper, the manager is a representative of the entire decision-making body of the firm. The speed of the manager's reaction to information is further constrained by business processes of corporate decision-making. Therefore, a manager is not able to react to each atomic information that is generated in an interval $\Delta\tau$. Rather, a manager will react to the aggregation of a large number of atomic information that have been generated from t to $t + 1$.

Now I examine stock return distribution. From equations (13) and (14), the stock return r_{t+1} from t to $t + 1$ is:

$$r_{i,t+1} = \frac{1}{V_{i,t}} \delta_{i,t+1} + \frac{D_{i,t+1}}{V_{i,t}} c_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^n \quad (15)$$

where $c_i = \frac{2\alpha}{g_i^{1+\beta}}$, $n = \frac{2\alpha m}{1+\beta}$ are positive. In equation (15), random variable $\delta_{i,t+1}$ is the direct impact on firm value by the new information that has been generated during the corporate decision-making period from t to $t + 1$. Given the model assumption 3,

¹⁹ My empirical work uses daily stock returns. To be consistent with my empirical results, without loss of generality, I will assume Δt be one business day.

$\Delta t \gg \Delta \tau$, therefore, $\delta_{i,t+1}$ will be a sum of a large number of atomic shocks ε_τ during Δt . As previously defined, random variable ε_τ is the direct impact of each atomic information on firm value, and ε_τ is independent, identically distributed. According to the central-limit theorem, $\delta_{i,t+1}$ will be normally distributed.

Although $\delta_{i,t+1}$ is normally distributed, the second term $\frac{D_{i,t+1}}{V_{i,t}} c_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^n$ in equation (15) will skew the distribution of stock return. Specifically, this term will lengthen the right tail of return distribution, but shorten its left tail. The economic intuition is as follows (illustrated by Figure 4): when a firm manager learn good news about the firm, the manager’s reactive effort will make the good news marginally better, an effect of “icing on the cake”. When the manager learn bad news, the manager’s reactive effort will make the bad news marginally less bad, an effect of “damage control”. There is an intrinsic asymmetry in the managerial efforts as reacting to good news and to bad news respectively. When the market is efficient, managerial efforts will be reflected in stock price changes as a positive skewness.

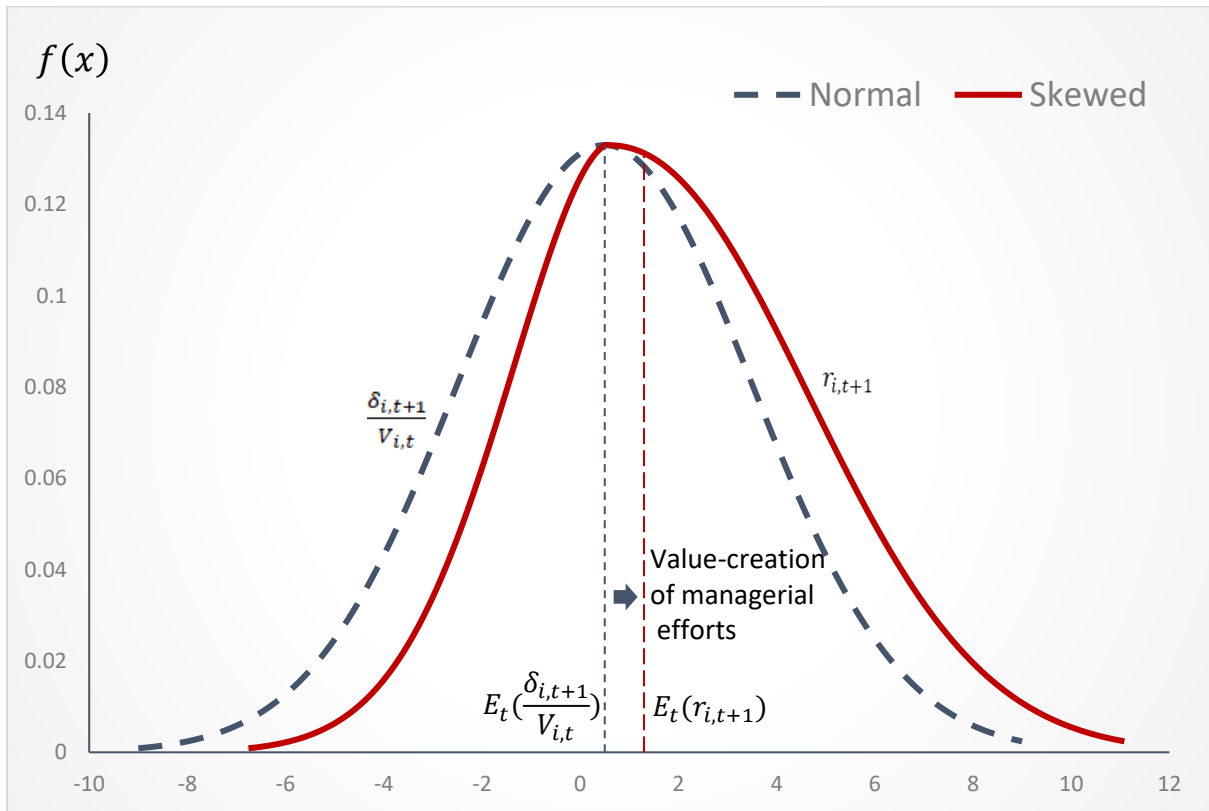


Figure 4 The economic intuition: An effect of “icing on the cake” on the right tail, along with an effect of “damage control” on the left tail, generate a positive skewness and shift the mean.

The model assumption 3, $\Delta t \gg \Delta \tau$, is the most critical element in my model. To see this, let's assume that, in an alternative world, firm managers are so fast that they are able to learn information and make decisions at the speed of milliseconds or faster, analogic to “algorithmic managing”. If a manager is able to react to each atomic information, $\Delta t = \Delta \tau$, then the managerial efforts will directly act on each atomic price change. The central limit theorem will takes effect after the firm manager has taken a large number of actions. If this is the case, the central limit theorem will eliminate the asymmetry caused by managerial efforts and generate a normal distribution instead.

Equation (15) allows me to calculate the skewness of stock return: If the power $n \leq 5$, I have the skewness $\gamma_{i,t+1}$ (mathematical deduction in Appendix A):²⁰

$$\gamma_{i,t+1} = \frac{E_t \left[(r_{i,t+1} - \bar{r}_{i,t+1})^3 \right]}{E_t \left[(r_{i,t+1} - \bar{r}_{i,t+1})^2 \right]^{\frac{3}{2}}} \approx \lambda_i \sigma_{i,t+1}^{n-1} D_i \quad (16)$$

where $\lambda_i = 3n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} c_i \Gamma\left(\frac{n+1}{2}\right)$ is a positive constant, and σ_i is the standard deviation of the shock $\delta_{i,t+1}$. According to (16), a higher information productivity D_i will lead to a higher skewness $\gamma_{i,t+1}$. Moreover, equation (16) implies that in the periods of large information shocks, i.e. large $\sigma_{i,t+1}$, skewness will be high.

The value creation of managerial efforts, as reflected by skewness, is an intrinsic component in the expected return. To better see this, I take expectation on both sides of equation (15),

$$E_t(r_{i,t+1}) = \frac{1}{V_{i,t}} E_t(\delta_{i,t+1}) + \frac{D_i c_i}{V_{i,t}} E(|\delta_{i,t+1} - E_t(\delta_{i,t+1})|^n) \quad (17)$$

where $\delta_{i,t+1}$ follows a normal distribution with mean $E_t(\delta_{i,t+1})$ and standard deviation $\sigma_{i,t+1}$; and $|\delta_{i,t+1} - E_t(\delta_{i,t+1})|$ follows a half-normal distribution. Hence, I have:

²⁰ If $n > 5$, the model still implies a positive relationship between $\gamma_{i,t+1}$ and D_i , but more complex than the equation (13). I deem very high power $n > 5$ unlikely.

$$E_t(r_{i,t+1}) = \frac{1}{V_{i,t}} E_t(\delta_{i,t+1}) + \frac{1}{V_{i,t}} 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} c_i \Gamma\left(\frac{n+1}{2}\right) D_i \sigma_{i,t+1}^n \quad (18)$$

Equation (18) indicates that $E_t(r_{i,t+1})$ will deviate from the point $\frac{E_t(\delta_{i,t+1})}{V_{i,t}}$, where the peak in the return distribution locates. As Figure 4 illustrates, managerial efforts will not only skew the distribution, but also shift the mean of $r_{i,t+1}$ to the right of the distribution peak. This shift reflects the value-creation by managerial labor.

Admittedly, not all managerial works are about learning and reacting to new information. Sometimes, a firm manager may just do daily routine works not much different from assembly line workers. Nevertheless, this paper believes that consciously learning new information and creatively marking new decisions captures the very essence of managerial labor in value creation.

I can rewrite the equations (18) with skewness. From equations (16) and (18), I have:

$$E_t(r_{i,t+1}) = \frac{1}{V_{i,t}} E_t(\delta_{i,t+1}) + \frac{1}{3n} \frac{\sigma_{i,t+1}}{V_{i,t}} \gamma_{i,t+1} \quad (19)$$

Nevertheless, one will need to take care to interpret either equation (18) or (19). The equation (18) and (19) are not outcomes from an equilibrium model of asset pricing. My model is silent about the trade-off between return and risk. The equations (18) and (19) can't determine the expected return $E_t(r_{i,t+1})$, because the present firm value $V_{i,t}$ is undetermined in the model. Hence, the equations (18-19) imply neither a cross-sectional relation between the expected return and productivity D_i , nor a cross-sectional relation between the expected return and skewness $\gamma_{i,t+1}$. The message of equation (19) is simply that conscious managerial efforts contribute to expected stock return, and this contribution is reflected by the positive skewness (the third moment).

3.2 An Extended Model with Information Asymmetry

In this section, I will extend the basic model to incorporate a practical element: the information asymmetry between firm manager and the market.

I incorporate the information asymmetry by relaxing the model assumption 2, while leaving other model settings unchanged. In section 3.1, when the market is fully efficient, the model assumption 2 was: the magnitude of new information's shock on firm value approximates the amount of new information. Now, I relax it as follows:

Model Assumption 2: The magnitude of new information's shock on firm value is significantly related to the amount of new information:

$$I_{i,t+1} = I_{i,0} + g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m + e_{i,t+1} \quad (20)$$

where $I_{i,t+1}$ is the amount of new information received by the firm manager, $I_{i,0}$ is a constant, $\delta_{i,t+1}$ is the direct impact of the new information on firm value, and e_i is an extra term independent to the shock $\delta_{i,t+1}$. The extra term e_i captures the information asymmetry between the firm manager and the market. I consider information asymmetry in the following three scenarios. In the first scenario, both a firm manager and the market know the new information immediately. Hence, there is no information asymmetry, and the term $e_i = 0$. In the second scenario, for some new information, a firm manager may immediately know, but the market does not, so market price does not immediately reflect the information. Consequently, in (20) the term e_i becomes non-negligible: On the day when the firm manager is informed, the information $I_{i,t+1}$ is large while the market shock $|\delta_{i,t+1} - E_t(\delta_{i,t+1})|$ is zero, so the term $e_i > 0$; On the day when the information is finally released to the market, market shock $|\delta_{i,t+1} - E_t(\delta_{i,t+1})|$ is large while $I_{i,t+1}$ is small, so the term $e_i < 0$. In both cases, the magnitude of e_i captures the degree of information asymmetry. In the third scenario, for some new information, the market first knows, but the firm manager does not. Assuming the firm manager can quickly learn this information from the market, this scenario converges to the first scenario, where $e_i = 0$.

From equations (13) and (20), the stock return r_{t+1} from t to $t + 1$ will be:

$$r_{i,t+1} = \frac{1}{V_{i,t}} \delta_{i,t+1} + \frac{D_{i,t+1}}{V_{i,t}} (I_{i,0} + g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m + e_{i,t+1})^{\frac{2\alpha}{1+\beta}} \quad (21)$$

The same as in section 3.1, when $\Delta t \gg \Delta \tau$, central limit theorem implies that the first term $\frac{\delta_{i,t+1}}{V_{i,t}}$ will be normally distributed. The second term in (21) will skew the return distribution by lengthening the right tail but shortening its left tail.

Equation (21) will not allow a close-form solution for skewness, unless I impose additional constraints on the parameters. Without loss of generality, I assume $\frac{2\alpha}{1+\beta} = 1$ in (21). I have:

$$r_{i,t+1} = \frac{1}{V_{i,t}} \delta_{i,t+1} + \frac{D_{i,t+1}}{V_{i,t}} (I_{i,0} + g_i |\delta_{i,t+1} - E_t(\delta_{i,t+1})|^m + e_{i,t+1}) \quad (22)$$

Equation (22) allows me to examine stock return distribution. Compared with equation (15), the $I_{i,0}$ term in (22) will cause a parallel shift to return distribution, and the error term e_i will cause an additional symmetric variance. The skewness is:

$$\gamma_{i,t+1} = \frac{E_t \left[(r_{i,t+1} - \bar{r}_{i,t+1})^3 \right]}{E_t \left[(r_{i,t+1} - \bar{r}_{i,t+1})^2 \right]^{\frac{3}{2}}} \approx \frac{\sigma_{i,t+1}^3}{(\sigma_{i,t+1}^2 + \sigma_{e,t+1}^2)^{\frac{3}{2}}} \times \lambda_i \sigma_{i,t+1}^{n-1} D_i \quad (23)$$

where σ_i is the standard deviation of the shock $\delta_{i,t+1}$, and σ_e is the standard deviation of the extra term e_i . Compared with (16), the skewness in (23) becomes smaller with a ratio factor $\frac{\sigma_{i,t+1}^3}{(\sigma_{i,t+1}^2 + \sigma_{e,t+1}^2)^{\frac{3}{2}}}$. The symmetric variance caused by the term e_i blurs the asymmetric distribution of returns, and therefore reduces the skewness. According to (22), if the market has complete information, then $\sigma_{e,t+1}^2 = 0$ and the skewness in (23) will return to that in (16). Conversely, when there is a high degree of information asymmetry between firm managers and the market, then $\sigma_{e,t+1}^2$ is relatively large, and skewness in (23) will become less positive.

Though the close-form solution in equation (23) results from a special setting of assuming $\frac{2\alpha}{1+\beta} = 1$, the model implication is qualitatively valid for more general cases. The information asymmetry between a firm manager and the market will cause a symmetric variance e_i . As long as stock price changes are still significantly related to

new information, the term e_i will only blur the skewness pattern, but not eliminate it. Moreover, a higher degree of information asymmetry, denoted by a larger $\sigma_{e,t+1}^2$, will lead to a lower skewness.

3.3 Model predictions and empirical evidences

My model predictions provide explanations for the empirical evidences presented in Section 2.

First, the skewness in my model does not rely on a correlation between the first moment and the second moment. My model attributes the positive skewness to managerial efforts. As managerial efforts are conditional as well as unconditional, my model implies that stock returns are skewed both conditionally and unconditionally. Moreover, as the mean of conditional managerial efforts is not different from unconditional efforts, my model implies that the mean skewness measured with a short time window should be same to that measured with a long time window. This prediction is consistent with the empirical observation that skewness across different measurement windows varies very little, as shown by figure 3 in section 2.

Secondly, my model provides a simple explanation for the otherwise puzzling observation that skewness is significantly higher in FOMC-announcement periods, shown by table 1 in section 2. According to the model, a large bulk of external information in FOMC announcement dates will induce greater managerial efforts, which leads to a higher skewness. In addition, monetary policy announcements are external information shocks, about which firm managers are no better informed than the market. This implies a low degree of information asymmetry between a firm manager and the market. According to equation (23), a smaller $\sigma_{e,t+1}^2$ will lead to a higher skewness.

Thirdly, in contrast to the FOMC-announcement effect, skewness is significantly lower in earnings-announcement periods, as shown by table 2 in section 2. My model also provides an explanation for the decrease of skewness. Earnings announcements are

information shocks to the market, but not shocks to the firm manager. A firm manager may have already known the information well before the earnings announcement dates. The information shocks will not induce additional managerial efforts on the earnings announcement dates. Instead, the information shocks only cause symmetric variance that blurs the skewness pattern. According to equation (23), a larger $\sigma_{e,t+1}^2$ will lead to a lower skewness. This model prediction is consistent with the empirical observation shown by table 2.

4. Conclusion

Stock investors would earn no returns in the absence of labor efforts on the production side. However, in classic asset pricing theories, when the market is fully efficient, the first moment of returns (the expected return) will be irrelevant to managerial efforts, as stock prices already reflect all information about managerial productivity. This paper, instead, suggests that the third moment of returns (the skewness) reflects the managerial value-creation efforts. Stock returns may exhibit positive skewness even in a perfectly efficient market. To demonstrate this, I construct a simple model that bases on three assumptions: First, new information is an indispensable input in a firm manager's value production. Second, the magnitude of new information's shock on firm value approximates the amount of new information. Third, corporate decision-making is at a speed much slower than information generating speed. My mode is consistent with the empirical evidences: Stock returns exhibit positive skewness conditionally as well as unconditionally. Skewness is significantly more positive on FOMC-announcement dates, but significantly less positive on earnings-announcement dates.

The most critical element in my theory is its third model assumption, i.e. the limit of reaction speed of a conscious mind. This speed limit shines light on the nature of skewness: Skewness is the footprint of conscious value-creation on the unconscious flow of information.

Appendix A

I rewrite the equation (15), a firm's stock return r_t from $t - 1$ to t :

$$r_{t+1} = \frac{1}{V_t} \delta_{t+1} + \frac{D_{t+1} c_i}{V_t} |\delta_{t+1} - E_t(\delta_{t+1})|^n \quad (\text{A1})$$

where δ_{t+1} follows a normal distribution with positive mean $E_t(\delta_{t+1})$ and standard deviation σ_{t+1} , and therefore $|\delta_{t+1} - E_t(\delta_{t+1})|$ follows a half-normal distribution; And $0 < n \leq 5$. The information productivity $D_{t+1} > 0$, and $c_i > 0$.

The variation in firm's real capital from $t - 1$ to t is small and predictable. Hence I assume that the firm's real capital K_{t+1} is predictable at time t , then no uncertainty in D_{t+1} . In the follows, I will rewrite D_{t+1} simply as D .

Hence, the expected return is:

$$E_t(r_{t+1}) = \frac{E_t(\delta_{t+1})}{V_t} + \frac{D c_i}{V_t} E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n) \quad (\text{A2})$$

in which

$$\begin{aligned} E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n) &= \frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \int_0^{\infty} [\delta_{t+1} - E_t(\delta_{t+1})]^n e^{-\frac{[\delta_{t+1} - E_t(\delta_{t+1})]^2}{2\sigma_{t+1}^2}} d(\delta_{t+1} - E_t(\delta_{t+1})) \\ &= \frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+1}{2})}{2\left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+1}{2}}} \\ &= 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \sigma_{t+1}^n \end{aligned} \quad (\text{A3})$$

Hence, from (A2, A3), I have:

$$E_t(r_{t+1}) = \frac{E_t(\delta_{t+1})}{V_t} + 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \times \left(\frac{D c_i \sigma_{t+1}^n}{V_t}\right) \quad (\text{A4})$$

From (A1), the variance of return is:

$$\begin{aligned}
Var(r_{t+1}) &= \frac{1}{V_t^2} Var(\delta_{t+1} - E_t(\delta_{t+1})) + \frac{D^2 c_i^2}{V_t^2} Var(|\delta_{t+1} - E_t(\delta_{t+1})|^n) \\
&\quad + \frac{2D c_i}{V_t^2} Cov\left((\delta_{t+1} - E_t(\delta_{t+1})), |\delta_{t+1} - E_t(\delta_{t+1})|^n\right) \\
&= \frac{\sigma_{t+1}^2}{V_t^2} + \frac{D^2 c_i^2}{V_t^2} \left\{ E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^{2n}) - (E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n))^2 \right\} \\
&= \frac{\sigma_{t+1}^2}{V_t^2} + \frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) - \Gamma^2\left(\frac{n+1}{2}\right) \right] \times \left(\frac{D c_i \sigma_{t+1}^n}{V_t}\right)^2
\end{aligned} \tag{A5}$$

The skewness of return is:

$$\gamma = \frac{E_t([r_{t+1} - E_t(r_{t+1})]^3)}{Var(r_{t+1})^{\frac{3}{2}}} \tag{A6}$$

Also from (A1), I have:

$$\begin{aligned}
&E_t([r_{t+1} - E_t(r_{t+1})]^3) \\
&= \frac{1}{V_t^3} E_t\left([\delta_{t+1} + D c_i |\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(\delta_{t+1} + D c_i |\delta_{t+1} - E_t(\delta_{t+1})|^n)]^3\right) \\
&= \frac{1}{V_t^3} E_t\left([\delta_{t+1} - E_t(\delta_{t+1})]^3\right) \\
&\quad + \frac{3D c_i}{V_t^3} E_t\left([\delta_{t+1} - E_t(\delta_{t+1})]^2 \times [|\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n)]\right) \\
&\quad + \frac{3D^2 c_i^2}{V_t^3} E_t\left([\delta_{t+1} - E_t(\delta_{t+1})] \times [|\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n)]^2\right) \\
&\quad + \frac{3D^3 c_i^3}{V_t^3} E_t\left([|\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n)]^3\right) \\
&= \frac{3D c_i}{V_t^3} E_t\left([\delta_{t+1} - E_t(\delta_{t+1})]^2 \times [|\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n)]\right) \\
&\quad +
\end{aligned}$$

$$\begin{aligned}
& + \frac{3D^3 c_i^3}{V_t^3} E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n - E_t(|\delta_{t+1} - E_t(\delta_{t+1})|^n))^3 \\
& = \frac{3D c_i}{V_t^3} \left(\frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+3}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+3}{2}}} - \frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+1}{2}}} \times \sigma_{t+1}^2 \right) \\
& \quad + \frac{3D^3 c_i^3}{V_t^3} \left(\frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{3n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{3n+1}{2}}} - 3 \times \frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+1}{2}}} \times \frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{2n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{2n+1}{2}}} + 3 \times \right. \\
& \quad \left. \left(\frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+1}{2}}} \right)^3 - \left(\frac{1}{\sigma_{t+1}} \sqrt{\frac{2}{\pi}} \frac{\Gamma(\frac{n+1}{2})}{2 \left(\frac{1}{2\sigma_{t+1}^2}\right)^{\frac{n+1}{2}}} \right)^3 \right) \\
& = \frac{3D c_i}{V_t^3} \left(n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \sigma_{t+1}^{n+2} \right) \\
& \quad + \frac{3D^3 c_i^3}{V_t^3} 2^{\frac{3n}{2}} \pi^{-\frac{3}{2}} \sigma_{t+1}^{3n} \left(\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \Gamma\left(\frac{n+1}{2}\right)^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \right)
\end{aligned} \tag{A7}$$

Then, from (A5-7), I have the skewness of return:

$$\begin{aligned}
\gamma & = \frac{E_t([r_{t+1} - E_t(r_{t+1})]^3)}{\text{Var}(r_{t+1})^{\frac{3}{2}}} \\
& = \frac{\left\{ \frac{3D c_i \sigma_{t+1}^{n+2}}{V_t^3} \left(n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right) + \frac{3D^3 c_i^3 \sigma_{t+1}^{3n}}{V_t^3} 2^{\frac{3n}{2}} \pi^{-\frac{3}{2}} \left(\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \Gamma\left(\frac{n+1}{2}\right)^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \right) \right\}}{\left[\frac{\sigma_{t+1}^2}{V_t^2} + \frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right] \times \left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2 \right]^{\frac{3}{2}}}
\end{aligned} \tag{A8}$$

A quantitative approximation is innocuous. In equation (A4):

$$E_t(r_{t+1}) = \frac{E_t(\delta_{t+1})}{V_t} + 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \times \frac{D c_i \sigma_{t+1}^n}{V_t}$$

where $\frac{E_t(\delta_{t+1})}{V_t}$ is positive, hence:

$$\frac{D c_i \sigma_{t+1}^n}{V_t} < \frac{E_t(r_{t+1})}{2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \tag{A9}$$

Hence, in equation (A5), the second term on the right:

$$\frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right] \times \left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2 < \frac{\frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right]}{\left[2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right]^2} \times E_t^2(r_{t+1}) \quad (\text{A10})$$

In the inequality (A10), a typical value of daily expected return $E_t(r_{t+1}) < 5 \times 10^{-4}$, Hence $E_t^2(r_{t+1}) < 2.5 \times 10^{-7}$. The constant $\frac{\frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right]}{\left[2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right]^2} < 16$, when $n \leq 5$.

Plugging into these numbers, I have the right side of the inequality (A10):

$$\frac{\frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right]}{\left[2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right]^2} \times E_t^2(r_{t+1}) < 4 \times 10^{-6} \quad (\text{A11})$$

From (A10) and (A11), the second term in the variance equation (A5):

$$\frac{2^n}{\pi} \left[\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1+2n}{2}\right) - \Gamma^2\left(\frac{1+n}{2}\right) \right] \times \left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2 < 4 \times 10^{-6} \quad (\text{A12})$$

A typical value of the variance of daily stock return is: $Var(r_t) > 5 \times 10^{-4}$. Hence, from (A12), the second term in the variance equation (A5) contributes only less than 1% of total variance. More than 99% of return variance comes from the first. Hence, I can rewrite the equation (A5) approximately as:

$$Var(r_{t+1}) \approx \frac{\sigma_{t+1}^2}{V_t^2} \quad (\text{A13})$$

Hence, the denominator in skewness equation (A8) approximates $Var(r_{t+1})^{\frac{3}{2}} \approx \frac{\sigma_{t+1}^3}{V_t^3}$.

Then I check the relative magnitude of the two terms in the numerator in skewness equation (A8). From the numerator in (A8), the ratio of the two terms is:

$$Ratio = \frac{D^3 c_i^3 \sigma_{t+1}^{3n} 2^{\frac{3n}{2}} \pi^{-\frac{3}{2}} \left(\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \left[\Gamma\left(\frac{n+1}{2}\right) \right]^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \right)}{D c_i \sigma_{t+1}^n \sigma_{t+1}^2 n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right)}$$

$$= \left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2 \times \frac{2^n \pi^{-1} \left(\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \left[\Gamma\left(\frac{n+1}{2}\right) \right]^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \right)}{n \Gamma\left(\frac{n+1}{2}\right)} \quad (\text{A14})$$

And from (A9), I have:

$$\left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2 < \frac{1}{\left[\frac{n}{2^2} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right]^2} \times E_t^2(r_{t+1}) \quad (\text{A15})$$

From (A13) and (A15), I have:

$$\left(\frac{D c_i \sigma_{t+1}^n}{\sigma_{t+1}} \right)^2 = \frac{\left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2}{\left(\frac{\sigma_{t+1}^2}{V_t^2} \right)} \cong \frac{\left(\frac{D c_i \sigma_{t+1}^n}{V_t} \right)^2}{\text{Var}(r_{t+1})} < \frac{1}{\left[\frac{n}{2^2} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right]^2} \times \frac{E_t^2(r_{t+1})}{\text{Var}(r_{t+1})} \quad (\text{A16})$$

Hence, from (A14) and (A16), I have:

$$\text{Ratio} < \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \left[\Gamma\left(\frac{n+1}{2}\right) \right]^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right)}{n \left[\Gamma\left(\frac{n+1}{2}\right) \right]^3} \times \frac{E_t^2(r_{t+1})}{\text{Var}(r_{t+1})} \quad (\text{A17})$$

In the inequality (A17), a typical value of $E_t^2(r_{t+1}) < 2.5 \times 10^{-7}$ and $\text{Var}(r_{t+1}) > 5 \times 10^{-4}$. Hence,

$$\frac{E_t^2(r_{t+1})}{\text{Var}(r_{t+1})} < 5 \times 10^{-4} \quad (\text{A18})$$

The right side of (A17) is an increasing function of n . From (A17) and (A18), I have $\text{Ratio} < 0.192$ if $n \leq 5$; $\text{Ratio} < 0.092$ if $n \leq 4$; $\text{Ratio} < 0.01$ if $n \leq 3$. Hence, when $n \leq 5$, the first term in the numerator of the skewness equation (A8) is far larger than the second term:

$$\frac{3Dc_i\sigma_{t+1}^{n+2}}{V_t^3} \left(n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right) \gg \frac{3D^3c_i^3\sigma_{t+1}^{3n}}{V_t^3} 2^{\frac{3n}{2}} \pi^{-\frac{3}{2}} \left(\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{3n+1}{2}\right) + 2 \left[\Gamma\left(\frac{n+1}{2}\right) \right]^3 - 3 \Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{n+1}{2}\right) \Gamma\left(\frac{2n+1}{2}\right) \right) \quad (\text{A19})$$

Hence, from (A8), (A13) and (A19), the skewness of return is:

$$\gamma \approx \frac{\frac{3Dc_i\sigma_{t+1}^{n+2}}{V_t^3} \left(n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \Gamma\left(\frac{n+1}{2}\right) \right)}{\left[\frac{\sigma_{t+1}^2}{V_t^2} \right]^{\frac{3}{2}}} = 3Dc_i n 2^{\frac{n}{2}} \pi^{-\frac{1}{2}} \sigma_{t+1}^{n-1} \Gamma\left(\frac{n+1}{2}\right) \quad (\text{A20})$$

References

Acharya, V. Acharya, DeMarzo Peter, and Kremer Ilan, 2011, Endogenous Information Flows and the Clustering of Announcements, *The American Economic Review* 101, 2955-2979.

Albuquerque, Rui, 2012, Skewness in Stock Returns: Reconciling the Evidence on Firm Versus Aggregate Returns, *Review of Financial Studies* 25, 1630-1673.

Ang, A., and J. Chen, 2002, Asymmetric Correlations of Equity Returns, *Journal of Financial Economics* 63, 443-494.

Black, Fisher, 1976, Studies of Stock Price Volatility Changes, In *Proceedings of the American Statistical Association, Business and Economics Statistics Section*, 177-81.

Boyer, B., T. Mitton, and K. Vorkink, 2010, Expected Idiosyncratic Skewness, *Review of Financial Studies* 23, 169-202.

Chen, J., H. Hong, and J. C. Stein, 2001, Forecasting Crashes: Trading Volume, Past Returns, and Conditional Skewness in Stock Prices, *Journal of Financial Economics* 61, 345-381.

Christie, Andrew A., 1982, The Stochastic Behavior of Common Stock Variances: Value, Leverage, and Interest Rate Effects, *Journal of Financial Economics* 10, 407-32.

Coase, Ronald H., 1937, The Nature of the Firm, *Economica* 4, 386-405.

Damodaran, Aswath, 1985, Economic Events, Information Structure, and the Return-Generating Process, *The Journal of Financial and Quantitative Analysis* 20, 423-434

Daniel, K., and T. J. Moskowitz, 2015, Momentum Crashes, *Journal of Financial Economics* 122, 221-247.

Duffee, G. R. 1995. Stock Returns and Volatility: A Firm-level Analysis. *Journal of Financial Economics* 37:399–420.

Duffee, G. R. 2002. Balance Sheet Explanations for Asymmetric Volatility. Working Paper, University of California, Berkeley.

Dye, R. 1990. Mandatory versus Voluntary Disclosures: The Cases of Financial and Real Externalities. *Accounting Review* 65:1–24.

French, K. R., G. W. Schwert, and R. F. Stambaugh, 1987, Expected Stock Returns and Volatility, *Journal of Financial Economics* 19, 3–29.

Hayek, Friedrich A., 1945, The Use of Knowledge in Society, *The American Economic Review* 35, 519-530.

Hong, H., and J. C. Stein. 2003. Differences of Opinion, Short-sales Constraints, and Market Crashes. *Review of Financial Studies* 16:487–525.

Levy, Haim, 1978, Equilibrium in an imperfect market: A constraint on the number of securities in a portfolio, *American Economic Review* 68, 643-658.

Mayshar, Joram, 1979, Transaction costs in a model of capital market equilibrium, *Journal of Political Economy* 87, 673-700.

Mayshar, Joram, 1981, Transaction costs and the pricing of assets, *Journal of Finance* 36, 583-597.

McNichols, Maureen, 1988, A Comparison of the Skewness of Stock Return Distributions at Earnings and Non-earnings Announcement Dates, *Journal of Accounting and Economics* 10, 239-273.

Merton, Robert C., 1987, A Simple Model of Capital Market Equilibrium with Incomplete Information, *Journal of Finance* 42, 483-510.

Simkowitz, Michael A., and William L. Beedles, 1978, Diversification in a three-moment world, *Journal of Financial and Quantitative Analysis* 13, 927-941.

Veronesi, P. 2004. The Peso Problem Hypothesis and Stock Market Returns. *Journal of Economic Dynamics and Control* 28:707–25.