Does Innovation Explain the Skewness of Stock Returns?

Ahmed Baig
Boise State University, Idaho, U.S.

Hassan Anjum Butt
Missouri Southern State University, Missouri, U.S.

Abrar Fitwi
Saint Mary's College, Notre Dame, Indiana, U.S.

Joey Smith
College of William & Mary, Virginia, U.S.

Does Innovation Explain the Skewness of Stock Returns?

Ahmed S. Baig[a], Hassan Butt[b], Abrar Fitwi[c], and Joey Smith[d]

ABSTRACT

This paper investigates the impact of firm-level innovation on the skewness of stock returns. Using data on a broad sample of equities from the major US stock exchanges, we find that innovative companies exhibit strong positive skewness. Our results are robust to both input and output measures of innovation as we find that increases in both firm-level research and development expenditure (R&D), as well as the number of patents, are positively associated with future stock return skewness. Our results hold using both systematic and idiosyncratic measures of skewness while controlling for various stock characteristics, time, and industry-fixed effects.

KEYWORDS

R&D, Skewness, Innovation, Patents

JEL classification: G11, G12, G14

INTRODUCTION

Recent academic research provides evidence that investors have a preference for positive skewness. More specifically, studies have shown that investors prefer stocks that resemble lotteries (Kraus and Litzenberger, 1976). For example, Barberis and Huang (2008) use prospect theory and show that some investors exhibit a strong preference for positive skewness and in doing so tend to overweight the tails in return distributions. In a return distribution characterized by positive skewness, investors expect a few large gains and frequent small losses from an investment. Generally, investors find positively skewed distribution of returns more desirable because even with a small probability of huge gains, all frequent but small losses can be covered. Barberis and Huang (2008) also show that stocks with a positive skewness exhibit price premiums and subsequent underperformance. More recent research provides support to these results. In the context of initial public offerings (IPOs), Green and Hwang (2012) show that IPOs with higher expected skewness exhibit significant first-day positive returns followed by significant negative abnormal returns over a one to five-year period. This supports the idea that higher first-day returns are due to investors' preference for positive skewness. Similarly, Blau and Whitby (2018) argue stocks with positive skewness are less efficient and investors' preference for positive skewness might explain the inefficiency of such stocks.

A separate stream of research finds that innovation is favorably valued by the financial market. It is well documented that research and development (R&D) expenditures, a key predictor of innovation,

---

[a] Boise State University, Idaho, U.S.
[b] Missouri Southern State University, Missouri, U.S.
[c] Saint Mary's College, Notre Dame, Indiana, U.S.
[d] College of William & Mary, Virginia, U.S.

Corresponding Author:
Butt (butt-h@mssu.edu)

affect a firm's market value, operating performance, and cash flows (see Hall, Jaffe, and Trajtenberg, 2000 and 2005; VanderPal, 2015; Farre-Mensa, Hegde, and Ljungqvist, 2020). In an efficient market, investors evaluate a firm based on expected cash flows: therefore, investments in R&D should be reflected in stock prices. Prior literature provides some support to this notion. For example, Eberhart, Maxwell, and Siddique (2004) find that an increase in R&D expenditures is associated with significant positive long-term abnormal operating performance and stock returns. In their recent paper Fitzgerald, Balsmeier, Fleming, and Manso (2021) find that innovative firms focusing on exploitation rather than exploration tend to generate superior subsequent short-term operating performance which translates into positive stock returns for these firms as well.

This study attempts to bridge the gap between these two streams of literature and examines the association between innovation and the skewness of stock returns. Positive skewness can intuitively be thought of as a deviation from the normal distribution with a long right tail and a higher probability of extremely positive gains. Based on prior evidence, investors tend to prefer positive skewness, and investors also value innovative firms. Accordingly, it is worth testing if stock returns of innovative firms exhibit positive skewness. We use a broad sample of equities from the major US stock exchanges and test two hypotheses. First, we investigate whether R&D expenditure leads to the positive skewness of stock returns. The first hypothesis is motivated by the findings of Eberhart, Maxwell, and Siddique (2004) where they show that firms experience significantly positive long-term abnormal operating performance following an increase in their R&D expenditures. The second hypothesis examines the association between patent registration and the skewness of stock returns. The motivation of the second hypothesis comes from the seminal work of Hall, Jaffe, and Trajtenberg (2005) where they show that patents and citations significantly contribute to the market value of the firms.

If the increase in R&D expenditure, as well as the number of patents, indeed increases the skewness of stock returns of innovative firms, our hypotheses predict a positive relationship between future skewness and R&D expenditure as well as the number of patents. Consistent with our prediction, we find that R&D expenditures and future skewness are positively and significantly associated in a multivariate regression framework. As a robustness check, we test our hypothesis using future idiosyncratic skewness of stock returns and find similarly significant results. Furthermore, we also find a positive relationship between the number of patents and future skewness: controlling for various firm characteristics, an increase in the number of patents is significantly associated with an increase in both our future skewness and idiosyncratic skewness measures. Finally, our analysis by industry shows that the healthcare sector contributes the most to this relation of innovation and skewness of stock returns. Our results are also consistent with Kapadia (2006), who highlights that the proxies for technological changes can forecast the future innovations in cross-sectional skewness.

The rest of this article proceeds as follows: Section 2 presents related literature. Section 3 describes our data. Section 4 presents the results of the empirical tests and Section 5 concludes.

LITERATURE REVIEW

In this section, we describe the literature related to our research questions. We first present the literature on investors' preference for lottery-like stocks. Second, we highlight several studies that describe the significance of innovation in relation to various aspects of a firm's performance.

We find empirical evidence that investors prefer positively skewed stocks. Mitton and Vorkink (2007) show that individual investors under diversify their portfolio at the expense of mean-variance efficiency to achieve higher skewness of expected returns. Similarly, Barberis and Huang (2008) present evidence that apart from having a preference for positively skewed stock returns, investors pay a premium for these positively skewed stocks which ultimately leads to contemporaneous overpricing and subsequent underperformance for such securities. Moreover, Boyer, Mitton, and

Vorkink (2010) provide empirical evidence that stocks with higher idiosyncratic skewness are more likely to suffer from lower-than-expected returns. Byun and Kim (2016) show that call options associated with positively skewed stocks underperform similar call options associated with less positively skewed stocks. Kumar (2009) shows a significant positive correlation between the preference for lottery-like stocks and preference for gambling in investors. He further shows that demand for the lottery-like stocks increases during economic downturns. Similarly, Kumar, Page, and Spalt (2011) show that regions with higher Catholic concentrations exhibit a stronger preference to hold stocks with positive skewness, and this preference correlates with their gambling preferences. Blau, Hsu, and Whitby (2019) find a higher preference for skewed stocks by investors and negative return premia associated with such preferences in countries where gambling is legalized. In a similar study, Yao et al (2019) show that Chinese stock markets present evidence of a preference for assets with higher idiosyncratic skewness which adversely impacts the efficiency of these markets. They argue that the efficiency of such financial markets can be improved by relaxing some arbitrage restrictions and better managing retail investors.

On the other hand, the financial literature contains several studies that show the significance of innovation at the macroeconomic as well as firm-level. Griffith, Redding, and Reenen (2004) study the aggregate R&D at the country level and find that R&D expenditures are associated with innovation and growth. Moreover, by using the newly created data on patents issued to U.S. firms from the 1926 to 2010 period, Kogan et al (2017) show that in the medium-term, technological innovation accounts for significant fluctuations in the total factor of production and aggregate economic growth. At the firm level, the financial literature shows that the current R&D spending levels, as well as recent changes in R&D spending, are positively related to subsequent excess stock returns. Chan, Martin, and Kensinger (1990) show that stock price response to an increase in R&D spending announcements are on average significantly positive, even when announcements occur during periods of earnings decline. Moreover, they show that positive abnormal returns following an increase in R&D announcement are more pronounced for high-tech firms as compared to low-tech firms. Eberhart, Maxwell, and Siddique (2004) find that when firms increase R&D expenditure unexpectedly and significantly, shareholders of the firm earn significant positive abnormal stock returns while the firms also experience positive abnormal long-term operating performance. Hall, Jaffe, and Trajtenberg (2000) use comprehensive patent data from 4,800 U.S. manufacturing firm to examine the contributions and the economic meaning of R&D spending, patents, and citation-weighted patents for the firms. They find that citation-weighted patents carry more economic value for the firms in terms of market valuation as firms with highly cited patents get more attention from investors. Furthermore, Hall, Jaffe, and Trajtenberg (2005) present similar empirical evidence that an additional citation per patent increases a firm's market valuation by almost 3%. They further show that self-citations are more valuable and carry stronger weight towards the market valuation. Farre-Mensa, Hegde, and Ljungqvist (2020) show that young firms that win patents earlier experience increased sales growth, employment growth, pursue higher-quality follow-on innovative activities, and attract more funding opportunities over the next five years. Scherer, Harhoff, and Kukies (2000) and Mata and Woerter (2013) present empirical evidence that the returns from the innovative activities are highly skewed. This means that only a handful of firms are successful in the innovation activities and relish the gains on R&D spending. Similarly, in the context of the Dutch economy, Marsili and Salter (2005) also find a highly skewed distribution of returns to innovation and that the degree of novelty of the innovation significantly increases the skewed distribution of return to innovative activities. Finally, Kapadia (2006) in a comprehensive study of cross-sectional skewness, idiosyncratic volatility and expected returns highlights that the proxies for technological changes such as lagged patent growth or R&D investment can forecast the future innovations in cross-sectional skewness.

_____

In short, we find two different strands of literature: one focused on the innovative activities of the firms and their consequences on the firm's performance, the other focused on investors' preference for positively skewed stocks and its undesirable consequences. This study contributes to both the literature on firm innovation activities as well as investors' preferences for higher moments. First, we bridge the gap between both pieces of literature by studying the impact of innovative activities of the firms on distributional characteristics of stock returns. Second, we add to the literature that highlights the factors that contribute to an increase in the positive stock-level skewness of returns.

DATA DESCRIPTION

The data utilized in this study are obtained from three primary sources. We obtain the daily stock level information for all common stocks from the Center of Research in Security Prices (CRSP). We obtain annual financial statement data from Compustat. Finally, we gather yearly data on firm-level patents from Kogan et al. (2017).[1] The analysis period for CRSP and Compustat data in our study range from 1993 to 2017, while the data period for Patent information is from 1993 to 2010. At the time of writing this paper, patent information after 2010 is not available from our data source. We further note that patent data is not available for all firms and accordingly patent data has missing observations. Using these stock characteristics, patent, and financial information we construct a series of variables that are utilized throughout the study. Our main dependent variables are skewness and idiosyncratic skewness. SKEWNESS is computed as the third moment of daily raw returns for each stock during each year. IDIOSKEWNESS is the third moment of daily residuals obtained from the Fama and French (1993) three-factor model regressions.

To capture innovation we follow Chen, Leung, and Evans (2016) and utilize both input and output-oriented firm-level innovation measures as our main independent variables of interest. As our first (input-based) innovation measure we utilize R&D which represents the research and development expenditure for firms in USD millions for the year. As our second (output-based) measure for innovation we use PATENTS which is the number of registered patents for the firm-year. For our control variables, we use the following: BM is the book -to-market ratio for the year. ROA is the return on assets for the year. DEBT/TA is the ratio of total debt to total assets. VOLATILITY is the annual standard deviation of daily returns. PRICE represents the closing stock price. Following prior finance literature, we remove stocks with average price of less than $2 from our sample (Baig et al, 2020; Baig et al., 2021). MARKETCAP represents the market capitalization of the firm on the last trading day of the year. SPREAD represents the bid-ask spread averaged over the year. TURNOVER represents the annual average ratio of daily trading volume scaled by share outstanding. ILLIQUIDITY represents the annual average of daily Amihud (2002) illiquidity measure trimmed at 1st and 99th percent to control for outliers. NASDAQ is a dichotomous variable that equals 1 for NASDAQ listed stocks and 0 otherwise. Table 1 presents the summary statistics for our sample. An average stock has a skewness of about 40% and idiosyncratic skewness of about 44%. Average R&D expenditures are close to $29 million. An average firm in our sample has around 19 registered patents. Book-to- market ratio averages about 64%. Similarly, the ratio of debt-to-total assets for an average firm is about 19%. An average stock in our sample has a volatility of 3.3%, price of $39, market capitalization of about $2.1 billion, the spread of about 2.1%, turnover of about 0.7%, illiquidity of about 1.49. Nearly 61% of our sample stocks are listed on NASDAQ.

---

[1] The data is available at: https://kelley.iu.edu/nstoffma/

Table 1. Summary Statistics

This table provides the statistics that summarize our sample. SKEWNESS is the third moment of daily raw returns for each stock during each year. IDIOSKEWNESS is the third moment of daily residuals obtained from the Fama and French (1993) three-factor model regressions. R&D represents the research and development expense for the firm in USD millions for the year. PATENTS is the number of patents for the firm-year. BM is the Book to Market ratio for the year. ROA is the return on assets for the year. DEBT/TA is the ratio of total debt to total assets. VOLATILITY is the annual standard deviation of daily returns. PRICE represents the closing stock price. MARKETCAP represents the market capitalization of the firm on the last trading day of the year. SPREAD represents the bid-ask spread averaged over the year. TURNOVER represents the annual average ratio of daily trading volume scaled by share outstanding. ILLIQUIDITY represents the annual average of daily Amihud (2002) illiquidity measure. NASDAQ is a dichotomous variable that equals 1 for NASDAQ listed stocks and 0 otherwise.

	OBS	MEAN	STANDARD DEVIATION	25th PERCENTILE	MEDIAN	75th PERCENTILE
	[1]	[2]	[3]	[4]	[5]	[6]
SKEWNESS	101840	0.404	1.330	-0.062	0.305	0.755
IDIOSKEWNESS	101840	0.436	1.476	-0.059	0.352	0.857
R&D	101840	29.068	209.475	0.000	0.000	6.570
PATENTS	18293	19.454	73.717	1.000	3.000	11.000
BM	101840	0.644	0.556	0.302	0.530	0.835
ROA	101840	0.069	0.228	0.023	0.096	0.164
DEBT/TA	101840	0.192	0.186	0.022	0.147	0.311
VOLATILITY	101840	0.033	0.020	0.020	0.029	0.042
PRICE	101840	39.440	1683.463	7.600	16.200	29.875
MARKETCAP	101840	2124742	7317660	77044	288839	1205510
SPREAD	101840	0.021	0.027	0.002	0.011	0.028
TURNOVER	101840	0.007	0.010	0.002	0.005	0.009
ILLIQUIDITY	101840	1.494	5.970	0.003	0.031	0.465
NASDAQ	101840	0.611	0.488	0.000	1.000	1.000

EMPIRICAL RESULTS

In this section we carry out a series of multivariate regressions using different econometric methods to analyze the relationship between firm-level innovation and skewness of stock returns. We utilize robust standard errors clustered at the firm level and include year fixed effects in all our regression specifications following Petersen (2009). This allows us to control for any potential serial and cross dependence issues. We also include industry fixed effects based on Fama-French 48 industry classifications in our regression models.

SKEWNESS AND R&D EXPENDITURE

In our first set of tests, we run the following OLS specification.

$$SKEWNESS_{i,t+1} \text{ or } IDIOSKEWNESS_{i,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} +$$

$$\beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} +$$

$$\beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \tag{1}$$

Table 2 presents the results from the estimation of equation (1). The dependent variable is skewness in the following year in column [1] while it is idiosyncratic skewness in the following year in column [2]. The coefficient on LNR&D is positive and statistically significant at 1% level in column [1]. More specifically, the coefficient is 0.023 (t-statistic=5.259) after controlling for various firm characteristics. Similarly, the coefficient on LNR&D is positive and statistically significant at 1% level in column [2]. The coefficient is 0.031 (t-statistic=6.111). These results again are robust to the control of various firm characteristics, industry factors and any potential time trends. These findings lend strong support to our hypothesis and suggest that our input measure of innovation (the research and development (R&D) expenditure) is indeed positively   related to skewness and idiosyncratic skewness. Regarding the control variables, book-to-market ratio and stock return volatility are positively and significantly associated with future skewness measures. Return on assets, natural logs of price and size, spread, turnover and Nasdaq dummy are negatively and significantly associated with future skewness measures. Debt-to-total assets ratio is negatively associated with both skewness measures while the Amihud illiquidity measure is positively associated with skewness and unrelated to the idiosyncratic skewness.

───────────────────────────────

Table 2. Skewness and R&D Regressions

This table provides the results from the estimation of the following OLS regression specification.

$$SKEWNESS_{i,t+1} \text{ or } IDIOSKEWNESS_{i,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is one year ahead SKEWNESS in column [1] and one year ahead IDIOKEWNESS in column [2]. The independent variable LNR&D represents the is natural log of R&D. For the variable definitions please refer to table 1. All specifications include industry fixed effects based on Fama-French 48 Industry classifications. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEWNESS $_{i,t+1}$	[2] IDIOSKEWNESS $_{i,t+1}$
LNR&D	0.023***	0.031***
	(5.259)	(6.111)
BM	0.079***	0.092***
	(6.677)	(7.226)
ROA	-0.481***	-0.535***
	(-12.569)	(-13.061)
DEBT/TA	-0.000	0.027
	(-0.011)	(0.824)
VOLATILITY	4.292***	4.353***
	(6.611)	(6.538)
LNPRICE	-0.111***	-0.118***
	(-10.768)	(-10.246)
LNSIZE	-0.090***	-0.090***
	(-18.029)	(-16.005)
SPREAD	-1.182***	-1.448***
	(-3.592)	(-4.269)
TURNOVER	-1.258*	-1.522*
	(-1.646)	(-1.844)
ILLIQUIDITY	0.003***	0.002
	(2.972)	(1.369)
NASDAQ	-0.116***	-0.128***
	(-9.080)	(-8.820)
Constant	1.618***	1.643***
	(24.556)	(22.836)
YEAR FE	YES	YES
INDUSTRY FE	YES	YES
ROBUST SE	YES	YES
Observations	84,651	84,651
R-squared	0.060	0.048

SKEWNESS AND PATENTS

In our second set of tests, we run the following OLS specification.

$$SKEWNESS_{i,t+1} \text{ or } IDIOSKEWNESS_{i,t+1} = \beta_0 + \beta_1 LN(PATENT)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} +$$
$$\beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} +$$
$$\beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \tag{2}$$

Table 3 presents the results from the estimation of equation (2). The dependent variable is skewness in the following year in column [1] and idiosyncratic skewness in the following year in column [2]. The independent variable LNPATENT represents the natural log of patents where patents represent the number of patents for the firm during the year. The coefficient on LNPATENT is positive and statistically significant at 1% level in column [1]. We find that, the coefficient on LNPATENT is 0.037 (t-statistic=3.650). Similarly, the coefficient on LNPATENT is also positive and statistically significant at 1% level in column [2]. More specifically it is 0.041 (t-statistic=3.362). These results are robust to the control of various firm and industry factors and any potential time trends. These results further affirm our hypothesis and suggest that our output measure of innovation (patents) is positively associated with future skewness and idiosyncratic skewness. As for our control variables, stock return volatility is positively and significantly associated with future skewness measures. Return on assets, natural logs of price and size, spread, and Nasdaq dummy are negatively and significantly associated with future skewness measures. Book-to-market ratio, debt-to-total assets ratio, turnover and illiquidity are unrelated to the skewness measures. The signs on these control variables are generally consistent with the previous table with some decrease in significance which could be explained by the lower number of observations.

Overall, the results from both these tables lend support to our hypothesis. Our empirical evidence suggests that both the input and output measures of innovation captured in the form of research and development expenditure and the number of patents for the firm are positively linked to the future skewness of stock returns.

Table 3. Skewness and Patents Regressions

This table provides the results from the estimation of the following OLS regression specification.

$$SKEWNESS_{i,t+1} \text{ or } IDIOSKEWNESS_{i,t+1} = \beta_0 + \beta_1 LN(PATENTS)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} +$$

$$\beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} +$$

$$\beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is one year ahead SKEWNESS in column [1] and one year ahead IDIOKEWNESS in column [2]. The independent variable LNPATENT represents the natural log of number of Patents for the firm-year. For the variable definitions please refer to table 1. All specifications include industry fixed effects based on Fama-French 48 Industry classifications. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEWNESS $_{i,t+1}$	[2] IDIOSKEWNESS $_{i,t+1}$
LNPATENT	0.037***	0.041***
	(3.650)	(3.362)
BM	0.018	0.029
	(0.668)	(0.941)
ROA	-0.352***	-0.395***
	(-4.636)	(-4.833)
DEBT/TA	0.084	0.111
	(1.121)	(1.295)
VOLATILITY	4.426***	4.168***
	(4.195)	(3.518)
LNPRICE	-0.157***	-0.177***
	(-7.367)	(-6.883)
LNSIZE	-0.095***	-0.100***
	(-8.116)	(-7.295)
SPREAD	-1.760**	-2.477***
	(-1.995)	(-2.619)
TURNOVER	-0.116	-1.288
	(-0.074)	(-0.676)
ILLIQUIDITY	0.003	-0.000
	(0.683)	(-0.037)
NASDAQ	-0.148***	-0.161***
	(-5.048)	(-4.673)
Constant	1.885***	2.046***
	(12.237)	(11.865)
YEAR FE	YES	YES
INDUSTRY FE	YES	YES
ROBUST SE	YES	YES
Observations	15,968	15,968
R-squared	0.069	0.059

ROBUSTNESS ANALYSIS

LOGISTIC REGRESSIONS

In this section we carry out a robustness analysis of our main tests using a series of logistic regressions. To do so we define a dummy variable for positive skewness that takes on the value one if yearly return skewness (or idiosyncratic skewness) for the stock is greater than +2, and zero otherwise. This definition follows George (2011) who defines acceptable value of skewness for a normal distribution to be 2. We run the following logistic model:

$$OSITIVESKEWNESS\ DUMMY_{i,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t\ or}\ \beta_1 LN(PATENT)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} +$$
$$\beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} +$$
$$\beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \tag{3}$$

Table 4 presents the results from the estimation of equation (3). The dependent variable is dummy variable for positive skewness in the following year in column [1] and [2] while it is dummy variable for positive idiosyncratic skewness in the following year in column [3] and [4]. The independent variables are the natural log of R&D in columns [1] and [3] or natural log of PATENT in columns [2] and [4]. All the independent variables, including the control variables, are defined the same way as in the previous tables. The coefficient on LNR&D is positive and statistically significant at 1% level in column [1] as well as column [3]. In particular, the coefficient in column [1] is 0.112 (t-statistic=7.125) while the coefficient in column [3] is 0.120 (t-statistic=9.270) suggesting that increases in research and development costs is associated with a significant increase in the probability of one year ahead positive skewness and idiosyncratic skewness. These results are robust to the control of various firm and industry factors and any potential time trends. Moreover, these results are also statistically and qualitatively consistent with results in table 2 lending further support to our main tests.

However, we find weak results for patents and skewness relation. The coefficient on LNPATENT in column [2] is significant at 10% level, while the coefficient in column [4] is statistically insignificant. More precisely, the results suggest that an increase in patent registration is significantly positively associated with the probability of positive skewness in the following year but is not significantly associated with idiosyncratic skewness in the following year. It shows that natural log of patents has relatively lower explanatory power to explain the probability of skewness and the results are only weakly consistent with our results in table 3. This could perhaps be due to the smaller dataset for the patent regressions. In the following section we employ one more test to verify the robustness of our main results.

Table 4. Logistic Regressions of Skewness and R&D
This table provides the results from the estimation of the following logistic regression specification.

$$SKEWNESS\ DUMMY_{i,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is SKEW (IDIOSKEW) DUMMY that takes on a value of one if return skewness (idiosyncratic skewness) for the stock is greater than +2, zero otherwise. We define this dummy following George (2011) who defines acceptable value of skewness for a normal distribution to be 2. We use the year ahead skewness dummy in our regression specifications. The independent variable LNR&D represents the is natural log of R&D. The independent variable LNPATENT represents the natural log of number of Patents for the firm-year. For the variable definitions please refer to table 1. All specifications include industry fixed effects based on Fama-French 48 Industry classifications. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEW DUMMY $_{i,t+1}$	[2] SKEW DUMMY $_{i,t+1}$	[3] IDIOSKEW DUMMY $_{i,t+1}$	[4] IDIOSKEW DUMMY $_{i,t+1}$
LNR&D	0.112***		0.120***	
	(7.125)		(9.270)	
LNPATENT		0.083*		0.056
		(1.753)		(1.506)
BM	0.082***	0.077	0.086***	0.068
	(2.772)	(0.851)	(3.104)	(0.830)
ROA	-0.513***	-0.390**	-0.451***	-0.350**
	(-6.518)	(-2.144)	(-6.295)	(-2.138)
DEBT/TA	0.171*	0.218	0.174**	0.282
	(1.720)	(0.840)	(1.967)	(1.269)
VOLATILITY	5.873***	3.943	5.295***	3.551
	(5.434)	(1.348)	(5.144)	(1.258)
LNPRICE	-0.071**	-0.067	-0.030	-0.036
	(-2.297)	(-0.841)	(-1.108)	(-0.520)
LNSIZE	-0.324***	-0.388***	-0.200***	-0.215***
	(-16.514)	(-7.752)	(-11.670)	(-5.312)
SPREAD	-3.225***	-3.793	-1.302	0.184
	(-2.826)	(-1.202)	(-1.189)	(0.057)
TURNOVER	5.398***	11.161**	5.680***	7.397*
	(3.964)	(2.514)	(4.175)	(1.887)
ILLIQUIDITY	0.002	-0.007	-0.003	-0.030
	(0.679)	(-0.490)	(-0.980)	(-1.423)
NASDAQ	-0.369***	-0.421***	-0.307***	-0.297***
	(-8.784)	(-4.077)	(-8.063)	(-3.267)
Constant	0.715*	1.463**	-1.085***	-0.789
	(1.916)	(2.213)	(-2.900)	(-1.375)
YEAR FE	YES	YES	YES	YES
INDUSTRY FE	YES	YES	YES	YES
ROBUST SE	YES	YES	YES	YES
Observations	84,651	15,788	84,651	15,914

COMMON SAMPLE REGRESSIONS

In this section, we carry out a robustness analysis with the common observations between the patents and R&D data sets. The idea behind this test is that in table 2 (R&D regressions) we have roughly 85,000 observations while in table 3 (patents regressions) we have roughly 16,000 observations. So, our results in table 2 could be driven by the uncommon observations. Accordingly, to alleviate this concern we run the following OLS specification using the common 15, 968 observations.

$$SKEWNESS\ MEASURE_{i,t} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \quad (4)$$

The dependent and independent variables are our skewness measures and the natural log of R&D, respectively. All the variables, including the control variables, are defined the same way as in the previous tables.

Table 5 presents the results from the estimation of equation (4). The dependent variable is skewness in the following year in column [1] while it is idiosyncratic skewness in the following year in column [2]. The coefficient on LNR&D is positive and statistically significant at 1% level in both columns [1] and [2]. More specifically the coefficient in column [1] is 0.018 (t-statistic=2.373) while in column [2] is 0.021 (t-statistic=2.226). These results are again robust to the control of various firm and industry factors and any potential time trends. These results are statistically and qualitatively consistent with table 2, with some difference in economic significance. These robustness regressions yet again confirm our hypothesis and suggest that our input measure of innovation (R&D expenditure) is positively related with future skewness and idiosyncratic skewness. Furthermore, these results are not driven by uncommon observations. The signs on control variables are generally consistent with the previous tables and with the prior literature.

_____

Table 5. Skewness and R&D Regressions for Common Sample

This table provides the results from the estimation of the following OLS regression specification.

$$SKEWNESS\ MEASURE_{i,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is one year ahead SKEWNESS in column [1] and one year ahead IDIOKEWNESS in column [2]. The independent variable LNR&D represents the is natural log of R&D. For the variable definitions please refer to table 1. All specifications include industry fixed effects based on Fama-French 48 Industry classifications. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEWNESS $_{i,t+1}$	[2] IDIOSKEWNESS $_{i,t+1}$
LNR&D	0.018**	0.021**
	(2.373)	(2.226)
BM	0.020	0.030
	(0.716)	(0.987)
ROA	-0.354***	-0.397***
	(-4.652)	(-4.847)
DEBT/TA	0.088	0.116
	(1.180)	(1.354)
VOLATILITY	4.304***	4.036***
	(4.090)	(3.418)
LNPRICE	-0.162***	-0.182***
	(-7.670)	(-7.224)
LNSIZE	-0.089***	-0.094***
	(-7.731)	(-6.962)
SPREAD	-1.574*	-2.271**
	(-1.791)	(-2.408)
TURNOVER	-0.257	-1.452
	(-0.162)	(-0.749)
ILLIQUIDITY	0.003	0.000
	(0.761)	(0.042)
NASDAQ	-0.153***	-0.166***
	(-5.203)	(-4.826)
Constant	1.849***	2.008***
	(11.977)	(11.517)
YEAR FE	YES	YES
INDUSTRY FE	YES	YES
ROBUST SE	YES	YES
Observations	15,968	15,968
R-squared	0.069	0.059

INDUSTRY ANALYSIS

In this study, we use the entire universe of common stocks from our data sources during our sample period. This allows us to draw interesting aggregate inferences on the relation of innovation and future skewness. However, it is also interesting to analyze at the industry-level in order to identify sectors and industries where the relation between innovation and skewness is more or less pronounced. To conduct this analysis, we categorize our stocks according to the Fama-French 5

industry classifications. The classification of 5 broad industries is as follows: [1] Consumer products (Consumer Durables, Nondurables, Wholesale, Retail, and Some Services (Laundries, Repair Shops)), [2] Manufacturing (Manufacturing, Energy and Utilities), [3] HiTech (Business Equipment, Telephone and Television Transmission), [4] Health (Healthcare, Medical Equipment, and Drugs) and [5] all other (Mines, Construction, Building Materials, Transport, Hotels, Entertainment, Finance).

SKEWNESS AND R&D REGRESSIONS BY FAMA-FRENCH 5 INDUSTRY CLASSIFICATIONS

To carry out this analysis, we run the following OLS regression specification for each of the 5 industry classifications and report the results in Table 6.

$$SKEWNESS\ MEASURE_{i,j,t+1} = \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \quad (5)$$

The dependent variable is the skewness in the following year in columns [1], [3], [5], [7] and [9] while it is the idiosyncratic skewness in the following year in columns [2], [4], [6], [8] and [10]. The main independent variable of interest in this table is the natural log of R&D. We find that for Consumer and HiTech industries, the relation between R&D expenditure and positive skewness is statistically insignificant. However, for the Manufacturing, Health, and all other industries we do observe a positive and significant association between R&D expenditure and positive skewness. Note that we find the strongest association in the healthcare industry. Healthcare industry is often viewed as a very innovative industry, so these findings appear consistent with intuition. Surprisingly, however, we do not find a significant association between innovation and future skewness of returns for the technology industry.

Table 6. Skewness and R&D Regressions for Fama-French 5 Industry Classifications

This table provides the results from the estimation of the following OLS regression specification.

$$SKEWNESS\ MEASURE_{i,t+1}$$
$$= \beta_0 + \beta_1 LN(R\&D)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t}$$
$$+ \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is one year ahead SKEWNESS in in columns [1], [3], [5], [7] and [9] and one year ahead IDIOKEWNESS in columns [2], [4], [6], [8] and [10] respectively. The independent variable LNR&D represents the is natural log of R&D. The classification of 5 broad industries is as follows: (FFI=1) Consumer products, (FFI=2) Manufacturing, (FFI=3) HiTech, (FFI=4) Health, and (FFI=5) all other. For the variable definitions please refer to table 1. All specifications include year fixed effects. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEW $_{i,t+1}$	[2] IDIOSKEW $_{i,t+1}$	[3] SKEW $_{i,t+1}$	[4] IDIOSKEW $_{i,t+1}$	[5] SKEW $_{i,t+1}$	[6] IDIOSKEW $_{i,t+1}$	[7] SKEW $_{i,t+1}$	[8] IDIOSKEW $_{i,t+1}$	[9] SKEW $_{i,t+1}$	[10] IDIOSKEW $_{i,t+1}$
	FFI=1		FFI=2		FFI=3		FFI=4		FFI=5	
LNR&D	0.005	0.016	0.028***	0.036***	-0.006	-0.001	0.088***	0.101***	0.033**	0.041**
	(0.561)	(1.493)	(5.045)	(5.359)	(-0.766)	(-0.132)	(5.915)	(6.047)	(2.276)	(2.379)
BM	0.091***	0.098***	0.044*	0.053**	0.028	0.034	0.130**	0.114	0.118***	0.144***
	(3.948)	(4.083)	(1.881)	(2.068)	(0.978)	(1.046)	(1.983)	(1.609)	(5.747)	(6.382)
ROA	-0.551***	-0.603***	-0.430***	-0.487***	-0.456***	-0.520***	-0.274***	-0.269***	-0.555***	-0.614***
	(-5.366)	(-5.583)	(-5.493)	(-5.534)	(-7.539)	(-8.079)	(-3.170)	(-2.902)	(-6.640)	(-6.810)
DEBT/TA	0.032	0.079	0.063	0.111	0.140**	0.187**	0.043	0.051	-0.014	-0.000
	(0.495)	(1.094)	(0.941)	(1.476)	(2.133)	(2.503)	(0.365)	(0.385)	(-0.309)	(-0.002)
VOLATILITY	2.116*	2.328*	2.456**	3.226***	3.872***	3.770***	4.475**	4.568**	4.180***	4.320***
	(1.821)	(1.868)	(2.403)	(2.880)	(4.099)	(3.790)	(2.447)	(2.387)	(3.510)	(3.564)
LNPRICE	-0.149***	-0.159***	-0.082***	-0.076***	-0.182***	-0.193***	-0.142***	-0.161***	-0.069***	-0.073***
	(-5.898)	(-5.742)	(-4.066)	(-3.278)	(-8.535)	(-7.587)	(-3.244)	(-3.334)	(-4.005)	(-3.882)
LNSIZE	-0.066***	-0.068***	-0.108***	-0.106***	-0.077***	-0.076***	-0.119***	-0.125***	-0.084***	-0.082***
	(-5.469)	(-5.040)	(-10.940)	(-9.599)	(-6.356)	(-5.419)	(-4.765)	(-4.583)	(-11.116)	(-9.673)
SPREAD	-0.551	-0.865	-1.241	-1.360*	-0.484	-0.721	0.238	-0.409	-1.381***	-1.870***
	(-0.644)	(-0.979)	(-1.644)	(-1.763)	(-0.500)	(-0.737)	(0.221)	(-0.355)	(-2.850)	(-3.727)
TURNOVER	-1.853	-2.418	-0.544	-0.352	0.720	0.083	-1.607	-1.312	-2.441	-2.551
	(-0.818)	(-0.988)	(-0.536)	(-0.317)	(0.525)	(0.057)	(-0.602)	(-0.437)	(-1.606)	(-1.570)
ILLIQUIDITY	0.005	0.004	0.005	0.003	0.005	0.002	-0.002	-0.003	0.003***	0.002*
	(1.088)	(0.800)	(1.330)	(0.819)	(1.464)	(0.687)	(-0.424)	(-0.596)	(2.710)	(1.678)
NASDAQ	-0.098***	-0.116***	-0.106***	-0.119***	-0.092***	-0.082**	-0.134**	-0.142**	-0.102***	-0.122***
	(-3.432)	(-3.500)	(-4.335)	(-4.142)	(-3.156)	(-2.381)	(-2.526)	(-2.403)	(-4.903)	(-5.239)
Constant	1.596***	1.674***	1.949***	1.935***	1.805***	1.847***	1.953***	2.111***	1.535***	1.554***
	(10.689)	(10.242)	(15.006)	(13.835)	(12.758)	(11.765)	(6.888)	(6.848)	(14.605)	(13.483)
YEAR FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
ROBUST SE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	15,399	15,399	16,333	16,333	16,630	16,630	7,538	7,538	28,751	28,751
R-squared	0.060	0.051	0.063	0.044	0.089	0.067	0.060	0.054	0.051	0.040

SKEWNESS AND PATENTS REGRESSIONS FOR INDUSTRY CLASSIFICATION

In this subsection we run the following OLS regression specification for each of the 5 industry classifications and report the results in Table 7.

$$SKEWNESS\ MEASURE_{i,j,t+1} = \beta_0 + \beta_1 LN(PATENTS)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} +$$
$$\beta_5 VOLATILITY_{i,t} + \beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} +$$
$$\beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t} \tag{6}$$

The dependent variable is the skewness in the following year in columns [1], [3], [5], [7] and [9] while it is the idiosyncratic skewness in the following year in columns [2], [4], [6], [8] and [10]. The main independent variable of interest in this table is the natural log of patents. We find that for the "Consumer Products" and the "All Other" industry classifications the association between patent registration and positive skewness is statistically insignificant. On the other hand, for the "Manufacturing", "Health", and "HiTech" industries the association between R&D expenditure and positive skewness is statistically significant. Similar to the previous table, the relation of innovation (measured by patents) and skewness measures is the strongest for the healthcare sector.

Table 7. Skewness and Patents Regressions for Fama-French 5 Industry Classifications

This table provides the results from the estimation of the following OLS regression specification.

$$SKEWNESS\ MEASURE_{i,t+1} = \beta_0 + \beta_1 LN(PATENTS)_{i,t} + \beta_2 BM_{i,t} + \beta_3 ROA_{i,t} + \beta_4 \left(\frac{DEBT}{TA}\right)_{i,t} + \beta_5 VOLATILITY_{i,t} +$$

$$\beta_6 LNPRICE_{i,t} + \beta_7 LNSIZE_{i,t} + \beta_8 SPREAD_{i,t} + \beta_9 TURNOVER_{i,t} + \beta_{10} ILLIQUIDITY_{i,t} + \beta_{11} NASDAQ_{i,t} + \varepsilon_{i,t}$$

Dependent variable is one year ahead SKEWNESS in in columns [1], [3], [5], [7] and [9] and one year ahead IDIOKEWNESS in columns [2], [4], [6], [8] and [10] respectively. The independent variable LNPATENTS represents the is natural log of PATENTS. The classification of 5 broad industries is as follows: (FFI=1) Consumer products, (FFI=2) Manufacturing, (FFI=3) HiTech, (FFI=4) Health, and (FFI=5) all other. For the variable definitions please refer to table 1. All specifications include year fixed effects. Robust standard errors are clustered at the firm level. T-stats are in parentheses below the coefficient estimates. *, **, *** denote statistical significance at the 0.1, 0.05, and the 0.01 levels, respectively.

	[1] SKEW $_{i,t+1}$	[2] IDIOSKEW $_{i,t+1}$	[3] SKEW $_{i,t+1}$	[4] IDIOSKEW $_{i,t+1}$	[5] SKEW $_{i,t+1}$	[6] IDIOSKEW $_{i,t+1}$	[7] SKEW $_{i,t+1}$	[8] IDIOSKEW $_{i,t+1}$	[9] SKEW $_{i,t+1}$	[10] IDIOSKEW $_{i,t+1}$
	FFI=1		FFI=2		FFI=3		FFI=4		FFI=5	
LNR&D	0.016	0.023	0.032**	0.039**	0.034**	0.040**	0.085**	0.086*	0.025	0.067
	(0.550)	(0.671)	(2.354)	(2.238)	(2.346)	(2.174)	(2.182)	(1.874)	(0.571)	(1.216)
BM	0.033	0.006	0.067	0.104*	-0.038	-0.045	-0.009	-0.017	0.012	0.030
	(0.509)	(0.084)	(1.465)	(1.963)	(-0.836)	(-0.847)	(-0.076)	(-0.129)	(0.146)	(0.314)
ROA	-0.504*	-0.490*	-0.429**	-0.501***	-0.398***	-0.467***	-0.139	-0.126	-0.520**	-0.642***
	(-1.883)	(-1.791)	(-2.498)	(-2.751)	(-3.107)	(-3.441)	(-0.984)	(-0.827)	(-2.379)	(-2.639)
DEBT/TA	0.238	0.275	0.160	0.235*	0.029	0.020	-0.217	-0.244	0.229	0.399
	(1.261)	(1.286)	(1.243)	(1.650)	(0.262)	(0.151)	(-0.935)	(-0.939)	(1.034)	(1.408)
VOLATILITY	5.192	4.836	1.198	2.025	3.009**	2.800*	2.930	2.247	7.945**	7.645**
	(1.185)	(1.021)	(0.558)	(0.787)	(2.109)	(1.741)	(0.874)	(0.633)	(2.528)	(2.067)
LNPRICE	-0.158**	-0.201***	-0.103***	-0.094**	-0.204***	-0.229***	-0.276***	-0.323***	-0.006	0.012
	(-2.496)	(-2.877)	(-2.712)	(-2.048)	(-6.413)	(-5.710)	(-3.758)	(-3.790)	(-0.099)	(0.165)
LNSIZE	-0.044	-0.042	-0.100***	-0.110***	-0.103***	-0.113***	-0.081**	-0.085*	-0.089**	-0.087*
	(-1.560)	(-1.302)	(-5.578)	(-5.321)	(-5.536)	(-4.917)	(-1.986)	(-1.808)	(-2.277)	(-1.830)
SPREAD	0.982	0.331	-2.779*	-3.346**	-1.714	-2.951*	1.245	-0.011	-0.149	1.164
	(0.340)	(0.110)	(-1.726)	(-1.981)	(-1.018)	(-1.665)	(0.475)	(-0.004)	(-0.055)	(0.361)
TURNOVER	-1.703	-2.236	0.524	-1.565	1.111	-0.480	5.881	7.246	-9.035	-7.670
	(-0.210)	(-0.255)	(0.162)	(-0.417)	(0.580)	(-0.191)	(1.103)	(1.191)	(-1.060)	(-0.668)
ILLIQUIDITY	0.005	0.003	0.014	0.010	0.000	-0.003	-0.025**	-0.026**	0.006**	0.005
	(0.251)	(0.133)	(1.492)	(1.084)	(0.000)	(-0.353)	(-2.054)	(-2.043)	(2.117)	(1.601)
NASDAQ	-0.125*	-0.151*	-0.062	-0.070	-0.143***	-0.149**	-0.204**	-0.222**	-0.149	-0.166
	(-1.654)	(-1.715)	(-1.331)	(-1.304)	(-2.959)	(-2.577)	(-2.274)	(-2.184)	(-1.422)	(-1.382)
Constant	1.198***	1.358***	1.940***	2.034***	2.225***	2.511***	2.111***	2.373***	1.414***	1.287**
	(2.934)	(3.002)	(7.751)	(7.613)	(9.169)	(8.882)	(4.203)	(4.260)	(2.843)	(2.214)
YEAR FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
ROBUST SE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	1,971	1,971	4,478	4,478	5,803	5,803	2,809	2,809	907	907
R-squared	0.061	0.054	0.065	0.055	0.090	0.074	0.61	0.057	0.095	0.072

CONCLUSION

In this study we develop and test the hypothesis that firm level innovation activities are associated with the positive skewness of the stock returns. Our hypothesis is motivated by the notion found in the literature that increases in R&D expenditures and patent registrations increase the future probability of innovative firms enjoy benefits such as increase in market values, improvement in operating performance and outperformance of stock returns in the long run. Consistent with our

hypothesis, our results show a significant positive association between measures of innovation and future skewness as well as idiosyncratic skewness of stock returns.

The results of this study provide important contributions to the literature. As mentioned earlier, we find two separate strands of literature: (1) firm's innovative activities and its aftermath, and (2) investors' preference for lottery-like stocks. However, the literature lacks research that examines the impact of innovative activity measures such as R&D expenditures and patent registrations on the future skewness of the stock returns. In this study, we bridge the gap between these two strands of literature and provide important insights to investors as well as the policymakers.

Finally, since it is well reported that investors' preference for positively skewed stock leads to contemporaneous overpricing and subsequent underperformance of stocks over the long run, it is important to determine the factors that contribute to the increase in skewness of stock returns. Our results identify one such factor in the form of firm innovation.

REFERENCES

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

Barberis, N., & Huang, M. (2008). Stocks as lotteries: The implications of probability weighting for security prices. *American Economic Review, 98*(5), 2066-2100.

Baig, A., Blau, B. M., & Griffith, T. G. (2020). Firm Opacity and the Clustering of Stock Prices: The Case of Financial Intermediaries. *Journal of Financial Services Research*, 1-20.

Baig, A. S., Blau, B. M., & DeLisle, R. J. (2021). Does mutual fund ownership reduce stock price clustering? Evidence from active and index funds. *Review of Quantitative Finance and Accounting*, 1-33.

Blau, B. M., & Whitby, R. J. (2018). Skewness, short interest and the efficiency of stock prices. *Applied Economics, 50*(20), 2229-2242.

Blau, B. M., Hsu, J., & Whitby, R. J. (2019). Skewness preferences and gambling cultures. *Pacific-Basin Finance Journal, 58*, 101206.

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *The Review of Financial Studies, 23*(1), 169-202.

Byun, S. J., & Kim, D. H. (2016). Gambling preference and individual equity option returns. *Journal of Financial Economics, 122*(1), 155-174.

Chan, S. H., Martin, J. D., & Kensinger, J. W. (1990). Corporate research and development expenditures and share value. *Journal of Financial Economics, 26*(2), 255-276.

Chen, J., Leung, W. S., & Evans, K. P. (2016). Are employee-friendly workplaces conducive to innovation?. *Journal of Corporate Finance*, 40, 61-79.

Eberhart, A. C., Maxwell, W. F., & Siddique, A. R. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *The Journal of Finance, 59*(2), 623-650.

Farre-Mensa, J., Hegde, D., & Ljungqvist, A. (2020). What is a patent worth? Evidence from the US patent "lottery". *The Journal of Finance, 75*(2), 639-682.

Fitzgerald, T., Balsmeier, B., Fleming, L., & Manso, G. (2021). Innovation search strategy and predictable returns. *Management Science, 67*(2), 1109-1137.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds." *Journal of Financial Economics*, 33(1), 3-56.

George, D. (2011). *SPSS for windows step by step: A simple study guide and reference, 17.0 update, 10/e.* Pearson.

Green, T. C., & Hwang, B. H. (2012). Initial public offerings as lotteries: Skewness preference and first-day returns. *Management Science, 58*(2), 432-444.

Griffith, R., Redding, S., & Reenen, J. V. (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of Economics and Statistics, 86*(4), 883-895.

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2000). Market value and patent citations: A first look. *NBER Working Paper # w7741*

Hall, B. H., Jaffe, A., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics*, 16-38.

Kapadia, N. (2006). The next Microsoft? Skewness, idiosyncratic volatility, and expected returns. Skewness, Idiosyncratic Volatility, and Expected Returns. *SSRN 3606637 Working Paper*

Kogan, L., Papanikolaou, D., Seru, A., & Stoffman, N. (2017). Technological innovation, resource allocation, and growth. *The Quarterly Journal of Economics, 132*(2), 665-712.

Kumar, A. (2009). Who gambles in the stock market? *The Journal of Finance*, 64(4), 1889-1933.

_____

Kumar, A., Page, J. K., & Spalt, O. G. (2011). Religious beliefs, gambling attitudes, and financial market outcomes. *Journal of Financial Economics, 102*(3), 671-708.

Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. The *Journal of Finance, 31*(4), 1085-1100.

Marsili, O., & Salter, A. (2005). 'Inequality'of innovation: Skewed distributions and the returns to innovation in Dutch manufacturing. *Economics of Innovation and New Technology, 14*(1-2), 83-102.

Mata, J., & Woerter, M. (2013). Risky innovation: The impact of internal and external R&D strategies upon the distribution of returns. *Research Policy, 42*(2), 495-501.

Mitton, T., & Vorkink, K. (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies, 20*(4), 1255-1288.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies, 22*(1), 435-480.

Scherer, F. M., Harhoff, D., & Kukies, J. (2000). Uncertainty and the Size Distribution of Rewards from Innovation. *Journal of Evolutionary Economics, 10*(1-2), 175-200.

VanderPal, G. A. (2015). Impact of R&D expenses and corporate financial performance. *Journal of Accounting and Finance, 15*(7), 135-149.

Yao, S., Wang, C., Cui, X., & Fang, Z. (2019). Idiosyncratic skewness, gambling preference, and cross-section of stock returns: Evidence from China. *Pacific-Basin Finance Journal, 53*, 464-483.