Skewness in Stock Returns:  
Reconciling the Evidence on Firm versus Aggregate Returns

Rui Albuquerque

Boston University, CEPR, and ECGI

March 29, 2011

Abstract

Aggregate stock market returns display negative skewness. Firm stock returns display positive skewness. The large literature that tries to explain the first stylized fact ignores the second. This paper provides a unified theory that reconciles the two facts by explicitly modeling firm-level heterogeneity. I build a stationary asset pricing model of firm announcement events where firm returns display positive skewness. I then show that cross-sectional heterogeneity in firm announcement events can lead to conditional asymmetric stock return correlations and negative skewness in aggregate returns. I provide evidence consistent with the model predictions.

Key words: Skewness, market returns, firm returns, earnings announcements, cross-sectional heterogeneity.

JEL Classifications: G12, G14, D82

*I would like to thank Adlai Fisher, Marcin Kacperczyk, Emilio Osambela, Lukasz Pomorski, Kevin Sheppard, Pietro Veronesi and Grigory Vilkov for comments and suggestions, as well as seminar participants at Boston University, ESSEC, Michigan Ross School of Business, Said School of Business and Oxford-Man Institute of Oxford University, University of Miami and at the 2010 EFA and the 2011 UBC Winter Finance conferences. Address at BU: Finance Department, Boston University School of Management, 595 Commonwealth Avenue, Boston, MA 02215. Email: ralbuque@bu.edu. The usual disclaimer applies.
1 Introduction

Aggregate stock market returns display negative skewness, the propensity to generate negative returns with greater probability than suggested by a symmetric distribution. A large body of literature has aimed to explain this stylized fact about the distribution of aggregate stock returns (e.g., Fama, 1965, Black, 1976, Christie, 1982, Blanchard and Watson, 1982, Pindyck 1984, French et al., 1987, Hong and Stein, 2003). The evidence on aggregate returns contrasts with another stylized fact, namely, that firm-level returns are positively skewed. For this reason, theories of negative skewness that model single-firm stock markets necessarily depict an incomplete picture. In this paper I provide a unified theory for both stylized facts by explicitly modeling firm-level heterogeneity and present evidence consistent with the theory.

The implications from the disconnect between firm-level return skewness and aggregate return skewness are best illustrated using the definition of sample skewness of a portfolio return. Skewness of a portfolio return is the sum of mean firm-return skewness and co-skewness terms. Because mean firm skewness is positive, negative portfolio-return skewness must be caused by negative co-skewness terms. The co-skewness terms capture the average co-movement in one firm’s return with the variance of the portfolio that comprises the remaining firms. Thus, co-skewness depends on the cross-sectional heterogeneity of firm co-movement, which makes the observed negative skewness in aggregate returns a cross-sectional phenomenon.

This paper argues that the behavior of stock prices around certain firm announcement events is consistent with the existence of positive skewness in firm returns and that cross-sectional heterogeneity in these events can account for the negative skewness in aggregate returns.

The paper provides a stationary asset pricing model of cash payout and earnings announcement events that captures the basic stylized facts on volatility and mean returns around such events. When cash payouts are periodic, cash flow news is discounted according to the time remaining until the next payout. The impact of news on the conditional return volatility is thus greater for news released closer to the payout. This gives rise to a pattern of increasing conditional return volatility, despite homoskedastic news shocks. In addition, discounting also implies that the conditional return volatility increases at an increasing rate. The presence of a risk-return trade-off in the model implies that these properties apply to con-
ditional mean returns and induces positive skewness in conditional mean returns. Similarly, the model predicts conditionally higher return volatility and mean returns around earnings announcement events due to large contemporaneous information flows. Firm returns may thus display sporadic and short-lived periods of high volatility and high mean returns around earnings announcements and positive skewness in conditional mean returns.

The simplicity of the model allows for a complete characterization of the conditional and unconditional distributional properties of equilibrium returns. I show that the unconditional distribution of equilibrium returns is a mixture of normals distribution. Under a mixture of normals distribution, skewness in stock returns is given by two components. The first component is skewness in conditional mean returns. The second component captures the association between expected returns and conditional return variance and is positive given the risk-return trade-off imbedded in the model. With both terms positive, the model can generate positive skewness in firm-level stock returns.

To study market return skewness, I introduce heterogeneity in firms’ announcement events. When firms have different announcement event dates, the high mean return and return volatility of some firms around their event date contrasts with the low return volatility of the portfolio of the remaining firms and may generate negative co-skewness in the market portfolio. In the model, sharp stock market downturns are likely to occur during an announcement season in which a significant fraction of firms displays high return volatility and strong co-movement, while the rest display low expected returns and low return volatility. These periods generate conditional asymmetry in stock correlations: Stocks become more strongly correlated with the market return in a market downturn than in a market upturn.

The paper provides evidence consistent with the above model predictions. Using CRSP daily stock returns to compute skewness over six-month periods from 1973 to 2009, I document two stylized facts. First, firm-level skewness is higher than aggregate skewness 96% of the time. Second, firm-level return skewness is always positive (except in the second half of 1987), whereas market skewness is almost always negative.

The evidence that the cross-sectional dispersion in event dates can produce the correct sign for aggregate return co-skewness uses data on earnings announcement events. As in the model, earnings announcements are associated with brief periods of high volatility and high mean returns (Beaver, 1968, Ball and Kothari, 1991, and Cohen et al., 2007). I use earnings announcement dates over the 1973 to 2009 period from the merged CRSP/Compustat quarterly file. I construct two experiments, both of which use daily return data over six-month
periods. In the first experiment, I form portfolios of firms based on the calendar week of their first earnings announcement in each semester. I then group the firms in the first portfolio (first-week announcers) with those firms announcing \( k \) weeks later and report the six-month portfolio return skewness. I show that, as in the model, there is a symmetric U-shaped pattern in skewness: The portfolio of firms that announce in weeks 1 and 2 has similar return skewness to the portfolio of firms that announce in weeks 13 and 1, and their skewness is higher than the skewness in any other portfolio configuration.

In the second experiment, I form portfolios of firms that announce in weeks 1 through \( k \) in the quarter for \( k = 2, \ldots, 13 \), and report the respective portfolio return skewness. This experiment constructs stock markets with announcement seasons. I show that, consistent with the model, there is a negative relationship between skewness and the increased heterogeneity that results from adding dispersion in event dates. I also show that portfolio skewness in the model can be negative if sufficient heterogeneity in event dates is allowed.

The predictive power of the model hinges on information flowing to the market in the form of announcement seasons. Consistent with other studies (e.g., Chambers and Penman, 1984, and Kross and Schroeder, 1984), I show that firms in the U.S. tend to announce between weeks two and eight in each quarter, giving rise to an earnings announcement season. The beginning of an announcement season is also the period in the model that most contributes to the overall negative skewness in the market. Consistent with this model prediction, I split aggregate skewness into its weekly components and document that aggregate skewness is particularly negative around the beginning of an earnings announcement season.

An alternative explanation for why market skewness differs in sign from firm skewness is the existence of a negatively skewed return factor (Duffee, 1995). Following Duffee (1995), I remove the market return—a negatively skewed factor—from firm returns to obtain “idiosyncratic” returns. I show that while some results are weaker when CAPM-based idiosyncratic returns are used, the evidence is still broadly consistent with the model. Ideally, the use of structural models that nest various theories of negative aggregate skewness can provide for more statistically powerful identification strategies.

The model is related to the literature that analyzes the flow of information in the stock market (e.g., He and Wang, 1995), and the literature that studies properties of stock returns around public news events (e.g., Kim and Verrecchia, 1991, 1994). Especially relevant is the work of Acharya et al. (2010). They study the optimal release of information and the clustering of announcements upon public news releases. In their model, as in Dye (1990), firms
delay the release of bad news, which gives rise to positively skewed firm values. In addition, they show that when firms can preempt the release of public industry news there is clustering of bad news upon the announcement which, they argue, could give rise to conditional negative aggregate skewness. There are two main differences between the their setting and mine. First, the mechanism in my paper does not rely on the endogeneity of the decision to release information. Acharya et al. is a model of voluntary disclosures, which are rare and difficult to predict (e.g. Bhojraj et al., 2010). In this paper, I model and present evidence based on earnings announcements, which are mandatory and predictable (e.g., Givoly and Palmon, 1982, and Chambers and Penman, 1984). Second, Acharya et al. present a result about conditional skewness whereas my result and the data presented speak to unconditional skewness in market returns.

Many studies have focused on asymmetric volatility as an explanation for negative skewness in aggregate stock returns. Black (1976) and Christie (1982) posit the existence of a leverage effect, whereby a low price leads to increased market leverage, which in turn leads to high volatility (see also Veronesi, 1999). Pindyck (1984), French et al. (1987), Campbell and Hentschel (1992), Bekaert and Wu (2000), Wu (2001), and Veronesi (2004) further propose the existence of a volatility feedback effect, whereby high volatility is associated with a high risk premium and a low price. Blanchard and Watson (1982) show that negative skewness can result from the bursting of stock price bubbles. Hong and Stein (2003) hypothesize that short sales constraints limit the market’s ability to incorporate bad news. According to their model, when more bad news arrives in the market, the price responds to the cumulative effect of news and falls at a time when volatility may be high (see also Bris et al., 2007). These papers have made important contributions to our understanding of the dynamics of return volatility and skewness, but they do not address the disconnect between firm skewness and market skewness. The current paper contributes to this literature by providing a bottom-up theory for negative skewness in aggregate stock returns that explicitly models positive skewness in firm-level returns and firm-level heterogeneity. This paper also contributes to the literature by documenting empirically the sources of negative skewness in aggregate returns:

Asymmetric correlations, as opposed to asymmetric volatility, explains the negative skewness

---

1 There is also a literature that documents that skewness is priced; total skewness (e.g., Arditti, 1967); co-skewness (e.g., Kraus and Litzenberger, 1976, and Harvey and Siddique, 2000); or, idiosyncratic skewness (e.g., Boyer et al., 2010). For models of positive skewness at the firm level see Acharya et al. (2010), Dye (1990), Duffee (2002), Grullon et al. (2010), Hong et al. (2008), and Xu (2007). Hong et al. (2007) develop a model that predicts negatively skewed returns for glamour stocks and positively skewed returns for value stocks.
in market returns in this paper. This prediction is consistent with the conditional asymmetry in stock correlations found in Longin and Solnik (2001), Ang and Chen (2002) where market downturns are shown to be associated with higher stock correlations.

The model in this paper is consistent with the evidence from dividend and earnings announcements. Aharony and Swary (1980), Kalay and Loewenstein (1985), and Amihud and Li (2006) show that dividend announcements are associated with high returns and high volatility of stock returns. Beaver (1968), Givoly and Palmon (1982), Ball and Kothari (1991) and Cohen et al. (2007) and others show that the high expected returns around earnings announcements are also associated with high volatility. Patton and Verardo (2010) document an economically and statistically significant increase in firm beta on days of earnings announcements. Finally, there is evidence that firm-level stock returns are well described by a mixture of normals distribution (see Kon, 1984, Zangari, 1996, and Haas et al., 2004).

The paper is organized as follows. Section 2 presents several facts about skewness and discusses the need to model cross-sectional heterogeneity. Section 3 describes the basic model and presents the stock market equilibrium. Section 4 extends the model to incomplete information and earnings announcement events. Section 5 analyzes the skewness properties of aggregate stock returns. Section 6 presents evidence on the paper’s main hypotheses and Section 7 concludes. The Appendix contains the proofs of the propositions and results on the correlated cash flow model.

2 Some Skewness Facts

This section starts by documenting several well-known facts about firm-level and aggregate return skewness. Figure 1 plots the time series of the mean firm stock return skewness and of skewness in the equally weighted market return computed using six-months of daily data. The return is the holding period arithmetic return from CRSP, inclusive of dividends. The data are further described in Section 6 below. Four salient stylized facts emerge from the figure. First, firm-level skewness is always positive, except in the second half of 1987. Second, skewness in market returns is almost always negative, representing 77% of the observations. Third, and as a combination of the two facts above, most semesters of large negative skewness in market returns are not accompanied by negative skewness in firm-level returns. Fourth, firm skewness is higher than aggregate skewness in 96% of the semesters. Because skewness is generally lower and more often negative for larger firms, I reproduce the same statistics using
value-weighted mean (or median) firm skewness and value-weighted market return skewness. Not surprisingly, the value-weighted mean (or median) of firm skewness is lower, but the general gist of the results above is unaffected. These results and those below are robust to using logarithmic returns and are available upon request.

To better understand these results and the need for cross-sectional heterogeneity in a model-free way, it is useful to write the expression for sample non-standardized skewness for a market composed of \( N \) firms (i.e., the sample estimate of the third-centered moment of returns). Assuming equal weights for simplicity, let \( r_{pt} = N^{-1} \sum_{i=1}^{N} r_{it} \) be the time-\( t \) market return, \( \bar{r}_i = T^{-1} \sum_{t=1}^{T} r_{it} \) be the mean sample return for firm \( i \), and \( \bar{r}_p = T^{-1} \sum_{t=1}^{T} r_{pt} \) be the mean sample market return. Then, sample non-standardized skewness is,

\[
T^{-1} \sum_{t} (r_{pt} - \bar{r}_p)^3 = \frac{1}{N^3} \sum_{i=1}^{N} \frac{1}{T} \sum_{t} (r_{it} - \bar{r}_i)^3 + \frac{3}{TN^3} \sum_{t} \sum_{i=1}^{N} (r_{it} - \bar{r}_i) \sum_{i' \neq i}^{N} (r_{i't} - \bar{r}_{i'})^2 + \frac{6}{TN^3} \sum_{t} \sum_{i=1}^{N} (r_{it} - \bar{r}_i) \sum_{i' > i}^{N} \sum_{l > i'}^{N} (r_{l't} - \bar{r}_{i'}) (r_{lt} - \bar{r}_i) .
\]

The first term in (1) is the mean of firm skewness and, as Figure 1 shows, it is positive. The second and third terms in (1) are the co-skewness terms. I label these terms \( \text{co-vol} \) and \( \text{co-cov} \), respectively. Together, they must be negative for skewness in market returns to be negative.

Loosely speaking, the co-skewness terms capture the average co-movement in one firm’s return with the variance of the portfolio that comprises the remaining firms. Thus, co-skewness depends on the cross-sectional heterogeneity of firm co-movement, implying that the negative skewness in aggregate returns is a cross-sectional phenomenon. Specifically, the \( \text{co-vol} \) term describes how one firm’s return co-moves with the return variance in the other firms in the portfolio. The \( \text{co-cov} \) term describes how one firm’s return behaves at times of greater or smaller co-movement in other stocks.

Next I show that the \( \text{co-cov} \) term dominates the sum in (1). The number of firms in a portfolio does not directly affect the calculation of sample skewness. Inspection of equation (1) reveals that \( N^{-3} \) multiplies every term. At the same time \( N^{-3} \) also multiplies every term in \( [T^{-1} \sum_t (r_{pt} - \bar{r}_p)^2]^{3/2} \), cancelling off in the calculation of normalized skewness. Where the number of firms matters is in the weights placed in the various terms. Observe that there
are $N$ firm-level skewness terms, $N(N - 1)$ terms in co-vol, and $N!/ [3! (N - 3)!]$ terms in co-cov. Hence, as the number of firms increases, the number of terms associated with co-cov increases faster than the number of terms associated with any other component of skewness. This does not immediately imply that the co-cov terms dominate the sum, because it may be the case that their component terms cancel each other out. In Figure 2, I plot the ratio of the standardized co-cov term to the sample skewness of market returns. With a ratio close to 100%, on average, the figure suggests that it is the co-cov term that drives negative skewness at the market level.

What determines the sign of the co-cov term is the presence of conditional asymmetries in stock correlations. Take a market downturn characterized by the average firm experiencing a return below the mean. If the pair-wise correlations of $r_{lt} - \tilde{r}_l$ and $r_{pt} - \tilde{r}_p$ for all $l$ are higher in downturns, then the typical term in co-cov, $(r_{lt} - \tilde{r}_l) (r_{lt} - \tilde{r}_p) (r_{lt} - \tilde{r}_l)$, not only is negative in downturns, but is larger in absolute value relative to market upturns, implying negative co-skewness. Hence, negative aggregate co-skewness is consistent with the evidence of higher stock correlations in downturns in Longin and Solnik (2001) and Ang and Chen (2002).

As a final remark, the skewness measure I report in the figures is the standardized skewness equal to $T^{-1} \sum_t (r_{pt} - \tilde{r}_p)^3 / [T^{-1} \sum_t (r_{pt} - \tilde{r}_p)^2]^{3/2}$. Formally, standardizing the third-centered moment introduces a discrepancy between mean firm skewness and the component of market skewness related to firm skewness. When normalized skewness is used, the first term in (1) becomes the volatility-weighted average of normalized firm skewness (with weights $\omega_i = \frac{\sum_t (r_{it} - \tilde{r}_i)^2}{\sum_t (r_{pt} - \tilde{r}_p)^2}]^{3/2}$). Because small firms tend to be more volatile and also have returns with more positive skew, this term is also positive, and negative normalized skewness can only arise from negative normalized co-vol and co-cov terms.

3 The Model

I construct a simple model that captures the observed changes in volatility and mean returns around dividend announcement events. I use the model to show that these patterns in the conditional mean and volatility of returns lead to positive skewness in firm-level returns. In Section 4, I study a model of incomplete information with earnings announcement events and find similar results.
3.1 Investment opportunities

Time is discrete and indexed by \( t = 1, 2, \ldots \). There is a risk-free asset with perfectly elastic supply that can be traded at the gross rate of return of \( R > 1 \). For now consider a stock market with one stock only that has a fixed supply of one share. Each share of the stock is infinitely divisible and trades competitively at time \( t \) at the ex-dividend price \( P_t \). A dividend is announced (and simultaneously paid) every \( K + 1 \) periods,

\[
D_t = F_t + \sum_{j=0}^{K} \varepsilon_{t-K+j}^D.
\]

If \( t \) corresponds to a non-dividend period, then \( D_t = 0 \).

To keep track of the time to the next dividend announcement, trading periods are further identified by event time using the superscript \( k = 0, \ldots, K \), where \( k = 0 \) refers to a dividend-paying period, and \( k > 0 \) refers to a non-dividend-paying period. It helps to think of a trading period as one week and of \( K + 1 \) periods as one quarter: Week \( k \) in the quarter is \( k \) weeks since the last dividend payment and \( K + 1 - k \) weeks to the next dividend payment.

The dividend can be decomposed into a persistent component,

\[
F_t = \rho_F F_{t-1} + \varepsilon_t^F, \quad 0 \leq \rho_F \leq 1,
\]

with \( \varepsilon_t^F \sim N(0, \sigma_F^2) \), and a transitory component, \( \sum_{j=0}^{K} \varepsilon_{t-K+j}^D \), with \( \varepsilon_t^D \sim N(0, \sigma_D^2) \). Note that dividend shocks are conditionally homoskedastic and thus any conditional heteroskedasticity in equilibrium returns is generated endogenously.

Denote by \( P_{t}\) and \( Q_{t}\) the stock price and return, respectively, that occur in period \( t, k \) periods after the last dividend payout. The excess return in a dividend-paying period is

\[
Q_{t}^0 \equiv P_{t}^0 + D_t - R P_{t-1}^K,
\]

and in a non-dividend-paying period is

\[
Q_{t}^k \equiv P_{t}^k - R P_{t-1}^{k-1}.
\]

3.2 Investors’ problem

There is a continuum of identical investors with unit mass. Investors choose their time \( t \) asset allocation, \( \theta_t \), to maximize utility over next period wealth, \( W_{t+1} \),

\[
-E \left[ \exp^{-\gamma W_{t+1}} | I_t \right], \quad (3)
\]
where $\gamma > 0$ is the coefficient of absolute risk aversion. The maximization is subject to the budget constraint

$$W_{t+1} = Q_{t+1}^{k+1} \theta_t + RW_t,$$

and the information set

$$\mathcal{I}_t = \{P_{t-s}, D_{t-s}, F_{t-s}, \varepsilon_{t-s}^D\}_{s \geq 0}.$$

For simplicity, I adopt the short-hand notation $E_t[.] = E[.|\mathcal{I}_t]$.

### 3.3 Stock market equilibrium

Investors trade competitively in the stock market, making their asset allocation while taking prices as given. In equilibrium, the stock price is consistent with market clearing:

$$\theta_t = 1.$$

In the Appendix, I show that:

**Proposition 1** The equilibrium price function is

$$P_t^k = p^k + \Gamma_k F_t + R^{-(K+1-k)} \sum_{j=0}^{k-1} \varepsilon_{t-j}^D,$$

for $\Gamma_k = \frac{(\rho_F/R)^{k+1-k}}{1-(\rho_F/R)^{K+1}}$ and any $k = 0, ..., K$. The constants $p^k < 0$ are given by

$$p^k = -\frac{1}{R^{K+1-1}} \sum_{j=0}^{K} R^{K-j} E_t \left[ Q_{t+1}^{k+1+j} \right],$$

where for any $k$, $E_t \left[ Q_{t+1}^{k+1+K} \right] = E_t \left[ Q_{t+1}^k \right]$.

The stock price at $k$ reflects the present value of dividends conditional on all available information. The present value accounts for the fact that at time $t$—after $k$ periods have elapsed since the last dividend payment—it will take another $K + 1 - k$ periods until dividends are paid again. Consider first the coefficient associated with $F_t$. With $k = 0$, the coefficient is, $\left[ (R/\rho_F)^{K+1} - 1 \right]^{-1}$, and the stock resembles a perpetuity discounted at rate $(R/\rho_F)^{K+1} - 1$. This is because the next payment arises in $K + 1$ periods and is discounted by $R^{K+1}$ and by that time $F_t$ will have decreased in expectation by $\rho_F^{K+1}$. $K + 1$ periods later, another payment occurs, which is also discounted at the same rate, and so on.
The transitory shock \( \varepsilon^D \) enters the stock price function because investors learn about it before it is paid as a dividend: \( \varepsilon^D_t \) enters the price function at time \( t \) with a coefficient of \( R^{-(K-k)} \), whereas \( \varepsilon^D_{t+1} \) enters the price function at time \( t+1 \) with a coefficient of \( R^{-(K-(k-1))} > R^{-(K-k)} \). Despite being transitory, \( \varepsilon^D_t \) has de facto persistence of one until the next dividend payment and persistence of zero thereafter.

### 3.4 Conditional distribution of stock returns

Define the conditional mean return as 
\[
\mu_k = E_t \left[ Q_{t+1}^{k+1} \right]
\]
and the conditional volatility of returns as 
\[
\sigma^2_k = E_t \left[ Q_{t+1}^{k+1} - E_t \left( Q_{t+1}^{k+1} \right) \right]^2.
\]
The investors’ first-order condition together with the stock market clearing condition requires that
\[
\mu_k = \gamma \sigma^2_k. \tag{9}
\]

To solve for the equilibrium values of \( \{\mu_k, \sigma^2_k\}_k \), use the price function above to express excess returns as
\[
Q^k_t = p^k - Rp^{k-1} + \Gamma_k \varepsilon^F_t + R^{-(K+1-k)} \varepsilon^D_t,
\]
for any \( k \). In this expression, \( Q^0_t \) is recovered by replacing \( k \) with \( K+1 \) and noting that \( p^{K+1} = p^0 \) and \( Q^{K+1}_{t+1} = Q^0_{t+1} \). Therefore,

**Corollary 1** *The conditional distribution of stock returns is normal,*

\[
Q_{t+1}^{k+1} | t \sim N \left( \mu_k, \sigma^2_k \right),
\]
with \( \mu_k \) given by equation (9) and \( \sigma^2_k \) given by

\[
\sigma^2_k = \Gamma^2_k \sigma^2_F + R^{-2(K+1-k)} \sigma^2_D. \tag{11}
\]

*The conditional mean and volatility of the stock return increase monotonically and are convex in \( k \), all else equal.*

The corollary states that the conditional stock return volatility increases with \( k \) despite the fact that the shocks \( \varepsilon^F_t \) and \( \varepsilon^D_t \) are conditionally homoskedastic. The intuition is that news that occurs farther away from the dividend payment is more highly discounted and contributes less to risk than news that occurs closer to the dividend payment. Further, discounting penalizes news asymmetrically (i.e., conditional mean and volatility of stock returns are convex in \( k \)), which yields distributions of conditional mean return and conditional return volatility that are positively skewed.
Quantitatively, the effect of discounting on conditional heteroskedasticity via the persistent shocks can be very large even for small interest rates. Consider the impact of \( k \) on the coefficient associated with \( \sigma_F^2 \) in equation (11). Specifically, evaluate the difference in coefficients at \( k = 0 \) and \( k = K \) and take the limit as \( \rho_F/R \to 1 \). Applying L’Hôpital’s rule,

\[
\lim_{\rho_F/R \to 1} \frac{(\rho_F/R)^2 \left(1 - (\rho_F/R)^{2K}\right)}{\left(1 - (\rho_F/R)^{K+1}\right)^2} = +\infty.
\]

Intuitively, a lower interest rate (and higher persistence \( \rho_F \)) reduces the impact of discounting associated with news that is released before the next payout, but increases the value of the perpetuity associated with the news. The second effect is stronger than the first producing the result. Because transitory shocks lack the second effect, when \( R \to 1 \) the discounting effect through transitory shocks disappears.

The result in the Corollary shows that the model is consistent with the evidence that dividend announcements are associated with both higher mean returns and higher volatility (e.g., Aharony and Swary, 1980, and Kalay and Loewenstein, 1985). More recently, Amihud and Li (2006) show evidence of a declining, but still significant, dividend announcement effect.

### 3.5 Unconditional distribution of stock returns

Corollary 1 shows that the firm’s stock return is conditionally normally distributed with mean \( \mu_k \) and variance \( \sigma_k^2 \). The unconditional distribution of the firm’s stock return is not normal because the mean and variance of a randomly drawn return observation depend on \( k \). In fact, because a \( k \)-period stock return is drawn from a normal density \( \phi(Q; \mu_k, \sigma_k^2) \) and such observations occur with frequency \( 1/(K+1) \), the unconditional distribution of returns is a mixture of normals distribution. Formally,

**Proposition 2** For \( K \geq 1 \), the unconditional distribution of stock returns is a mixture of normals distribution with density

\[
f(Q) = \frac{1}{K+1} \sum_{k=0}^{K} \phi(Q; \mu_k, \sigma_k^2),
\]

where \( \phi(.) \) is the normal density function. For \( K = 0 \), returns are unconditionally normally distributed.

The periodicity of dividends – by generating time-varying conditional volatility in stock returns – leads to the derived mixture of normals distribution for stock returns for \( K \geq 1 \).
This result provides a theoretical justification for attempting to fit a mixture of normals distribution to stock returns (e.g., Fama, 1965, Granger and Orr, 1972, Kon, 1984, and Tucker 1992).

In the Appendix, I prove the following corollary.

**Corollary 2** The unconditional mean and variance of stock returns are

\[
E(Q_{t+1}) = \frac{1}{K+1} \sum_{k=0}^{K} \mu_k, \\
Var(Q_{t+1}) = \frac{1}{K+1} \sum_{k=0}^{K} \left[ \sigma_k^2 + (\mu_k - E(Q_{t+1}))^2 \right].
\]

The unconditional (non-standardized) skewness in stock returns is

\[
E[(Q - E(Q_{t+1}))^3] = \frac{1}{K+1} \sum_{k=0}^{K} (\mu_k - E(Q_{t+1}))^3 + \frac{3}{(K+1)^2} \sum_{k=0}^{K} \sum_{j<k} (\sigma_k^2 - \sigma_j^2) (\mu_k - \mu_j).
\]

The unconditional mean return is simply the mean of the \(k\)-conditional expected returns. The unconditional mean variance is the mean of the \(k\)-conditional variances plus the variance of the \(k\)-conditional means.

Skewness in stock returns can be decomposed into two terms. The first term in (13) is the level of skewness in expected returns, \(\mu_k\). For \(K \leq 3\), it is possible to show that this term is non-negative because of the monotonicity and convexity of \(\mu_k\). For larger values of \(K\), it is not possible to sign this term, but numerically it is always found to be positive. Intuitively, this term is positive because an increasing and convex \(\mu_k\) means that a small number of event periods display high expected returns relative to the larger number of event periods with low expected returns. The second term describes the impact on skewness of the co-movement between return volatility and expected returns. The risk-return trade-off implied by equation (9) guarantees that the second term in (13) is positive: Periods of high expected returns are associated with periods of high volatility. In summary, stock returns display positive skewness.

---

2The proof is quite lengthy and is ommitted but is available upon request.
3.6 Discussion

The stochastic discount factor that is implicit in the equilibrium problem formulated above is

\[ m_{t+1}^{k+1} = \gamma \exp \left[ -\gamma \mu_{t+1}^{k+1} - \gamma \Gamma_{k+1}^{F} \varepsilon_{t+1}^{F} - \gamma R^{-(K+1-k)} \varepsilon_{t+1}^{D} \right]. \]

It can be derived directly from the first-order conditions if written as \( E_t \left[ m_{t+1}^{k+1} Q_{t+1}^{k+1} \right] = 0, \) and imposing the market clearing condition, \( \theta_t = 1. \) The stochastic discount factor changes with both calendar time as well as event time, reflecting the fact that shocks to dividends carry a higher risk premium the closer they are to a payout period. Formulating the problem as partial equilibrium and assuming an exogenous stochastic discount factor, as opposed to specifying preferences and budget constraints, is less restrictive and offers a simple and general approach to modeling the effects described in this paper, but lacks microfoundations. This paper provides a microfoundation for event-time variation in the stochastic discount factor.

The model generates skewness in firm-level stock returns by making use of the time-series patterns in volatility that arise from having cash payouts spread out over time. While these patterns in conditional volatility are consistent with the evidence, there could be other explanations for the same facts. For example, it could be the case that the resolution of uncertainty afforded by earnings announcements also results in greater volatility and higher expected returns. I explore this idea next by modeling earnings announcements.

The model takes the cash payout dates as fully predictable, which eliminates considerations about strategic timing of events. While this assumption is made for tractability it finds support in the data (Kalay and Loewenstein, 1985). Likewise, earnings announcements -to be discussed next- are very predictable (e.g., Givoly and Palmon, 1982, Chambers and Penman, 1984, Kross and Schroeder, 1984) and this predictability arises mostly from past earnings announcement behavior, which is often attributed to tradition (e.g. Givoly and Palmon, 1982). I return to this issue in subsection 6.4.

Positive skewness arises despite the fact that prices and returns are conditionally normally distributed. The source of skewness in the model is thus distinct from that which affects arithmetic returns mechanically due to truncation at zero. This benefit, due to exponential utility and normal shocks, comes at the cost of having negative prices with positive probability. To minimize this probability, it is customary to add a positive long-run mean dividend to the process in equation (2). Because all main results (i.e., patterns in conditional
volatility and expected returns in event time) are unchanged, I have assumed away this constant for simplicity of presentation. Nevertheless, one can never rule out the possibility of negative prices in this setting, which is why the model should be understood as an approximation to reality. Another cost of the present setup is that it describes properties of dollar returns. To characterize the properties of simple, percent returns, and for comparability with the empirical analysis, I resort to numerical simulations of a model that allows for a mean dividend. The results in this paper appear robust to these considerations as well (available upon request).

4 A Model with Earnings Announcements

Modeling earnings announcement events is important for two reasons. First, earnings announcements are firm events with similar return and volatility properties to dividend announcements in the data. Second, the empirical analysis in this paper is carried out using earnings announcement events.

I allow for an intermediate earnings announcement event at event date $1 < K_a < K$. For the earnings announcement to be informative, I introduce incomplete information in the model. To do this with minimal deviation from the model above, I assume that for any $1 \leq k \leq K_a - 1$, investors learn

$$S_t^F = \varepsilon_t^F + \varepsilon_t^{SF},$$
$$S_t^D = \varepsilon_t^D + \varepsilon_t^{SD},$$

with the information noise $\varepsilon_t^{SF} \sim N(0, \sigma_{SF}^2)$ and $\varepsilon_t^{SD} \sim N(0, \sigma_{SD}^2)$ independent of each other and of all other shocks. It is assumed that the earnings announcement at event date $K_a$ reveals all current and past shocks. Also, for simplicity, shocks are known with certainty after $K_a$. This gives rise to the following information structure. Let $t$ be any trading period and $k$ be the corresponding date in event time. For any $k = 0$ or $k > K_a - 1$,

$$\mathcal{I}_t^k = \{P_{t+k-s}, D_{t+k-s}, F_{t+k-s}, \varepsilon_{t+k-s}^D\}_{s \geq 0},$$

and for any $1 \leq k \leq K_a - 1$,

$$\mathcal{I}_t^k = \{P_{t+k-s}, S_{t+k-s}^F, S_{t+k-s}^D, \varepsilon_{t+k-s}^D\}_{s=0,\ldots,k-1}.$$

The Appendix shows the following proposition:
**Proposition 3** The equilibrium price function is

\[ P^k_t = p^k + \Gamma_k E_t (F_t) + R^{-(K+1-k)} \sum_{j=0}^{k-1} E_t (\varepsilon^{D}_{t-j}) , \]

for any \( k = 0, \ldots, K \).

The stock price function takes the same form as before with the actual values of the random variables replaced by their conditional expectations. After \( K_a \), the expectations operators drop out because the shocks are in the investors’ information set. With the equilibrium prices, it is possible to derive the equilibrium stock return. For any period \( 1 \leq k \leq K_a - 1 \),

\[ Q^k_t = p^k - Rp^{k-1} + \Gamma_k E_t (\varepsilon^F_t) + R^{-(K+1-k)} E_t (\varepsilon^D_t) . \]

When the signals that investors get are infinitely precise and \( \sigma_{SD}^2 = \sigma_{SF}^2 = 0 \), equation (10) is recovered. For \( k = K_a \),

\[ Q^k_t = p^k - R^k + \Gamma_k \varepsilon^F_t + R^{-(K+1-k)} \varepsilon^D_t + \rho_F \Gamma_k [E_t (F_{t-1}) - E_{t-1} (F_{t-1})] + R^{-(K+1-k)} \sum_{j=0}^{k-2} [\varepsilon^D_{t-1-j} - E_{t-1} (\varepsilon^D_{t-1-j})] . \]

The resolution of uncertainty with the earnings announcement implies that the stock return at \( K_a \) responds to the unanticipated realizations of the past shocks. Finally, for \( k > K_a \), returns take the same form with the same conditional moments as before.

To conclude the derivation of the equilibrium, use the return process above to get the conditional stock return variance, and equation (9) to obtain the conditional mean stock return. It is straightforward to show that for any period \( 1 \leq k \leq K_a - 1 \),

\[ Var_{t-1} \left( Q^k_t \right) = \Gamma^2_k \frac{\sigma^2_F}{\sigma^2_F + \sigma^2_{SF}} + R^{-2(K+1-k)} \frac{\sigma^4_D}{\sigma^2_D + \sigma^2_{SD}} , \]

and for period \( k = K_a \),

\[ Var_{t-1} \left( Q^k_t \right) = \Gamma^2_k \sigma^2_F + R^{-2(K+1-k)} \sigma^2_D + \Gamma^2_k \rho_F Var_{t-1} (F_{t-1}) + R^{-2(K+1-k)} \sum_{j=0}^{k-2} Var_{t-1} (\varepsilon^D_{t-1-j}) . \]

The process for the conditional variance of firm returns is increasing and convex up to \( K_a \). At \( K_a \), the conditional variance may drop so that

\[ Var_{t-1} \left( Q^K_a \right) > Var_t \left( Q^{K_a+1}_{t+1} \right) . \]
This case arises for sufficiently low precision of the signals prior to the earnings announcement, which generates significant resolution of uncertainty at $K_a$. This pattern resembles that of the non-stationary event model of He and Wang (1995).

The patterns in conditional volatility and mean returns described here are consistent with the evidence in Beaver (1968), Givoly and Palmon (1982), and Ball and Kothari (1991). Studying a more recent sample, Cohen et al. (2007) report persistent, significant earnings announcement premia, albeit a smaller one in the later part of the sample. They associate the more recent lower premia with increased voluntary disclosures, which is also consistent with the model above.

In summary, it is possible to have the conditional return variance, and thus also the conditional mean return, displaying two distinct periods of convexity in the event time from 0 to $K$ (one for the earnings announcement and another for the cash payout). By making the periods of high conditional mean returns more likely, returns become less positively skewed. By itself this feature cannot generate negative skewness in aggregate returns, but may contribute to more negative skewness in market returns relative to the benchmark model. Overall, the results with the earnings announcement model are qualitatively similar to those in the model with dividend announcements.3 For this reason, the analysis below considers the simpler setting with dividend announcements only.

5 Skewness in Aggregate Stock Returns

For the remainder of the paper, I consider a stock market composed of firms with i.i.d. cash flows that differ only with respect to the timing of their cash payouts. Together with the assumptions of negative exponential utility and normal shocks, the assumption of i.i.d. cash flows guarantees that stock returns are independent and that the equilibrium firm returns share the properties of the equilibrium returns in the single-stock case studied above. While the independence of stock returns is an unrealistic result, it is useful for two reasons. First, it isolates the effect of cross-sectional heterogeneity in cash payout dates on aggregate skewness: With uncorrelated returns, market skewness can only arise from the cross-sectional heterogeneity in cash payout dates. Second, it gives rise to a simpler presentation with less notation. In the Appendix, I show that the results follow through in the general case of correlated cash flows and discuss implications for systematic risk.

3Numerical examples were part of an earlier version of this paper and are available upon request.
I start by presenting the unconditional distribution of aggregate stock returns and computing skewness in aggregate returns.

5.1 The unconditional distribution of aggregate returns

Let the stock market be composed of \(N\) firms, each with fixed supply of one share. The stock market dollar return is the return from buying and selling the stock on all \(N\) firms. The purchase price is \(P_{\text{it}}\) and the sale price plus the dividend is \(P_{\text{it}} + D_{\text{it}}\). Thus, the per share dollar excess return is \(Q_{Mt} = \frac{1}{N} (Q_{1t} + \ldots + Q_{Nt})\). The unconditional distribution of the stock market return is therefore a mixture of normals distribution:

\[
f(Q^k_M) = \frac{1}{K + 1} \sum_{k=0}^{K} \phi \left( Q^k_M; \mu^M_k, \sigma^2_{M,k} \right).
\]

(14)

Cross sectional heterogeneity is introduced in the following way. Each firm makes a dividend announcement at equidistant periods and with equal frequency. Firms are assumed to differ at most by \(K\) periods in their announcements, which limits the amount of heterogeneity with respect to announcement dates to \(K + 1\) possible dates. A firm of type \(k = 0, 1, \ldots, K\) is identified in the following manner. I arbitrarily assign firm-type 0 to a group of firms announcing in the same period. All other firm types are identified using the distance of their announcement date to that of firms of type 0. Therefore, a firm’s type is set vis-à-vis firm-type 0’s event time. To track the entire cross-section of firms it is thus enough to track event time for one type of firms. I arbitrarily assign the index \(k\) in \(Q^k_M\) to track event time for firm-type 0.

The Appendix shows that (non-standardized) skewness in aggregate stock returns is

\[
E \left[ (Q_{Mt} - E(Q_{Mt}))^3 \right] = \frac{1}{N^3} \sum_{i=1}^{N} E \left[ (Q_{it} - E(Q_i))^3 \right] \]

\[
+ \frac{3}{K + 1} \frac{1}{N^3} \sum_{k=0}^{K} \sum_{i=1}^{N} (\mu^i_k - E(Q)) \sum_{i' \neq i}^{N} \left[ \sigma^2_{k,i'} + \left( \mu^i_k - E(Q) \right)^2 \right] \]

\[
+ \frac{6}{K + 1} \frac{1}{N^3} \sum_{k=0}^{K} \sum_{i=1}^{N} (\mu^i_k - E(Q)) \sum_{i' > i}^{N} \sum_{i'' > i'}^{N} \left( \mu^i_k - E(Q) \right) \left( \mu^i_k - E(Q) \right).
\]

(15)

Skewness in aggregate stock returns is the sum of average firm skewness (first term on the right-hand side of equation (15)) and the co-skewness terms (remaining two terms). The first of the co-skewness terms describes the co-movement of one firm’s stock with other firms’ volatility and is the theoretical equivalent to the \(\text{co-vol}\) term. The second co-skewness term
describes the co-movement of one firm’s stock with the covariance between any two other firms and is equivalent to the co-cov term. Note that it requires \( N \geq 3 \) in the stock market to be non-zero.

Because firm-level skewness is positive in this model, negative aggregate skewness must come from the co-skewness terms: Negative stock market skewness becomes a cross-sectional phenomenon. The portfolio return becomes negatively skewed when a low return for one firm is associated with high volatility in the remaining firms in the portfolio. One way in which this is achieved is via conditional asymmetric correlations. If stock return correlations increase in market downturns, then the co-cov term is negative. Indeed, I show below that the model can generate negative co-skewness and that its main cause is the presence of conditional asymmetric correlations.

5.2 Skewness and cross-sectional heterogeneity in announcement events

To evaluate the effect of cross-sectional heterogeneity in payout dates on co-skewness, I conduct two numerical experiments that simulate a variety of stock market configurations. In all experiments and for simplicity, I assume one firm per firm type. I use dollar returns because the model provides closed-form solutions for all relevant moments, but model simulations show that the results hold for simple, percent returns as well.

In the first experiment, each stock market is composed of two types of firms with cash payouts separated by \( k \) periods, where \( k \in \{0, 1, ..., K\} \). By varying \( k \), the two firms start off similar, become increasingly dissimilar, and end up similar again. I choose \( K = 12 \) so that each trading period represents one week and the time from 0 to \( K \) corresponds to one calendar quarter. Because \( N = 2 \), this experiment explores the effect of cross-sectional heterogeneity ignoring the co-cov term.

Panel A of Figure 3 plots mean firm skewness (dashed line) and market skewness (solid line), for the various stock market configurations. For comparability with the empirical analysis, skewness is the third centered moment of returns normalized by the standard deviation cubed. Mean firm skewness is constant because with i.i.d. cash flows firm skewness does not depend on a firm’s payout date. Market skewness is symmetric because having the second firm pay out \( k \) periods after the first firm or \( k \) periods before the first firm results in identical cross-sectional heterogeneity. Co-skewness can be very large and negative but never sufficiently so in order to offset the individual skewness terms. Co-skewness is particularly negative when the two firms pay out at dates that are farthest apart because then
the high volatility of the announcing firm contrasts the most with the contemporaneously low expected return of the non-announcing firm. In summary, the experiment suggests that the \textit{co-vol} terms can significantly reduce market skewness relative to firm-level skewness, but cannot generate negative market skewness. This result is confirmed with many other parameterizations.

In the second experiment, I allow a role for the \textit{co-cov} term by having the number of firms in the stock market grow as heterogeneity across firms also changes. Each stock market is indexed by \( k \), meaning it consists of \( k+1 \) firm types with cash payout dates at periods \( 0, 1, \ldots, k \). The period from 0 to \( k \) thus denotes an announcement season during the window of time \( 0, \ldots, K \). Panel B of Figure 3 depicts mean firm skewness (dashed line) and market skewness (solid line) in each of the stock market configurations. As in panel A, mean firm skewness is constant because it does not depend on a firm’s payout date. Market skewness displays a flipped J-curve with respect to \( k \). For \( k = 0 \) there is only one firm type in the stock market, and mean firm and market skewness are identical. For \( k = 1 \), the stock market has two firm types, one announcing at 0 and the other at 1. This case is also present in panel A of the figure. For \( k > 1 \) skewness drops faster than it did in panel A because of a negative \textit{co-cov} term. As more firm types are added and the range of cash payout dates is widened, market skewness becomes negative. The negative market skewness occurs despite the fact that mean firm skewness is positive. Market skewness remains negative until the stock market consists of one firm of each type. When the stock market consists of one firm of each type, skewness is zero because every period looks the same with equal aggregate stock market conditional mean and volatility of returns.

The possibility that the \textit{co-cov} term is responsible for the negative skewness in the stock market is investigated further in Figure 4. This figure plots market skewness (solid line) in each of the stock market configurations under experiment two as well as the respective \textit{co-cov} term (dashed line) also normalized by market volatility. A common property of the numerical examples studied, and of this one in particular, is that the \textit{co-cov} term is the main driver of negative skewness in the stock market consistent with evidence presented in Figure 2. The symmetry of events in the model implies that as \( k \) approaches \( K \) and market skewness goes to zero, the \textit{co-cov} term turns positive and the \textit{co-vol} terms turn negative. The \textit{co-vol} terms are negative for large \( k \) because almost every period \( t \) consists of an event period with one firm with the highest conditional volatility (the one with an event at \( t+1 \)) and all the others with low volatility possibly below their respective unconditional means.
A negative co-cov term arises from asymmetric stock correlations in market upturns versus market downturns. To show this, I follow Longin and Solnik (2001) and Ang and Chen (2002), and compute exceedance correlations defined as the correlation between a firm’s stock return and the market return in market upturns (i.e., market return above its unconditional mean) and the correlation between a firm’s stock return and the market return in market downturns (i.e., market return below its unconditional mean). Figure 5 depicts the model simulated average pair-wise exceedance correlation in market upturns, and the ratio of the average pair-wise correlation in downturns to that in upturns, across the various stock market configurations. The figure shows two main properties of the model. First, stock markets with more firm types have lower exceedance correlations because firm cash flows are uncorrelated and each firm represents a smaller weight on the market portfolio. Second, and more importantly, there is a strong conditional asymmetry in correlations. When the stock market is composed of firms announcing in periods 0, 1, ..., 7, and market skewness is negative, the exceedance correlation in market downturns is roughly 60% higher than in market upturns. This prediction is consistent with the evidence in Ang and Chen (2002) who document higher correlations for U.S. stocks in market downturns.

It is also interesting to analyze which trading periods in the quarter contribute most toward overall skewness. Specifically, I am interested in the properties of skewness with respect to the timing of the announcement season. Figure 6 presents a decomposition of the negative skewness for the stock market consisting of eight firm types, each firm type announcing at a different period $k$, with $k = 0, ..., 7$. The figure shows that most negative skewness occurs around the start of the announcement season when some firms’ volatility spikes vis-à-vis that of others.

In the numerical examples above, I assume that $K = 12$ so that there are always 13 periods between any two events for the same firm. While the choice is meant to identify each period as one week and each set of 13 periods as one quarter to match the regularity of the events studied, this choice is not innocuous. Taking $K = 0$ means that payouts occur at every period and in the model returns become unconditionally normally distributed with zero skewness. More generally, $K$ controls the amount of firm heterogeneity in payout dates. Small values of $K$ imply that there cannot be much heterogeneity and make it harder to generate negative aggregate skewness. For example, consider a stock market that consists of two firm types and $K = 2$. When one firm-type has a payout event, the other will either have one next period or the period after. Because of the regularity of the payout events, both
configurations would imply the same level of market skewness. Because of the closeness of the announcements, market skewness would generally be positive.\footnote{Moreover, empirically, a large $K$ may affect the precision of the skewness estimates. In addition, two facts about the timing of earnings announcements suggest looking at weekly periods. First, earnings announcements are fairly predictable \citep[e.g. ][]{chambers1984, givoly1982}. For quarterly announcements, Chambers and Penman estimate that for the representative firm the standard deviation of the actual earnings date minus the estimated date is three to four calendar days. Letting $K = 12$ eliminates some of the concern that investors cannot predict the announcement date as well as they can in the model. Second, firms tend to announce bad news on Fridays \citep[e.g., ][]{damodaran1989, penman1987}. Letting $K = 66$ adds a concern for special week days that is absent in the model.}

6 Empirical Evidence

This section presents evidence on the three main predictions of the model. For reasons that will become clearer, I focus on earnings announcements. First, earnings announcement events are neither uniformly distributed on average in a quarter nor concentrated in one week in the quarter. If the former were true, the model would predict zero unconditional skewness. If the latter were true, the model would predict positive skewness in aggregate returns because of the clustering in volatility in the same week for all firms. Second, cross-sectional dispersion in earnings announcement events can generate large enough negative co-skewness and negatively skewed market returns. I demonstrate this by replicating experiments one and two developed above. I also show that skewness is most negative around the start of an earnings announcement season. Third, negative skewness arises due to co-skewness and in particular the $co-cov$ term. In addition, in a robustness exercise, I repeat the analysis allowing for a negatively skewed factor in returns.

I use daily return data on AMEX/NASDAQ/NYSE stocks from CRSP for the period between 1/1/1973 and 12/31/2009. I use the arithmetic holding period total return from CRSP, inclusive of dividends. I also obtain from CRSP dividend distribution information. I use variable DCLRDT to retrieve the date the board declares a distribution and variable DISTCD to select ordinary dividends and notation of issuance. Information about earnings announcement events is from the merged CRSP/Compustat quarterly file for the period 1/1/1973 through 6/30/2009 (variable RDQ). Below, skewness is estimated using six months of daily return data. Firms are required to have complete return data within each semester to be included in the sample.
6.1 Cross-sectional heterogeneity in event dates

I start by describing the cross-sectional dispersion in cash payout announcements and in earnings announcements. I am interested in the calendar week of the announcement within the quarter. Figure 7 plots the histograms of the announcement week for cash payouts (Panel A) and of the announcement week for earnings announcements (Panel B).\(^5\) Cash payouts are close to uniformly distributed across the quarter. In contrast, and consistent with other studies (e.g., Chambers and Penman, 1984, and Kross and Schroeder, 1984), earnings announcements are on average concentrated between weeks two and eight in the calendar quarter, leaving the other half of the quarter with less than 20% of the announcements. These patterns are consistent across various subsamples and also across the various quarters. This evidence suggests that cross-sectional dispersion in payout dates may not be able to explain the negative skewness in aggregate returns, but that cross-sectional dispersion in earnings announcement events may explain the negative skewness in aggregate returns.\(^6\) I use data on earnings announcements below.

Next, I use data to reproduce the experiments that give rise to Figure 3. For every semester, I group firms by week of first earnings announcement in the semester. This gives rise to 13 portfolios, labeled \(P_1\) through \(P_{13}\), one for each of the weeks in the first quarter of the semester. The portfolios vary greatly in the number of firms that comprise them because of the concentration of earnings announcement events during the quarter (see Figure 7). To keep a constant number of firms across portfolios, I randomly drop firms from portfolios to match the number of firms in the smallest portfolio. It is important to note that it is not possible to replicate in the data the absolute symmetry that exists in the model because firms do not consistently announce in the same week in every quarter. Forcing firms in portfolio \(P_k\) to contain only firms that announce in week \(k\) in both quarters in the semester would lead to a significant loss of observations. I consider two samples: (i) the full sample since 1973; and, (ii) the subsample with data from 1/1/1988, because the earlier years in the full sample have fewer firms. The results below have been replicated when performed over a quarter of data.

Figure 8 replicates experiment one above and Figure 9 replicates experiment two. Figure 8 plots the sample skewness in the equally weighted portfolio return for the portfolios consisting

---

\(^5\)For earnings announcements, observations with an announcement date before the end of the quarter are dropped.

\(^6\)In addition, many firms do not pay dividends, which results in a much smaller sample relative to the earnings announcement sample with a consequent decrease in the precision of estimates.
of the firms in $P1$ and $Pk$ against the index $k = 1, 2, ..., 13$. I also plot the corresponding 10% confidence bands constructed using the sample standard deviation of the estimated skewness measures. The figure shows that portfolio return skewness displays a U-shaped pattern in both samples, consistent with the symmetric U-shaped pattern in Figure 3.

Figure 9 plots the sample skewness in the equally weighted portfolio return for the portfolios that result from the unions $P1UP2U...UPk$ against the index $k = 1, 2, ..., 13$, and the corresponding 10% confidence bands. In both sample periods, there is a negative relationship between skewness and the increased heterogeneity that results from adding dispersion in earnings announcement dates into the portfolio. This evidence is also consistent with the model prediction as depicted in Figure 3.

An alternative interpretation of the results in Figure 9 is that investors with a preference for positive skewness prefer to remain underdiversified (see Mitton and Vorkink, 2007). Using a large dataset of individual investor accounts, Mitton and Vorkink (2007) find that less diversified investor portfolios tend to be more positively skewed because they are composed of firms with more positively skewed stock returns. While this alternative interpretation is plausible, it does not apply to the announcement-week portfolios constructed here. Figure 10 shows that the larger (and more diversified) portfolios in Figure 9 have approximately the same mean firm skewness as the smaller portfolios.

Lastly, I present evidence of how the earnings announcement season is related to skewness. I decompose market skewness computed using six months of data into its weekly components. The decomposition guarantees that adding up the weekly components yields the market skewness for the six-month period. Recalling Panel B of Figure 7, an earnings announcement season starts shortly after the beginning of every quarter. Figure 11 shows that an earnings announcement season is also when skewness has its most negative components during the quarter, consistent with the model prediction illustrated in Figure 6.

There are two main caveats regarding the evidence presented and the model predictions. First, in Figure 9, skewness strictly declines with $k$ whereas in the model, when all firm types are allowed, skewness becomes zero. The result in the model relies on the artifact of symmetry where a firm always announces in the same calendar week in every quarter. This assumption is not validated in the data. Second, in both Figures 8 and 9, point estimates of portfolio return skewness are negative. One possible explanation for the negative portfolio skewness is that even the firms in the same portfolio $Pk$ differ in the week of earnings announcement in the second quarter of the semester. Another explanation is that the cross-sectional heterogeneity
in events is not subsumed in the cross-sectional heterogeneity of earnings announcement events. Finally, it could be the case that firm returns are exposed to a common factor that is negatively skewed. I return to this last point below.

6.2 The number of firms in a portfolio

Section 2 discusses how the number of firms in a portfolio affects the weights placed on each of the various terms that compose the skewness of a portfolio. There are two additional facts about how the number of firms in a portfolio, \( N \), relates to skewness in the portfolio return.

The first is that the co-skewness terms are important and negative even when portfolios are composed of a small number of firms. The second is that the co-skewness terms appear to be monotonically decreasing in \( N \). To show these two facts, I construct equally weighted portfolios of size \( N = 25,625 \) in the following way. First, I assign a random number to each firm and rank firms accordingly. Second, non-overlapping portfolios are formed by taking each consecutive group of \( N \) firms according to their ranking. This procedure guarantees that if two firms are in the same portfolio for \( N = 25 \) they are also in the same portfolio for \( N = 625 \) – a property that is needed to capture the effect of increasing \( N \). Finally, mean portfolio skewness is computed across the \( N \)-firm portfolios. The procedure is then repeated for every semester.

The upshot of the exercise is Figure 12 where I repeat the plots of mean firm skewness and market skewness from Figure 1. The figure shows that the co-skewness terms are important even for small \( N \) and that they appear to be monotonic in \( N \). Using median skewness produces a similar observation. The observed monotonicity pattern can be fully attributed to monotonicity in co-skewness to \( N \) because the mean return skewness across portfolios is the same no matter how many firms are in a portfolio (provided the variance of the portfolio does not change much). This evidence is consistent with the model, but is also consistent with the existence of a negatively skewed common factor in returns. If returns follow \( r_{it} = \beta_t f_t + \epsilon_{it} \) where \( f_t \) is the common factor, it can be shown that, as \( N \to \infty \), non-normalized sample skewness converges to \( \bar{\beta}^3 T^{-1} \sum_i (f_t - \bar{f})^3 \), where \( \bar{\beta} \) is the average exposure to the common factor.

6.3 Negatively skewed factor in returns

Duffee (1995) proposes that the discrepancy in measured skewness in firm and market returns can be accounted for by the existence of a negatively skewed factor in returns. Duffee suggests
looking at the market return as that factor, but does not try to explain the negative skewness in the market return. While my model can explain the skewness in market returns from the cross-sectional heterogeneity in firm announcements, it is possible that market returns are negatively skewed due to other factors, such as jumps in the cash flow process. Separating these different hypotheses is important but difficult because the inclusion of factors, especially those driven by statistical validation, introduces the possibility of “throwing the baby out with the bath water”, that is, of a false rejection of the paper’s null hypothesis. In fact, this is the limitation of the analysis in this subsection. By removing the market factor from firm returns it assumes that the skewness in the market factor is unrelated to the mechanism proposed in the paper.

I remove one common factor from returns for the following reasons. First, the Appendix shows that the model is a one-factor model. Second, the market factor may capture the effect of peso problems or jumps in (common factors in) cash flows that would arise in a more general model, whereas a second factor may capture the skewness induced by cross-sectional heterogeneity in firm events. Third, Engle and Mistry (2007) suggest that the ICAPM is inconsistent with priced risk factors that do not display asymmetric volatility or for which time aggregation changes the sign of skewness. In their paper, the market factor is negatively skewed across all frequencies. The size and momentum factors are negatively skewed at high frequencies but positively skewed at lower frequencies and the book-to-market factor is positively skewed across all frequencies. The results in the subsample from 1988 are especially interesting, because they coincide with the period of study in Engle and Mistry.

To remove the market factor, I run a regression of firm-level daily returns on market returns,

\[ q_{it} = a_i + b_{i1}q_{Mt} + b_{i2}q_{Mt-1} + b_{i3}q_{Mt-2} + \varepsilon_{it}, \]

over the largest possible sample period from 1963 to 2009 for each firm \( i \), from which I obtain the estimated “idiosyncratic” returns, \( \hat{\varepsilon}_{it} \). I use logarithmic returns, \( q_{it} \) and \( q_{Mt} \), as opposed to arithmetic returns and allow for two lags of the market return because of microstructure effects such as non-synchronous trading (Duffee, 1995). The use of logarithmic returns eliminates the positive skewness that arises mechanically because prices are bounded below at zero (see Duffee, 1995, and Chen et al., 2001). However, the overall impact on the level of skewness is unclear because removing the market factor acts in the opposite direction to increase the level of skewness.
Having obtained the residuals $\hat{\epsilon}_{it}$, I proceed as in subsection 6.1, creating portfolios of firms according to the calendar week of their first earnings announcement in each semester. Again, I label these portfolios $P_1$ through $P_{13}$, one for each of the weeks in the first quarter of the semester. I then repeat experiments one and two using the estimated residuals $\hat{\epsilon}_{it}$.

Figure 13 depicts skewness across the various portfolios for experiment one. Removal of the market factor contributes to less negative portfolio skewness as compared to Figure 8. In the full sample, portfolio skewness is always insignificant though the point estimate for $P_1$ is positive. However, the symmetry present in Figure 8 is lost. In the subsample with post 1988 data, not only is the skewness in $P_1$ significantly positive, but there also is a more symmetric relation in the point estimates. It is possible that the greater number of firms listed in this subsample is contributing to the better results. Figure 14 depicts the results for experiment two. Again, compared to Figure 9, portfolio skewness is higher after the removal of the market factor. Consistent with the model there are now several positive point estimates for portfolio skewness and there also is a more pronounced flattening of the skewness curve as $k$ increases.

### 6.4 Timing of announcements

Acharya et al. (2010) propose a model of endogenous information releases. In their model, if firms can preempt the release of public industry, or economy-wide news, then there can be clustering of bad firm news at the time of the public news release. This clustering of bad news may give rise to negative conditional aggregate skewness and thus potentially explain the disconnect between unconditional firm-level and aggregate return skewness. Their model also predicts the existence of conditional asymmetric correlations in market upturns and downturns. Moreover, if the timing of earnings announcements is influenced by the release of public news, then the prediction in Acharya et al. would be consistent with the evidence presented in Figure 11.

Despite the similarity with some of the predictions, the information events that Acharya et al. study are very different in nature from earnings announcements modeled here. Theirs is a model of voluntary disclosures, which are infrequent and tend to be unpredictable. In a recent study of voluntary disclosure practices, Bhojraj et al. (2010) document that the median number of disclosures per quarter is 0.00 and that the mean is 0.529 in a sample of 21,880 firm-quarters. In addition, the pseudo R-squares in the regressions explaining this frequency are under 20%. In contrast, earnings announcements are mandatory and predictable,
thus less subject to discretion. Several papers have shown that earnings announcements can be well predicted using information from past choices of earnings announcement dates and industry patterns (e.g., Chambers and Penman, 1984, and Givoly and Palmon, 1982). Further, Chambers and Penman (1984) estimate that for quarterly announcements the standard deviation of the actual earnings date minus the estimated date from their model is three to four calendar days.

To further distinguish the two hypotheses, I construct the distribution of earnings announcements conditioning on market movements. According to Acharya et al., bunching of earnings announcements is to be expected after market downturns, relative to market upturns. I take the monthly CRSP value-weighted return (including distributions) prior to every quarter and label market upturns as those quarters that follow a return above the median and market downturns as those quarters that follow a return below the median. I also repeat the analysis using as cut-offs the 75th percentile and the 25th percentile, respectively. Conditional on the market return, I then construct the transition matrix of earnings announcements by counting the firms that announce in weeks $k_1$ in the previous quarter and $k_2$ in the current quarter and dividing by the number of firms that announce in week $k_1$ in the previous quarter, for $k_1, k_2 = 1, \ldots, 13$. These transition matrices have no zero elements and thus have a unique stationary distribution (see Ljungqvist and Sargent, 2000). The stationary conditional cumulative distributions are depicted in Figure 15. The distribution depicted using a dashed line conditions on market upturns and the distribution depicted using a solid line conditions on market downturns. There is no apparent bunching after market downturns and instead a larger fraction of firms tend to announce at the beginning of the quarter after market upturns. The results are robust to using the lagged quarterly return to determine market upturns and downturns, or the current monthly or quarterly return.

7 Conclusion

The main contribution of this paper is to model and provide evidence on a new source of negative skewness in market returns. This source consists of the cross-sectional heterogeneity in the timing of earnings announcement events. The paper develops a simple model to capture the observed changes in volatility and mean returns around cash payout and earnings announcement events. The model shows that periodicity in these events gives rise to conditional heteroskedasticity and positive skewness in firm returns consistent with the data.
The model also shows that heterogeneity in the timing of these events can lead to negative skewness in market returns despite the positive skewness in firm returns. The negative skewness in market returns in the model is further shown to be caused by stock correlations that are asymmetrically higher in market downturns. These model predictions are consistent with evidence based on the cross-sectional dispersion of earnings announcement events.

The results in this paper can be informative to the literature on rare disasters that tries to explain the equity premium puzzle and also predicts negative skewness in aggregate stock returns (e.g., Rietz, 1988, and Barro, 2006). Like this paper, Chang et al. (2009) present evidence suggestive that aggregate skewness does not appear to be related to jump risk. Future research should develop structural models nesting several hypotheses to better identify the sources of negative skewness in aggregate returns.

The results in this paper are also pertinent to the large literature that tries to model the dynamics of aggregate return volatility. The model predicts that aggregate return volatility is partly explained by the cross-sectional heterogeneity of firm-level volatility. Testing this prediction is left for future research.
Appendix A: Proofs

This appendix collects the proofs of the propositions in the text.

Proof of Proposition 1: Guess that equilibrium stock returns are conditionally normal with

\[ Q_{t+1}^{k+1} | t \sim N (\mu_k, \sigma_k^2). \]

The representative investor solves (3) subject to equations (4) and (5). The problem yields the familiar first-order necessary and sufficient condition

\[ \theta_t = \frac{\mu_k}{\gamma \sigma_k^2}. \]

Imposing the market clearing condition that the representative investor holds all shares, \( \theta_t = 1 \), gives equation (9), \( \mu_k = \gamma \sigma_k^2 \). Using equation (9), and assuming without loss of generality that time \( t+1 \) corresponds to a payout period, it is possible to write the following set of equilibrium conditions:

\[
\begin{align*}
P_t^K & = R^{-1} \left[ -\gamma \sigma_k^2 + E_t \left[ P_{t+1}^0 + D_{t+1} \right] \right] \\
P_{t-1}^{K-1} & = R^{-1} \left[ -\gamma \sigma_k^2 + E_{t-1} \left[ P_t^K \right] \right] \\
\vdots \\
P_0^{t-K} & = R^{-1} \left[ -\gamma \sigma_k^2 + E_{t-K} \left[ P_{t-K+1}^1 \right] \right] \\
P_{t-K-1}^{K-1} & = R^{-1} \left[ -\gamma \sigma_k^2 + E_{t-K-1} \left[ P_{t-K} + D_{t-K} \right] \right].
\end{align*}
\]

Assuming a stationary solution to this system of stochastic difference equations, recursive substitution yields equation (7) in the proposition.

After constructing equilibrium returns from the price function (see equation (10)), it is straightforward to show that the values for \( p_k \) obey the recursion

\[ \mu_k \equiv E_t \left[ Q_{t+1}^{k+1} \right] = p^{k+1} - Rp^k, \]

which can be solved to yield equation (8). Stationarity implies that for any \( k \), \( E_t \left[ Q_{t+1}^{k+1+K} \right] = E_t \left[ Q_{t+1}^k \right]. \]

Proof of Corollary 1: Given equation (7), construct returns (10). It is then straightforward to derive the conditional variance of stock returns. The conditional variance is increasing and convex in \( k \), because \( R^{-(K+1-k)} \) is increasing and convex in \( k \) and, with \( \rho_F / R < 1 \), \( \Gamma_k \) is also increasing and convex in \( k \). The conditional mean return is proportional to the conditional return variance (see equation (9)) and thus is also increasing and convex in \( k \).
Proof of Corollary 2: Using the definition of $f(Q)$, the unconditional mean stock return is

$$E(Q_{t+1}) = \frac{1}{K+1} \sum_{k=0}^{K} E_k(Q_{t+1}) = \frac{1}{K+1} \sum_{k=0}^{K} \mu_k.$$ 

The unconditional variance in stock returns is

$$Var(Q_{t+1}) = \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - E(Q_{t+1}))^2 \phi(Q; \mu_k, \sigma_k^2) dQ$$

$$= \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - \mu_k + \mu_k - E(Q_{t+1}))^2 \phi(Q; \mu_k, \sigma_k^2) dQ$$

$$= \frac{1}{K+1} \sum_{k=0}^{K} \left( \sigma_k^2 + (\mu_k - E(Q_{t+1}))^2 \right).$$

Finally, unconditional skewness is

$$E\left[ (Q - E(Q_{t+1}))^3 \right] = \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - E(Q_{t+1}))^3 \phi(Q; \mu_k, \sigma_k^2) dQ$$

$$= \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - \mu_k + \mu_k - E(Q_{t+1}))^3 \phi(Q; \mu_k, \sigma_k^2) dQ$$

$$= \frac{1}{K+1} \sum_{k=0}^{K} \left[ (\mu_k - E(Q_{t+1}))^3 + 3\sigma_k^2 (\mu_k - E(Q_{t+1})) \right]. \quad (A4)$$

The third equality uses $\int (Q - \mu_k) \phi(Q; \mu_k, \sigma_k^2) dQ = 0$ and the fact that skewness is zero for a normal variable, $\int (Q - \mu_k)^3 \phi(Q; \mu_k, \sigma_k^2) dQ = 0$. The second term under the summation sign in (A4) can be manipulated to yield the expression in the corollary by noting that

$$\mu_k - E(Q_{t+1}) = \frac{1}{K+1} \sum_{j=0, j \neq k}^{K} (\mu_k - \mu_j),$$

and grouping terms together under the last summation sign.$\blacksquare$

Proof of Proposition 3: Guess prices to be

$$P^k_t = p^k + \Gamma_k E_t(F_t) + R^{-(K+1-k)} \sum_{j=0}^{k-1} E_t(z^D_{t-j}),$$
for all $k$. Obviously for $k \geq K_a$, the expectations operators drop out because the shocks are in investors’ information set. Excess stock returns are

$$Q^k_t = P^k_t - RP^k_{t-1}$$

$$= p^k + \Gamma_k E_t(F_t) + R^{-(K+1-k)} \sum_{j=0}^{k-1} E_t(\varepsilon^D_{t-j})$$

$$- R \left( p^{k-1} + \frac{(\rho_F/R)^{K+2-k}}{1 - (\rho_F/R)^{K+1}} E_{t-1}(F_{t-1}) + R^{-(K+2-k)} \sum_{j=0}^{k-2} E_{t-1}(\varepsilon^D_{t-1-j}) \right),$$

for any period $1 \leq k \leq K_a - 1$. Because

$$E_t(F_t) = \rho_F E_{t-1}(F_{t-1}) + E_t(\varepsilon^F_t)$$

$$= \rho_F E_{t-1}(F_{t-1}) + \frac{\sigma^2_F}{\sigma^2_F + \sigma^2_{SF}} S^F_t,$$

the expression for returns reduces to

$$Q^k_t = p^k - Rp^{k-1} + \Gamma_k \frac{\sigma^2_F}{\sigma^2_F + \sigma^2_{SF}} S^F_t + R^{-(K+1-k)} \frac{\sigma^2_D}{\sigma^2_D + \sigma^2_{SD}} S^D_t.$$

Above, I used

$$E_t(F_{t-1}) = E_{t-1}(F_{t-1}), E_t(\varepsilon^D_{t-1}) = E_{t-1}(\varepsilon^D_{t-1}), ..., E_t(\varepsilon^D_{t-k+1}) = E_{t-1}(\varepsilon^D_{t-k+1}),$$

knowing that time $t$ signals are not informative about $t-n$ shocks for any $n > 0$. For period $k = K_a$,

$$Q^k_t = P^k_t - RP^k_{t-1}$$

$$= p^k + \Gamma_k F_t + R^{-(K+1-k)} \sum_{j=0}^{k-1} \varepsilon^D_{t-j}$$

$$- R \left( p^{k-1} + \frac{(\rho_F/R)^{K+2-k}}{1 - (\rho_F/R)^{K+1}} E_{t-1}(F_{t-1}) + R^{-(K+2-k)} \sum_{j=0}^{k-2} E_{t-1}(\varepsilon^D_{t-1-j}) \right),$$

or rearranging,

$$Q^k_t = p^k - Rp^{k-1} + \Gamma_k \rho_F [F_{t-1} - E_{t-1}(F_{t-1})] + \Gamma_k \varepsilon^F_t$$

$$+ R^{-(K+1-k)} \left\{ \varepsilon^D_t + \sum_{j=0}^{k-2} [\varepsilon^D_{t-1-j} - E_{t-1}(\varepsilon^D_{t-1-j})] \right\}.$$

Finally, for $k > K_a$, returns take the same form with the same conditional moments as in Corollary 1.
It is now easy to construct conditional return moments. For variance, and for any period \(1 \leq k \leq K_a - 1\),
\[
\text{Var}_{t-1} \left( Q^k_t \right) = \Gamma_k^2 \frac{\sigma_F^4}{\sigma_F^2 + \sigma_{SF}^2} + R^{-2(K+1-k)} \frac{\sigma_D^4}{\sigma_D^2 + \sigma_{SD}^2},
\]
which is increasing and convex in \(k\). For period \(k = K_a\),
\[
\text{Var}_{t-1} \left( Q^K_t \right) = \Gamma_k^2 \rho_F^2 \text{Var}_{t-1} \left[ F_{t-1} - E_{t-1} \left( F_{t-1} \right) \right] + \Gamma_k^2 \sigma_F^2
\]
\[+ R^{-2(K+1-k)} \left\{ \sigma_D^2 + \sum_{j=0}^{k-2} \text{Var}_{t-1} \left[ \varepsilon^{D}_{t-1-j} - E_{t-1} \left( \varepsilon^{D}_{t-1-j} \right) \right] \right\}.
\]
In addition,
\[
E_{t-1} \left( \varepsilon^D_{t-1} \right) = \frac{\sigma_D^2}{\sigma_D^2 + \sigma_{SD}^2} S^D_{t-1},
\]
and
\[
\text{Var}_{t-1, K_a-1} \left[ F_{t-1} - E_{t-1} \left( F_{t-1} \right) \right] = \text{Var}_{t-1, K_a-1} \left[ \varepsilon^F_{t-1} - E_{t-1} \left( \varepsilon^F_{t-1} \right) + \ldots + \rho_F^{K_a-2} \left( E_{t-K_a} \left( \varepsilon^F_{t-K_a+1} \right) - E_{t-K_a+1} \left( \varepsilon^F_{t-K_a+1} \right) \right) \right]
\]
\[= \frac{\sigma_F^2 \sigma_{SF}^2}{\sigma_F^2 + \sigma_{SF}^2} \left\{ 1 + \rho_F^2 + \ldots + \rho_F^{K_a-2} \right\}.
\]
For \(k \leq K_a\),
\[
\text{Var}_{t-1} \left( Q^K_t \right) > \text{Var}_{t-2} \left( Q^{k-1}_{t-1} \right).
\]
Furthermore,
\[
\text{Var}_{t-1} \left( Q^{K_a}_t \right) > \text{Var}_t \left( Q^{K_a+1}_{t+1} \right)
\]
is possible if the arrival of information from past shocks is relevant enough. In that case the path of conditional variance displays two distinct periods of convexity. Finally, knowing that \(\mu_k = \gamma \text{Var}_k \left( Q^{k+1}_{t+1} \right)\), it is possible to recover the constants \(p^k\) verifying that the price function above is an equilibrium price.

**Calculations in the many firm case:** Here I derive several unconditional moments of aggregate returns including skewness, which is given in the main text in equation (15). Using the definition of \(f \left( Q \right)\), for a stock market composed of \(N\) firms, the unconditional mean stock return is
\[
E \left( Q_{Mt+1} \right) = \frac{1}{K+1} \sum_{k=0}^{K} E_k \left( Q_{Mt+1} \right) = \frac{1}{K+1} \sum_{k=0}^{K} \frac{1}{N} \sum_{i=1}^{N} \mu^i_k.
\]
The unconditional variance in stock returns is

\[
\text{Var}(Q_{Mt+1}) = \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - E(Q_{t+1}))^2 \phi(Q; \mu_k^M, \sigma_{M,k}^2) \, dQ
\]

\[
= \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - \mu_k^M + \mu_k^M - E(Q_{t+1}))^2 \phi(Q; \mu_k^M, \sigma_{M,k}^2) \, dQ
\]

\[
= \frac{1}{K+1} \sum_{k=0}^{K} \left( \sigma_{M,k}^2 + (\mu_k^M - E(Q_{t+1}))^2 \right).
\]

Unconditional skewness is

\[
E \left[ (Q_{Mt} - E(Q_{Mt}))^3 \right] = \frac{1}{K+1} \sum_{k=0}^{K} \int (Q - E(Q))^3 \phi(Q; k) \, dQ
\]

\[
= \frac{1}{K+1} \sum_{k=0}^{K} \int \left( (Q - \mu_k^M)^3 + (\mu_k^M - E(Q))^3 \right) \phi(Q; k) \, dQ
\]

\[
+ \frac{3}{K+1} \sum_{k=0}^{K} \int (Q - \mu_k^M)^2 (\mu_k^M - E(Q)) \phi(Q; k) \, dQ
\]

\[
+ \frac{3}{K+1} \sum_{k=0}^{K} \int (Q - \mu_k^M) (\mu_k^M - E(Q))^2 \phi(Q; k) \, dQ,
\]

or

\[
E \left[ (Q_{Mt} - E(Q_{Mt}))^3 \right] = \frac{1}{K+1} \sum_{k=0}^{K} (\mu_k^M - E(Q))^3 + \frac{3}{K+1} \sum_{k=0}^{K} (\mu_k^M - E(Q)) \sigma_{M,k}^2.
\]

Expressing market returns as a sum of firm-level returns leads to

\[
= \frac{1}{N^3} \sum_{i=1}^{N} E \left[ (Q_{it} - E(Q_{it}))^3 \right]
\]

\[
+ \frac{3}{K+1} \frac{1}{N^3} \sum_{k=0}^{K} \sum_{i=1}^{N} \sum_{i' \neq i} (\mu_k^i - E(Q_i))^2 \left( \mu_k^{i'} - E(Q_{i'}) \right) + \frac{3}{K+1} \frac{1}{N^3} \sum_{k=0}^{K} \sum_{i=1}^{N} \sum_{i' \neq i} (\mu_k^i - E(Q_i)) \sum_{i' \neq i} \sigma_{i,k}^2
\]

\[
+ \frac{6}{K+1} \frac{1}{N^3} \sum_{k=0}^{K} \sum_{i=1}^{N} \sum_{i' \neq i} \sum_{l \neq i'} (\mu_k^i - E(Q_i)) \left( \mu_k^{i'} - E(Q_{i'}) \right) \left( \mu_k^l - E(Q_l) \right).
\]
Appendix B: Model with Correlated Cash Flows

In the correlated cash flow case, firm $i$’s persistent dividend factor is $F_{it} = \rho_{Fi} F_{i,t-1} + \varepsilon^F_{it}$, $0 \leq \rho_{Fi} \leq 1$, with $\varepsilon^F_{it} \sim N(0, \sigma^2_{Fi})$, and the transitory component is $\varepsilon^D_{it} \sim N(0, \sigma^2_{Di})$. For any two firms $i$ and $i'$, $E[\varepsilon^D_{it} \varepsilon^D_{i',t-s}] = \sigma^D_{ii'}$ and $E[\varepsilon^F_{it} \varepsilon^F_{i',t-s}] = \sigma^F_{ii'}$ when $s = 0$, and zero otherwise. I am interested in the case in which shocks have one or more common components that affect the cash flows of all firms in the economy in the same direction, $\sigma^D_{ii'}, \sigma^F_{ii'} \geq 0$. For simplicity, $E[\varepsilon^D_{it} \varepsilon^F_{it-s}] = 0$ for any two firms $i$ and $i'$ and any $s$. As in the main text, all dividend shocks are homoskedastic.

Denote by $Q_t^k = (Q_{1t}^k, \ldots, Q_{Nt}^k)^T$ the column vector of time $t$ stock returns. The superscript $k$ indicates that firms of type $k$ (if there are any) announce at time $t$. Again, with fixed heterogeneity in firm announcements, $k$ is a sufficient statistic for the heterogeneity in firm announcements at $t$.

Assuming that returns are jointly conditionally normal

$$Q_{t+1}^k | t \sim N(\mu_k, \Sigma_k),$$

investors’ problem yields the first-order necessary and sufficient condition

$$\theta_t = \gamma^{-1} \Sigma_k^{-1} \mu_k.$$

Imposing the equilibrium condition that the representative investor holds all shares in the market, $\theta_t = 1$, gives $\mu_k = \gamma \Sigma_k 1$.

Following the steps of the proof of Proposition 1 and assuming stationarity yields the equilibrium price function for firm $i$:

$$P_{it}^k = p_i^k + \Gamma_{ki} F_{it} + R^{-(K+1-k_i)} \sum_{j=0}^{k_i-1} \varepsilon^D_{i,t-j}.$$ 

The expression for the constants $p_i^k < 0$ is the same as in equation (8).

To solve for the equilibrium values of $\{\mu_k, \Sigma_k\}_{k}$, use the price function above to express excess returns as

$$Q_{it}^k = p_i^k - Rp_i^{k_i-1} + \Gamma_{ki} \varepsilon^F_{it} + R^{-(K+1-k_i)} \varepsilon^D_{it},$$

for any $k_i$. Then, the elements of $\Sigma_k$ are

$$\sigma^2_{ik} \equiv Var_t \left[ Q_{it+1}^k \right] = \Gamma_{k_i+1}^2 \sigma^2_{Fi} + R^{-2(K-k_i)} \sigma^2_{Di},$$

$$\sigma_{ii',k} \equiv Cov_t \left[ Q_{it+1}^k, Q_{i't+1}^{k_i'} \right] = \Gamma_{k_i+1} \Gamma_{k_{i'}+1} \sigma^F_{ii'} + R^{-2(K-k_i-k_{i'})} \sigma^D_{ii'}.\]
For each firm $i$, the conditional mean and volatility of the stock return increase monotonically and are convex in $k_i$, all else equal. As in the uncorrelated cash flow case, the conditional stock return variance increases with $k_i$, all else equal.

In this model, as in the uncorrelated cash flow case, the stock market equilibrium has a conditional CAPM representation. Let $\alpha \equiv 1/N$ and write $Q_{Mt}^k = \alpha^i Q^i_t$. Then, $\mu^M_k \equiv E_t \left[ Q_{Mt+1}^{k+1} \right] = \alpha^i \mu_k$ and $\sigma_{M,k}^2 \equiv E_t \left[ \left( Q_{Mt+1}^{k+1} - \mu^M_k \right)^2 \right] = \alpha^i \nu_k \alpha$. Then, using $\mu_k = \gamma \nu_k 1$, gives:

$$\mu_k = \beta_k \mu^M_k,$$

where $\beta_k \equiv \text{Cov}_k \left( Q_t^k, Q_{Mt}^k \right) / \sigma_{M,k}^2$ and $\alpha^i \beta_k = 1$.

If firm $i$ has a high expected return around its announcement event, it must also be that $\beta_k$ is high around the event. This systematic risk is driven by the volatility associated with the information flow in common factors. For example, if $F_{it} = F_t$ for all $t$ and $i$, and $\sigma^2_{it'} = 0$, then the economy has only one common factor, which is persistent. Shocks to this common factor, $\varepsilon_t^F$, affect stock returns of firms differently depending on how far each firm is from its respective payout event. This timing explains the dynamics in the conditional stock return moments because proximity to a payout event determines the impact of (systematic) information on returns. Consistent with this model prediction, Patton and Verardo (2010) show that daily firm betas increase by an economically and statistically significant amount around earnings announcement events.

After deriving the equilibrium conditional distribution of stock returns, it is straightforward to derive the equilibrium unconditional distribution following the same steps as in the main text. For $K \geq 1$, the unconditional distribution of stock returns for firm $i$ is a mixture of normals distribution with density

$$f \left( Q^i \right) = \frac{1}{K+1} \sum_{k=0}^K \phi \left( Q^i; \mu^i_k, \sigma^2_{ik} \right),$$

where $\phi(\cdot)$ is the normal density function, and for $K = 0$, returns are unconditionally normally distributed. The expressions for the unconditional mean and variance of stock returns and the unconditional skewness in stock returns are the same as in the main text. With correlated cash flows it is not possible to sign skewness because when $k_i$ changes other firms’ event time, say $k_{i'}$, also changes which may lead to non-monotonicity in the conditional return covariance between $i$ and $i'$ and hence in conditional mean returns for firm $i$. However, in all numerical examples studied this effect is dominated and firm-level skewness is positive.
The unconditional distribution of aggregate market returns is a mixture of normals distribution with
\[
f(Q_M) = \frac{1}{K+1} \sum_{k=0}^{K} \phi(Q_M; \mu_k^M, \sigma^2_{M,k}),
\]
and skewness in aggregate stock returns is as in equation (15) plus the following term:
\[
\frac{6}{K+1} \frac{1}{N^3} \sum_{k=0}^{K} (\mu_k^M - E(Q_M)) \sum_{i=1}^{N} \sum_{i' > i} \sigma_{i'i',k}.
\]
This term is likely to be positive because the return covariance is likely to be highest at event dates \(k\), where mean returns \(\mu_k^M\) are also likely to be higher. Hence in the correlated cash flow case, market skewness tends to be higher than in the uncorrelated case and there tends to be less conditional asymmetry in stock return correlations. However, numerical examples show that qualitatively the results in the main text apply also to this more general model.
References


Barro, R. J., 2006, Rare Disasters and Asset Markets in the Twentieth Century, Quarterly Journal of Economics 121, 823-866.


Chang, B. Y., Christoffersen, P., and Jacobs, K., 2009, Market Skewness Risk and the Cross-Section of Stock Returns, working paper, McGill University.


Figure 1: **Skewness in firm-level and aggregate stock returns.** The figure plots mean skewness in daily firm-level returns (dashed line) and skewness in the equally weighted market return (solid line), both computed using six months of trading data. Data comprise all firms in CRSP with complete daily return data by semester. Period of analysis is 1/1/1973 through 31/12/2009.
Figure 2: **Skewness decomposition.** The figure plots co-cov as a fraction of overall market skewness. Skewness is computed using equally weighted portfolios in six months of daily returns. Data comprise all firms in CRSP with complete daily return in the specific year and semester. The sample period is January 1, 1973 to December 31, 2009.
Figure 3: **Stock market skewness in various stock market configurations.** In Panel A, each stock market consists of two types of firms with cash payout dates separated by $k$ periods, where $k \in \{0, 1, ..., K\}$. In Panel B, each stock market consists of $k + 1$ different types of firms with cash payout dates of $0, 1, ..., k$. Each panel depicts market skewness (solid line) and mean firm skewness (dashed line). Parameters are: $K = 12$, $\sigma_D^2 = \sigma_F^2 = 1$, $\rho_F = 0.9$, $\gamma = 5$ and $R = 1.0025$. 
Figure 4: Decomposing Stock Market Skewness. Each stock market consists of $k + 1$ different types of firms with cash payout dates of 0, 1, ..., and $k$ as in Panel B of Figure 3. The figure depicts market skewness (solid) and its co-skewness component term $co-cov$ (dashed). Parameters are: $K = 12$, $\sigma_D^2 = \sigma_F^2 = 1$, $\rho_F = 0.9$, $\gamma = 5$, and $R = 1.0025$. 


Figure 5: **Exceedance correlations in various stock market configurations.** Each stock market consists of $k + 1$ different types of firms with cash payout dates of 0, 1, ..., and $k$ as in Panel B of Figure 3. The figure depicts the exceedance correlation in market upturns (solid line) and the ratio of exceedance correlation in downturns to the exceedance correlation in upturns (dashed line). Parameters are: $K = 12$, $\sigma_D^2 = \sigma_F^2 = 1$, $\rho_F = 0.9$, $\gamma = 5$ and $R = 1.0025$. 


Figure 6: **Contribution of each trading period to stock market skewness.** The stock market consists of firms with cash payout dates of 0, 1, ..., and 7. The figure plots the component of normalized skewness, $E(Q_t - E(Q_t))^3 / [E(Q_t - E(Q_t))^2]^{3/2}$, due to each trading period. Parameters are: $K = 12$, $\sigma_D^2 = \sigma_F^2 = 1$, $\rho_F = 0.9$, $\gamma = 5$ and $R = 1.0025$. 
Figure 7: **Histogram of announcement week.** The figure plots the empirical frequency by calendar week of cash payouts (Panel A) and earnings (Panel B) announcements. Data come from the merged Compustat/CRSP quarterly files. The sample period is 1973:Q1 to 2009:Q2. Observations with announcement date before the end of the quarter are dropped.
Figure 8: **Skewness and announcement portfolios.** The figure plots portfolio return skewness with 10% confidence bands. Portfolios are constructed by grouping firms that announce in the first week of the first quarter in the semester ($P_1$) with firms that announce in week $k$ of the first quarter in the semester ($P_k$), $k = 2, \ldots, 13$. Skewness is calculated using daily returns over six months. Portfolio returns are equally weighted. Portfolios are constrained to have the same number of firms, which is done by randomly dropping firms from the larger portfolios. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the merged Compustat/CRSP quarterly file and the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 9: **Skewness and announcement portfolios.** The figure plots portfolio return skewness with 10% confidence bands. Portfolios are constructed by grouping firms that announce between the first week of the first quarter in the semester ($P1$) and week $k$ of the first quarter in the semester ($Pk$), $k = 2, ..., 13$. Skewness is calculated using daily returns over six months. Portfolio returns are equally weighted. Portfolios are constrained to have the same number of firms, which is done by randomly dropping firms from the larger portfolios. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the merged Compustat/CRSP quarterly file and the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 10: **Mean firm skewness and announcement portfolios.** The figure plots the mean firm return skewness with 10% confidence bands. Portfolios are constructed by grouping firms that announce between the first week of the first quarter in the semester (P1) and week k of the first quarter in the semester (Pk), k = 2, ..., 13. Firm skewness is calculated using daily returns over six months. Portfolios Pk are constrained to have the same number of firms as is done in Figure 9. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the merged Compustat/CRSP quarterly file and the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 11: **Skewness and calendar week.** The figure plots the weekly component of market skewness with 10% confidence bands. Skewness is calculated using daily returns over six months. Portfolio returns are equally weighted. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 12: **Skewness in portfolios of varying size.** The figure plots mean skewness in daily returns from portfolios of size $N$. Skewness is computed using equally weighted portfolio returns and six months of daily data. The portfolios are constructed by randomly ranking the firms and then grouping them. If two firms are in the same portfolio when $N = 25$, then they will also be in the same portfolio for $N = 625$. The dash-dotted line plots firm-level skewness. The solid line, labeled *Market*, plots skewness of equally weighted returns of all firms in CRSP. The sample period is January 1, 1973 to December 31, 2009.
Figure 13: **Skewness and announcement portfolios using CAPM residuals.** The figure plots portfolio return skewness with 10% confidence bands. Portfolios are constructed by grouping firms that announce in the first week of the first quarter in the semester ($P_1$) with firms that announce in week $k$ of the first quarter in the semester ($P_k$), $k = 2, ..., 13$. Skewness is calculated using daily idiosyncratic returns over six months. Portfolio returns are equally weighted. Portfolios are constrained to have the same number of firms, which is done by randomly dropping firms from the larger portfolios. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the merged Compustat/CRSP quarterly file and the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 14: **Skewness and announcement portfolios using CAPM residuals.** The figure plots portfolio return skewness with 10% confidence bands. Portfolios are constructed by grouping firms that announce between the first week of the first quarter in the semester ($P_1$) and week $k$ of the first quarter in the semester ($P_k$), $k = 2, ..., 13$. Skewness is calculated using daily idiosyncratic returns over six months. Portfolio returns are equally weighted. Portfolios are constrained to have the same number of firms, which is done by randomly dropping firms from the larger portfolios. Confidence bands use the sample standard deviation of the estimated skewness values. Data are obtained from the merged Compustat/CRSP quarterly file and the CRSP daily return file. The sample period is January 1, 1973 to December 31, 2009.
Figure 15: **Stationary conditional cumulative distributions of earnings announcements.** Panel A classifies market upturns (downturns) as quarters preceded by a CRSP value-weighted return above (below) the historical median. Panel B classifies market upturns (downturns) as quarters preceded by a CRSP value-weighted return above the historical 75th percentile (below the 25% percentile). The distribution depicted using a dashed line conditions on market upturns and the distribution depicted using a solid line conditions on market downturns. Data on earnings announcements come from the merged Compustat/CRSP quarterly files. The sample period is 1973:Q1 to 2009:Q2. Observations with announcement date before the end of the quarter are dropped.