4
Information Flow, Disposition Effect, and the Skewness of Stocks

4.1. Introduction

Frequently risk is referred to as the standard deviation of returns (also called volatility), both in finance literature and in market jargon. Risk-return trade-off analyses are usually based on this single parameter of returns distribution. Options prices, for instance, are listed with implied volatility. But this single factor is a poor representation of returns distribution, at least because volatility is known to change over time. Also, if we take bank regulation (Basle Accord), risk is related to potential losses, as emphasized by Poon and Granger (2003). Thus, higher order moments of returns distributions, especially skewness, are also relevant to assess risk.

Not surprisingly, skewness[1] has been shown by Harvey and Siddique (2000) to be relevant for asset pricing. Chung, Johnson and Shill (2006) have shown that skewness and kurtosis may even substitute Fama and French's (1993) SMB and HML factors. Boyer, Mitton, Vorkink (2009) reported that even idiosyncratic skewness affects expected returns. They attribute this to a preference for positive skewness, in the same way as people make bets in lotteries. But it has also been shown that skewness is not persistent for individual stocks. Particularly, Singleton and Wingender (1986) found that "positively-skewed equity portfolios in one period are not likely to be positive skewed in the next time period."

The objective of this essay is to identify determinants of skewness.[2] Harvey and Siddique (2000) have shown that skewness decreases (or becomes more negative) with size3 and past returns. Analysis by Chen, Hong and Stein

---

[1] We consider that it is implicit in any reference to skewness or kurtosis that it refers to the skewness or kurtosis of the distribution of returns of individual stocks, unless it is specified otherwise.

[2] A positive skewness indicates that extreme positive returns are more probable than extreme negative returns, and vice-versa.

[3] In the present study, as well as in referred studies, firm size is measured by market capitalization, which is computed as the product of the price of each share of a stock by the number of shares

(2001) and Xu (2007) confirm these findings and add additional variables to test specific asset pricing models based on divergence of opinion. The novelty of the present study is to assume that the distribution of returns is composed by a superposition of two distributions, one related to news and the other unrelated to news. We conjecture which are the relevant driving factors of skewness in each of these distributions, and test whether proxies for these factors indeed affect skewness.[4]

To explain our conjectures we start by the part of returns distribution which is not related to public news (either idiosyncratic or systematic news). Some investors may, independently from public news, decide to buy or sell shares of a specific stock.[5] We will refer to trades that are not motivated by public news as “demand shocks”. We postulate that price changes which are not related to news arise to accommodate demand shocks. Our contribution is to provide empirical evidence supporting that a behavioral bias labeled “disposition effect”, by Shefrin and Statman (1985), contributes to negative skewness of the distribution of returns related to demand shocks. The “disposition effect” is the tendency of investors to hold stocks with negative cumulative returns (“losers”) for too long and sell stocks with positive cumulative returns (“winners”) too early. In the previous chapter we provided evidence that the disposition effect plays a role in the distribution of returns, particularly in the “asymmetric volatility” (the negative autocorrelation between volatility and lagged returns). In the presence of demand shocks for a stock, investors subject to disposition effect are more prone to sell the stock. This dampens positive returns for stocks with positive cumulative returns. It makes large positive returns be less frequent than large negative returns. Thus skewness should be more negative or less positive when cumulative returns are positive. Conversely, Shefrin and Stataman (1985) state

outstanding for that stock.

[4] Indeed, Chen, Hong and Stein (2001) use three different theories to explain the effect of the three main independent variables in their estimation of conditional skewness. Xu (2007) develops a model that explains positive skewness, but refers to other theories that might explain negative skewness.

[5] These decisions may be taken at random or based on private information, as in Kyle (1985). Traders that trade randomly were labeled “noise traders” by Kyle (1985), an expression that was widely used in finance literature afterwards. Trade decisions that are taken independently from news may also be motivated by idiosyncratic shocks to private investment opportunities (assuming financial markets are incomplete), as in Wang (1994). The essence of both models is that other traders (e.g. market makers) can not know whether they are trading with privately informed investors, because there are also investors without private information continuously

that to prevent from huge losses due to this propensity to hold stocks with negative cumulative returns, investors develop self-control mechanisms, such as a stop losses limit, at which they send sell orders at market price. This increases the possibility that investors coordinate selling for stocks with negative cumulative returns, originating a demand shocks that amplifies price drops. The consequences in both cases, positive and negative cumulative returns, are consistent with the stylized fact that skewness is negatively correlated with past returns. We use Grinblatt and Han's (2005) measure of capital gain as a proxy for the intensity of disposition effect to be more specific on testing its impact on skewness.

We turn now to the component of returns distribution related to news. Our contribution is to show empirical evidence supporting that the rate of information arrival concerning news about a stock is relevant to the stock's distribution of returns. Although there is frequent news affecting the results of a firm, its effects on results are disclosed at a lower frequency. Thus, the impact on returns at disclosure is the cumulative impact of the pieces of news that have arrived so far. This makes the distribution of returns has different characteristics than the distribution of the impact of each independent piece of news on stocks prices. Our approach is based on the jump-diffusion model developed by Press (1967). In this model the cumulative firm return over a fixed period of time is composed by the information arrivals within the period. The number of pieces of news that arrive at each period is assumed to be distributed as a Poisson process.[6] Press has shown that the expected skewness of returns distribution has the same sign of the mean return related to information arrivals. This is also consistent with the finding of Xu (2007) that skewness is positively correlated with mean current returns. Since expected returns are positive, otherwise investors would not pay prevailing prices, expected skewness of news is also positive. Because firm returns represent the cumulative effect of news, the distribution of returns will get more skewed as the information disclosure gets less frequent. This is consistent with a negative partial correlation between firm size and skewness, as reported by Harvey and Siddique (2000), since we may expect smaller firms to be followed by less institutional

trading.

[6] This approach of assuming that the returns we observe are composed by pieces of news was labeled "Mixture of Distribution Hypothesis", by Epps and Epps (1976). Differences in this approach stem from the assumptions about the distribution of the number of pieces of news that arrive at each fixed period of time, which in the case of Press (1967) follows a Poisson

investors and analysts, and thus incorporates news to prices less frequently. As Chen, Hong and Stein (2001) and Xu (2007), we associate volume with divergence of opinion. Divergence of opinion stems from different evaluations of the impact of news on the future cash flow of the firm. Thus, extremely low turnover for a stock indicates either absence of news, or that news is not being fully incorporated into prices.[7] To be more specific on testing the rate of news updating, we will use the number of days with extremely low turnover within a year as a proxy for low frequency of incorporation of news to prices. Our analysis shows a negative partial correlation between skewness and this proxy.

In the next section we describe determinants of the skewness of the underlying distributions that we are assuming to compound returns. In Section 4.3 we describe the sources of data, sample selection and procedures to determine secondary variables. In Section 4 we present the pooled OLS regressions, splitting the sample into sub-periods and into groups formed by size. Section 4.5 concludes.

4.2.
Skewness related to news and unrelated to news

Empirical evidence suggests that news arrival does not fully explain variations in stock returns. Roll (1988) shows that even after controlling for a market index return and the average return of firms in the same industry, and after removing from the sample of returns the days with disclosure of idiosyncratic news relative to the firm, most of the variation in returns remains unexplained. (The R^2 of the regression of returns of a stock on market and industry returns remains as low as 20% even in the sample that excludes idiosyncratic news).[8]

Even if there is very frequent news, the actual impact on a firm may be unclear until information is disclosed, for instance by a formal announcement. This is consistent with the increase in trading volume around announcement days

process.

[7] The works of Chen, Hong and Stein (2001) and Xu (2007) state that if there are short selling constraints, investors may stop trading, because the optimists hold the stock and the pessimists do not. If divergence of opinion is great, only very significant news, or the composition of many small pieces of news, will lead to increase in trading.

[8] Consistent with Roll's analysis, Shiller (1981) and West (1988) show that prices volatility in the stock market are much greater than the ex-post volatility of fundamental value calculated with the discount value of dividends.

reported by Chae (2005). The return on the announcement day then comprises the composed effect of many pieces of news in the preceding days. Even if the impact on price of each piece of news is a sample of the same non-skewed distribution, the distribution of returns in the announcement day may be skewed. This way to approach return as the total impact of independent events was labeled by Epps and Epps (1976) as the Mixture of Distributions Hypothesis. In the way Press (1967) formulates this approach, the number of news comprised in each measure of return follows a Poisson process with mean number of pieces of news in each measure equal to λ (the single parameter of the Poisson distribution). The first three moments of the distribution of returns that follows this process are given by:

$$m_1 = \mathrm{E}[r_{j,d}] = \lambda\,\theta \tag{15a}$$

$$m_2 = E\left[\left(r_{j,d} - \bar{r}_j\right)^2\right] = \lambda\left(\theta^2 + \sigma_n^2\right) \tag{15b}$$

$$m_3 = E\left[\left(r_{j,d} - \bar{r}_j\right)^3\right] = \lambda\theta\left(\theta^2 + 3\sigma_n^2\right) \tag{15c}$$

where θ is the mean and σ_n^2 is the variance of the distribution of the impacts of news on returns, which is assumed to be normal. These are not necessarily days of official announcements. Idiosyncratic information may be conveyed, for instance, through trading, when an informed investor accepts to pay a higher price to buy more shares. We see in the set of equation (15a) – (15c) that all moments increase with parameters λ and θ.

The mean impact of news (θ) should be positive, otherwise risk averse or risk neutral investors would not buy or hold stocks at prevailing price, when there is a risk free asset with positive return as an alternative investment opportunity.[9] For the same rate of news arrival, if prices incorporate news less frequently, λ tends to be greater. This means that the lower the rate of information arrival, the more skewed is the distribution of returns. Because size is considered a proxy for the inverse of information asymmetry, as in Llorente, Michaely, Saar and Wang (2002), stocks of small firms are good candidates for stocks with prices that less

[9] This is a generalization. People may engage into financial transactions even when expected return is negative, if this transaction provides more stable cash flow over time. This is clearly the case of insurance. However stocks returns are almost always positively correlated with the business cycle, and thus can not provide a compensation for periods of lower income. That is why the generalization we use is valid taking a large sample of stocks.

frequently incorporate news into prices. Then, the smaller the firm, the more skewed is the distribution of the part of returns related to news. Because expected skewness is positive, the smaller the firm, the more positive is expected skewness. This is consistent with the negative partial correlation between size and skewness reported by Harvey and Siddique (2000), Chen, Hong and Stein (2001), and Xu (2007).

From equation (15c), skewness is positively correlated with expected return, represented by θ. Variables that affect the expected return of a stock will affect its skewness in the same direction. De Bondt and Thaler (1985) argue that investors tend to overreact to recent performance, and this causes a reversal of returns within periods of three years. That is, due to a behavioral bias, we may expect higher returns in the next years if the previous years have been bad and vice-versa. De Bondt and Thaler (1987) and Lakonishok, Shleifer and Vishny (1994) provide additional empirical evidence to this long run reversal, and do not find evidence that higher risk might explain the higher return of past losers (stocks with lower return in the past).[10] This is consistent with the negative correlation between skewness and past returns reported by Harvey and Siddique (2000), Chen, Hong and Stein (2001), and Xu (2007).

So far we have argued that the Mixture of Distributions Hypothesis, more specifically the model developed by Press (1967), is consistent with previous works. Now, to be more specific on evaluating the effect of the frequency of incorporation of news into prices (λ), we will argue that days with extremely low turnover are days when no information is conveyed through prices. By one side, we know, from Chae (2005) that associated to public announcements, which are undoubtfully days when news are incorporated into prices, turnover is high. By another side, trading volume has increasingly been associated with divergence of

[10] It is said that there is a long run reversal because the reversal occurs for stocks that presented poor performance for periods of three to five years. Jegadeesh and Titman (1993) show that, taking periods of three to twelve months, instead of reversal, there is a continuation of past performance, the "momentum effect". Fama and French (1996) argue that the long term reversal in compatible with three-factor model of returns, but there is no clear reason why covariance with the factor that explains this long term reversal (the HMB factor) represents greater risk to investors. Besides, this model fails to explain the "momentum effect". Substituting in the three factor model the HMB factor by its inverse or other negatively correlated variable, one could explain the "momentum effect", leaving the long run reversal unexplained. Anyway, being a behavioral bias or a rational equilibrium, the long run reversal, together with equation (15c) indicates why skewness might be negatively correlated with past returns.

opinion. Chen, Hong and Stein (2001) and Xu (2007), to take only articles related to skewness, assume this relation is true.[11] Divergence of opinion stems from different evaluations of the impact of news on the future cash flow of the firm. Thus, extremely low turnover for a stock indicates either absence of news or that news is not being fully incorporated into prices.[12] We will use the number of days with extremely low turnover within a year as a proxy for λ. That is, the more days when news is not incorporated into prices, the greater is the number of pieces of news that are incorporated into prices when this incorporation occurs. Thus, we expect a positive partial correlation between skewness and the number of days with extremely low turnover. Because it is quite an arbitrary task to determine what is an extremely low turnover, we use two different thresholds, and define them in a way that a significant fraction of our sample has at least one day with extremely low turnover per year.

We turn now to the contribution of the part of returns unrelated to news. This part of return arises from private motivation of investors. We will call the trades that comprise these private motivations demand shocks. In Kyle (1985), demand shocks may arise from private information or random trading by "noise traders". Noise trading may stem from behavioral bias, as in DeLong, Shleifer, Summers and Waldmann (1990). Demand shocks may also arise under rational expectations equilibrium, from a shock to private investment opportunities in incomplete markets, as in Wang (1994). In the presence of a demand shock, prices change, because investors that do not have a private motivation, will charge a premium to engage into trading. In a competitive market, the premium should, at first sight, tend to zero. However, because there may be informed traders, as in both Kyle (1985) and Wang (1994), there is a problem of adverse selection. So, even in a competitive market, investors charge a strictly positive premium, in equilibrium, to account for the potential loss of engaging into trading with an informed investor.

In the previous chaoter we argued that in the presence of demand shocks, a behavioral bias labeled "disposition effect", by Shefrin and Statman (1985),

---

[11] An early reference on the association between volume and divergence of opinion is Harris and Raviv (1993). A review on this literature is provided by Hong and Stein (2007).

[12] The works of Chen, Hong and Stein (2001) and Xu (2007) state that if there are short selling constraints, investors may stop trading, because the optimists hold the stock and the pessimists do not. If diverge of opinion is great, only very significant news, or the composition of many

affects conditional volatility. The “disposition effect” is the propensity of some investors to sell stocks with high past returns (“winners”) too soon and hold stocks with low past returns (“losers”) for too long. Also according to Shefrin and Stataman (1985), to avoid huge losses due to the propensity to hold stocks with negative cumulative returns, investors develop self-control techniques. A widely used technique of this kind is the stop-losses price limit, at which investors send sell orders at market price. So, investors subject to disposition effect start selling stocks with positive cumulative return when prices are rising, due to demand shocks, damping positive returns. Conversely, when prices are dropping and cumulative returns are negative, they start selling because stop-losses price limits are reached, amplifying prices drop.

If, positive returns are damped by disposition effect, for stocks with positive cumulative returns, and negative returns are amplified for stocks with negative returns, then skewness should be negatively correlated with cumulative past returns, which is a stylized fact in finance literature However, for the disposition effect, stocks have to present cumulative returns from the point of view of the investors subject to disposition effect. This not necessarily coincides with having cumulative positive or negative return in the previous year, or other fixed period of time. For stocks with high turnover the reference price for investors will be more recent than for stocks with low turnover. To be more specific on testing the impact of disposition effect on skewness, we use Grinblatt and Han’s (2005) measure of capital gain (g) as a proxy for the intensity of disposition effect. In this measure, daily turnover is taken into account to compute an average reference price. We expect a negative partial correlation between skewness and capital gain.

4.3. Data

We use daily returns, prices, trading volumes (number of shares traded) and shares outstanding from CRSP database, for the period beginning in January 1963, and ending in December 2007. Only common stocks listed at NYSE are considered. Volatility (standard deviation) and skewness of daily returns distributions are calculated for each year and each stock. Thus, one observation in

small pieces of news, will lead to increase in trading.

our analyses corresponds to one stock-year. The choice of periods of one year follows Xu (2007), and allows a good balance between sample size (number of observations, i.e. stock-years), and accuracy of the estimates of volatility and skewness. Besides, periods of one year are preferable to periods of one semester, as used by Chen, Hong and Stein (2001), because it neutralizes the January effect,[13] since all observations contain January returns. To improve accuracy of these estimates, we remove from our sample any observation (stock-year) with missing data (even with one day of missing data in a whole year). This procedure also has the benefit of removing from our sample periods with IPOs or when firms are delisted, which may present idiosyncrasies in the distributions of returns that we do not want to affect our analyses. Because simple returns are obviously more positively skewed, daily returns are taken as logarithm changes in prices:

$$r_{j,d} = ln\ (P_{j,d}) - ln\ (P_{j,d-1}) = ln\ (1 + x_{j,d}), \tag{16}$$

where $P_{j,d}$ and $x_{j,d}$ are, respectively, the price and the simple return of stock j on day d.

We are interested in the analysis of the drivers of skewness in the cross section of stocks. To emphasize the contribution of explicative variables to the skewness of individual stocks, we calculate idiosyncratic returns. Chen, Hong and Stein take the individual part of returns either as excess returns over market return, or as beta adjusted returns, that is, the excess return over the expected return from the CAPM. However these procedures do not remove expected returns related to higher order moments. Because we motivate this work arguing that higher order moments do play a role in asset pricing, we take idiosyncratic returns from an asset pricing model consistent with this argument. As Harvey and Siddique (2001), we base our empirical analyses on the conditional form of Kraus and Litzenberger's (1976) model.[14] In their model the skewness of returns is relevant for the utility function. Expected returns are then written as:

$$E_t[r_{j,t}] = A_{j,t}\ E_t[r_{M,t}] + B_{j,t}\ E_t[r^2_{M,t}] \tag{17}$$

[13] A recent analysis of the January effect is provided by Chen and Singal (2004).
[14] The model on which Harvey and Siddique (2001) base their empirical analysis is the conditional version of Karus and Litzenberger's (1976) model.

where $r_{j,t}$ is the expected return of stock j in period t, E_t is the expectation operator, M denotes the market index, and A_t and B_t are conditional parameters of the model. We take idiosyncratic returns $r_{j,t,i}$ as the difference between actual and expected return for each period t:

$$r_{j,t,i} = r_{j,t} - E_t[r_{j,t}] \tag{18}$$

where $E[r_{j,t}]$ is given by equation (17). We will compute $A_{j,t}$ and $B_{j,t}$ for each year, from a linear regression of $r_{j,t}$ on $r_{M,t}$ and $r^2_{M,t}$.

The mean $\bar{r}_{j,y}$, the standard deviation $\sigma_{j,y}$, and the skewness $sk_{j,y}$ for the distribution of the daily returns of stock j in year y are estimated as:

$$\bar{r}_{j,y} = \frac{1}{D_{j,y}} \sum_{d=1}^{D_{j,y}} r_{j,d,i} \tag{19}$$

$$\hat{\sigma}_{j,y} = \left[\frac{\sum_{d=1}^{D_{j,y}} \left(r_{j,d,i} - \bar{r}_{j,y}\right)^2}{\left(D_{j,y} - 1\right)} \right]^{\frac{1}{2}} \tag{20}$$

$$sk_{j,y} = \frac{\frac{D_{j,y}}{\left(D_{j,y}-1\right)\left(D_{j,y}-2\right)} \sum_{d=1}^{D_{j,y}} \left(r_{j,d,i} - \bar{r}_{j,y}\right)^3}{\hat{\sigma}^3_{j,y}} \tag{21}$$

where $r_{j,d,i}$ is the idiosyncratic return (logarithm change in price) of stock j on day d, and $D_{j,y}$ is the number of non missing daily returns for stock j in year y. To keep consistency, cumulative return in a year ($R_{j,y}$) is the logarithm of the price changes between the last day of the previous year and the last day of the current year. That is:

$$R_{j,y} = \sum_{d=1}^{D_{j,y}} r_{j,d,i} \tag{22}$$

To evaluate the impact of disposition effect on skewness, we use a measure $g_{j,d}$ of capital gain, similar to the one proposed by Grinblatt and Han (2005):[15]

$$g_{j,d} = \sum_{k=1}^{250} w_{j,d-k} \left[\prod_{n=0}^{k-1} \left(1 + r^{*}_{j,d-k}\right) - 1 \right] \tag{23a}$$

$$w_{j,t} = \frac{V_{j,t-k} \prod_{i=0}^{k-1} \left(1 - V_{j,t-i}\right)}{\sum_{k=1}^{250} \left[V_{j,t-k} \prod_{i=0}^{k-1} \left(1 - V_{j,t-i}\right) \right]} \tag{23b}$$

where $V_{j,d}$ is the turnover of stock *j* on day *d*, computed as the ratio of total shares traded on that day by shares outstanding, and $r^{*}_{j,d}$ is the ex-dividend return of stock *j* on day *d*. For each stock-year we compute the mean capital gain as:

$$\bar{g}_{j,y} = \frac{1}{D_{j,y}} \sum_{d=1}^{D_{j,y}} g_{j,d} \tag{24}$$

This will be our proxy for the intensity of disposition effect. Because $\bar{g}$ is a linear combination of daily returns, one might argue that this variable captures the same effect as cumulative returns *R*, from equation (22). But the *g* of the first day of a year has only returns from the past year, and the *g* of the last day of a year has only returns from the same year. However, partial correlation between skewness and the cumulative return of previous and current year are negative and positive, respectively. Thus, if mean $\bar{g}$ captures the effects of *R*, then its sign should be

[15] There are three differences from our measure to the measure proposed by Grinblatt and Han (2005). The first is the use of compounded ex-dividend returns, instead of price differences. We do so because in our database prices are the nominal closing prices on each day. Any split or grouping of shares would cause problems to our measure. The ex-dividend returns are adjusted for splits, groupings and other price changes that do not impact the wealth of shareholders (except dividend distribution). The second is the use of daily data (returns and turnovers), instead of weekly data. This was done for simplicity, since our database comprises daily data. The third difference is the use of 250 past days (approximately 1 year), instead of 250 past weeks (approximately 5 years). We chose this option because we are using compounded ex-dividend returns to calculate effective (instead of nominal) price changes. To improve the accuracy of this procedure we discard measurements with any missing data in the period over which capital gain is calculated. Using 5 years of data would reduce our sample and remove from it new firms.

ambiguous. One might otherwise argue that the weights w, from equation (23b), are increasing as the day of a daily return gets closer to the day for which g is calculated according to equation (23a). Consequently mean $\overline{g}$, from equation (24), should represent better the returns of the contemporaneous year than the returns of the previous year. (Indeed the correlation between $\overline{g}_{j,y}$ and $R_{j,y}$, over all observations, is 0.79, while the correlation between $\overline{g}_{j,y}$ and $R_{j,y-1}$ is 0.28.) However we predict that the partial correlation between $sk_{j,y}$ and $\overline{g}_{j,y}$ is negative, while previous works show that the partial correlation between skewness and contemporaneous cumulative return is positive. Thus, if there is an unambiguous negative partial correlation between skewness and contemporaneous $\overline{g}$, it can not be attributed to the same effect as cumulative returns.

Our proxy for firm size is the mean market capitalization in a year[16] $mc_{j,y}$ calculated as:

$$\overline{mc}_{j,y} = \frac{1}{D_{j,y}} \sum_{d=1}^{D_{j,y}} P_{j,d} O_{j,d} \tag{25}$$

where $P_{j,d}$ and $O_{j,d}$ are, respectively, the price and the number of shares outstanding for stock j on day d.

Finally, the proxy used for frequency of news updating is the number of days with extremely low turnover ($f_{j,y}$). The definition of extremely low turnover is quite arbitrary and for this reason we use two different thresholds: 0.001% and 0.01% daily turnover, which corresponds to approximately 0.25% and 2.5% yearly turnover, respectively. A lower threshold would make most observations have the proxy variable $f_{j,y}$ equal to zero. Variable f is not the same as turnover. Indeed, size (mc) explains more of the variance of f in the cross-section of stocks than turnover.[17]

[16] Some researchers prefer to use as proxy of size the market capitalization at the end of the previous year (when working with frequencies equal to or higher than one observation per year). Because prices and skewness are auto-correlated and there is no obvious causality relation, we do not see any benefit in using a previous measure of market capitalization. Conversely, using contemporaneous data better reflects the partial correlation between the two variables. Also, the use of a mean value instead of the market capitalization of a single day improves the accuracy of the estimate. Our choice follows Xu (2007).

[17] We determine the percentage of the variation in the cross section of f that is explained by market

Table 11 - Summary Statistics (sub-samples selected by period)

All variables are calculated by stock per year. Only observations (stock-years) with data for all trading days in the year are considered. Mean values over all observations in the period are presented in the table, with corresponding standard deviations in parenthesis. *R* is yearly idiosyncratic return, $\hat{\sigma}$ and *sk* are respectively the standard deviation and the skewness of daily idiosyncratic returns, *mc* is the mean market capitalization, *turnover* is mean daily turnover, $\overline{g}$ is the mean capital gain, where daily capital gain is determined through equations (23a) and (23b), similarly to Grinblatt and Han (2005), *f* is the number of days with extremely low volume (either *turnover* < 0.01%, or *turnover* < 0.001%), R_m is the cumulative return of the market (CRSP) index, and sk_m is the skewness of daily market returns.

Period	1965 – 2007	1965 – 1979	1980 – 1993	1994 – 2007
R	0.001 (0.457)	-0.001 (0.411)	0.006 (0.511)	-0.008 (0.457)
$\hat{\sigma}$	0.020 (0.012)	0.020 (0.009)	0.021 (0.014)	0.020 (0.012)
Sk	0.149 (1.075)	0.270 (0.695)	0.226 (1.003)	0.017 (1.289)
ln(*mc*)	12.9 (1.8)	11.9 (1.5)	12.8 (1.7)	13.5 (1.8)
Turnover	78% (95%)	32% (34%)	63% (58%)	115% (121%)
$\overline{g}$	0.033 (0.128)	0.017 (0.137)	0.040 (0.136)	0.039 (0.116)
f (*tu*< 0.01%)	10.0 (26.2)	16.1 (29.7)	9.7 (24.8)	6.7 (24.1)
f (*tu*< 0.001%)	2.8 (11.9)	3.4 (10.6)	2.8 (11.8)	2.4 (12.7)
R_m	0.073 (0.158)	0.024 (0.182)	0.101 (0.115)	0.085 (0.160)
sk_m	-0.301 (0.781)	-0.005 (0.348)	-0.642 (0.101)	-0.261 (0.085)
No. of observs.	69,845	18,598	19,737	31,510

capitalization and turnover, through regressing *f* on both variables and then on each one separately. We run panel data regression with random effects and consider the R^2 of the between estimation. For a reference on this type of analysis see Wooldridge (2002). For *f* defined with threshold daily turnover of 0.01%, the between R^2 when we regress *f* on the logarithm of market capitalization is 6.7%, while the between R^2 when we regress *f* on detrended turnover is 0.6%. Detrended turnover is determined similarly to Chen, Hong and Stein (2001), by subtracting from the turnover of a year the mean turnover of the two preceding years. They use this procedure because they “want to eliminate any component of turnover that can be thought of as a relatively fixed firm characteristic.” Indeed, if we use turnover instead of detrended turnover, the between R^2 is 6.8%. If we use the definition of *f* with threshold daily turnover of 0.001%, the between R^2 when we regress *f* on the logarithm of market capitalization, increases to 10.8%. However, the between R^2 of the regressions on detrended turnover and on turnover decrease to 0.2% and 1.7%, respectively.

We present in Tables 11 and 12 summary statistics for all variables specified in equations (20), (21), (22), (24), and (25), plus variable *f*, turnover and the yearly cumulative return and yearly skewness for the distribution of returns of the aggregate market index (the CRSP value weighed index in our case). In Table 11 we divide our sample in three periods. The first period has one year more than the others and corresponds to a period used in Xu (2007). In Table 12 we classified each observation (that is, each stock-year), by size, within the cross-section of each year. Stocks were classified into five equal sized groups.[18] The proxy for size was mean market capitalization computed as in equation (25). Except for the number of observations, values presented in Tables I and II are mean values for respective variable, with standard deviation in parentheses, calculated with all observations (stock – years).

As can be seen in Table 11, to determine variable *f*, a turnover threshold greater than 0.01%, corresponding to a 2.5% yearly turnover, barely could be classified as extremely low turnover. Taking for instance 0.1%, corresponding to 25% yearly turnover, it would approach the mean yearly turnover. In Table 11 we can also see that mean yearly turnover is increasing over time, raising the question whether threshold should be fixed over time. Indeed, as we can see in Figure 3, the mean value of *f* is decreasing over time. However, Figure 3 also shows that mean skewness also presents a downward trend. This makes us comfortable to use fixed thresholds.

It is also worth of note in Table 11 that the skewness of returns distribution is negative for the market index, while the mean skewness of idiosyncratic returns is positive. The occurrence of opposite signs for the skewness of the market index and the mean skewness of individual stocks had already been reported, for instance, by Campbell, Lo and MacKinlay (1997). We notice in Figure 3 a downward trend of the mean skenwess of individual stocks. This induces the thought that whatever makes individual stocks have positive skewness, besides varying in the cross-section of stocks, probably has also changed over time. This is the case of the frequency at which news are incorporated into prices, which we try to measure through variable *f*, also depicted in Figure 3. In Table 12 it gets

[18] The groups with smaller firms had fewer observations, whenever the number of stocks in our sample, for a specific year, was not a multiple of 5.

clear that skewness and *f*, besides seeming to be correlated in the time series, also seem to be correlated in the cross-section of stocks.

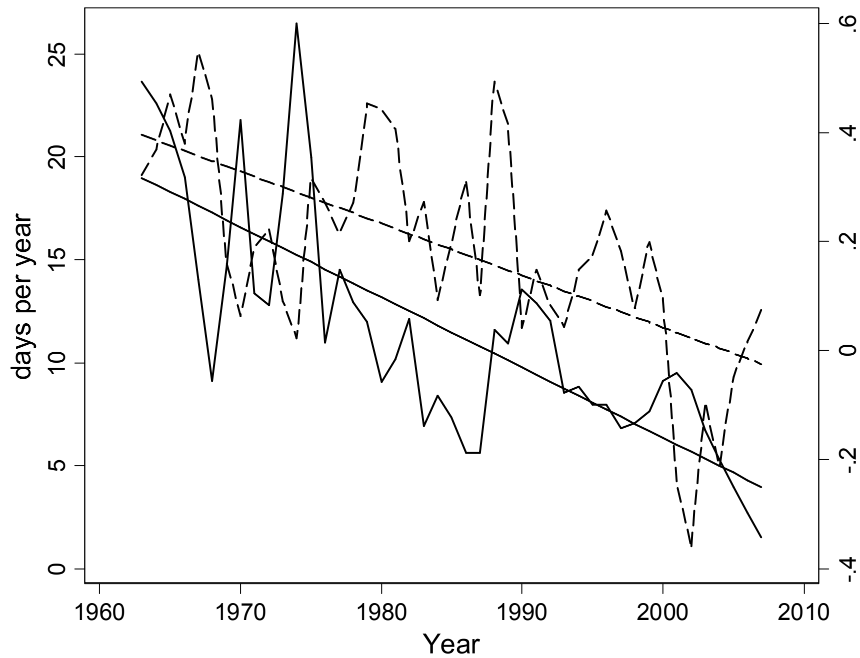

Figure 3. Evolution of mean *f* and mean *sk*. Variable *f*, i.e. the number of days with extremely low volume (daily turnover below 0.01%), is calculated for each stock-year. Then, for each year, the mean *f* of all stocks is calculated and plotted in the solid line, with a trend line. Values of *f* are presented in the left vertical axis. The skewness of idiosyncratic daily returns is calculated for each stock-year. Then, for each year, the mean idiosyncratic skewness of all stocks is calculated and plotted in the dashed line with a trend line. Values of idiosyncratic skewness are presented in the right vertical axis.

An interesting observation in Table 12 is that mean idiosyncratic yearly return is increasing with size. This seems contradictory with the size anomaly, first reported by Banz (1981). However, we classify stocks based on mean market capitalization in a calendar year and use yearly return in the same year to calculate mean idiosyncratic yearly return. This means that idiosyncratic yearly return is related to market capitalization: the lower the return, the higher the probability that a stock will be classified in a lower quintile, and vice-versa. Indeed Schwert (2003) reports that funds that try to mimic the strategy described by Banz (1981) did not present higher return than CAPM's prediction, between 1982 and 2002. In the Appendix we analyze this issue and provide empirical support to these two arguments.

Table 12 - Summary Statistics (sub-samples selected by size)

The sample period is from January 1965 to December 2007. All variables are calculated by stock per year. Only observations (stock-years) with data for all trading days in the year are considered. Mean values over all observations classified in the group are presented in the table, with corresponding standard deviations in parenthesis. *R* is yearly idiosyncratic return, $\hat{\sigma}$ and *sk* are respectively the standard deviation and the skewness of daily idiosyncratic returns, *mc* is the mean market capitalization, *turnover* is mean daily turnover, β is the CAPM's beta calculated using stock and market (CRSP index) daily returns within each year, $\overline{g}$ is the mean capital gain, where daily capital gain is determined through equations (23a) and (23b), similarly to Grinblatt and Han (2005), *f* is the number of days with extremely low volume (either *turnover* < 0.01%, or *turnover* < 0.001%). Size quintiles are determined within the sample, for each year, based on mean market capitalization (*mc*).

Size Group	1 (small)	2	3	4	5 (big)
R	-0.134 (0.642)	-0.007 (0.459)	0.026 (0.412)	0.050 (0.358)	0.071 (0.312)
$\hat{\sigma}$	0.027 (0.019)	0.020 (0.011)	0.019 (0.009)	0.018 (0.007)	0.016 (0.007)
Sk	0.281 (0.980)	0.183 (1.078)	0.132 (1.168)	0.093 (1.137)	0.059 (0.982)
ln(*mc*)	10.7 (0.9)	11.9 (0.8)	12.8 (0.8)	13.7 (0.8)	15.2 (1.3)
Turnover	61% (81%)	70% (88%)	82% (105%)	89% (107%)	81% (89%)
$\overline{g}$	-0.003 (0.151)	0.025 (0.130)	0.040 (0.123)	0.049 (0.114)	0.056 (0.105)
f (*tu*< 0.01%)	19.7 (30.7)	10.6 (24.0)	9.1 (24.5)	6.7 (23.5)	6.0 (27.4)
f (*tu*< 0.001%)	9.2 (21.2)	2.6 (9.9)	1.3 (6.9)	0.8 (7.4)	0.4 (4.8)
No. observs.	14,357	14,376	14,370	14,376	14,392

4.4.Regressions

We run the following regression, using pooled OLS:

$$sk_{j,y} = a + \sum_{k=0}^{3} b_k R_{j,y-k} + c_1 \ln\left(mc_{j,y}\right) + c_2 \overline{g}_{j,y} + c_3 f_{j,y} + \sum_{s=1}^{Y-1} e_s I_s + \varepsilon_{j,y} \quad (26)$$

where a, b_k, c_1, c_2, c_3, c_4 and e_s are the coefficients to be determined, I_s are individual dummies for each of the Y years in the sample (except one to avoid linear dependence of the dummies), j indexes stocks, y indexes years, sk is the estimate of skewness, from equation (21), when indexed by m, sk is the skewness of the distribution of CRSP index returns, R is the yearly return, from equation (22), $\overline{g}$ is the capital gain (proxy for intensity of disposition effect), from equation (24), f is the number of days with extremely low turnover (proxy for the inverse of frequency of news updating into prices), and mc is market capitalization, from equation (25). The indexes j and y identify individual stocks and years, respectively.

Differently than previous works by Chen, Hong and Stein (2001) and Xu (2007) we do not use volatility neither detrended turnover as explanatory variables in equation (26), because they do not play a clear role in our argument. Xu (2007) makes no argument for the use of volatility in the regression and finds ambiguity in the sign of the coefficient of volatility, depending on the period and on the set of explanatory variables. Chen, Hong and Stein (2001) include volatility because they are not making a contemporaneous analysis, but trying to forecast volatility, and thus they want to make sure that their explanatory variables are not forecasting volatility, which might be correlated with contemporaneous skewness. Both referred articles included detrended turnover because they made specific hypothesis concerning the relation between turnover and skewness that they wanted to test. Anyway, we ran regressions including both, volatility and detrended turnover, and the results obtained (not reported here) are qualitatively the same, with minor quantitative changes. Indeed, if we departure from the explanatory variables reported as relevant for skewness by Harvey and Siddique (2000), that is contemporaneous return, past returns and size, the increase in the R^2 of the regression is much lower when we include detrended turnover and volatility, than when we include capital gain ($\overline{g}$) and the number of days with extremely low turnover (f). (These regressions are also omitted in this report.)

Because we use two lags of yearly cumulative returns, our sample effectively starts in 1965, despite we have daily data since January of 1963. For robustness check, we split the sample period in three: 1965 through 1979, 1980 through 1993, and 1994 through 2007. The first subsample coincides with one of the sub-samples used by Xu (2007), allowing comparison of results, even though our sample selection is not exactly the same. Table 13 shows the results of regressing equation (26), with a column for each period considered, and the t statistics in parenthesis. (The estimated values for coefficients e_s of the year dummies I_s are not reported to save space.) Standard errors used in t statistics are robust to heteroskedasticity. The daily turnover threshold used in the definition of f is 0.01%. Results (not reported here) are very similar when we use the threshold 0.001%. We see in Table 13 that all coefficients present expected signs in all periods, being statistically significant at less than 1% level (except for the coefficient of variable f, in the period 1965 – 1979, which is statistically significant at 10% level).

Table 13 -
Determination of skewness (sub-samples selected by year): pooled regression

The dependent variable is idiosyncratic skewness calculated from 1 year's worth of each stock idiosyncratic daily returns if there is data for all trading days in the calendar year for the stock. R_y, R_{y-1}, R_{y-2} , R_{y-3} are idiosyncratic returns of the current year and the 3 previous years, *ln*(*mc*) is the logarithm of mean market capitalization, $\overline{g}$ is the mean capital gain, where daily capital gain is determined through equations (23a) and (23b), similarly to Grinblatt and Han (2005), and *f* is the number of days with extremely low volume (either *turnover* < 0.01%). All regressions include time dummies (unreported) for each year; *t*-statistics, in parentheses, are adjusted for heteroskedasticity.

Periods	1965 – 2007	1965 – 1979	1980 – 1993	1994 – 2007
R_y	0.844 (25.9)	0.764 (17.7)	0.652 (14.7)	1.036 (20.5)
R_{y-1}	-0.239 (-16.7)	-0.100 (-4.9)	-0.227 (-10.5)	-0.265 (-11.1)
R_{y-2}	-0.175 (-12.3)	-0.119 (-6.3)	-0.115 (-5.1)	-0.233 (-9.7)
R_{y-3}	-0.150 (-10.4)	-0.087 (-4.9)	-0.099 (-4.4)	-0.186 (-7.9)
ln(*mc*) (x 10^3)	-42.8 (-15.6)	-97.4 (-25.8)	-43.5 (-8.3)	-20.2 (-4.8)
$\overline{g}$	-0.556 (-6.7)	-0.817 (-7.4)	-0.272 (-2.5)	-0.757 (-5.4)
f (x 10^3)	2.20 (10.7)	0.44 (1.9)	3.52 (7.4)	2.84 (7.6)
Intercept	0.844 (20.9)	1.577 (30.1)	0.841 (10.9)	0.586 (9.6)
R^2	11.7%	15.5%	9.8%	11.3%
No. of observs.	53,496	13,713	15,673	24,110

In Table 14 we take sub-samples by size. Each year, observations are classified in five size quintiles. We then regress equation (24) five times considering only observations classified in each of the quintiles. Again the daily turnover threshold used in the definition of *f* is 0.01%, and results are very similar when we use the threshold 0.001%.

Table 14 –
Determination of skewness (sub-samples selected by size): pooled regression

The dependent variable is idiosyncratic skewness calculated from 1 year's worth of each stock idiosyncratic daily returns if there is data for all trading days in the calendar year for the stock. R_y, R_{y-1}, R_{y-2} , R_{y-3} are idiosyncratic returns of the current year and the 3 previous years, *ln*(*mc*) is the logarithm of mean market capitalization, $\bar{g}$ is the mean capital gain, where daily capital gain is determined through equations (23a) and (23b), similarly to Grinblatt and Han (2005), and *f* is the number of days with extremely low volume (either *turnover* < 0.01%). All regressions include time dummies (unreported) for each year; *t*-statistics, which are in parentheses, are adjusted for heteroskedasticity. Size quintiles are determined within the sample, for each year, based on mean market capitalization (*mc*).

Size Group	1 (small)	2	3	4	5 (big)
R	0.492 (13.3)	0.855 (12.1)	1.171 (12.2)	1.506 (14.2)	1.156 (16.5)
R_{t-1}	-0,208 (-9.0)	-0.230 (-6.5)	-0.177 (-4.5)	-0.030 -(0.7)	0.004 (0.1)
R_{t-2}	-0,139 (-5.6)	-0.172 (-5.1)	-0.176 (-4.7)	-0.104 (-3.2)	-0.086 (-2.5)
R_{t-3}	-0,134 (-5.6)	-0.160 (-4.8)	-0.168 (-4.7)	-0.031 (-0.8)	-0.015 (-0.4)
Ln(*mc*) (x 10^3)	-45.1 (-2.4)	-66.0 (-1.7)	-20.1 (-0.5)	-34.1 (-1.0)	-39.1 (-4.2)
$\bar{g}$	-0.150 (-1.3)	-0.648 (-3.6)	-1.122 (-4.6)	-1.963 (-7.6)	-1.850 (-8.4)
f (x 10^3)	2.34 (5.1)	2.57 (5.2)	2.43 (4.4)	1.22 (2.3)	0.73 (2.3)
Intercept	1.094 (5.6)	1.065 (2.1)	0.856 (1.7)	0.834 (2.0)	0.856 (6.3)
R^2	13.0%	13.8%	13.8%	12.3%	15.2%
No. observs.	9,683	9,712	10,251	11,298	12,552

The first point worth of note in Table 14 is that past returns, particularly the first lag, lose significance from group 1 to group 5. Indeed the first and third lags are not statistically significant for groups 4 and 5 (large firms).[19] When we run regressions without $\bar{g}$ and f in the right side of equation (26), we find that the first lag of returns is statistically significant, at less than 1% level, for all size groups.[20] (The results of these regressions are not reported here.) Indeed, we find that it is the inclusion of variable $\bar{g}$ as explanatory variable that reduces the negative impact on skewness of the lags of yearly returns. Our interpretation is that $\bar{g}$ is best suited to classify winners and losers, from the perspective of investors subject to the disposition effect, than lags of returns. And the main contribution of past returns to skewness is through the disposition effect. As a matter of fact, the inclusion of $\bar{g}$ in the regressions reduces the absolute value of the coefficient of the first lag (R_{y-1}) for all size groups.

The second point worth of note in Table 14 is that size, that is $ln(mc)$, is statistically significant, at 5% level, to only two size groups. And these are the extreme groups (of smaller and larger firms), the ones with greater variance of size, as we see in Table 12. When we run regressions without $\bar{g}$ and f in the right side of equation (26) we find that the absolute value of its coefficient increases for all size groups and is statistically significant at 5% level to three size groups (groups 1, 2, and 5). We find that it is the inclusion of variable f as explanatory variable that reduces the negative impact of size on skewness. This supports our argument that, first, information flow is relevant for skewness, and second, that the correlation between size and skewness occurs because size is a proxy for information flow. This is in agreement with Xu (2007), who finds that institutional ownership and the breadth of institutional ownership, both usual proxies for (the inverse of) information asymmetry, have negative effect on skewness, and make size lose significance.[21]

4.5. Conclusion

We interpreted the distribution of returns of individual stocks as composed by a superposition of distributions. We assume there are two main fundamental

[19] Xu (2007) does not run separate regressions by size group to allow comparison.

[20] The only exceptions were the lags of order 3 for groups 4 and 5, which presented the negative sign but were not statistically significant.

[21] We replicated the analyses which results were presented in Tables 13 and 14 with simple returns, instead of idiosyncratic returns. All results were qualitatively the same.

distributions: one relate to news and other unrelated to news. For each of these fundamental distributions, we proposed one variable that should be relevant for the skewness, and evaluated how they affect the skewness. For price changes related to news, we proposed that the distribution of related returns would be more skewed if news was incorporated less frequently to prices, according to the jump diffusion model proposed by Press (1967). We defined a variable, other than size, that should proxy for this frequency of incorporation of news into prices. We associated price changes unrelated to news with behavioral biases, specifically the disposition effect, labeled by Shefrin and Stataman (1985). As shown in the previous chapter, the disposition effect affects volatility. We extended this reasoning to skewness and used a measure of the intensity of disposition effect, the measure of capital gain proposed by Grinblatt and Han (2005), as an explanatory variable for skewness.

We ran regressions using common stocks listed at NYSE from January 1963 through December 2007. As previous works by Harvey and Siddique (2000), Chen, Hong and Stein (2001), and Xu (2007), we find that contemporaneous return contributes positively to skewness, past returns contribute negatively, and size (the logarithm of market capitalization) contributes negatively. Additionally, we find that for the two variables we added as explanatory variables, the coefficients are statistically significant and present the expected sign. We split our sample in three continuous and non-overlapping periods of time, and find the results are robust across the subsamples.

Besides splitting our sample in periods of time, we also split it into five groups formed with the quintiles of size distribution in the cross-section of each year. Again, the two variables that we added, as explanatory variables in the regression to determine skewness, had coefficients statistically significant and with the expected sign. We also find that the use of capital gain makes the first lag of returns loses its statistical significance as explanatory variable to skewness of larger firms. We interpret this as an indicative that the impact of past returns in skewness was due to disposition effect, and the variable capital gain is better suited to capture this effect. Finally we find that size also loses significance in some size groups. Our interpretation of this result is that size is linked to skewness through the frequency of incorporation of news to prices. When we add a variable that is a better proxy for this frequency the relevance of size decreases. This is consistent with Xu (2007) for whom the introduction of institutional ownership, which is a usual proxy for information asymmetry, as an explanatory variable, also reduces the relevance of size.