R&D Information Quality and Stock Returns*

Tao Huang† Junye Li‡ Fei Wu§ Ning Zhu¶

First version: January 2016; This version: December 2016

Abstract

Investors demand higher premiums from firms whose future performance in R&D is difficult to evaluate. We construct an R&D information quality (IQ) measure by linking a firm's historical innovation input (R&D expenditures) and innovation outcome (sales) and find statistically and economically significant evidence that expected excess returns decrease with R&D IQ. The high-minus-low R&D IQ hedge portfolio earns excess returns of about −39 (−48) basis points per month in value-weighted (equal-weighted) returns. The R&D IQ effect is weakly correlated with commonly used risk factors, is stronger for firms with greater uncertain business environment, and exhibits incremental pricing power.

Keywords: Research and Development; Ambiguity Aversion; Information Quality; Return Predictability; Factor Models.

JEL Classification: G12, G14, O32

*We would like to thank Yu Yuan, Jennifer Huang, Hung Wan Kot, David Ng, Matt Richardson, Yan Xu, and seminar participants at 2015 International Conference on Systemic Risk, 2016 Asian Finance Association Annual Meeting, and 2016 China International Conference in Finance for helpful comments.

†International Business School Suzhou, Xi'an Jiaotong-Liverpool University, Suzhou, China. Email: tao.huang@xjtlu.edu.cn

‡ESSEC Business School, Paris-Singapore, Singapore. Email: li@essec.edu

§Shanghai Advanced Institute of Finance, Shanghai Jiaotong Univeristy, Shanghai, China. Email: fwu@saif.sjtu.edu.cn

¶Shanghai Advanced Institute of Finance, Shanghai Jiaotong Univeristy, Shanghai, China. Email: nzhu@saif.sjtu.edu.cn

1. Introduction

As an essential component for gauging a firm's competitive advantages, research and development (R&D) information plays a critical role in guiding investors' evaluations of a firm's future prospects. Numerous works have investigated whether a firm's R&D information is fully impounded in its stock prices (Chan, Lakonishok, and Sougiannis, 2001; Eberhart, Maxwell, and Siddique, 2004; Li, 2011; Cohen, Diether, and Malloy, 2013). However, relatively little attention has been paid to investors' reactions to R&D information quality. R&D information is difficult to evaluate because R&D is usually featured with future-oriented long-term activities in science and technology, whose information is hard to process and whose outcomes are difficult to predict. Moreover, the lack of accounting disclosure suggests that investors may not be fully informed of all information related to firms' R&D activities, creating problems of asymmetric information (Aboody and Lev, 2000). All of these facts make it critical to understand how investors resolve uncertainties inherent in R&D from conceptualization to commercialization when making investment decisions based on publicly available information.

In the theoretical camp, inconsistent predictions on how information quality/uncertainty affects asset returns have been presented. Veronesi (2000) considers a pure exchange economy with power utility preferences and shows that the equity risk premium increases with information quality. However, Brevik and d'Addona (2010) introduce Epstein-Zin recursive preferences into Veronesi's model and find an opposite result. In a production-based long-run risk model, Ai (2010) also finds that high information quality decreases equity premiums. Furthermore, Epstein and Schneider (2008) present a model in which ambiguity-averse agent follows a recursive multiple-priors utility and behaves as if he maximizes expected utility every period under a worst-case belief that is chosen from a set of conditional probabilities and show that investors require compensation for holding assets with low quality information. In this paper, we empirically explore the relationship between R&D information quality and expected stock returns. We examine whether investors demand higher risk premiums from firms whose future performance in R&D is more difficult to evaluate and whether such behaviors affect firms' future stock returns.

However, the measurement of R&D information quality is a delicate issue. Traditional information quality measures such as firm size, firm age, analyst coverage, dispersion in analyst forecasts, return volatility, and/or cash flow volatility (e.g., Zhang, 2006) may not be appropriate in such a context. In this paper, we present a measure of R&D information quality (IQ) as how much variation of a firm's fundamentals can be explained by its R&D expenditures. More specifically, a firm's R&D IQ is measured based on the R-square generated from a regression of its sales growth on its realized R&D capital. When constructing this measure of R&D IQ, we take into account the fact that different firms may have different R&D lifespans and the fact that each year's R&D expenditures may have different effects on future sales growth. Using all firms listed on NYSE, AMEX, and NASDAQ with valid accounting and returns data, we find that R&D IQ is generally persistent, as its one-year-apart persistence is as high as 0.57 ($t = 32.9$), and is not correlated with other firm-specific variables (e.g., size, book-to-market, cash holdings, return of assets, return of equity, and certain innovation-related variables). This suggests that R&D IQ is distinct from well-known firm characteristics and contains different information and that the historical measure of R&D IQ is a good predictor of future R&D information quality.

We hypothesize that when the firm's past track record indicates low R&D information quality (a small R-square), investors face a high degree of uncertainty in evaluating its R&D information and are unwilling to make investments in its future R&D activities: they would require high premiums to make such an investment. To test our hypothesis, we conduct a portfolio analysis similar to Fama and French (1996). At the end of June of each year, we sort all firms into three R&D IQ portfolios (low, middle, and high) based on the 30th and 70th percentiles of R&D IQ in the previous year and construct a hedge portfolio that longs the high IQ portfolio and shorts the low IQ portfolio. We hold these portfolios over the next 12 months and compute their value/equal-weighted monthly returns. We find that average excess portfolio returns decrease with R&D IQ. For example, the low IQ portfolio earns 127 basis points ($t = 4.23$) per month in value-weighted returns and 126 basis points ($t = 4.14$) per month in equal-weighted returns, whereas the high IQ portfolio earns only 78 basis points ($t = 2.56$) per month in value-weighted excess returns and 88 basis points ($t = 2.99$) per month in equal-weighted excess returns. Furthermore, the monthly return on the hedge portfolio is economically substantial and statistically significant, yielding -48 basis points ($t = -4.69$) in value-weighted excess returns and -39 basis points ($t = -4.16$) in equal-weighted excess returns. The same pattern also holds for characteristic- and industry-adjusted returns.

The alphas of factor models also decrease with R&D IQ. More specifically, in the Fama-French three-factor model (Fama and French, 1993), the alpha for the high-minus-low IQ hedge portfolio is -40 basis points ($t = -4.04$) per month in value-weighted excess returns and is -52 basis points ($t = -4.73$) per month in equal-weighted excess returns. The pattern for the estimated alphas for the low, middle, and high IQ portfolios and the hedge portfolio is the same in the Carhart four-factor model (Carhart, 1997). We further investigate risk-adjusted returns using the recently developed q-factor (Hou, Xue, and Zhang, 2015) and M-factor (Stambaugh and Yuan, 2016) models. Again, we find that in both models the alphas for the hedge portfolio is economically substantial and statistically significant in both value- and equal-weighted returns. These results suggest that investors are less certain about the prospects of low IQ firms' future R&D activities and therefore require higher premiums when making such investments.

We further perform Fama-MacBeth cross-sectional regressions that allow us to control for a large number of variables, including size, book-to-market, momentum, leverage, idiosyncratic volatility, illiquidity, and innovation-related variables. Despite such extensive controls, the coefficient on R&D IQ is always negative and statistically significant. This finding provides further evidence in support of our hypothesis that expected excess returns decrease with R&D IQ.

To examine whether high expected excess returns due to low R&D IQ are related to firms' fundamentals, we conduct Fama-MacBeth regressions of firms' future operating performance as measured by return on assets (ROA), cash flows (CF), and performance (PM) on R&D IQ. Even after controlling standard variables in the regressions such as size, book-to-market, leverage, idiosyncratic volatility, illiquidity, and certain innovation-related variables, we find that for all three proxies of fundamentals, the coefficient on IQ is insignificant, whereas coefficients on lagged fundamentals and changes in fundamentals are significant. These findings indicate that our R&D IQ measure is not driven by the undervaluation or overvaluation of fundamental information.

Our hypothesis is that investors are ambiguity-averse and require high premiums to invest in firms with low R&D IQ. We therefore posit that the R&D IQ-return relationship should be stronger for firms with higher information uncertainty, such as firms with smaller market capitalization, younger firms, firms facing more financial constraints, and firms with higher fundamental volatility. These firms may have more uncertain business environments and investors are more ambiguous about their future prospects. To test this hypothesis, we perform independent double sorts on R&D IQ and size, age, the KZ index, and cash-flow uncertainty. We find that the high-minus-low IQ hedge portfolio earns -76 ($t = -2.91$) basis points per month for firms with small size, whereas it earns only -28 ($t = -2.34$) basis points per month for firms with large size. Alphas from the Fama-French three-factor model, the Carhart four-factor model, the q-factor model, and the M-factor model for the hedge portfolio are -99 ($t = -3.02$), -110 ($t = -3.23$),

-121 ($t = -3.31$), and -107 ($t = -2.82$) basis points per month, respectively, for small firms, whereas they become small and marginally significant, -24 ($t = -1.94$), -25 ($t = -1.96$), -30 ($t = -2.25$), and -25 ($t = -1.88$) basis points per month, respectively, for large firms. Our tests on age, the KZ index, and cash-flow uncertainty present the same implications.

To further explore the relationship between R&D IQ and future stock returns and to examine whether the IQ effect reflects commonalities in returns that are not captured by existing factors, we construct a factor-mimicking portfolio for R&D information quality following the similar methodology used in Fama and French (1993). At the end of June of each year, we sort firms independently into two groups based on size (small "S" and big "B") and into three IQ groups (low "L", middle "M", and high "H"). The intersection of these portfolios forms six size-IQ portfolios (S/L, S/M, S/H, B/L, B/M, and B/H). The IQ factor (IQF) is constructed as (S/L + B/L)/2 - (S/H + B/H)/2. We find that the IQF is not highly correlated with commonly used factors such as the market factor (MKT), the size factor (SMB), the value factor (HML), and the momentum factor (MOM). For example, the monthly correlation between IQF and MKT is only about 3%; its monthly correlations with SMB, HML, and MOM are 22%, -13%, and 19%, respectively. We also find that IQF captures a different pricing factor that is distinct from the existing factors through constructions of tangency portfolios. For example, adding IQF to the Fama-French three factors improves the ex post Sharpe ratio of the tangency portfolio by 14% with the weight on IQF being 42%, which is larger than weights on MKT (26%), SMB (25%), and HML (7%).

Our study relates and contributes to two strands of literature. On the one hand, many works have examined whether asset prices fully impound information contained in the innovation process. Chan, Lakonishok, and Sougiannis (2001) find that R&D intensity measured as R&D expenditures relative to the market value of equity has the ability to predict future returns. However, its predictability power disappears when the ratio of R&D expenditures to sales is used. Eberhart, Maxwell, and Siddique (2004) empirically report significantly positive long-term abnormal stock returns following unexpected and economically significant increases in R&D and argue that R&D increases are beneficial investments, but the market underreacts to this benefit. Li (2011) argues that the positive relationship between R&D intensity and stock returns exists only in financially constrained firms, and this relationship is robust to measures of R&D intensity. Cohen, Diether, and Malloy (2013) demonstrate that firm-level innovation is persistent and predictable, but the market appears to ignore the publicly available information in R&D when valuing future innovation. Gu (2005) finds that changes in patent citations relative to total assets are positively related with firm's future earnings and stock returns. Pandit, Wasley, and Zach (2011) show that firm's patent citations positively associate with its future operating performance. Hirsleifer, Hsu, and Li (2013, 2015) construct an innovative efficiency (IE) measure and an innovative originality (IO) measure, respectively, using the number of patents and patent citations of a firm and find that both IE and IO positively predict the future stock returns. They mainly attribute this positive IE/IO-return relationship to limited investor attention. However, unlike the above works, our paper focuses on information quality/uncertainty related to the innovation process by relating innovation input (R&D) and innovation outcome (sales).

On the other hand, how information quality/uncertainty affects asset returns has attracted considerable attention. Veronesi (2000) finds that the equity risk premium increases with information quality, whereas Brevik and d'Addona (2010) and Ai (2010) find an opposite result. Chen and Epstein (2002) show in a theoretical model that excess return should be composed of a risk premium and a premium for Knightian uncertainty (ambiguity), and Epstein and Schneider (2008) make a further refinement and present a model in which investors demand a premium for holding assets with low quality information. Zhang (2006) implements an empirical investigation on the relationship between information uncertainty and stock returns. He finds that greater information uncertainty leads to higher expected excess returns following good news but lower returns following bad news. In this paper, we focus on this seemingly contentious issue using information contained in a firm's R&D activities and empirically investigate the relationship between R&D information quality and future stock returns. We find robust empirical results that expected excess returns contain a premium for R&D information quality and that the higher information quality is, the smaller the future excess returns will be.

The rest of the paper is organized as follows. Section 2 introduces the data and provides summary statistics. Section 3 investigates R&D information quality and return predictability using portfolio analysis and Fama-MacBeth cross-sectional analysis. Section 4 implements several robustness checks. Section 5 provides further evidence on the return predictability power of R&D information quality. Section 6 constructs a R&D information quality factor. And Section 7 concludes the paper.

2. Data and Summary Statistics

In this section, we present the data used for empirical analysis and construct our measure of R&D information quality by connecting innovation input (R&D) and innovation outcome (sales) in Subsection 2.1, and report summary statistics of the R&D information quality measure in Subsection 2.2.

2.1. Data and R&D Information Quality

The sample we use in this paper combines different data sources and covers the period ranging from July of 1980 to July of 2012. We obtain firm-specific accounting data such as R&D expenditures, sales, and book equity from Compustat, and monthly stock returns, shares outstanding, and volume capitalization from the Center for Research in Security

Prices (CRSP). All common stocks trading on the NYSE, AMEX, and NASDAQ with valid accounting and return data are included in the sample. Firms need to be listed on Compustat for two years before included in our sample. We exclude financial firms, which have four-digit standard industrial classification (SIC) codes between 6,000 and 6,999 (finance, insurance, and real estate sectors). Like Fama and French (1993), we discard closed-end funds, trusts, American Depository Receipts, Real Estate Investment Trusts, units of beneficial interest, and firms with negative book equity. For some of our tests, we also use the firm-level patent-related data, which are mainly drawn from the updated National Bureau of Economic Research (NBER) patent database originally developed by Hall, Jaffe, and Trajtenberg (2001). However, these data are only available to December of 2006.

In general, information quality is measured by signal precision. Suppose that there is a parameter θ that an investor wants to learn, and assume that a signal s is related to the parameter by $s = \theta + \epsilon$, where $\epsilon \sim (0, \sigma_s^2)$. Then the signal precision is given by $h_s = 1/\sigma_s^2$. In practice, σ_s^2 is unknown and must be estimated, $\hat{\sigma}_s^2 = Var(e)$, and $e = s - \theta$. However, this estimate is sensitive to outliers when the sample size used is small, and this is exactly the case in our study. We therefore rely on R-square: $R^2 = 1 - \frac{Var(e)}{Var(s)}$, which captures variation of s explained by θ and which is more robust to outliers. There is clearly a one-to-one relationship between signal precision and R-square. The smaller the R-square is, the lower information quality is.

This paper focuses on how investors resolve uncertainties inherent in R&D from conceptualization to commercialization. The main objective of a firm's R&D activities is to develop new products, services, and technologies, whose success is directly reflected in its sales. We therefore measure R&D information quality by assessing how much variation of a firm's sales growth can be explained by its R&D expenditures. Given the facts that different firms may have different R&D lifespans and each year's R&D expenditures may have different effects on firms' future sales growth, we regress sales growth separately on each of the past five-year's realized R&D capitals and take the largest resulting R-square as our measure of information quality. More specifically, our regression for any firm i for each year t takes the following form

$$\log\left(\frac{sales_{i,t}}{sales_{i,t-1}}\right) = \alpha_{i,j} + \beta_{i,j}\log(1 + RDC^k_{i,t-j}) + \epsilon_{i,t}, \tag{1}$$

where

$$RDC^k_{i,t-j} = \frac{k}{10}RD_{i,t-j-1} + (\frac{k}{10})^2 RD_{i,t-j-2} + (\frac{k}{10})^3 RD_{i,t-j-3} + (\frac{k}{10})^4 RD_{i,t-j-4}, \tag{2}$$

for $j = 1, 2, ..., 5$ and $k = 1, 2, ..., 9$, where RD stands for R&D expenditures and RDC for the realized R&D capital. When constructing their innovative ability measure, Cohen, Diether, and Malloy (2013) employ a similar regression to Equation (1) where R&D expenditures scaled by sales instead of R&D capital are used in the left-hand side. Our definition of R&D capital in Equation (2) is more flexible than that defined in Chan, Lakonishok, and Sougiannis (2001) who assume that the productivity of R&D spending declines linearly by 20 percent each year. The key point in Equations (1) and (2) is that simply applying the same timespan between R&D input and output and the same R&D productivity decay rate to all firms is too restrictive. Some firms can take longer time to materialize R&D spending (e.g., pharmacy) than other firms (e.g., utilities); and technologies and services in some industries (e.g., chemicals) can be utilized for longer time periods than those in other industries (e.g., machinery). To accommodate these concerns, we regress sales growth on R&D inputs in two dimensions: the timespan between R&D input and output (j) and the R&D capital decay rate (k). Equation (1) assumes that R&D capitals from year $t-5$ to $t-1$ are of relevance to sales of year t. In fact, we have tried other assumptions on R&D lifespan for up to ten years. We find that our results are virtually the same.

The regression is run for each firm for each fiscal year t on time series from year $t-7$ to year t. We require that there are at least 6 valid observations on R&D expenditures and that at least 4 RDCs are non-zero. As a result, for each firm-year, there are 45 regressions in total. R&D information quality is then defined as

$$IQ_{i,t} = \max\left\{R^2(j,k)\right\}, \tag{3}$$

where $R^2(j,k)$ is the R-square resulting from the regression in Equation (1). The selection of the largest value of R-squares in Equation (3) is for finding the most relevant R&D lifespan and decay rate for each firm and should not pose any problems to our study, as we hypothesize that the lower information quality is, the higher excess future returns will be. Taking maximum here in fact works against finding any significant results.

2.2. Summary Statistics

Panel A of Table 1 presents the pooled mean, standard deviation, median, and 25th and 75th percentiles of the R&D IQ measure for each industry according to Fama-French 17 industry classifications. It also presents the number of firms in each industry included in our sample, the market share of each industry in our sample, and the market share of each industry in the universal sample.

The mean (median) value of the R&D IQ measure varies across industries from 0.41 (0.37) for Automobiles to 0.67 (0.69) for Utilities. The standard deviation does not vary considerably across industries: its minimum is 0.20 for Transportation, and its maximum is 0.25 for Retail Stores. However, the number of firms in each industry included in our sample varies considerably. For example, the sample includes 583 Machinery and Business Equipment firms, 122 Drugs, Soap, Perfumes, and Tobacco firms, 11 Mining and Minerals firms, and only 4 Utilities firms. This indicates that it is crucial to control for industry effects when examining the R&D IQ-return relationship. A comparison of market share of each industry in our sample and that in the universal sample shows that for most industries, these two shares are similar. For example, the market share of Textiles, Apparel & Footwear in our sample is 0.9% while it is 0.8% in the universal sample; the market share of Food in our sample is 4.9% while it is 5.7% in the universal sample. However, there are a few exceptions: Utilities in our sample accounts for only 0.1% whereas it accounts for 5.3% in the universal sample; and Machinery accounts for 28.2% in our sample but only 13.8% in the universal sample. This suggests that while our sample is economically meaningful and representative overall, it is necessary to control for industry effect in our study.

In Panel B, we present average values of certain firm-specific variables for the three R&D IQ portfolios constructed according to the 30th and 70th percentiles of the lagged IQ (see Section 3 for a detailed discussion). These variables include (log) market equity (ME), the book-to-market ratio (BEME), return on assets (ROA, Income before extraordinary items plus interest expenses divided by lagged total assets), return on equity (ROE, Income before extraordinary items plus interest expenses divided by lagged common equity), leverage (DXA, long-term debt plus debt in current liabilities divided by total assets), cash holding, the industry concentration index (HHI, Hou and Robinson, 2006), and idiosyncratic volatility (IVOL, the standard deviation of Fama-French three-factor residuals for the past 12 months). We also construct some innovation-related variables: R&D intensity (RDA, R&D expenditures divided by total assets), innovation ability (InnAb, Cohen, Diether, and Malloy, 2013), and innovative efficiency (IE, Hirsleifer, Hsu, and Li, 2013).

We find that innovation ability increases with R&D information quality. It is 0.29, 0.57, and 0.80 for the low, middle, and high IQ portfolios, respectively, whereas the innovative efficiency of low and high IQ portfolios is quite similar (3.08) and is larger than that of the middle IQ portfolio (2.66). We note that size increases with R&D IQ: the high IQ portfolio has larger size than the low IQ portfolio (278 vs. 205). The low IQ portfolio presents a larger ROA but a smaller ROE than the high IQ portfolio. Book-to-market, R&D intensity, leverage, cash holding, industry concentration, and idiosyncratic volatility do not vary significantly among these three portfolios.

Panel C reports the time-series average of cross-sectional correlations between IQ and the above-mentioned firm characteristics. IQ is weakly correlated with these variables. Correlations range from -0.05 (with ROA) to 0.07 (with InnAb). Correlations with ROE, R&D intensity, and IE are the smallest, ± 0.01. The above findings indicate that our R&D IQ measure is distinct from well-known firm characteristics and may contain different information.

Panel A of Figure 1 presents the R&D IQ periodic cross-sectional persistence, which is computed for each time t as the cross-sectional correlation of R&D IQ between time t and time $t-1$. It is evident that the persistence varies from 0.39 (between years 2006 and 2007) to 0.73 (between years 1983 and 1984) over time. The time-series average of the periodic cross-sectional persistence is about 0.57 ($t = 32.9$), indicating that R&D IQ is a fairly persistent variable. Economically, this suggests that the historical measure of R&D IQ is a good predictor of future R&D information quality, making the historical data-based value a useful measure for analyses aimed at discerning the relation between R&D IQ and future stock returns.

3. R&D Information Quality and Return Predictability

In this section, we examine the relationship between R&D IQ and future stock returns. Our main hypothesis is that there is a premium for information uncertainty and that expected excess returns should decrease with R&D information quality. We first implement portfolio sorts in Subsection 3.1, then perform Fama-MacBeth cross-sectional regressions in Subsection 3.2, and finally investigate effects of R&D IQ on firms' subsequent operating performance in Subsection 3.3.

3.1. Portfolio Analysis

We first examine R&D information quality and return predictability using portfolio sorts. Similar to Fama and French (1996), we sort all firms into three R&D IQ portfolios (low, middle, and high) at the end of June of each year from 1981 to 2012. The low IQ portfolio contains all stocks below the 30th percentile in R&D IQ, and the high IQ portfolio contains all stocks above the 70th percentile in R&D IQ. Stocks between the 30th and 70th percentiles belongs to the middle IQ portfolio. We further form a hedge portfolio that longs the high IQ portfolio and shorts the low IQ portfolio.

3.1.1. Portfolio Returns

We hold these portfolios over the next twelve months and compute their value/equal-weighted monthly returns. Panel A of Table 2 presents both value- and equal-weighted average monthly returns in excess of one-month Treasury bill rates for these portfolios. We find that portfolio returns decrease with R&D IQ. This result holds for both value- and equal-weighted excess returns. For example, the low IQ portfolio earns 127 basis points ($t = 4.23$) per month in value-weighted excess returns and 126 basis points ($t = 4.14$) per month in equal-weighted excess returns. However, the high IQ portfolio only earns 88 basis points ($t = 2.99$) per month in value-weighted excess returns and 78 basis points ($t = 2.56$) in equal-weighted excess returns. More importantly, the monthly returns of the high-minus-low hedge portfolio are economically substantial and statistically significant. The hedge portfolio earns -39 basis points ($t = -4.16$) and -48 basis points ($t = -4.69$) per month in value- and equal-weighted excess returns, respectively.

To ensure that our results are robust for firm characteristics and industry effects, characteristic- and industry-adjusted returns are also reported. Characteristic-adjusted returns are computed following Daniel et al. (1997) as the difference between individual firms' returns and 125 size/book-to-market/momentum benchmark portfolios, and industry-adjusted returns are calculated as the difference between individual firms' returns and the returns of firms in the same industry according to the Fama-French 17 industry classifications. Characteristic- and industry-adjusted returns also indicate that portfolio returns decrease with R&D IQ. The low IQ portfolio earns 18 basis points ($t = 2.25$) and 30 basis points ($t = 2.46$) per month in value-weighted characteristic- and industry-adjusted returns, respectively, and 16 basis points ($t = 1.98$) and 14 basis points ($t = 0.91$) per month in equal-weighted characteristic- and industry-adjusted returns, respectively. However, the high IQ portfolio only earns -24 basis points ($t = -2.70$) and -15 basis points ($t = -1.37$) per month in value-weighted characteristic- and industry-adjusted returns, respectively, and -33 basis points ($t = -3.36$) and -36 basis points ($t = -2.08$) per month in equal-weighted characteristic- and industry-adjusted returns, respectively. Returns on the hedge portfolio are again economically substantial and statistically significant in characteristic- and industry-adjusted returns, -42 basis points ($t = -4.23$) and -45 basis points ($t = -4.80$) in value-weighted characteristic- and industry-adjusted returns, respectively, and -50 basis points ($t = -4.46$) and -50 basis points ($t = -4.90$) in equal-weighted characteristic- and industry-adjusted returns, respectively.

Panel B of Figure 1 presents the time series of annual equal-weighted and value-weighted excess returns on short position of the hedge portfolio for July of 1981 to July of 2012. We find that annual returns to this strategy are relatively stable over time. Volatility is about 7.9% for value-weighted returns and is about 8.5% for equal-weighted returns, whereas it is about 17.3% for excess market returns for the same period. The annual correlation between returns of this strategy and excess market returns is modest: it is about 28.3% for value-weighted returns and is about 30.3% for equal-weighted returns.

3.1.2. Risk-Adjusted Returns

We further examine whether R&D IQ portfolio excess returns can be explained by commonly used risk factors. We consider the Fama-French three-factor model (Fama and French, 1993),

$$r_i - r_f = \alpha_i + \beta_{i,MKT}MKT + \beta_{i,SMB}SMB + \beta_{i,HML}HML + e_i, \tag{4}$$

and the Carhart four-factor model (Carhart, 1997),

$$r_i - r_f = \alpha_i + \beta_{i,MKT}MKT + \beta_{i,SMB}SMB + \beta_{i,HML}HML + \beta_{i,MOM}MOM + e_i, \tag{5}$$

where $r_i - r_f$ denotes portfolio returns in excess of one-month T-bill rates, and MKT, SMB, HML, and MOM are the usually used market, size, value, and momentum factors, respectively. We also consider the two recently developed factor models: the q-factor model (Hou, Xue, and Zhang, 2015),

$$r_i - r_f = \alpha_i + \beta_{i,MKT}MKT + \beta_{i,SMB}SMB + \beta_{i,I/A}I/A + \beta_{i,ROE}ROE + e_i, \tag{6}$$

where I/A is the investment factor, which is constructed as the difference between the return on a portfolio of low investment stocks and the return on a portfolio of high investment stocks, and ROE is the profitability factor constructed as the difference between the return on a portfolio of high profitability stocks and the return on a portfolio of low profitability stocks; and the mispricing-factor model (Stambaugh and Yuan, 2016),

$$r_i - r_f = \alpha_i + \beta_{i,MKT}MKT + \beta_{i,SMB}SMB + \beta_{i,MGMT}MGMT + \beta_{i,PERF}PERF + e_i, \tag{7}$$

where $MGMT$ and $PERF$ are referred to as the mispricing factors, which aggregate information across 11 well-known anomalies by averaging rankings within two clusters exhibiting the greatest co-movement in long-short returns. The first cluster of anomalies represent quantities that firms' managements can affect directly, and the factor arising from it is MGMT. The second cluster is related more to performance and is less directly controlled by management, and the factor constructed from this cluster is PERF. There is some evidence that both the q- and mispricing-factor models perform better than the Fama-French three-factor model, the Carhart four-factor model, and the Fama-French five-factor model (Fama and French, 2015) in explaining most of anomalies (Hou, Xue, and Zhang, 2015, 2016; Stambaugh and Yuan, 2016).

Panels B, C, D, and E of Table 2 present alphas and factor loadings from regressing portfolio excess returns on the Fama-French three-factor model, on the Carhart four-factor model, on the q-factor model, and on the mispricing-factor model, respectively. The alpha estimates deliver the same implication as shown above. In all four models, the alpha is positive for the low IQ portfolio, whereas it is negative for the high IQ portfolio in both value- and equal-weighted returns. For example, in the q-factor model, it is 17 basis points ($t = 1.52$) per month in value-weighted returns and 28 basis points ($t = 2.61$) per month in equal-weighted returns for the low IQ returns; however, it is only -21 basis points ($t = -1.71$) per month in value-weighted returns and -24 basis points ($t = -1.89$) per month in equal-weighted returns for the high IQ portfolio.

More importantly, the hedge portfolio's alpha is negative, economically substantial, and highly statistically significant in the four models. For example, the alpha from the Fama-French three-factor model is -40 basis points ($t = -4.04$) per month in value-weighted returns and -52 basis points ($t = -4.73$) per month in equal-weighted returns; the alpha from the Carhart four-factor model is -38 basis points ($t = -3.71$) per month in value-weighted returns and -49 basis points ($t = -4.40$) per month in equal-weighted returns; the alpha resulted from the q-factor model is -38 basis points ($t = -3.50$) per month in value-weighted returns and -53 basis points ($t = -4.38$) per month in equal-weighted returns; and it is -33 basis points ($t = -2.96$) per month in value-weighted returns and -43 basis points ($t = -3.54$) per month in equal-weighted returns in the mispricing-factor model. These findings indicate that investors are uncertain about low IQ firms' future R&D activities and typically require higher compensation when making investments in their future R&D.

In the Fama-French three-factor model and the Carhart four-factor model, all three IQ portfolios load positively and significantly on market, size, and value factors but negatively and significantly on momentum. However, in the q- and mispricing-factor models, these three IQ portfolios load positively and significantly on market and size factors, but load nearly insignificant on investment, profitability, and mispricing factors. Factor loadings for the hedge portfolio are small and hardly significant in all these four models, indicating that returns on this portfolio do not covary with any of these well-known factors and suggesting that there may be important factor(s) missed in addition to these well-known factors.

3.2. Fama-MacBeth Cross-Sectional Analysis

In this section, we test the return predictive power of R&D IQ by employing monthly Fama-MacBeth cross-sectional regressions (Fama and MacBeth, 1973). This analysis allows for extensive controls of industry effects and of variables that have been found to have predictive power for stock returns. To be specific, we control for size (Banz, 1981), book-to-market ratio (Fama and French, 1992), and momentum (Carhart, 1997). We also consider leverage (Miller and Modigliani, 1958; Ozdagli, 2012), illiquidity (Amihud, 2002), idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang, 2006; Bali, Cakici, and Whitelaw, 2011), one-month lagged returns, turnover, capital expenditures (CapEX), and industry concentration (HHI) in our regressions. An industry dummy is also introduced to control for any industry-related effects that may drive our results. Definitions of these variables have been given in Section 2.

For each month from July of year t to June of year $t+1$, we regress monthly excess returns of individual stocks on our R&D IQ measure and the above control variables of year $t-1$. Table 3 presents the regression results, which confirm our hypothesis: the lower R&D information quality is, the higher excess returns we expect, as the coefficient on IQ in each regression we consider is negative and statistically significant. Model 1 in the table considers a simple regression in which we exclude all control variables and take R&D IQ as the only predictor. The coefficient on R&D IQ is -0.81 and is highly statistically significant ($t = -4.36$), and the adjusted R^2 is about 1.3% and is highly statistically significant ($t = 6.46$). When we introduce size, book-to-market, momentum, leverage, lagged returns, and turnover in Model 2, the slope estimate on IQ becomes slightly small, -0.69, but remains highly significant ($t = -3.67$). Coefficients on the other variables are not statistically significant. The adjusted R^2 increases to 5.4% ($t = 13.0$). Our sample mainly includes R&D-intensive firms, whose average size is larger than that of the universal sample. This may explain insignificance of the traditional control variables.

Our R&D information quality measure constructed from Equation (3) depends on R-square from regression (1), where volatility of sales growth is also an important input. In Model 3, except variables introduced in Model 2, we also control volatility of sales growth, which is measured as the standard deviation of sales growth for the period from year $t-7$ to year t. We find that the coefficient on IQ is still negative and statistically significant, -0.61 ($t = -3.39$). We also find that volatility of sales growth significantly and adversely affects the future stock returns, -1.02 ($t = -2.50$). The adjusted R^2 in this model is 6.3% ($t = 14.3$). We further introduce idiosyncratic volatility and illiquidity in Model 4 and find that the coefficient of our interest, IQ, is -0.53 ($t = -2.94$) and the adjusted R^2 is further increased to 7.2% ($t = 15.2$). The coefficient on turnover is negative and significant, -0.59 ($t = -2.91$), and the coefficient on IVOL is positive and marginally significant, 0.04 ($t = 1.91$), consistent with Bali, Cakici, and Whitelaw (2011). Model 5, which introduces capital expenditures to Model 2, and Model 6, which introduces industry concentration to Model 2, deliver similar implication that the coefficient on IQ is still negative and significant. We note that in both Models 5 and 6, the coefficient on turnover is negative and significant and that in Model 5, the coefficient on CapEx is negative and highly significant. The adjusted R^2s from these two models are 5.6% ($t = 13.5$) and 5.5% ($t = 13.2$), respectively.

Recently, several works reveal a positive R&D-return relationship. Chan, Lakonishok, and Sougiannis (2001) find that firms with high ratios of R&D expenditures to market equity earn high subsequent returns. Eberhart, Maxwell, and Siddique (2004) show that significant R&D increases predict positive future abnormal returns. Cohen, Diether, and Malloy (2013) construct an innovation ability measure and argue that firms that exhibit high ability in the past and that continue to spend a large amount of R&D outperform in the future. We therefore introduce variables of (i) R&D intensity: RDS (R&D expenditures scaled by sales) and RDA (R&D expenditures relative to total assets); (ii) significant R&D increases (RDG); and (iii) innovation ability in our regressions. The estimates in Model 7 show that after controlling for R&D intensity, RDA, and the variables used in Model 2, our estimate on IQ is still negative and significant, -0.64 ($t = -3.46$), and the coefficient on RDA is not significant. We find that in this model, the coefficient on turnover is negative and significant. Furthermore, the estimates in Model 8 show that even after controlling for RDS and innovation ability (InnAb), we obtain the similar result as obtained before: the estimate on IQ is negative and significant, -0.56 ($t = -3.04$). The coefficient on RDS is insignificant and the coefficient on InnAb is negative and significant, -0.28 ($t = -2.53$). Whenever we introduce RDG in Model 9, the coefficient on IQ is -0.73

($t = -3.49$), but the coefficient on RDG is not significant. The adjusted R^2s from Models 7, 8, and 9 are 6.0% ($t = 13.7$), 6.2% ($t = 14.1$), and 5.3% ($t = 11.6$), respectively.

Hirsleifer, Hsu, and Li (2013) show that firms' patents and patent citations contain rich information on future stock returns. We therefore construct their innovative efficiency measure (IE) and include it in our test. As the NBER patent database only runs to December of 2006, our sample here is from July of 1980 to July of 2006. Model 10 shows that when even IE is introduced into the Fama-MacBeth regression, the coefficient on IQ is still negative and statistically significant, -0.69 ($t = -2.78$), whereas the coefficient on IE is insignificant. The adjusted R^2 is about 5.7% ($t = 11.2$).

The results in Subsections 3.1 and 3.2 provide empirical support for some theoretical conclusions. Brevik and d'Addona (2010) find that high information quality decreases the equity premium in a pure exchange economy with Epstein-Zin recursive preferences. In a production-based long-run risk model, Ai (2010) also find that high information quality decreases the equity premium. Epstein and Schneider (2008) show that in markets with ambiguous information, expected excess returns decrease with future information quality, and ambiguity-averse investors require compensation to hold assets with low quality information. Furthermore, as R&D spending is usually regarded as a good signal of firms' future prospects, our results are also consistent with those in Zhang (2006), who implements an empirical investigation on information uncertainty and stock returns and who finds that greater information uncertainty leads to higher expected excess returns following good news but lower returns following bad news.

3.3. R&D IQ and Subsequent Operating Performance

If R&D IQ constructed in Equation (3) really captures information quality in a firm's R&D, it should have little effects on its fundamentals. We take return on assets (ROA), cash flows (CF), and performance (PM, operating income before depreciation scaled by lagged sales) as proxies for fundamentals and examine the relationship between IQ and subsequent operating performance via Fama-MacBeth regressions. As before, we control for size, book-to-market, leverage, idiosyncratic volatility, illiquidity, and some innovation-related variables such as R&D intensity and innovation ability. We also introduce the lagged values and changes in fundamental variables in the regressions.

Table 4 reports the Fama-MacBeth regression results. We find that for all three proxies of fundamentals, coefficients on IQ are insignificant. For example, the coefficient on IQ is -0.01 ($t = -1.36$) in the ROA regression, 0.02 ($t = 1.53$) in the PM regression, and -0.01 ($t = -1.47$) in the CF regression. For these three regressions, coefficients on the lagged fundamentals and changes in fundamentals are highly statistically significant except for changes in PM. We also find that the coefficient on size is highly statistically significant in all cases, indicating that the larger a firm's size is, the better its subsequent performance is. These findings suggest that our IQ measure is not related to undervaluation or overvaluation to public information already known to the market.

4. Robustness Checks

4.1. Long-Term Cumulative Returns

We further test our hypothesis by examining long-term cumulative portfolio returns. As stated above, at the end of June of each year, we construct three R&D IQ portfolios and a hedge portfolio and hold them over the next 12, 24, and 36 months. Table 5 reports the value-weighted excess returns and the three-, four-, q-, and M-factor alphas for these portfolios. Even though the low, middle, and high IQ portfolios present similar cumulative returns and the hedge portfolio's cumulative return is not significant for the past 12 months, the low IQ portfolio earns much higher cumulative return, 13.38% ($t = 4.77$), than the high IQ portfolio, 9.73% ($t = 2.99$), and the hedge portfolio's cumulative return is statistically significant, -3.64% ($t = -2.80$) for the future 12 months. A similar pattern is found for the three-, four-, q-, and M-factor alphas.

All returns and alphas for the three IQ portfolios and hedge portfolio keep increasing and are statistically significant for the future 24 and 36 months. The hedge portfolio's cumulative return is -5.84% ($t = -2.66$) for the future 24 months and is -8.60% ($t = -2.89$) for the future 36 months, and its three-, four-, q-, and M-factor alphas are -6.27% ($t = -2.64$), -5.09% ($t = -1.75$), -5.62% ($t = -2.66$), and -5.31% ($t = -2.03$), respectively, for the future 24 months, and are -11.3% ($t = -3.82$), -8.40% ($t = -2.31$), -8.45% (-2.43), and -8.49% ($t = -2.82$), respectively, for the future 36 months. We do not find any reversal, suggesting that our IQ measure does capture a premium for information quality rather than any form of overreaction.

4.2. R&D IQ Effect and Investor Sentiment

We also investigate whether the R&D IQ effect observed above is caused by behavioral bias. To do this, we regress the high-minus-low R&D IQ portfolio returns (Spread) on investor sentiment indices proposed by Baker and Wurgler (2006), who construct a composite index (SENT1) that captures the common component in the six proxies of investor sentiment, i.e., the close-end fund discount, turnover, the number of IPOs, the average first-day returns, equity share, and the dividend premium. Furthermore, to avoid effects of a common business cycle component, they construct a second sentiment index (SENT2) that explicitly removes business cycle variations.

Table 6 presents our regression results. It is evident that regardless of whether value- or equal-weighted R&D IQ spread returns are used, coefficients on SENT1, SENT2, ΔSENT1, and ΔSENT2, and on their corresponding lagged values are statistically insignificant. The adjusted R^2s from all regressions are almost zero. For example, for value-weighted spread returns, the coefficient on SENT1 is 0.022 with the $t-$statistics of

0.14 and the adjusted R^2 is about -0.003, and the coefficient on SENT2 is -0.021 with the $t-$statistics of -0.12 and the adjusted R^2 is about -0.003. For equal-weighted spread returns, the coefficient on SENT1 is 0.001 with the $t-$statistics of 0.00 and the adjusted R^2 is about -0.003, and the coefficient on SENT2 is -0.067 with the $t-$statistics of -0.36 and the adjusted R^2 is about -0.002. Regressions of R&D IQ spread returns on ΔSENT1, ΔSENT2, and the lagged values of the above variables present similar results, suggesting a lack of explanatory power of investor sentiment to abnormal returns on R&D IQ portfolios.

4.3. Other Robustness Checks

We argue that the success of a firm's R&D activities is directly reflected in its sales. We therefore use sales growth as a signal to construct R&D IQ. To have a robustness check that the R&D IQ-return relationship we have found is not a reflection of the negative relation of sale growths and future returns (Lakonishok, Shleifer, and Vishny, 1994), we replace sales growth by return-on-assets (ROA), which should also be affected by firms' R&D activities, in regression (1) to construct our R&D IQ measure. Following the same procedure, we find similar results as in Section 3. For instance, the analog of the hedge portfolio in Table 2 has monthly excess return of -22 basis points ($t = -2.34$) in equal-weighted returns and of -18 basis points ($t = -2.11$) in value-weighted returns.

As an alternative robustness check, we replace R&D capital by a more tangible variable, capital expenditure, in regression (1). In this case, we find that the results we have observed in Section 3 are completely disappeared. These two robustness checks suggest that the measure constructed in Equation (3) does capture R&D information quality and there exists a premium for R&D information quality.

5. Further Empirical Evidence

In this section, we provide further evidence on the relationship between R&D IQ and future stock returns. If our R&D IQ measure really captures information quality in firms' R&D, and there exists a premium for R&D IQ in excess returns, we conjecture that the relationship should become stronger and the premium should be larger in firms with smaller size, younger age, greater financial constraints, and higher return and fundamental volatility, as such firms generally operate in more uncertain business environments and investors are more ambiguous to their future prospects.

We perform independent double sorts on R&D IQ and firm size, firm age, financial constraint, return volatility, and fundamental volatility. At the end of June of each year, we first sort all firms into three portfolios based on each of the above conditioning variables and then sort each of these three portfolios into three subgroups based on R&D IQ and form a high-minus-low IQ hedge portfolio in each of these three portfolios. In his study, Zhang (2006) uses firm size, firm age, stock return volatility, and cash flow volatility (as well as analyst coverage) to measure information uncertainty. In what follows, we only report results based on value-weighted portfolio returns. The results in equal-weighted returns are similar and are available in an unreported appendix, in which we also implement monthly Fama-MacBeth cross-sectional regressions across subsamples split by the above conditioning variables, respectively, and find the same results as those presented below.

5.1. Firm Size

We measure firm size by its market capitalization. Small firms usually have more expensive access to external financial fundings, are more likely to be growing firms in rapidly developing and intrinsically volatile industries, are less diversified, and have more serious asymmetric information problems. Banz (1981) regards firm size a proxy for risk; Amihud and Mendelson (1986) and Liu (2006) find that the size effect is linked to liquidity risk; Zhang (2006) takes firm size as a proxy for information uncertainty.

Table 7 presents the double-sorting results in value-weighted returns, which strongly confirm our conjecture. From Panel A, we find that the hedge portfolio's returns and alphas are economically substantial and statistically significant for small firms, whereas they become smaller (though still significant) for big firms. For example, monthly excess returns and characteristic- and industry-adjusted returns are -96 basis points ($t = -2.91$), -100 basis points ($t = -3.02$), and -104 basis points ($t = -3.03$), respectively, for small firms, whereas they are only -28 basis points ($t = -2.34$), -32 basis points ($t = -2.57$), and -30 basis points ($t = -2.58$), respectively, for big firms. The Fama-French three-factor alpha is -99 basis points ($t = -3.02$) per month, the Carhart four-factor alpha is -110 basis points ($t = -3.23$) per month, the q-factor alpha is -121 basis points ($t = -3.31$) per month, and the M-factor alpha is -107 basis points ($t = -2.82$) per month, for small firms. However, these alphas are only -24 basis points ($t = -1.94$), -25 basis points ($t = -1.96$), -30 basis points ($t = -2.25$), and -25 basis points ($t = -1.88$), per month for big firms.

5.2. Firm Age

Young firms may face liability of newness (Stinchcombe, 1965). They are vulnerable to unexpected shocks, and their growth paths are difficult to predict. This makes their future prospects more ambiguous to investors. By constrast, old firms may have smoother growth paths with fewer bumps and surprises and usually have more easy-to-access information available to investors (Barry and Brown, 1985). Investors should become concerned when they observe low R&D quality information from young firms.

Panel B reports portfolio results based on firm age and R&D IQ. Firm age is defined as the number of years listed on Compustat with non-missing price data. Consistent to our conjecture, we find that for young firms, the low IQ portfolio always earns higher returns per month, which are always statistically significant, than the high IQ portfolio, whose returns are hardly significant. For example, for young firms, the excess return and characteristic- and industry-adjusted returns are 150 basis points ($t = 4.13$), 41 basis points ($t = 2.46$), 55 basis points ($t = 2.94$) per month, respectively, and the three-, four-, q-, and M-factor alphas are 42 basis points ($t = 2.27$), 43 basis points ($t = 2.22$), 64 basis points ($t = 3.07$), and 39 basis points ($t = 1.81$) per month, respectively, for the low IQ portfolio, whereas all three returns and alphas are smaller for the high IQ portfolio. However, the pattern that holds for young firms is hardly observed for old firms.

Furthermore, the high-minus-low IQ hedge portfolio earns much more substantial and significant returns and alphas per month in young firms than in old firms. The hedge portfolio earns -120 basis points of excess return per month, -112 basis points of characteristic-adjusted return per month, and -123 basis points of industry-adjusted return per month, all of which are statistically significant at 1% level, and its three-, four-, q-, and M-factor alphas are -110 basis points ($t = -3.98$), -91 basis points ($t = -3.22$), -103 basis points ($t = -2.98$), and -75 basis points ($t = -2.31$) per month, respectively. However, for older firms, the hedge portfolio's returns and alphas are very small and completely insignificant.

5.3. Firm Financial Constraints

Firms with financial constraints have limited capacities to fund their desired investments. Lamont, Polk, and Saá-Requejo (2001) show that financial constraints affect firm value and that the severity of constraints varies over time, but constrained firms surprisingly earn lower returns than unconstrained firms. However, Whited and Wu (2006) find that more constrained firms earn higher average returns than less constrained firms, but the difference is not significant. Livdan, Sapriza, and Zhang (2009) revisit the relationship between financial constraints and stock returns and find that more financially constrained firms are riskier and earn higher expected stock returns than less financially constrained firms. Campello and Chen (2010) find evidence suggesting that financially constrained firms have higher systematic risk and that the constraint risk is priced in the financial markets. Li (2011) finds that the positive R&D-return relationship only exists for financially constrained firms.

Financial constraint is alway related to firm size and firm age. Small firms and young firms are usually considered to be more financially constrained than larger firms and old firms. For example, Li (2011) takes firm size and firm age as two proxies for financial constraint. We show above that for small and young firms, the relationship between R&D IQ and future stock returns is much stronger than that for large and old firms, indirectly indicating that investors require higher premium for ambiguous R&D information quality. Here, we further investigate this implication by using a more formal measure of financial constraint: the KZ index (Kaplan and Zingales, 1997).

Panel C compares R&D IQ effect between financially constrained (high KZ index) firms and financially unconstrained (low KZ index) firms. We find that the IQ effect is much stronger for firms with high KZ index. The returns and alphas of the hedge portfolio are large and statistically significant for financially constrained firms, whereas they become small and always insignificant for financially unconstrained firms. For example, the monthly excess return, and characteristic- and industry-adjusted returns of the hedge portfolio are -42 basis points ($t = -3.18$), -41 basis points ($t = -3.02$), and -51 basis points ($t = -3.83$), respectively, for high KZ index firms, whereas they are only -23 basis ponts, -37 basis points, and -41 basis points, respectively, and are not statistically significant for low KZ index firms. The three-, four-, q-, and M-factor alphas exhibit the same pattern: they are -43 basis points ($t = -3.07$), -46 basis points ($t = -3.27$), -48 basis points ($t = -3.32$), and -39 ($t = -2.67$) basis points per month, respectively, for high KZ index firms, whereas they become small and insignificant for low KZ index firms.

5.4. Firm's Fundamental and Return Volatility

Zhang (2006) takes fundamental volatility and return volatility as two proxies for information uncertainty. In a theoretical model, Epstein and Schneider (2008) show that investors require more compensation for poor information quality when fundamentals are more volatile, whereas when fundamentals do not move much, investors do not care much about whether information quality is good or not. We empirically investigate this issue by using our R&D IQ measure.

Fundamental volatility is measured by cash flow uncertainty, which is defined as the standard deviation of return on asset (ROA) for the past three years. From Panel D, we do find that returns and alphas of the high-minus-low hedge portfolio are more economically substantial and statistically significant for firms with high fundamental volatility than for firms with low fundamental volatility. For example, the monthly excess return, and characteristic- and industry-adjusted returns of the hedge portfolio are -103 basis points ($t = -3.91$), -103 basis points ($t = -3.58$), and -110 basis points ($t = -4.43$), respectively, and its three-, four-, q-, and M-factor alphas are -98 basis points ($t = -3.54$), -104 basis points ($t = -3.42$), -121 basis points ($t = -3.74$), and -100 basis points ($t = -2.82$) per month, respectively, for high fundamental volatility firms. However, both returns and alphas of the hedge portfolio are much small and insignificant for low fundamental volatility firms.

We further examine the R&D IQ effect for high return volatility firms and low return volatility firms, where return volatility is calculated as the standard deviation of the Fama-French three-factor residuals for the past 12 months. We find exactly the same pattern as observed above.

6. An R&D Information Quality Factor

Table 2 shows that commonly used factor models such as the Fama-French three-factor model and the Carhart four-factor model cannot fully explain return dynamics. To further examine whether R&D IQ effect on future stock returns reflects commonality in returns that is not captured by the existing factors, we construct a factor-mimicking portfolio for R&D information quality following the same methodology as that in Fama and French (1993). Given that firm size increases with R&D IQ as reported in Table 1, we control for size in constructing the R&D IQ factor. At the end of June of year t from 1981 to 2012, we independently sort firms into two size portfolios (small "S" and big "B") based on NYSE median size breakpoint at the end of June of year t, and into three R&D IQ portfolios (low "L", middle "M", and high "H") based on the 30th and 70th percentiles of R&D IQ in year $t-1$. The intersection of these portfolios forms six size-IQ portfolios, namely, S/L, S/M, S/H, B/L, B/M, and B/H.

We hold these six portfolios for the next 12 months and compute their monthly value-weighted returns. The factor-mimicking portfolio for R&D IQ (IQF) is constructed as follows: (S/L + B/L)/2 - (S/H + B/H)/2. The IQF factor is thus size-adjusted and reflects the return comovement associated with R&D information quality. The IQF factor constructed from equal-weighted returns is quite similar and available upon request. Panel A of Table 8 reports the means, standard deviations, and ex post Sharpe ratios of IQF and the commonly used factors, i.e., the market factor (MKT), size factor (SMB), value factor (HML), and momentum factor (MOM). To compare with other innovation-related measures, we construct the following innovation factors: RDF (a factor based on R&D intensity), RDGF (a factor based on significant R&D growth), IEF (a factor based on Hirsleifer, Hsu, and Li's (2013) innovative efficiency), and NPF (a factor based on the number of patents scaled by market equity).

The average return of IQF is 30 basis points per month, which is smaller than that of MKT (60 basis points), HML (36 basis points), and MOM (60 basis points), but larger than average returns of SMB (10 basis points) and all other innovation-related factors. The standard deviation of IQF is 2.83%, which is smaller than those of nearly all of the factors considered except for that of NPF (2.67%). Furthermore, the ex post Sharpe ratios of these factors show that IQF offers a Sharpe ratio of 0.11, which is slightly lower than those for MKT (0.13), HML (0.12), and MOM (0.13) but larger than those for SMB (0.03) and all of the innovation-related factors.

Panel B of Table 8 presents the monthly correlations of all examined factors. We find that IQF is weakly correlated with and distinct from these factors. Its correlation with MKT is the smallest, 0.03, and its correlation with RDF is the strongest, 0.27. The average of absolute correlations between IQF and the other factors is about 0.17, which is smaller than those of the other factors except for MOM (0.13) and RDGF (0.15).

Figure 2 plots annual returns on the IQ factor (IQF) and on the market factor (MKT) from 1981 to 2012. The market factor is more volatile than the IQ factor. It can be as large as about 30% and as small as nearly -40%, and its standard deviation is about 17.32%. However, the IQ factor ranges from about -15% to about 20% and has a standard deviation of 8.45%. In the figure, we also highlight NBER recessions as gray areas. For the four recessions occurring in 1982, 1991, 2001, and 2008, the IQ factor performs better than the market factor in 1982, 2001, and 2008, and its outperformance is particularly striking during the Internet bubble burst of 2001 and during the recent global financial crisis of 2008. In 2001, the market factor has a return of -15.2%, whereas the IQ factor earns a positive return of 8.27%. In 2008, there is a severe market downturn: the return on the market factor reaches a historical low of -38.34%; however, the return on the IQ factor remains positive, 3.36%. The annual correlation between MKT and IQF is about 34.2%.

These findings indicate that IQF captures a different factor and that it may be beneficial to add IQF to existing factor models. For this purpose, similar to Hirshleifer, Hsu, and Li (2013), we construct different tangency portfolios using the above risk factors. Panel C presents optimal portfolio weights and ex post Sharpe ratios for these tangency portfolios. It is evident that when we only use the market factor (MKT), the monthly optimal Sharpe ratio is 0.13. When we introduce SMB together with MKT, the optimal weight on SMB is only 3%, whereas it is 97% on MKT. The optimal Sharpe ratio remains the same as above (0.13). When we use the Fama-French three factors (MKT, SMB, and HML) to construct the tangency portfolio, the optimal Sharpe ratio increases to 22% with mean of 0.40 and standard deviation of 1.79, and the largest weight is on HML (52%) followed by MKT (33%) and SMB (15%).

Upon applying our R&D IQ factor (IQF) and the Fama-French three factors, we find that the optimal Sharpe ratio further increases to 0.25 with mean of 0.39 and standard deviation of 1.54. The largest portfolio weight is now on IQF (42%) and the smallest weight is on HML (7%). When the momentum factor (MOM) is also available, the optimal Sharpe ratio reaches 0.31 with mean of 0.44 and standard deviation of 1.44. For this tangency portfolio, the largest weight is still on IQF (37%) followed by SMB (24%), MKT (15%), MOM (19%), and HML (5%).

From rows 6 to 9, we individually introduce the innovation-related factors (RDF, RDGF, IEF, and NPF) into the tangency portfolio together with IQF and the Fama-French three factors. We find that the Sharpe ratios of these tangency portfolios are nearly the same as that based on IQF and the Fama-French three factors, and the weights on these innovation-related factors are small, ranging from 2% (in row 6) to 5% (in row 7). When we put all factors together in row 10, the Sharpe ratio is 0.31, which is the same as that of the tangency portfolio in row 5, and the weights on these innovation-related factors are still very small (ranging from -3% for RDF to 5% for NPF). The largest weight in these tangency portfolios is again found for IQF, ranging from 35% to 42%.

The significant weight on IQF in these tangency portfolios and its role in improving the ex post Sharpe ratio are consistent with that shown in Panels A and B, where IQF has a relatively high mean and a small standard deviation, and its correlations with other factors are small. The above findings suggest that IQF does capture a pricing factor that is distinct from other well-known existing factors.

7. Conclusion

R&D investments are surrounded by a high degree of uncertainty due to the nature of R&D activities and a lack of accounting disclosure. We hypothesize that there exists a premium for ambiguous R&D information. Even though we cannot know its future information quality of a firm's R&D activities, past information on how much variation of its fundamentals can be explained by its R&D expenditures serves as a useful measure for evaluating its future R&D activities. We construct an R&D information quality (IQ) measure by connecting innovation input (R&D expenditures) and innovation outcome (sales). More specifically, R&D information quality is captured by the R-square from the regression of sales growth on the realized R&D capital.

We find strong evidence that expected excess returns decrease with R&D information quality. The high-minus-low IQ hedge portfolio earns excess return of about -39 basis point per month, characteristic-adjusted return of about -42 basis points per month, and industry-adjusted return of about -45 basis points per month in value-weighted returns. In value-weighted returns, the risk-adjusted monthly alpha of the hedge portfolio is about -40 basis points in the Fama-French three-factor model, about -38 basis points in the Carhart four-factor model, about -38 basis points in the q-factor model, and about -33 basis points in the M-factor model. All of these values are highly statistically significant. The same pattern is found in equal-weighted returns. Our Fama-MacBeth cross-sectional analysis shows that these results are robust in controlling for firm-specific variables that are known to have return predictability power and for some innovation-related variables.

The IQ-return relationship is even stronger in firms with smaller size, younger age, greater financial constraints, and higher fundamental and return volatility, as these firms usually have more uncertain business environments and investors are more ambiguous to their future prospects. Based on R&D IQ, we form a factor-mimicking portfolio (IQF), that is found to be weakly correlated with commonly used factors such as the market, size, value and momentum factors and with innovation-related factors proposed in the literature. Constructions of tangency portfolios show that adding IQF to the Fama-French three factors improves the ex post Sharpe ratio by 14% and that the weight on IQF dominates the other factors, indicating that IQF has incremental pricing effects relative to well-known pricing factors.

REFERENCES

Aboody, D. and Lev, B. (2000). "Information Asymmetry, R&D, and Insider Gains". *Journal of Finance* 55, 2747–2766.

Ai, H. (2010). "Information Quality and Long-Run Risk: Asset Pricing Implications". *Journal of Finance* 65, 1333–1367.

Amihud, Y. (2002). "Iliquidity and Stock Returns: Cross-Section and Time-Series Effects". *Journal of Financial Markets* 5, 31–56.

Amihud, Y. and Mendelson, H. (1986). "Asset Pricing and the Bid-Ask Spread". *Journal of Financial Economics* 17, 223–249.

Ang, A., Hodrick, R.J., Xing, Y., and Zhang, X. (2006). "The Cross-Section of Volatility and Expected Returns". *Journal of Finance* 61, 259–299.

Baker, M. and Wurgler, J. (2006). "Investor Sentiment and the Cross-Section of Stock Returns". *Journal of Finance* 61, 1645–1680.

Bali, T., Cakici, N., and Whitelaw, R. (2011). "Maxing Out: Stocks as Lotteries and the Cross-Section of Expected Returns". *Journal of Financial Economics* 99, 427–446.

Banz, R.W. (1981). "The Relationship between Return and Market Value of Common Stocks". *Journal of Financial Economics* 9, 3–18.

Barry, C.B. and Brown, S.J. (1985). "Differential Information and Security Market Equilibrium." *Journal of Financial and Quantitative Analysis* 20, 407–422.

Brevik, F. and d'Addona, S. (2010). "Information Quality and Stock Returns Revisited". *Journal of Financial and Quantitative Analysis* 45, 1419–1446.

Campello, M. and Chen, L. (2010). "Are Financial Constraints Priced? Evidence from Firm Fundamentals and Stock Returns". *Journal of Money, Credit, and Banking* 42, 1185–1198

Carhart, M. (1997). "On Persistence in Mutual Fund Performance". *Journal of Finance* 52, 57–282.

Chan, L., Lakonishok, J., and Sougiannis, T. (2001). "The Stock Market Valuation of Research and Development Expenditures". *Journal of Finance* 56, 2431–2456.

Chen, Z. and Epstein, L. (2002). "Ambiguity, Risk, and Asset Returns in Continuous Time". *Econometrica* 70, 1403–1443.

Cohen, L., Diether, K., and Malloy, C. (2013). "Misvaluing Innovation". *Review of Financial Studies* 26, 635–666.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. (1997). "Measuring Mutual Fund Performance with Characteristic-based Benchmarks". *Journal of Finance* 52, 1035–1058.

Eberhart, A.C., Maxwell, W.F., and Siddique, A.R. (2008). "An Examination of Long-Term Abnormal Stock Returns and Operating Performance Following R&D Increases". *Journal of Finance* 59, 623–650.

Epstein, L. and Schneider, M. (2008). "Ambiguity, Information Quality, and Asset Pricing". *Journal of Finance* 63, 197–228.

Fama, E. and French, K. (1992). "The Cross-Section of Expected Stock Returns". *Journal of Finance* 47, 427–465.

Fama, E. and French, K. (1993). "Common Risk Factors in the Returns on Stocks and Bonds". *Journal of Financial Economics* 33, 3–56.

Fama, E. and French, K. (1996). "Multifactor Explanation of Asset Pricing Anomalies". *Journal of Finance* 51, 55–84.

Fama, E. and French, K. (2015). "A Five-Factor Asset Pricing Model". *Journal of Financial Economics* 116, 1–22.

Fama, E. and MacBeth, J. (1973). "Risk, Return, and Equilibrium: Empirical Tests". *Journal of Political Economy* 71, 607–636.

Gu, F. (2005). "Innovation, Future Earnings, and Market Efficiency". *Journal of Accounting, Auditing and Finance* 20, 473–496.

Hall, B.H., Jaffe, A., and Trajtenberg, M., (2001). "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools". *NBER Working Paper* 8498.

Hirshleifer, D., Hsu, P.H., and Li, D. (2013). "Innovative Efficiency and Stock Returns". *Journal of Financial Economics* 107, 632–654.

Hirshleifer, D., Hsu, P.H., and Li, D. (2015). "Don't Hide Your Light Under a Bushel: Innovative Originality and Stock Returns". Working Paper.

Hou, K. and Robinson, D. (2006). "Industry Concentration and Average Stock Returns". *Journal of Finance* 61, 1927–1956.

Hou, K., Xue, C., and Zhang, L. (2015). "Digesting Anomalies: An Investment Approach". *Review of Financial Studies* 28, 650–705.

Hou, K., Xue, C., and Zhang, L. (2016). "A Comparison of New Factor Models". Working Paper, Ohio State University.

Kaplan, S.N. and Zingales, L. (1997). "Do Financial Constraints Explain Why Investment Is Correlated with Cash Flow?" *Quarterly Journal of Economics* 112, 169–216.

Lakonishok, J., Shleifer, A., and Vishny, R. (1994). "Contrarian Investment, Extrapolation, and Risk". *Journal of Finance* 49, 1541–1578.

Lamont, O., Polk, C., and Saá-Requejo, J. (2001). "Financial Constraints and Stock Returns". *Review of Financial Studies* 14, 529–554.

Li, D. (2011) "Financial Constraints, R&D Investment, and Stock Returns". *Review of Financial Studies* 25, 1033–1069.

Liu, W. (2006). "A Liquidity-Augmented Capital Asset Pricing Model". *Journal of Financial Economics* 82, 631–671.

Livdan, D., Sapriza, H., and Zhang, L. (2009). "Financially Constrained Stock Returns". *Journal of Finance* 60, 1827–1862.

Modigliani, F. and Miller, M. (1958). "The Cost of Capital, Corporation Finance and the Theory of Investment". *American Economic Review* 48, 261–297.

Ozdagli, A.K. (2012) "Financial Leverage, Corporate Investment, and Stock Returns". *Review of Financial Studies* 24, 2974–3007.

Pandit, S., Wasley, C.E., and Zach, T. (2011). "The Effect of R&D Inputs and Outputs on the Relation between the Uncertainty of future Operating Performance and R&D Expenditures". *Journal of Accounting, Auditing, and Finance* 26, 121–144.

Stambaugh, R. and Yuan, Y. (2016). "Mispricing Factors". *Review of Financial Studies*, forthcoming.

Veronesi, P. (2000). "How Does Information Quality Affect Stock Returns". *Journal of Finance* 55, 807–837.

Whited, T.M. and Wu, G. (2006). "Financial Constraints Risk." *Review of Financial Studies* 19, 531–559.

Zhang, X.F. (2006). "Information Uncertainty and Stock Returns". *Journal of Finance* 61, 105–136.

Table 1: **Summary Statistics**

Panel A reports the pooled mean, standard deviation, 25th percentile, median, and 75th percentile of the R&D information quality (IQ) measure across industries according to the Fama-French 17 industry classifications for the period from 1980 to 2012. The number of firms in each industry included in the sample (NFirm), the market share of each industry in the sample (SShare), and the market share of each industry in Compustat (MShare) are also reported. Panel B reports the average values of some selected firm-specific variables, including (log) market equity (ME), book-to-market ratio (BEME), return on assets (ROA, Income before extraordinary items plus interest expenses divided by lagged total assets), return on equity (ROE, Income before extraordinary items plus interest expenses divided by lagged common equity), leverage (DXA, long-term debt plus debt in current liability divided by total assets), cash holding, industry concentration index (HHI, Hou and Robinson, 2006), idiosyncratic volatility (IVOL, standard deviation of Fama-French three-factor residuals for the past 12 months), R&D intensity (RDA, R&D expenditures divided by total assets), innovation ability (InnAb, Cohen, Diether, and Malloy, 2013), and innovative efficiency (IE, Hirsleifer, Hsu, and Li, 2013) for the three R&D IQ portfolios constructed according to the 30th and 70th percentiles of the lagged IQ. Panel C reports the time-series correlations between R&D IQ and the above-mentioned firm characteristics.

Panel A: R&D Information Quality across Industries

	Mean	STD	Q25	Median	Q75	NFirms	SShare	MShare
Cars	0.41	0.21	0.24	0.37	0.56	43	1.5	2.0
Chems	0.43	0.23	0.23	0.38	0.62	74	3.2	2.4
Clths	0.47	0.22	0.28	0.46	0.62	45	0.9	0.8
Cnstr	0.46	0.22	0.29	0.42	0.61	83	2.8	2.6
Cnsum	0.47	0.23	0.26	0.44	0.67	122	19.2	12.5
Durbl	0.46	0.23	0.25	0.42	0.65	94	2.0	1.0
FabPr	0.47	0.23	0.27	0.45	0.62	28	0.6	0.3
Food	0.44	0.24	0.24	0.40	0.61	63	4.9	5.7
Machn	0.45	0.23	0.26	0.42	0.61	583	28.2	13.8
Mines	0.45	0.21	0.29	0.43	0.57	11	0.6	0.8
Oil	0.55	0.23	0.34	0.59	0.73	29	4.9	9.1
Other	0.49	0.23	0.29	0.47	0.67	693	27.6	30.5
Rtail	0.54	0.25	0.34	0.51	0.77	22	0.2	7.9
Steel	0.47	0.23	0.26	0.46	0.63	42	1.2	0.8
Trans	0.49	0.20	0.33	0.47	0.67	52	2.0	4.4
Utils	0.67	0.22	0.54	0.69	0.85	4	0.1	5.3

Panel B: Summary Statistics across IQ Portfolios

	ME	BEME	ROA	ROE	DXA	Cash	HHI	Ivol	RDA	IE	InnAb
Low IQ	205	0.68	0.03	-0.04	0.18	0.43	0.18	10.5	0.06	3.08	0.29
Mid IQ	248	0.73	0.02	-0.09	0.19	0.43	0.19	10.8	0.05	2.66	0.57
High IQ	278	0.71	0.01	0.09	0.20	0.52	0.19	11.3	0.06	3.08	0.80

Panel C: Correlation Matrix

	IQ	ME	BEME	ROA	ROE	DXA	Cash	HHI	Ivol	RDA	IE	InnAb
IQ	1.00											
ME	0.03	1.00										
BEME	0.02	-0.12	1.00									
ROA	-0.05	0.12	-0.14	1.00								
ROE	0.01	0.01	0.00	0.07	1.00							
DXA	0.05	-0.01	0.06	-0.11	0.01	1.00						
Cash	0.02	-0.02	-0.05	-0.19	-0.02	-0.13	1.00					
HHI	-0.02	-0.04	0.10	0.10	0.02	0.12	-0.10	1.00				
Ivol	0.06	-0.11	0.09	-0.38	-0.01	-0.01	0.09	-0.14	1.00			
RDA	-0.01	-0.06	-0.16	-0.45	-0.05	-0.17	0.26	-0.24	0.29	1.00		
IE	-0.01	0.32	-0.05	0.06	0.01	-0.01	-0.02	0.01	-0.05	-0.04	1.00	
InnAb	0.07	-0.04	0.04	-0.03	-0.01	0.05	-0.00	0.04	0.06	-0.06	-0.04	1.00

Table 2: **R&D Information Quality and Return Predictability**

This table presents average monthly portfolio returns (in %) based on single sort using R&D IQ. Each month stocks with non-missing lagged IQ are sorted into three groups based on the 30%/40%/30% breakpoints of R&D IQ. When forming portfolios, we impose the restriction that lagged price must be greater than $5 (breakpoints are computed before imposing the lagged price restriction). We hold these portfolios over the next 12 months and compute both their equal-weighted and value-weighted returns. In Panel A, we report excess returns, characteristic-adjusted returns, and industry-adjusted returns. Excess return is the difference between portfolio returns and the one-month Treasury bill rate. Characteristic-adjusted returns are computed by adjusting returns using 125 ($5 \times 5 \times 5$) size/book-to-market/momentum portfolios (Daniel et al., 1997), and industry-adjusted returns are computed by adjusting returns using 17 industry portfolios (Fama and French, 1997). In Panel B, C, D and E, we report the alphas and factor loadings from regressing portfolio excess returns on the Fama-French three factors (Fama and French, 1993), on the Carhart four factors (Carhart, 1997), on the q factors (Hou, Xue, and Zhang, 2015), and on the mispricing factors (Stambaugh and Yuan, 2016), respectively. The sample period is from July 1981 to June 2012.

	Value-Weighted Returns				Equal-Weighted Returns			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Panel A: Portfolio Returns								
Excess Returns	1.27	1.23	0.88	-0.39	1.26	1.22	0.78	-0.48
	(4.23)	(4.32)	(2.99)	(-4.16)	(4.14)	(4.14)	(2.56)	(-4.69)
Char-Adj Returns	0.18	0.14	-0.24	-0.42	0.16	0.11	-0.33	-0.50
	(2.25)	(1.84)	(-2.70)	(-4.23)	(1.98)	(1.45)	(-3.36)	(-4.46)
Ind-Adj Returns	0.30	0.23	-0.15	-0.45	0.14	0.07	-0.36	-0.50
	(2.46)	(1.96)	(-1.37)	(-4.80)	(0.91)	(0.47)	(-2.08)	(-4.90)
Panel B: Alphas and Loadings from the Three-Factor Model (Fama and French, 1993)								
Alpha	0.16	0.14	-0.24	-0.40	0.20	0.14	-0.32	-0.52
	(1.59)	(1.58)	(-2.32)	(-4.04)	(2.02)	(1.52)	(-2.94)	(-4.73)
MKT	1.06	1.04	1.06	-0.01	1.02	1.02	1.03	0.01
	(44.9)	(40.7)	(42.3)	(-0.29)	(47.6)	(36.7)	(41.4)	(0.43)
SMB	0.45	0.44	0.41	-0.04	0.58	0.55	0.51	-0.07
	(6.33)	(6.06)	(6.07)	(-1.28)	(9.90)	(7.72)	(7.78)	(-1.82)
HML	0.15	0.14	0.21	0.06	0.07	0.12	0.16	0.09
	(2.97)	(2.81)	(3.68)	(1.58)	(1.51)	(2.48)	(2.81)	(2.14)
Panel C: Alphas and Loadings from the Four-Factor Model (Carhart, 1997)								
Alpha	0.25	0.22	-0.13	-0.38	0.29	0.20	-0.20	-0.49
	(2.70)	(2.30)	(-1.29)	(-3.71)	(3.08)	(2.02)	(-2.01)	(-4.40)
MKT	1.04	1.02	1.02	-0.01	0.99	1.00	1.00	0.00
	(44.2)	(42.3)	(35.7)	(-0.46)	(45.7)	(38.6)	(36.3)	(0.16)
SMB	0.46	0.45	0.42	-0.04	0.59	0.56	0.52	-0.07
	(7.68)	(7.09)	(7.89)	(-1.23)	(12.3)	(8.74)	(10.2)	(-1.80)
HML	0.12	0.11	0.17	0.05	0.03	0.10	0.12	0.09
	(2.44)	(2.55)	(3.53)	(1.46)	(0.83)	(2.25)	(2.63)	(2.04)
MOM	-0.11	-0.09	-0.13	-0.02	-0.11	-0.07	-0.14	-0.03
	(-3.71)	(-2.75)	(-3.58)	(-0.90)	(-4.25)	(-2.06)	(-3.92)	(-0.98)

Panel D: Alphas and Loadings from the q-Factor Model (Hou, Xue, and Zhang, 2015)								
Alpha	0.17	0.19	-0.21	-0.38	0.28	0.20	-0.24	-0.53
	(1.52)	(1.79)	(-1.71)	(-3.50)	(2.61)	(1.89)	(-1.89)	(-4.38)
MKT	1.05	1.02	1.04	-0.01	0.99	0.99	1.01	0.01
	(38.6)	(35.2)	(35.0)	(-0.39)	(41.5)	(32.9)	(36.1)	(0.46)
SMB	0.45	0.45	0.38	-0.04	0.55	0.56	0.47	-0.08
	(7.19)	(6.73)	(5.64)	(-2.18)	(11.4)	(8.61)	(7.32)	(-1.99)
I/A	0.10	0.03	0.20	0.10	-0.01	-0.00	0.12	0.13
	(1.79)	(0.38)	(2.80)	(1.81)	(-0.14)	(-0.01)	(1.52)	(1.96)
ROE	0.01	0.01	-0.05	-0.06	-0.07	-0.00	-0.09	-0.02
	(0.25)	(0.16)	(-0.85)	(-1.63)	(-1.69)	(-0.05)	(-1.63)	(-0.42)
Panel E: Alphas and Loadings from the Mispricing-Factor Model (Stambaugh and Yuan, 2016)								
Alpha	0.14	0.20	-0.19	-0.33	0.21	0.20	-0.22	-0.43
	(1.40)	(2.01)	(-1.85)	(-2.96)	(2.12)	(1.99)	(-2.12)	(-3.54)
MKT	1.04	0.99	1.01	-0.03	0.99	0.96	0.97	-0.02
	(46.1)	(40.1)	(32.4)	(-0.85)	(42.1)	(35.9)	(32.2)	(-0.49)
SMB	0.52	0.51	0.45	-0.06	0.62	0.61	0.54	-0.08
	(8.09)	(8.03)	(6.96)	(-1.78)	(13.6)	(10.9)	(9.64)	(-1.97)
MGMT	-0.00	-0.07	0.02	0.02	-0.10	-0.11	-0.06	0.04
	(-0.05)	(-1.73)	(0.40)	(0.64)	(-2.53)	(-2.98)	(-1.21)	(0.91)
PERF	-0.04	-0.07	-0.09	-0.05	-0.06	-0.08	-0.12	-0.05
	(-1.51)	(-2.43)	(-2.24)	(-1.56)	(-2.50)	(-2.56)	(-3.11)	(-1.63)

Table 3: **Fama-MacBeth Cross-Sectional Regressions**

This table presents monthly Fama-MacBeth (1973) regressions of returns on R&D IQ. Control variables include: size (Banz, 1981), book-to-market ratio (Fama and French, 1992), momentum (Carhart, 1997), leverage (Miller and Modigliani, 1958; Ozdagli, 2012), illiquidity (Amihud, 2002), volatility of sales growth (VolSale), idiosyncratic volatility (Ang, Hodrick, Xing, and Zhang, 2006; Bali, Cakici, and Whitelaw, 2011), one-month lagged returns, turnover, capital expenditures (CapEX), and industry concentration (HHI). Some innovation-related variables are also taken into consideration: R&D expenditures scaled by sales, R&D expenditures scales by total assets, significant increases of R&D expenditures (Eberhart, Maxwell, and Siddique, 2004), innovation ability (Cohen, Diether, and Malloy, 2013), and innovative efficiency (Hirsleifer, Hsu, and Li, 2013). All regressions includes industry dummies (using Fama and French (1997) 17-industry classification scheme). The regressions only include stocks with lagged price greater than \$5. The sample period is from July 1981 to June 2012. *t*-statistics are in parenthesis.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
IQ	-0.81	-0.69	-0.61	-0.53	-0.68	-0.66	-0.64	-0.56	-0.73	-0.69
	(-4.36)	(-3.67)	(-3.39)	(-2.94)	(-3.61)	(-3.46)	(-3.46)	(-3.04)	(-3.49)	(-2.78)
log(ME)		0.04	-0.00	0.08	0.05	0.04	0.04	-0.02	0.05	0.02
		(0.85)	(-0.01)	(2.05)	(1.01)	(0.82)	(0.88)	(-0.35)	(1.00)	(0.25)
BEME		0.04	-0.02	0.08	-0.02	0.03	0.03	-0.02	0.19	0.26
		(0.22)	(-0.12)	(0.50)	(-0.08)	(0.17)	(0.17)	(-0.11)	(1.00)	(1.23)
MOM		0.08	0.08	0.10	0.09	0.08	0.06	0.05	0.09	0.12
		(0.53)	(0.52)	(0.71)	(0.59)	(0.51)	(0.41)	(0.35)	(0.63)	(0.71)
DXA		-0.13	-0.30	-0.27	-0.21	-0.08	-0.10	-0.31	0.03	-0.13
		(-0.38)	(-0.87)	(-0.78)	(-0.61)	(-0.22)	(-0.29)	(-0.90)	(0.07)	(-0.29)
R_{-1}		-0.00	-0.00	-0.01	-0.00	-0.00	-0.00	-0.00	0.00	-0.00
		(-0.49)	(-0.41)	(-1.26)	(-0.45)	(-0.48)	(-0.53)	(-0.40)	(0.09)	(-0.20)
turnover		-0.38	-0.21	-0.59	-0.42	-0.39	-0.44	-0.29	-0.14	-0.28
		(-1.92)	(-1.07)	(-2.91)	(-2.12)	(-1.98)	(-2.26)	(-1.46)	(-0.63)	(-1.08)
VolSale			-1.02							
			(-2.50)							
IVOL				0.04						
				(1.91)						
ILLIQ				-0.00						
				(-0.91)						
CapEx					-3.69					
					(-2.93)					
HHI						-0.34				
						(-1.13)				
RDS								-0.54		
								(-0.61)		
RDA							1.24			
							(1.03)			
InnAb								-0.28		
								(-2.53)		
RDG									-0.00	
									(-0.00)	
IE										-0.01
										(-0.89)
Intercept	1.40	0.95	1.61	0.01	1.06	0.99	0.94	1.71	0.60	1.16
	(4.29)	(1.25)	(2.15)	(0.02)	(1.39)	(1.31)	(1.32)	(2.22)	(0.74)	(1.18)
Adj R^2	1.3	5.4	6.3	7.2	5.6	5.5	6.0	6.2	5.3	5.7
	(6.46)	(13.0)	(14.3)	(15.2)	(13.5)	(13.2)	(13.7)	(14.1)	(11.6)	(11.2)
Industry	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 4: **R&D Information Quality and Subsequent Operating Performance**

This table reports the average slopes (in percent) and their time series t-statistics in parentheses from annual Fama-MacBeth (1973) cross-sectional regressions of individual stocks' operating performance measures in year $t+1$ on R&D IQ and other control variables in year t. We measure operating performance by return on assets (ROA), cash flows (CF), and performance (PM, operating income before depreciation scaled by the lagged sales). We control for size, book-to-market, leverage, idiosyncratic volatility, illiquidity, and some innovation-related variables such as R&D intensity and innovation ability. We also introduce the lagged values and the changes of fundamental variables in the regressions. Industry dummies are also introduced based on the Fama and French (1997) 17 industry classification.

	ROA_{t+1}	PM_{t+1}	CF_{t+1}
IQ	-0.01	0.02	-0.01
	(-1.36)	(1.53)	(-1.47)
ROA	0.61		
	(13.67)		
ΔROA	-0.13		
	(-3.64)		
PM		0.59	
		(4.96)	
ΔPM		0.06	
		(1.08)	
CF			0.58
			(15.2)
ΔCF			-0.22
			(-9.66)
log(ME)	0.00	0.02	0.01
	(4.46)	(3.37)	(6.06)
BEME	-0.02	-0.01	-0.00
	(-3.44)	(-1.08)	(-0.63)
DXA	-0.02	0.04	0.08
	(-1.95)	(1.23)	(6.22)
IVOL	-0.00	-0.00	-0.00
	(-6.24)	(-0.36)	(-3.18)
ILLIQ	0.00	-0.00	-0.00
	(1.39)	(-0.09)	(-2.56)
RDS	-0.02	-0.04	-0.05
	(-2.38)	(-0.27)	(-2.94)
InnAb	-0.00	-0.00	-0.00
	(-0.85)	(-0.46)	(-0.96)
Intercept	0.03	-0.11	0.01
	(2.26)	(-1.92)	(0.23)
Adj R^2	43.6	55.6	39.4

Table 5: **R&D Information Quality and Long-Term Future Returns**

This table presents long-term portfolio cumulative returns (in %) based on single sort using R&D IQ. At the end of June of each year, stocks with non-missing lagged IQ are sorted into three groups based on the 30%/40%/30% breakpoints of R&D IQ. We then hold these portfolios over the next 12, 24, and 36 months and compute value-weighted cumulative returns of these IQ portfolios. We report excess returns, the three-factor (Fama and French, 1993), four-factor (Carhart, 1997), q-factor (Hou, Xue, and Zhang, 2015), and M-factor (Stambaugh and Yuan, 2016) alphas. When forming portfolios, we also impose the restriction that lagged price must be greater than \$5 (breakpoints are computed before imposing the lagged price restriction). We also report the past 12-month portfolio returns. The sample period is from July 1981 to June 2012.

	Past 12-Month Returns				Future 12-Month Returns			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	25.66	27.82	24.68	-0.98	13.38	13.49	9.73	-3.64
	(7.55)	(7.23)	(6.71)	(-0.36)	(4.77)	(5.44)	(2.99)	(-2.80)
FF3F Alphas	23.13	24.30	21.93	-1.20	16.70	15.78	12.54	-4.16
	(7.11)	(5.87)	(4.94)	(-0.38)	(3.72)	(3.83)	(2.51)	(-2.48)
Carhart4F Alphas	24.26	25.25	24.10	-0.16	14.13	13.33	10.92	-3.21
	(5.10)	(4.97)	(4.09)	(-0.05)	(3.68)	(3.48)	(2.49)	(-1.80)
q-Factor Alphas	26.07	25.28	24.95	-1.12	13.09	13.10	10.68	-2.41
	(6.31)	(6.27)	(4.98)	(-0.31)	(2.82)	(2.76)	(2.12)	(-1.80)
M-Factor Alphas	27.37	28.79	26.15	-1.12	14.56	13.81	11.46	-3.11
	(5.71)	(5.08)	(4.96)	(-0.20)	(3.03)	(3.16)	(2.89)	(-2.65)
	Future 24-Month Returns				Future 36-Month Returns			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	24.46	26.42	18.62	-5.84	37.83	39.62	29.24	-8.60
	(5.15)	(5.48)	(3.53)	(-2.66)	(4.82)	(5.38)	(4.43)	(-2.89)
FF3F Alphas	28.70	29.97	22.44	-6.27	43.26	44.36	31.96	-11.30
	(4.59)	(4.98)	(3.44)	(-2.64)	(4.67)	(5.84)	(4.26)	(-3.82)
Carhart4F Alphas	28.40	30.62	23.32	-5.09	39.53	42.09	31.13	-8.40
	(4.59)	(4.58)	(3.47)	(-1.75)	(4.57)	(5.41)	(4.13)	(-2.31)
q-Factor Alphas	28.15	30.37	22.88	-5.62	39.32	42.78	30.87	-8.45
	(3.66)	(3.86)	(3.21)	(-2.66)	(3.05)	(4.09)	(3.73)	(-2.43)
M-Factor Alphas	30.15	31.85	24.56	-5.31	39.18	44.04	30.69	-8.49
	(3.88)	(4.14)	(3.73)	(-2.03)	(3.18)	(4.62)	(3.83)	(-2.82)

Table 6: **R&D IQ Effect and Investor Sentiment**

This table presents results from regressions of R&D IQ spread returns on different measures of investor sentiment. SENT1 is the Baker and Wurgler (2006) investor sentiment index, which is constructed as the first principal component of the six proxies of investor sentiment, including the close-end fund discount, turnover, the number of IPOs, the average first-day returns, equity share, and the dividend premium. SENT2 is the investor sentiment index, which removes business cycle variation from SENT1. SENT1L (SENT2L) represents the first lag of SENT1 (SENT2). ΔSENT1 and ΔSENT2 (ΔSENT1L and ΔSENT2L) are the first difference of SENT1 and SENT2 (SENT1L and SENT2L), respectively. R&D IQ spread returns are high-minus-low R&D IQ portfolio returns. The sample period is from July 1981 to June 2012.

	Value-Weighted R&D IQ Spread							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SENT1	0.022							
	(0.14)							
SENT2		-0.021						
		(-0.12)						
ΔSENT1			-0.356					
			(-0.42)					
ΔSENT2				-0.013				
				(-0.02)				
SENT1L					0.037			
					(0.25)			
SENT2L						-0.020		
						(-0.12)		
ΔSENT1L							-0.068	
							(-0.11)	
ΔSENT2L								-0.022
								(-0.04)
Constant	-0.392	-0.378	-0.386	-0.385	-0.396	-0.379	-0.385	-0.385
	(-3.79)	(-3.44)	(-4.06)	(-4.05)	(-3.85)	(-3.50)	(-4.04)	(-4.05)
Adj R^2	-0.003	-0.003	-0.002	-0.003	-0.002	-0.003	-0.003	-0.003
	Equal-Weighted R&D IQ Spread							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SENT1	0.001							
	(0.00)							
SENT2		-0.067						
		(-0.36)						
ΔSENT1			0.037					
			(0.04)					
ΔSENT2				0.338				
				(0.41)				
SENT1L					-0.001			
					(-0.01)			
SENT2L						-0.091		
						(-0.50)		
ΔSENT1L							-0.116	
							(-0.16)	
ΔSENT2L								-0.072
								(-0.12)
Constant	-0.483	-0.460	-0.483	-0.482	-0.482	-0.452	-0.483	-0.483
	(-4.27)	(-3.82)	(-4.60)	(-4.060)	(-4.28)	(-3.82)	(-4.58)	(-4.59)
Adj R^2	-0.003	-0.002	-0.003	-0.002	-0.003	-0.002	-0.003	-0.003

Table 7: **Firm Characteristics and R&D IQ Effect**

This table presents monthly portfolio returns (in %) based on double sorts on firm characteristics and R&D IQ. At each month stocks with non-missing lagged firm characteristics and R&D IQ are firstly sorted into three portfolios at 30%/40%/30% breakpoints based on each firm's characteristics (firm size in Panel A, firm age in Panel B, firm's financial constraint in Panel C, and fundamental volatility in Panel D) and each of these portfolios is then sorted into three sub-groups at 30%/40%/30% breakpoints based on R&D IQ. Excess return is the difference between portfolio returns and the one-month Treasury bill rate. Characteristic-adjusted returns are computed by adjusting returns using 125 ($5 \times 5 \times 5$) size/book-to-market/momentum portfolios (Daniel et al., 1997), and industry-adjusted returns are computed by adjusting returns using the Fama-French 17 industry portfolios (Fama and French, 1997). When forming portfolios, we impose the restriction that lagged price must be greater than $5. The three-factor (Fama and French, 1993), four-factor (Carhart, 1997), q-factor (Hou, Xue, and Zhang, 2015), and M-factor (Stambaugh and Yuan, 2016) alphas are also reported. The sample period is from July 1981 to June 2012.

Panel A: Firm's Size

	Small Size				Big Size			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	1.39	0.91	0.44	-0.96	1.28	1.22	1.00	-0.28
	(4.34)	(2.57)	(1.05)	(-2.91)	(4.39)	(4.75)	(3.78)	(-2.34)
Char-Adj Returns	0.11	-0.34	-0.89	-1.00	0.21	0.19	-0.11	-0.32
	(0.56)	(-1.55)	(-2.84)	(-3.02)	(2.15)	(2.21)	(-1.11)	(-2.57)
Ind-Adj Returns	0.47	-0.11	-0.58	-1.04	0.25	0.26	-0.06	-0.30
	(1.65)	(-0.49)	(-1.72)	(-3.03)	(2.11)	(2.28)	(-0.51)	(-2.58)
FF3F Alphas	0.50	-0.12	-0.51	-0.99	0.13	0.15	-0.10	-0.24
	(2.24)	(-0.57)	(-1.57)	(-3.02)	(1.13)	(1.31)	(-0.80)	(-1.94)
Carhart4F Alphas	0.54	-0.13	-0.56	-1.10	0.23	0.30	-0.02	-0.25
	(2.27)	(-0.59)	(-1.71)	(-3.23)	(2.03)	(2.49)	(-0.18)	(-1.96)
q-Factor Alphas	0.72	-0.13	-0.49	-1.21	0.08	0.23	-0.23	-0.30
	(2.90)	(-0.59)	(-1.41)	(-3.31)	(0.58)	(1.54)	(-1.67)	(-2.25)
M-Factor Alphas	0.50	-0.16	-0.56	-1.07	0.13	0.32	-0.13	-0.25
	(1.99)	(-0.67)	(-1.72)	(-2.82)	(1.05)	(2.18)	(-0.95)	(-1.88)

Panel B: Firm's Age

	Young Age				Old Age			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	1.50	1.16	0.30	-1.20	1.07	1.14	1.09	0.01
	(4.13)	(3.60)	(0.84)	(-4.47)	(3.82)	(4.37)	(3.49)	(0.08)
Char-Adj Returns	0.41	0.12	-0.71	-1.12	0.14	0.14	0.06	-0.08
	(2.46)	(0.91)	(-3.63)	(-4.30)	(1.11)	(1.37)	(0.36)	(-0.44)
Ind-Adj Returns	0.55	0.16	-0.68	-1.23	0.15	0.23	0.02	-0.13
	(2.94)	(1.03)	(-3.36)	(-4.90)	(1.03)	(1.58)	(0.12)	(-0.74)
FF3F Alphas	0.42	0.11	-0.68	-1.10	-0.00	0.08	-0.00	0.00
	(2.27)	(0.79)	(-3.29)	(-3.98)	(-0.03)	(0.68)	(-0.02)	(0.01)
Carhart4F Alphas	0.43	0.24	-0.48	-0.91	0.07	0.16	0.04	-0.02
	(2.22)	(1.62)	(-2.32)	(-3.22)	(0.47)	(1.29)	(0.23)	(-0.13)
q-Factor Alphas	0.64	0.27	-0.39	-1.03	-0.19	-0.03	-0.19	0.00
	(3.07)	(1.46)	(-1.68)	(-2.98)	(-1.00)	(-0.19)	(-0.92)	(0.02)
M-Factor Alphas	0.39	0.40	-0.35	-0.75	-0.13	-0.02	-0.19	-0.06
	(1.81)	(2.39)	(-1.61)	(-2.31)	(0.76)	(-0.11)	(-0.98)	(-0.33)

Panel C: Firm's Financial Constraint								
	Low KZ Index				High KZ Index			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	1.16	1.17	0.93	-0.23	1.28	1.32	0.86	-0.42
	(2.94)	(3.32)	(2.92)	(-0.87)	(4.65)	(4.88)	(2.95)	(-3.18)
Char-Adj Returns	0.10	0.04	-0.26	-0.37	0.17	0.25	-0.24	-0.41
	(0.52)	(0.18)	(-1.38)	(-1.47)	(1.93)	(2.74)	(-2.13)	(-3.02)
Ind-Adj Returns	0.35	0.22	-0.06	-0.41	0.29	0.31	-0.22	-0.51
	(1.38)	(0.96)	(-0.29)	(-1.53)	(2.22)	(2.33)	(-1.78)	(-3.83)
FF3F Alphas	-0.12	0.01	-0.28	-0.16	0.23	0.28	-0.19	-0.43
	(-0.53)	(0.03)	(-1.30)	(-0.61)	(2.14)	(2.58)	(-1.56)	(-3.07)
Carhart4F Alphas	-0.05	0.08	-0.15	-0.10	0.33	0.38	-0.13	-0.46
	(-0.21)	(0.39)	(-0.75)	(-0.34)	(3.43)	(3.34)	(-1.04)	(-3.27)
q-Factor Alphas	0.04	0.04	-0.17	-0.21	0.25	0.37	-0.23	-0.48
	(0.15)	(0.19)	(-0.70)	(-0.67)	(2.12)	(2.60)	(-1.63)	(-3.32)
M-Factor Alphas	-0.06	0.02	-0.25	-0.19	0.20	0.38	-0.20	-0.39
	(-0.20)	(0.08)	(-1.22)	(-0.54)	(2.02)	(2.90)	(-1.62)	(-2.67)
Panel D: Firm's Fundamental Volatility								
	Low Volatility				High Volatility			
	IQ_L	IQ_M	IQ_H	H-L	IQ_L	IQ_M	IQ_H	H-L
Excess Returns	1.33	1.29	1.05	-0.27	1.20	0.96	0.17	-1.03
	(4.99)	(5.31)	(3.69)	(-1.82)	(3.25)	(2.95)	(0.47)	(-3.91)
Char-Adj Returns	0.21	0.20	-0.01	-0.22	0.09	-0.04	-0.94	-1.03
	(1.96)	(2.08)	(-0.06)	(-1.48)	(0.50)	(-0.23)	(-4.60)	(-3.58)
Ind-Adj Returns	0.35	0.24	-0.02	-0.37	0.24	-0.037	-0.86	-1.10
	(2.19)	(1.83)	(-0.17)	(-2.41)	(1.18)	(-0.38)	(-4.76)	(-4.43)
FF3F Alphas	0.23	0.24	-0.02	-0.25	0.06	-0.13	-0.92	-0.98
	(1.72)	(2.11)	(-0.12)	(-1.64)	(0.32)	(-0.84)	(-4.55)	(-3.54)
Carhart4F Alphas	0.31	0.31	0.11	-0.20	0.18	-0.02	-0.85	-1.04
	(2.50)	(2.82)	(0.67)	(-1.27)	(0.84)	(-0.09)	(-3.95)	(-3.42)
q-Factor Alphas	0.06	0.15	-0.02	-0.08	0.36	0.15	-0.85	-1.21
	(0.40)	(1.08)	(-0.11)	(-0.47)	(1.60)	(0.81)	(-3.84)	(-3.74)
M-Factor Alphas	0.10	0.22	0.16	0.06	0.01	0.14	-0.99	-1.00
	(0.77)	(1.85)	(0.94)	(0.27)	(1.60)	(0.81)	(-3.84)	(-2.82)

Table 8: **The Factor-Mimicking Portfolios**

At the end of June of year t from 1981 to 2012, we firstly sort firms into two size portfolios (small "S" and big "B") based on NYSE median size breakpoint at the end of June of year t, and then sort each size portfolio into three R&D IQ portfolios (low "L", middle "M", and high "H") based on the 30th and 70th percentiles of R&D IQ in year $t-1$. As a result, there are in total six size-IQ portfolios, namely, S/L, S/M, S/H, B/L, B/M, and B/H. We hold these six portfolios over the next 12 months and compute their monthly value-weighted returns. The factor-mimicking portfolio for R&D IQ (IQF) is constructed as follows: (S/L + B/L)/2 - (S/H + B/H)/2. Size is the market equity at the end of June of year t. We also construct four innovation-related factors based on R&D intensity (R&D expenditures scaled by sales), significant R&D growth (RDG), innovative efficiency (IE), and the number of patents scaled by market equity, respectively. MKT is the return on the value-weighted NYSE, Amex, and Nasdaq portfolio minus the one-month Treasury bill rate. SMB and HML are the returns on two factor-mimicking portfolios associated with the size effect and the book-to-market effect, respectively. MOM denotes the momentum factor. Panel A reports the mean, standard deviation, and ex post Sharpe ratio (SR) for these factors. Panel B reports the Pearson correlation coefficients among these factors. Panel C report the portfolio weights and monthly Sharpe ratios of ex post tangency portfolios based on investing in subsets of these factor-mimicking portfolios. All returns and standard deviations are in percentage.

Panel A: Summary Statistics

	IQF	MKT	SMB	HML	MOM	RDF	RDGF	IEF	NPF
Mean	0.30	0.60	0.10	0.36	0.60	0.06	0.12	0.06	0.03
Stdev	2.83	4.54	3.09	3.04	4.57	3.68	3.18	3.10	2.67
SR	0.11	0.13	0.03	0.12	0.13	0.02	0.04	0.02	0.01

Panel B: Correlation Matrix

	IQF	MKT	SMB	HML	MOM	RDF	RDGF	IEF	NPF
IQF	1.00								
MKT	0.03	1.00							
SMB	0.22	0.23	1.00						
HML	-0.13	-0.33	-0.34	1.00					
MOM	0.19	-0.18	0.05	-0.13	1.00				
RDF	0.27	0.28	0.38	-0.44	0.11	1.00			
RDGF	0.16	0.14	0.06	-0.19	0.12	0.38	1.00		
IEF	0.23	0.15	0.26	-0.29	-0.09	0.40	0.08	1.00	
NPF	0.09	0.14	0.16	-0.22	-0.14	0.34	-0.07	0.84	1.00

Panel C: Constructions of Tangency Portfolio

	Portfolio Weights									Sharpe Ratio		
	MKT	SMB	HML	IQF	MOM	RDF	RDGF	IEF	NPF	Mean	Stdev	SR
1.	1.00									0.60	4.54	0.13
2.	0.97	0.03								0.58	4.42	0.13
3.	0.33	0.15	0.52							0.40	1.79	0.22
4.	0.26	0.25	0.07	0.42						0.39	1.54	0.25
5.	0.15	0.24	0.05	0.37	0.19					0.44	1.44	0.31
6.	0.25	0.25	0.06	0.42		0.02				0.38	1.52	0.25
7.	0.24	0.24	0.07	0.40			0.05			0.37	1.47	0.25
8.	0.25	0.24	0.06	0.41				0.04		0.38	1.50	0.25
9.	0.25	0.24	0.06	0.41					0.03	0.38	1.50	0.25
10.	0.12	0.22	0.04	0.35	0.18	-0.03	0.03	0.04	0.05	0.41	1.31	0.31

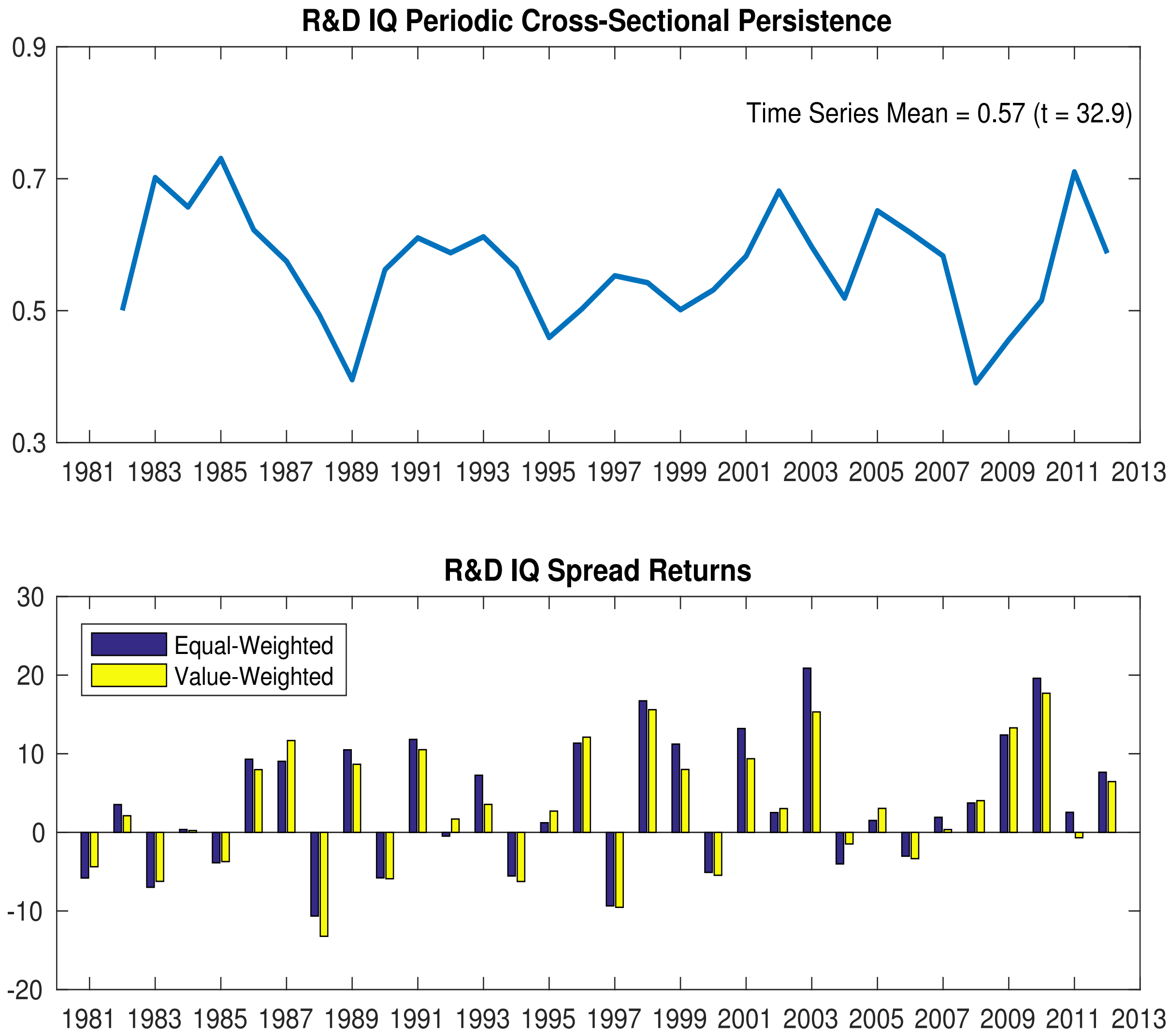

Figure 1: **R&D IQ Persistence and Returns on Spread Portfolios**

The upper panel presents the one-year apart periodic cross-sectional persistence of R&D IQ. Its time-series average is also computed. The lower panel presents the time series of annual equal-weighted and value-weighted excess returns on short position of the high-minus-low hedge portfolio over the period from July 1981 to July 2012. When computing portfolio excess returns, each month stocks with non-missing lagged IQ are sorted into three groups based on the 30%/40%/30% breakpoints of R&D IQ, and these portfolios are then held over the next 12 months.

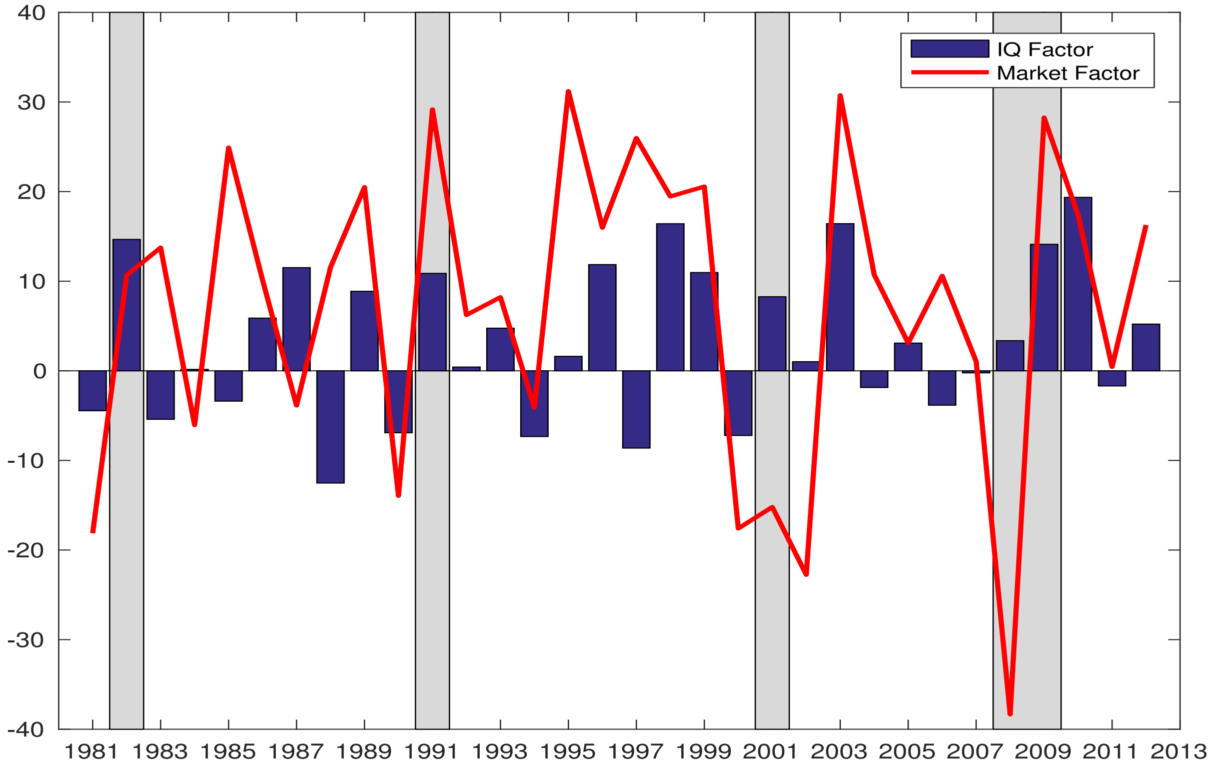

Figure 2: **Annual Returns on the IQ Factor and the Market Factor**
This figure plots returns (on a per annum basis) for the IQF factor and the market factor from 1981 to 2012. MKT is the return on the value-weighted NYSE, Amex, and Nasdaq portfolio minus the one-month Treasury bill rate. At the end of June of year t from 1981 to 2012, we firstly sort firms into two size portfolios (small "S" and big "B") based on NYSE median size breakpoint at the end of June of year t, and then sort each size portfolio into three R&D IQ portfolios (low "L", middle "M", and high "H") based on the 30th and 70th percentiles of R&D IQ in year $t-1$. As a result, there are in total six size-IQ portfolios, namely, S/L, S/M, S/H, B/L, B/M, and B/H. We hold these six portfolios over the next 12 months and compute their monthly value-weighted returns in excess of the one-month Treasury bill rates. The factor-mimicking portfolio for R&D IQ (IQF) is constructed as follows: (S/L + B/L)/2 - (S/H + B/H)/2. The gray areas represent NBER recessions.