INNOVATIVE EFFICIENCY AND STOCK RETURNS*

David Hirshleifer[a] Po-Hsuan Hsu[b] Dongmei Li[c]

March 2011

* We thank James Ang, Frederick Bereskin, Joao Gomes, Bronwyn Hall, Danling Jiang, Xiaoji Lin, Alfred Liu, Siew Hong Teoh, Allan Timmermann, Sheridan Titman, Neng Wang, Lu Zhang, Lu Zheng, and seminar participants at the CRSP Forum 2010 (Chicago, IL) and Florida State University for valuable discussions and comments.

[a] Paul Merage School of Business, University of California, Irvine.
[b] School of Business, University of Connecticut.
[c] Rady School of Management, University of California, San Diego.

INNOVATIVE EFFICIENCY AND STOCK RETURNS

We find that innovative efficiency (IE), patents or citations scaled by R&D, is a strong positive predictor of future returns after controlling for firm characteristics and risk. The IE-return relation is associated with the loading on a mispricing factor, and the high Sharpe ratio of the Efficient Minus Inefficient (EMI) portfolio suggests that mispricing plays an important role. Further tests based upon attention and uncertainty proxies suggest that limited attention contributes to the effect. The high weight of the EMI portfolio return in the tangency portfolio suggests that IE captures incremental pricing effects relative to well-known factors.

Keywords: innovative efficiency, limited attention

JEL Classification: G12, G14, O32

1. Introduction

Recent studies have provided evidence suggesting that owing to limited investor attention, prices do not fully and immediately impound the arrival of relevant public information, especially when such information is less salient or arrives during a period of low investor attention (e.g., Klibanoff, Lamont, and Wizman 1998; Huberman and Regev 2001; DellaVigna and Pollet 2009; Hirshleifer, Lim, and Teoh 2009; Hou, Peng, and Xiong 2009). Several papers have therefore argued that limited attention results in underreaction and return predictability. Theoretical models also predict that limited investor attention affects stock prices and can cause market underreaction (Merton 1987; Hirshleifer and Teoh 2003; Peng and Xiong 2006).

These studies consider the processing of news about current performance such as earnings announcements. However, we would expect investors to have even greater difficulty processing information that is less tangible, and about firms whose future prospects are highly uncertain. For example, information about the prospects of new technologies or other innovations should be especially hard to process, because the significance of such news depends upon strategic options and other complex considerations. If so, there will on average be price drift after the arrival of non-salient public news about the prospects for firms' innovations. In other words, on average there will be positive (negative) abnormal returns after good (bad) news.

In this study, we examine the relation between innovative efficiency and subsequent operating performance as well as stock returns. By innovative efficiency, we mean a firm's ability to generate patents or citations per dollar of R&D (research and development) investment. Patents are a measure of innovative output since innovations are usually officially introduced to the public in the form of approved patents with detailed information. U.S. firms have increasingly recognized the necessity to patent their innovations and hence have been especially active in patenting activities since the early 1980s (Hall and Ziedonis 2001; Hall 2005) owing to the creation of the Court of Appeals for the Federal Circuit (CAFC) in 1982 and several well-documented patent lawsuits (e.g., the Kodak-Polaroid case). Patents are thus the most important measure of contemporary firms' innovative output (Griliches 1990); they are actively traded in intellectual property markets (Lev 2001).

A firm's past innovative efficiency is not necessarily as salient to investors as explicitly forward-looking information about the prospects for the particular R&D projects that the firm is examining. According to Kahneman and Lovallo (1993), people tend to consider the judgment or decision problem they are facing as unique, and in consequence, "neglect the statistics of the past in evaluating current plans." Kahneman and Lovallo call a focus on the uniqueness of the problem the 'inside view,' and a focus on relevant statistical performance data from previous trials the 'outside view.' An excessive focus on the inside view implies that people will tend to be overoptimistic about prospects for success when they neglect unfavorable non-salient statistical information; and tend to be less optimistic, and perhaps over-pessimistic about the prospects of success, when they neglect favorable statistical information.[1]

Furthermore, there is extensive evidence that individuals pay less attention to, and place less weight upon, information that is harder to process (see, e.g., the review of Song and Schwarz 2010). As argued above, information about innovations is hard to process, because it requires developing and applying a theory of how the economic fundamentals of a firm or its industry are changing. It also requires an analysis of the road from patents to final products on the market, the profit of which

[1] Lovallo and Kahneman (2003) emphasize that the inside view tends to promote overoptimism on the part of managers because they are required to weave scenarios, imagine events, or gauge their own levels of ability and control, all of which are susceptible to organizational pressure and cognitive biases such as overoptimism, anchoring, and competitor neglect. The argument for optimism of managers does not necessarily extend to investors, who have much less of a personal attachment to the firm's projects. (However, a possible example might be an analyst who has chosen to follow or recommend a firm based on a positive analysis, and thereby becomes prone to an optimistic inside view.) In any case, our focus is on how the degree of optimism/pessimism varies with statistical performance information, rather than the overall average degree of optimism.

can be highly uncertain and long deferred. We would expect such hard-to-process information to be underweighted unless there is some offsetting effect (such as high salience).

These considerations suggest that investors will underreact to the information content in innovative efficiency because of the difficulty evaluating the economic implications of patents granted. If so, then firms that are more efficient in innovations may be undervalued, whereas firms that are less efficient in innovations may be overvalued. Therefore, we expect a positive relation between innovative efficiency and future stock returns and operating performance.

An alternative argument for why innovative efficiency should predict higher future returns derives from the *q*-theory (Cochrane 1991, 1996; Liu, Whited, and Zhang 2009). Firms with higher innovative efficiency tend to be more profitable and have higher return on assets. All else equal, the *q*-theory~~ies~~ implies~~y~~ that higher profitability predicts higher returns, because a high return on assets indicates that these assets were purchased by the firm at a discount (i.e., that they carry a high risk premium.).

Specifically, suppose that the market for capital being purchased by a firm is competitive and efficient. When a firm makes an R&D expenditure to purchase innovative capital, the price it pays will be appropriately discounted for risk. For concreteness, we can think for an example of a firm that acquires a high-tech target at a competitive market price.[2] In this scenario, a firm will on average achieve higher 'return' (large number of patents, resulting in high cash flows) on its innovative expenditures as fair compensation if its purchased innovative capital is highly risky, and will receive low return if capital is relatively low-risk. Past innovative efficiency is therefore a proxy for risk, so firms that have high past innovative efficiency should subsequently be productive

[2] When firms make acquisitions, under appropriate circumstances it can book part of the expenditure as 'in process R&D.'

in patenting (Dierickx and Cool 1989) and earn higher profits and stock returns.[3]

To test our key hypothesis that innovative efficiency (IE) predicts operating performance and stock returns, we use three proxies for IE in year t: patents granted in year t scaled by R&D capital in year $t-2$ (Patents/RDC), patents granted in year t scaled by R&D expenses in year $t-2$ (Patents/RD), and adjusted patent citations received in year t by patents granted in years $t-1$ to $t-5$ scaled by the sum of R&D expenses in years $t-3$ to $t-7$ (Citations/RD). The lag between the innovative input (R&D) and output (patents) reflects the average two-year application-grant lag (Section 2.1 provides more details). These IE measures are not highly correlated with other innovation-related return and operating performance predictors, such as R&D intensity, R&D growth, and change in adjusted patent citations scaled by average total assets (ΔAPC).[4] Therefore, the IE measures potentially contain useful incremental information.

Fama-MacBeth (1973) regressions show that higher IE measures predict significantly higher return on assets (ROA) and cash flows (CF) over the next year, controlling for ROA (CF), R&D intensity, R&D growth, ΔAPC, and industry. We also find a significantly positive relation between the three IE measures and future stock returns through Fama-MacBeth (1973) regressions controlling for industry and other return predictors, such as R&D intensity, R&D growth, ΔAPC, size, book-to-market, momentum, capital investment intensity, ROA, asset growth, net stock issues, and institutional ownership.

Portfolio analysis confirms the significantly positive relationship between IE and future stock

[3] There are other possible rational risk arguments consistent with a positive relation between past innovative efficiency and future stock returns. A high level of innovative activity, even if successful in the past, is likely to be associated with greater economic uncertainty and real options, and therefore high risk and expected return. See, e.g., Greenwood and Jovanovic (1999), Berk, Green, and Naik (2004), Hsu (2009), Pastor and Veronesi (2009), Garleanu, Kogan, and Panageas (2011), and Garleanu, Panageas, and Yu (2011).

[4] Adjusted patent citations in year t is patent citations received in year t by patents granted in years $t-1$ to $t-5$ adjusted by patent subcategory, citing year, and cited year. The procedure of adjustment is detailed in Section 2.1. The average total assets is averaged over years t and $t-1$. Gu (2005) denotes the same measure by ΔPCI (change in patent citation impact).

returns. The monthly value-weighted size-adjusted returns of the high-minus-low IE portfolios for Patents/RDC, Patents/RD, and Citations/RD are 41 (t = 3.91), 42 (t = 3.83), and 28 (t = 2.71) basis points respectively.

Standard risk factor models do not explain these returns. For example, the monthly alphas of the high-minus-low IE portfolios estimated from the Carhart (1997) four-factor model range from 38 basis points (t = 3.28) for Citations/RD to 42 basis points (t = 3.91) for Patents/RDC. In fact, the high-minus-low IE portfolios load negatively on the market, size, and momentum factors, implying high IE firms are less risky than low IE firms if these factor models are interpreted as capturing the market pricing of risk.

The alternative risk factor model motivated by investment-based asset pricing includes the market, investment (INV) and ROA factors (Chen, Novy-Marx, and Zhang 2010; henceforth CNZ). The alphas estimated from this model are smaller due to the positive loadings of the hedge portfolios on the ROA factor. However, the alphas remain large, ranging from 30 to 33 basis points per month and are significant at the 1% level.

To examine if mispricing helps explain the IE effect and is incremental to the above risk factors, we add a financing-based mispricing factor UMO (Undervalued Minus Overvalued; Hirshleifer and Jiang 2010) to the Carhart four-factor model and the CNZ model. We find the UMO factor improves the explanatory power for both models. Furthermore, the positive loadings of the high-minus-low IE portfolios on UMO suggest that high IE firms are undervalued relative to low IE firms. However, the alphas estimated from these augmented models remain high at 22-32 basis points per month and are statistically significant at the 1% or 5% level.

If the IE-return relation represents a market inefficiency driven by psychological constraints such as limited attention, we expect to observe greater return predictability among stocks with low investor attention and among hard-to-value stocks. To test this hypothesis, we use size and analyst coverage as proxies for attention to a stock (Hong, Lim, and Stein 2000; Hirshleifer and Teoh 2003) and firm age, turnover, and idiosyncratic volatility as proxies for valuation uncertainty (Kumar 2009).[5] We expect a stronger IE effect among firms with small market capitalization, low analyst coverage, young age, high turnover, and high idiosyncratic volatility.

The Fama-MacBeth subsample regressions provide supporting evidence. The average IE slopes among low attention and hard-to-value stocks are all significant at the 1% level and are substantially larger than those among high attention and easy-to-value stocks, which often are insignificant. For example, the IE slopes range from 0.08% to 0.14% among small firms, but only 0.03% to 0.04% among big firms. Similarly, the slopes range from 0.10% to 0.16% among firms with low analyst coverage, but are only 0.05% among firms with high analyst coverage. Although the cross-subsample differences in the IE slopes are not always statistically significant, their magnitudes are economically substantial.

Portfolio analysis also confirms a stronger IE-return relation among smaller firms. For example, for Patents/RDC, the monthly value-weighted return of the hedge portfolio formed among small firms is 55 basis points and significant at the 1% level. The monthly alphas estimated from factor models with or without the mispricing factor UMO are significant at the 1% level and range from 32 to 59 basis points. Furthermore, UMO dominates the risk factors in the augmented models and improves the explanatory power substantially. In contrast, the return of the hedge portfolio formed among large firms is only 27 basis points and significant only at the 10% level. The alphas estimated from the CNZ model and models augmented with UMO are statistically insignificant. The explanatory power of these models mainly derives from the ROA factor and the mispricing factor

[5] Other measures of investor attention are related to these variables. For example, Fang and Peress (2009) report that media coverage increases in firm size and analyst coverage.

UMO.

To further examine if the IE-based return predictability is driven by risk, mispricing, or both, we construct a factor-mimicking portfolio for innovative efficiency, EMI (Efficient Minus Inefficient), following the procedure of Fama and French (1993). We focus on the IE measure, Patents/RDC. By combining long positions on firms with high IE and short positions on firms with low IE within different size categories, we form a portfolio that captures any incremental return comovement associated with innovative efficiency.[6] The EMI factor is negatively correlated with the market factor (–0.10), the size factor (–0.20), and the momentum factor (–0.02), and is positively correlated with the profitability factor (0.21), the mispricing factor (0.21), the investment factor (0.08), and the value factor (0.13).

The average monthly return of the EMI factor is 0.41%, which is higher than that of the size factor (0.06%), the value factor (0.39%), and the investment factor (0.22%). Furthermore, the EMI factor offers an ex post Sharpe ratio, 0.22, which is higher than that of all the above factors except the mispricing factor (0.28). The high level of the equity premium is a well-known puzzle for rational asset pricing theory (Mehra and Prescott 1985). Therefore, on the face of it, the high ex post Sharpe ratio associated with the EMI factor also suggests that the IE-return relation may be too strong to be entirely explained by rational risk premia.

Adding EMI to the Fama-French three factors increases the ex post Sharpe ratio of the tangency portfolio from 0.29 to 0.38 with a weight of 38% on EMI. Even when all the above factors are included, the weight on EMI in the tangency portfolio is 20%, which is substantially higher than that of any of the other factors except the mispricing factor and the market factor. These findings imply the IE-return relation captures return predictability effects above and beyond those captured by the

[6] The EMI factor return is essentially the size-adjusted return of the high-minus-low IE portfolio based on Patents/RDC as discussed before.

other well-known factors.

Our study differs from previous literature in identifying a new innovation-related measure, IE, that predicts operating performance and stock returns. While existing related studies mainly focus on either the input or the output of innovation, we focus on the *productivity* of R&D investment as a ratio of innovative output (patents) to input (R&D). On the input side, several studies have shown that R&D expenses, R&D capital, R&D reporting bias, and R&D growth predict the level and variability of operating performance and stock returns (e.g., Lev and Sougiannis 1996; Chan, Lakonishok, and Sougiannis 2001; Kothari, Laguerre, and Leone 2002; Chambers, Jennings, and Thompson 2002; Lev, Sarath, and Sougiannis 2005; Eberhart, Maxwell, and Siddique 2004, 2008; Li 2010).

On the output side, several studies have shown that patents or citations predict the level and variability of earnings, cash flows, and/or stock returns. Lanjouw and Schankerman (2004) find that patent quality is positively associated with contemporaneous stock market value of firms. Gu (2005) finds a positive relation between ΔAPC and future ROA and stock returns. Pandit, Wasley, and Zach (2011) find that successful R&D efforts measured with patent data are associated with higher levels and lower variability of future operating performance. Matolcsy and Wyatt (2008) develop *aggregate* measures of technological innovation conditions based on patent data and provide empirical evidence on the relation between these aggregate measures and firms' contemporaneous market value of equity and future earnings.

We find that the ability of innovative efficiency to predict returns and operating performance is incremental to that of other innovation-related variables such as R&D intensity, change in patent citations, and R&D growth. We also examine whether the IE effect is driven by risk or mispricing, and explore the effects of proxies for limited attention and valuation uncertainty.

2. The Data, the Innovative Efficiency Measures, and Summary Statistics

2.1. The Data and the Innovative Efficiency Measures

Our sample consists of firms in the intersection of Compustat, CRSP (Center for Research in Security Prices), and the NBER patent database. We obtain accounting data from Compustat and stock returns data from CRSP. All domestic common shares trading on NYSE, AMEX, and NASDAQ with accounting and returns data available are included except financial firms, which have four-digit standard industrial classification (SIC) codes between 6000 and 6999 (finance, insurance, and real estate sectors). Following Fama and French (1993), we exclude closed-end funds, trusts, American Depository Receipts, Real Estate Investment Trusts, units of beneficial interest, and firms with negative book value of equity. To mitigate backfilling bias, we require firms to be listed on Compustat for two years before including them in our sample. For some of our tests, we also obtain analyst coverage data from the Institutional Brokers Estimate System (IBES) and institutional ownership data from the Thomson Reuters Institutional (13f) Holdings dataset.

Patent-related data are from the updated NBER patent database originally developed by Hall, Jaffe, and Trajtenberg (2001).[7] The database contains detailed information on all U.S. patents granted by the U.S. Patent and Trademark Office (USPTO) between January 1976 and December 2006: patent assignee names, firms' Compustat-matched identifiers, the number of citations received by each patent, the number of citations excluding self-citations received by each patent, application dates, grant dates, and other details. Patents are included in the database only if they are eventually granted by the USPTO by the end of 2006. It is noted that the NBER patent database contains two time placers for each patent: its application date and grant date. To prevent any

[7] The updated NBER patent database is available at https://sites.google.com/site/patentdataproject/Home/downloads.

potential look-ahead bias, we choose the grant date as the effective date of each patent, and measure firm i's innovation output in year t as the number of patents granted to firm i in year t ("patents").

We use three proxies for IE: patents granted scaled by R&D capital (Patents/RDC), patents granted scaled by R&D expenses (Patents/RD), and adjusted patent citations scaled by R&D expenses (Citations/RD). We allow a two-year gap between the innovative input and output as it takes, on average, two years for the USPTO to grant a patent application (Hall, Jaffe, and Trajtenberg 2001). Specifically, our primary IE measure, Patents/RDC, is defined as the ratio of firm i's patents granted in year t ($Patents_{i,t}$) scaled by its R&D capital (the five-year cumulative R&D expenses assuming an annual depreciation rate of 20% as in Chan, Lakonishok, and Sougiannis (2001)) in fiscal year ending in year $t-2$:

$$Patents_{i,t} \,/\, (R\&D_{i,t-2} + 0.8 * R\&D_{i,t-3} + 0.6 * R\&D_{i,t-4} + 0.4 * R\&D_{i,t-5} + 0.2 * R\&D_{i,t-6}),$$

where $R\&D_{i,t-2}$ denotes firm i's R&D expenses in fiscal year ending in year $t-2$, and so on. This IE measure is premised on R&D expenses over the preceding five years contributing to successful patent applications.

Following Lanjouw and Schankerman (2004), our second proxy of innovative efficiency, Patents/RD, is premised on only contemporaneous R&D expenses contributing to patent applications and is defined as:

$$Patents_{i,t} \,/\, R\&D_{i,t-2}.$$

Since forward citations may better reflect a patent's technological or economic significance (e.g., Trajtenberg 1990; Harhoff, Narin, Scherer, and Vopel 1999; Hall, Jaffe, and Trajtenberg 2005), the third proxy of IE, Citations/RD, is the ratio of forward citations received in year t by firm i's patents granted over the prior five year to matching R&D expenses in the following form:

$$\frac{\text{Citations}}{\text{RD}} = \frac{\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} \cancel{C_{ik}^{t-j}}\; C_{ik}^{t-j}}{\left(R\&D_{i,t-3} + R\&D_{i,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}\right)},$$

where C_{ik}^{t-j} $\cancel{C^{t-j}_{ik}}$ is the adjusted number of citations received in year t by patent k ($k = 1 \ldots N_{t-j}$), issued to firm i in year $t - j$ ($j = 1, 2, 3, 4, 5$), and N_{t-j} is the total number of patents issued to firm i in year $t - j$ that are cited in year t. The five-year period used in computing the firm-level adjusted citations is roughly consistent with prior findings that technology cycles measured by the duration of the benefits of R&D spending are approximately five years in most industries (Lev and Sougiannis 1996). Following Gu (2005) and Pandit, Wasley, and Zach (2011), the adjusted number of citations C^{t-j}_{ik} is the ratio of number of citations received by patent k to the mean number of citations received by patents of the same subcategory/grant year group cited in the same year.[8] This procedure helps adjust for citation propensity attributed to differences in technology fields, citing year (the year in which the citation takes place) and cited year (the year in which the cited patent is issued). To allow for the two-year application-grant lag, we use the sum of R&D expenses in years $t - 3$ to $t - 7$ in the denominator matching the cited patents granted in years $t - 1$ to $t - 5$ in the numerator.

We construct these three IE measures for each year from 1981 to 2006. The sample period starts in 1981 to ensure the quality of R&D expenditure data. Since the accounting treatment of R&D expenses reporting is standardized in 1975 (Financial Accounting Standards Board Statement No. 2), the estimate of R&D capital starts in 1979 to allow for a full five-year period with reliable R&D expenditure data. Furthermore, the two-year application-grant lag makes 1981 the first year for the IE measure, Patents/RDC. To make the results comparable, we use the same sample period for the other two IE measures.

2.2. Summary Statistics

[8] Patent subcategory is based on the technology classification system of the USPTO. A brief description of selected subcategories is as follows: 14 (organic compounds), 21 (communications), 22 (computer hardware and software), 23 (computer peripherals), and 24 (information storage). More details are provided in the updated NBER patent database.

Table 1 reports the pooled mean, standard deviation, 25th percentile, median, and 75th percentile of the three IE measures for selected innovation-intensive two-digit SIC industries. We observe significant variation in industrial innovative efficiency. Transportation equipment industry outperforms all the other industries. The average Patents/RDC, Patents/RD, and Citations/RD for this industry are 2.11, 1.64, and 10.66, respectively. In contrast, on average it costs more for biotech and pharmaceutical industry to create patents and citations. The average Patents/RDC, Patents/RD, and Citations/RD for this industry are 0.49, 0.86, and 2.57, respectively. These statistics suggest an important role of controlling for industry effect in examining the relation between IE and future operating performance and stock returns.

At the end of February of year *t*, we form three IE portfolios based on the 33th and 66th percentiles of Patents/RDC measured in fiscal year ending in year $t-1$.[9] We report average annual number of firms and median characteristics as they are highly skewed for the three IE portfolios in Panel A of Table 2.

The average annual number of firms in the low, middle, and high IE groups is 574, 274, and 424, respectively. Moreover, the median market equity of the low, middle, and high IE groups is $42.05 million, $630.59 million, and $208.54 million, respectively. We note that firms with non-missing IE cover 55% of total U.S. market equity. This is therefore an economically meaningful set of firms to study.

There are significant variations in the three IE measures across the IE groups formed on Patents/RDC. The median Patents/RDC, Patents/RD, and Citations/RD for the high IE group are 0.41, 0.97, and 1.14, respectively. In contrast, the counterparts of these statistics are zero for the low IE group and are 0.08, 0.20, and 0.30, respectively, for the middle IE group. Table 2 also reports

---

[9] All characteristics are for the year prior to the ranking year except market equity and momentum, which are measured at the end of February of the ranking year.

median future citations, future citations excluding self-citations, and future self-citations (all scaled by R&D capital). These measures also increase across the low, middle, and high IE groups. Note future citations are citations received from the grant year till the end of the sample period, 2006. Therefore, these measures are unobservable to investors at the time of portfolio formation and cannot be used to predict future returns and operating performance.

Table 2 also reports other characteristics including R&D expenditure, R&D expenditure to sales (RDS), R&D growth (RDG), ΔAPC, ROA in current year and next year, book-to-market ratio, net stock issues, asset growth, capital expenditure scaled by total assets (CapEx/Assets), momentum, and institutional ownership. To measure R&D intensity, we scale R&D expenditure by sales rather than market equity (ME) because the latter is associated with the size effect.[10] RDG is the growth in annual R&D expenses. ΔAPC denotes change in adjusted patent citations scaled by average total assets (Gu 2005), and its detailed definition is provided in Footnote 4. ROA is income before extraordinary items divided by lagged total assets. Book-to-market ratio (BTM) is book equity over market equity. Net stock issues (NS) is the change in the natural log of split-adjusted shares outstanding. Asset growth (AG) is the change in total assets divided by lagged total assets. CapEx/Assets is capital expenditure divided by lagged total assets. Momentum is the prior six-month returns (with one-month gap between the holding period and the current month; see, e.g., Hou, Peng, and Xiong 2009). Institutional ownership (IO) is the fraction of a firm's shares outstanding owned by institutional investors.

The middle IE group spends most on R&D ($28.5 million) while the high (low) IE group spends $7.37 ($2.00) million. The R&D intensity (RDS) of the high IE group is 3.98%, which is lower than

[10] Past research indicates that market price can proxy for various sources of either systematic risk or mispricing (e.g., Fama and French 1993 and 1996; Berk 1995; Daniel, Hirshleifer, and Subrahmanyam 2001). So the use of ME in the denominator for an R&D variable can capture effects unrelated to innovation. Moreover, Eberhart, Maxwell, and Siddique (2008) suggest for other reasons that using ME to compute R&D intensity can bias inferences.

that of the low IE group (5.37%). These patterns suggest that greater R&D spending does not necessarily lead to higher innovative efficiency. R&D growth of the high IE group (10%) is slighter larger than that of the low IE group (9%). In addition, ΔAPC also increases across the three IE groups.

The median ROAs in the year before and after portfolio formation increase monotonically with IE. For example, the median ROA in the year after the portfolio formation is 2.15%, 4.56%, and 4.64% for the low, middle, and high IE groups, respectively. The Wilcoxon rank-sum test shows the difference in ROA between the low and high IE groups in the year after portfolio formation is significant at the 0.01% level. This evidence suggests a significantly positive relation between IE and future ROA. For the other characteristics, the high IE group has lower BTM, same NS, higher AG, higher CapEx/Assets, higher momentum, and higher IO than the low IE group.

Panel B of Table 2 reports the pairwise Pearson and Spearman rank correlation coefficients among the three IE measures and other characteristics. The three IE measures are positively and significantly correlated. The Pearson (Spearman) correlation between Patents/RDC and Patents/RD is 0.58 (0.99), and the Pearson (Spearman) correlation between Patents/RDC and Citations/RD is 0.09 (0.63). The three IE measures are highly correlated with Future citations/RDC. For example, the Pearson (Spearman) correlation between Patents/RDC and Future citations/RDC is 0.83 (0.92). This high correlation suggests the counts-based IE measures are indeed good proxies for the long-run citation performance of a firm's patents and contain value-relevant information. The correlations between the IE measures and other characteristics including R&D intensity, R&D growth, ΔAPC, ROA, size, and BTM are low in magnitude (though sometimes statistically significant). Therefore the IE measures potentially contain independent information.

3. Predictability of Operating Performance Based upon Innovative Efficiency

To further examine the effect of innovative efficiency on firms' future operating performance, we conduct annual Fama-MacBeth (1973) cross-sectional regressions in the following form:

$$OP_{i,t+1} = \alpha_0 + \alpha_1 \ln(1 + IE_{i,t}) + \alpha_2 \ln(1 + RDS_{i,t}) + \alpha_3 RDG_{i,t} + \alpha_4 \Delta APC_{i,t} + \alpha_5 OP_{i,t} + \sum_{j=1}^{48} \gamma_j Industry_j,$$

where $OP_{i,t+1}$ is firm *i*'s operating performance in year *t* + 1 (ROA or cash flows), ln(1+IE) is the natural log of one plus IE, ln(1+RDS) is the natural log of one plus R&D expenditure to sales, RDG is R&D growth, ΔAPC is change in adjusted patent citations scaled by average total assets of years *t* and *t* – 1, and Industry_j is a dummy variable that equals 1 for the industry that firm *i* belongs to and 0 otherwise based on Fama and French's (1997) 48 industry classifications.[11]

As the distributions of IE measures are highly skewed and are often zero, we use ln(1+IE) in the regression.[12] Since R&D intensity is sometimes zero as well, we also use ln(1+RDS) in the regression. Following Gu (2005) and Pandit, Wasley, and Zach (2011), we include lagged operating performance in the model to accommodate the persistence in operating performance. We winsorize all variables at the 1% and 99% levels, and after winsorization we standardize all independent variables to zero mean and one standard deviation to facilitate the comparison of the economic magnitudes of effects.

Table 3 reports the average slopes and corresponding time-series *t*-statistics from the annual cross-sectional regressions for ROA (Panel A) and cash flows (Panel B). Panel A presents a significantly positive relation between the three IE measures and future ROA controlling for all the

[11] Cash flows (CF) is defined as net income plus depreciation scaled by lagged total assets. The results for other definitions of cash flows generate similar results.
[12] The logarithmic linearization follows Lerner (1994), in which the natural log of one plus patent count is used as the main explanatory variable for biotechnology firm value.

above variables. The slopes on IE range from 0.23% for Patents/RDC to 0.41% for Citations/RD with *t*-statistics of 3.55 or greater. The larger magnitude of the slopes on citation-based IE measure is intuitive as influential inventions should benefit the firm's profitability to a larger degree. The slope on RDS is significantly negative, and the slope on RDG is positive but insignificant. These findings suggest that it is difficult to draw inferences on how R&D input promotes profitability due to potential inefficient investment.

The slope on ΔAPC is significantly negative, and the slope on ROA is significantly positive. The finding for ΔAPC contrasts with the finding of Gu (2005) of a positive slope on ΔAPC. The difference may be due to difference in model specifications as we control for R&D intensity, R&D growth, and industry effects.

In unreported tables, we conduct univariate regressions in which each of the three IE measures is the only explanatory variable, and find that IE plays an economically substantial role in explaining future ROA. The coefficients on Patents/RDC, Patents/RD, and Citations/RD are 5.80%, 6.76%, and 6.65%, respectively, with *t*-statistics over 10. Moreover, we also include the natural log of employees and the natural log of capital expenditure in multivariate regressions to control for labor and capital input contributing to future ROA, and obtain similar results.

Similarly, Panel B shows a significantly positive relation between IE and future cash flows (CF). The slopes on IE range from 0.27% for Patents/RDC to 0.44% for Citations/RD with *t*-statistics of 4.20 or greater. The pattern in the slopes on the other variables is similar to that in Panel A. Unreported univariate regression results indicate that the slopes on Patents/RDC, Patents/RD, and Citations/RD are 5.56%, 6.48%, and 6.39%, respectively, with *t*-statistics over 10. We also obtain similar results when the natural log of number of employees and the natural log of capital expenditure are included in multivariate regressions.

Together these findings suggest that the proposed IE measures contain incremental information about firms' operating performance, and that innovative efficiency matters for firm value.

4. Predictability of Returns Based upon Innovative Efficiency

4.1. Fama-MacBeth Regression Results

In this subsection, we examine the ability of IE to predict returns using monthly Fama-MacBeth (1973) cross-sectional regressions in order to control for other characteristics that can predict returns and make sure the relation between stock returns and IE measures, if any, does not simply reflect familiar effects or industry characteristics. For each month from July of year t to June of year $t + 1$, we regress monthly excess returns of individual stocks on ln(1+IE) of year $t - 1$ and other control variables. The control variables include ln(1+RDS), R&D growth (RDG), ΔAPC, size, book-to-market ratio, momentum, CapEx/Assets, ROA, asset growth, net stock issues, institutional ownership, and industry dummies based on Fama and French's (1997) 48 industries.[13] All control variables are measured in the fiscal year ending in year $t - 1$ except size and momentum. Size is the natural log of the market equity at the end of June of year t. ~~Momentum is the prior six-month returns (with the standard one-month gap between the holding period and the current month).~~ The definitions of the other control variables are detailed in Section 2. The minimum six-month lag between stock returns and the other independent variables ensures the accounting variables are fully observable. We winsorize all independent variables at the 1% and 99% levels, and after winsorization we standardize all independent variables to zero mean and one standard deviation to

[13] On the capital investment effect, see, e.g., Lyandres, Sun, and Zhang (2008) and Polk and Sapienza (2009). On the asset growth effect, see, e.g., Cooper, Gulen, and Schill (2008). On the net stock issuance effect, see, e.g., Daniel and Titman (2006), Fama and French (2008), and Pontiff and Woodgate (2008). Similar results are obtained in the Fama-MacBeth regressions without industry dummies.

facilitate the comparison of the economic magnitudes of effects.

Table 4 reports the average slopes and their time-series *t*-statistics from the monthly cross-sectional regressions. There is a significantly positive relation between IE and stock returns controlling for all the variables above. The slopes on the three IE measures range from 0.06% to 0.10% and are all statistically significant at the 1% level. The IE slopes are larger than the RDS slopes, which are insignificant and range from 0.00% to 0.04% with *t*-statistics between 0.04 and 0.81

The RDG slopes are positive but insignificant. In addition, we find that the slopes on ΔAPC are negative and often are significantly negative. This finding differs from the positive return predictability identified in Gu (2005). The difference may be due to the different model specification; we control for R&D growth, industry fixed effects, and many other characteristics known to predict returns.

The slopes on the other control variables are in general consistent with the literature. Firms with smaller size, higher book-to-market ratio, higher ROA, lower asset growth, and higher institutional ownership provide significantly higher future stock returns. The slopes on momentum, CapEx/Assets, and net stock issues are insignificant. In unreported tables, we find that the magnitude and statistical significance of the IE slopes are robust to alternative measures of R&D intensity based on market equity, total assets, CapEx, and number of employees. These results indicate that the explanatory power of IE is robust to and is distinct from that of the other commonly known return predictors.

4.2. Portfolio Tests

In this subsection, we confirm the return predictability of IE and examine whether the IE effect is explained by risk or mispricing through portfolio analysis. At the end of February of year *t*, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint and three IE groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of IE for each of the three IE measures. Size is the market equity at the end of February of year *t*, and IE is measured in year *t* – 1.[14] The intersection forms six size-IE portfolios (S/H, B/H, S/M, B/M, S/L, and B/L). We hold these portfolios over the next twelve months (March of year *t* to February of year *t* + 1) and compute value-weighted monthly returns of the six portfolios. We then calculate monthly size-adjusted returns of the IE portfolios by averaging the returns of the two size groups within the same IE group. Specifically, the size-adjusted returns of the low, middle, and high IE portfolios are (S/L + B/L)/2, (S/M + B/M)/2, and (S/H + B/H)/2, respectively. To examine the significance of the relation between IE and future stock returns, we also form a high-minus-low IE portfolio. The size-adjusted return of this hedge portfolio is (S/H + B/H)/2 – (S/L + B/L)/2.

Table 5 shows that average size-adjusted excess return, measured as the difference between monthly size-adjusted returns of the IE portfolio and the one-month Treasury bill rate, increases monotonically with IE for the three IE measures. For Patents/RDC (Panel A), the value-weighted monthly size-adjusted excess returns on the low, middle, and high IE portfolios are 42 basis points (t = 1.25), 79 basis points (t = 2.58), and 82 basis points (t = 2.61), respectively. The difference between the high and low IE portfolio returns is 41 basis points per month (t = 3.91), which is statistically and economically significant.

We also examine whether the significant return spread between the high and low IE portfolios can be explained by standard risk factor models, such as the CAPM and the Carhart (1997) four-

[14] The two-month lag between the granted year end and the time of portfolio formation is imposed to ensure patent information is known to the public.

factor model, by regressing the time-series of portfolio excess returns on corresponding risk factor(s) returns.[15] The CAPM contains the market factor (MKT), the return on the value-weighted NYSE/AMEX/NASDAQ portfolio minus the one-month Treasury bill rate. The Carhart four-factor model includes the market factor, the size factor (SMB), the value factor (HML), and the momentum factor (MOM). SMB, HML, and MOM are returns on the factor-mimicking portfolios associated with the size effect, the value effect, and the momentum effect.

Panel A shows that these models do not explain the return spread between the low and high IE portfolios. In fact, the risk adjustment increases rather than decreases the return spread. The monthly alphas (intercepts) estimated from the CAPM and the Carhart four-factor model for the high-minus-low IE portfolio are 44 and 42 basis points with *t*-statistics of 4.14 and 3.91, respectively. The high IE portfolio has lower loadings on the market and size factors, suggesting high IE firms are less risky.

The results for the other two IE measures are similar. For Patents/RD (Panel B), the difference in the value-weighted size-adjusted return between the high and low IE portfolios is 42 basis points per month ($t = 3.83$). The monthly alphas estimated from the CAPM and the Carhart four-factor model for the high-minus-low IE portfolio are 48 and 42 basis points with *t*-statistics of 4.34 and 3.76, respectively. For Citations/RD (Panel C), the difference in the returns between the high and low IE portfolios is 28 basis points per month ($t = 2.71$). The monthly alphas estimated from the CAPM and the Carhart four-factor model for the high-minus-low IE portfolio are 28 and 38 basis points with *t*-statistics of 2.53 and 3.28, respectively.

---

[15] Due to limited space, we do not report Fama-French (1993) three-factor model results but obtain consistent results in unreported tables. We obtain four-factor returns and the one-month Treasury bill rate from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

If the IE-return relation is caused by limited investor attention, it reflects market mispricing. Several authors have suggested that there is commonality in mispricing (e.g., Daniel, Hirshleifer, and Subrahmanyam 2001 and Barberis and Shleifer 2003). For example, if investors do not fully impound information about common movements in innovative efficiency (arising, e.g., from technological change), we would expect a degree of commonality in the mispricing of innovative efficiency.

To examine whether commonality in mispricing helps capture the IE effect, we augment the Carhart model by adding the mispricing factor (UMO, Undervalued Minus Overvalued) from Hirshleifer and Jiang (2010).[16] The UMO factor is the returns to a zero-investment portfolio that goes long on firms with debt repurchases or equity repurchases and short on firms with IPOs, SEOs, and debt issuances over the past 24 months. The authors interpret the new issue/repurchase activities as managers' response to overvaluation/undervaluation of their firms and provide evidence suggesting the UMO loadings proxy for the common component of a stock's mispricing.

Table 6 shows the addition of the UMO factor is a substantial improvement as it reduces the Carhart alphas of the high-minus-low IE portfolios by 12, 15, and 6 basis points per month for Patents/RDC, Patents/RD, and Citations/RD in Panels A, B, and C, respectively. Furthermore, the high-minus-low IE portfolios load positively on the UMO factor: 0.19 (t = 3.44), 0.23 (t = 4.41), and 0.09 (t = 1.27) for the three IE measures, respectively. This pattern suggests that the high IE firms are undervalued relative to the low IE firms. Nevertheless, this augmented model does not fully explain the IE-return relation as the alphas remain substantial (ranging from 27 to 32 basis points per month) and statistically significant at the 1% level.

These findings are consistent with the innovative efficiency effect being a consequence of market mispricing. Alternatively, innovative efficiency may load on other risk factors. We therefore

[16] We obtain UMO factor returns from Danling Jiang's website: http://mailer.fsu.edu/~djiang/.

study whether the investment-based three-factor model of Chen, Novy-Marx, and Zhang (2010) can explain the IE-return relation. This model is motivated by investment-based asset pricing and consists of the market factor, an investment factor (INV), and a return on assets factor (ROA).[17] The INV factor is the difference between the return to a portfolio of low-investment stocks and the return to a portfolio of high-investment stocks. The ROA factor is the difference between the return to a portfolio of stocks with high ROA and the return to a portfolio of stocks with low ROA. We construct the INV and ROA factors following the methodology detailed in Chen, Novy-Marx, and Zhang (2010).

Table 7 shows the investment-based three-factor model cannot fully explain the monthly return spread between the high and low IE portfolios as the monthly alphas range from 30 to 33 basis points and are significant at the 1% level. Moreover, the high-minus-low IE portfolios load (significantly) positively on the ROA factor, but insignificantly on the INV factor. This evidence confirms that high IE firms are more profitable than low IE firms. Adding the UMO factor to the investment-based three-factor model reduces the alphas of the hedge portfolios. However, the alphas are still significant at the 5% level and range from 22 to 29 basis points per month. The high-minus-low IE portfolios load positively on the ROA and UMO factors.

Among all the factors we consider thus far, only ROA and UMO help partially explain the IE-return relation, indicating that both rational risk premia and mispricing contribute to the relation. However, the fact that none of the models eliminate the IE effect suggests that we find a novel pattern calling for further analyses.

---

[17] As the authors explain, investment predicts returns because given expected cash flows, high costs of capital imply low net present values of new capital and low investment. ROA predicts returns because high expected ROA relative to low investment implies high discount rates. The high discount rates are necessary to offset the high expected ROA and induce low net present values of new capital and low investment. If the discount rates were not high enough to offset the high expected ROA, firms would observe high net present values of new capital and invest more. Balakrishnan, Bartov, and Faurel (2010) report that the level of profits predicts individual stock returns (even after controlling for earnings surprise). Hirshleifer, Lim, and Teoh (2009) provide a model in which profitability predicts returns because of imperfect rationality.

5. Limited Attention, Valuation Uncertainty, and the Strength of Return Predictability

5.1. Fama-MacBeth Regressions within Attention and Valuation Uncertainty Subsamples

To test our proposed hypothesis that limited investor attention leads to a positive IE-return relation, we conduct Fama-MacBeth regressions within subsamples split by size and analyst coverage as proxies for investor attention to a stock.[18] We expect firms with smaller size and lower analyst coverage to receive less attention from investors, and therefore potentially to have more sluggish short-term stock price reactions to the information contained in innovative efficiency, and greater predictability of stock returns. Hong, Lim, and Stein (2000) test the information-diffusion model of Hong and Stein (1999) and report that the profitability of momentum strategies decreases with size and analyst coverage. In their theoretical paper on limited attention and stock prices, Hirshleifer and Teoh (2003) propose both size and analyst coverage measures as proxies for investor attention. Evidence on stock return lead-lags suggests that information diffuses gradually across between large and small firms, and between firms that are followed by different numbers of analysts (Brennan, Jegadeesh, and Swaminathan 1993; Hong, Torous, and Valkanov 2007; Hou 2007; Cohen and Frazzini 2008).

Also, consistent with size being a proxy for investor attention, Chambers and Penman (1984) and Bernard and Thomas (1989) find that post-earning announcement drift, an anomaly in which market prices underreact to earnings surprises, is strongest among small firms. Furthermore, the accrual anomaly, wherein market prices do not seem to fully reflect the information contained in cash flows versus accruals, is also stronger among small firms (Mashruwala, Rajgopal, and Shevlin

[18] Size is market equity, while analyst coverage is the average monthly number of analysts providing current fiscal year earnings estimates, averaged over the previous year. The Pearson and Spearman correlation coefficients between size and analyst coverage are 0.37 and 0.75, respectively.

2006).

Several other papers also use measures of analyst stock coverage as proxies for analyst and/or investor attention. Irvine (2003) documents a rise in liquidity following initiation of analyst coverage of a stock. Bushee and Miller (2007) document that small firms that hire an investor relations firm subsequently experience increased analyst following, suggesting that investor relations expenditures succeed in increasing investor attention. Koester, Lundholm, and Soliman (2010) find that positive extreme earnings surprises are associated with low analyst following, and that the analysts who are following such firms are busier following a greater number of other firms. They further suggest that managers engineer such surprises to attract attention, and document that after such surprises analyst following increases. Andrade, Bian, and Burch (2010) propose that information dissemination mitigates equity bubbles, and find that lower analyst coverage leads to greater price drop in reaction to the securities transaction tax increase in China in 2007.

We construct two size subsamples based on the NYSE median size breakpoint at the end of February of year *t* and two analyst coverage subsamples based on the median of analyst coverage calculated at the end of year $t - 1$.[19] Within each subsample, we then regress monthly returns of individual stocks in each month from July of year *t* to June of year $t + 1$ on IE and the same set of control variables as in the full-sample regressions (Table 4). We winsorize all independent variables at the 1% and 99% levels and standardize all independent variables to zero mean and one standard deviation within each subsample to facilitate the comparison of slopes across subsamples. All regressions control for the industry effect. For simplicity, we only report the IE slopes in Table 8.

Panel A of Table 8 shows that the IE-return relation is stronger among firms with smaller size and lower analyst coverage for different measures of IE. The average IE slopes among low attention

[19] To be consistent with the portfolio sorts, we split the sample at the end of February. Splitting the sample at the end of June generates similar results in unreported tables.

stocks are always significant at the 1% level and are substantially larger than those among high attention stocks, which often are insignificant.

Specifically, the IE slopes range from 0.08% to 0.14% with *t*-statistics of 3.24 or above for small firms and from 0.03% to 0.04% with *t*-statistics of 1.49 or below for big firms. Similarly, the IE slopes range from 0.10% to 0.16% with *t*-statistics of 2.61 or above for firms with low analyst coverage and are 0.05% with *t*-statistics ranging from 1.46 to 1.77 for firms with high analyst coverage. Although the cross-subsample differences in the IE slopes are not always statistically significant, their magnitudes are economically substantial. These sharp contrasts thus support the hypothesis that limited investor attention leads to a positive IE-return relation.

To further evaluate the effect of limited attention on the IE-return relation, we perform Fama-MacBeth regressions within subsamples formed on valuation uncertainty (VU). Past literature has reported stronger behavioral biases among stocks or portfolios with higher VU.[20] We expect the IE-return relation to be stronger among firms with more valuation uncertainty. Following Kumar (2009), we use three measures of VU: firm age, turnover, and idiosyncratic volatility (IVOL).[21] Firm age is the number of years listed on Compustat with non-missing price data. Turnover is the average monthly turnover over the prior year, and the monthly turnover is the number of shares traded during a month divided by the number of shares outstanding at the end of the month.[22] IVOL

[20] According to Einhorn (1980), overconfidence is greater in decision tasks involving greater uncertainty and less reliable feedback. Chan, Lakonishok, and Sougiannis (2001) find that the value effect (which is often interpreted as a behavioral anomaly) is stronger among firms with high R&D, for which valuation uncertainty is likely to be higher. Mashruwala, Rajgopal, and Shevlin (2006) find that the accrual anomaly is stronger among firms with high idiosyncratic volatility. This is consistent with greater misperceptions about such firms, or with high volatility being a barrier to arbitrage. Teoh, Yang, and Zhang (2009) also report that four financial anomalies are stronger among firms with lower R-squares. In a test of the model of Daniel, Hirshleifer, and Subrahmanyam (2001), Kumar (2009) reports greater individual investor trading biases among stocks with greater valuation uncertainty.

[21] The turnover has alternatively been used as a proxy of investor attention in Hou, Peng, and Xiong (2009), so its interpretation is necessarily mixed. However, the other two measures of VU, firm age and idiosyncratic volatility, are not subject to the concern of dual-interpretations.

[22] Following the literature, e.g., LaPlante and Muscarella (1997) and Hou (2007), we divide the NASDAQ volume by a factor of two.

is the standard deviation of the residuals from regressing daily stock excess returns on market excess returns over a maximum of 250 days.

We interpret firms with younger age, higher turnover, or higher idiosyncratic volatility as having higher valuation uncertainty. We form two age and two IVOL subsamples based on the medians of these measures at the end of year $t - 1$, and two turnover subsamples based on the median of turnover at the end of February of year *t.* Within each VU subsample, we conduct the same Fama-MacBeth regressions as were performed in the attention subsamples (Panel A). We winsorize all independent variables at the 1% and 99% levels and standardize all independent variables to zero mean and one standard deviation within each subsample to facilitate the comparison of slopes across subsamples.

Panel B of Table 8 shows a stronger IE-return relation in subsamples with higher valuation uncertainty. For example, the average slopes of Patents/RDC are 0.11% and 0.00% with *t*-statistics of 3.54 and 0.12 in the young and old age subsamples, respectively. In untabulated results, the difference in slopes across the age subsamples is statistical significant at the 1% level. The average slopes of Patents/RDC are 0.10% and 0.01% with *t*-statistics of 3.17 and 0.34 in the high and low turnover subsamples, respectively, and the difference in slopes is statistical significant at the 5% level. The average slopes of Patents/RDC are 0.08 and 0.05 with *t*-statistics of 2.28 and 2.65 in the high and low IVOL subsamples, respectively. Although the difference in IE slopes across the IVOL subsamples are statistically insignificant, its magnitude is substantial. These sharp contrasts across VU subsamples are robust to various measures of IE, suggesting that the IE-return relation is more likely to be driven by behavioral biases. In unreported results, we find the sharp contrasts are robust to alternative measures of R&D intensity.

5.2. Double-Sorted Portfolios Based upon IE and Attention Measures

We also confirm a stronger IE-return relation in smaller firms through portfolio sorts for the three IE measures. For brevity, we only report the results for Patents/RDC. Tables 9 and 10 show the monthly value-weighted return of the high-minus-low IE portfolio formed among small firms is 55 basis points and significant at the 1% level. The monthly alphas estimated from factor models with or without the mispricing factor UMO are all significant at the 1% level and range from 32 basis points for the investment-based model plus UMO to 59 basis points for the CAPM.

Furthermore, UMO dominates the risk factors in the augmented models and improves the models' explanatory power substantially. For example, Table 9 shows adding UMO to the Carhart model reduces the alpha from 46 basis points to 34 basis points per month. The loading on UMO is positive and significant at the 1% level, while the loadings on the market, size, value, and momentum factors are insignificant and sometimes negative.

In contrast to the findings for small firms, the return of the hedge portfolio formed among large firms is 27 basis points and significant at the 10% level. The alphas estimated from the investment-based model and models augmented with UMO are statistically insignificant. The explanatory power of these models mainly derives from the risk factor ROA and the mispricing factor UMO. These findings empirically support the hypothesis that the market reacts sluggishly to news of firms with lower attention, leading to the return predictability associated with innovative efficiency.

6. The EMI (Efficient Minus Inefficient) Factor

To further examine if the IE-return relation is driven by risk or mispricing (or both) and whether IE reflects commonality in returns not fully captured by existing factors, we construct an EMI (Efficient Minus Inefficient) factor based on six size-IE portfolios detailed in Section 4.2 following

Fama and French (1993). As before, we focus on the IE measure, Patents/RDC. The EMI factor is the difference between the average of the value-weighted returns on the two high IE portfolios (S/H and B/H) and the average of the value-weighted returns on the two low IE portfolios (S/L and B/L). In other words, the EMI factor is (S/H + B/H)/2 – (S/L + B/L)/2. This return series tracks any factor return comovement associated with innovative efficiency, regardless of the underlying causes coming from systematic mispricing or rational risk.[23]

Figure 1 plots the EMI factor returns and the MKT factor returns on a per annum basis from 1982 to 2008. Surprisingly, the EMI factor returns are negative for only five years out of 27 years, while the MKT factor returns are negative for eight years. Moreover, the EMI factor also appears to be a good hedge against market downturns; the EMI factor returns are almost always positive in those years in which the MKT factor returns are negative. For example, the MKT factor returns in 1984, 1987, 1990, 1994, 2000, 2001, 2002, and 2008 are –6.12%, –3.52%, –13.00%, –4.50%, –16.10%, –14.63%, –22.15%, and –41.78%, respectively. In contrast, the corresponding EMI factor returns are 2.31%, 15.20%, 6.07%, 7.72%, 31.72%, –1.49%, 4.70%, and 15.36%, respectively. More interestingly, even though the internet bubble burst in 2000, the EMI factor performed extremely well, with a substantial return of 31.72%.

Table 11 reports the summary statistics for EMI and other well-known factor returns. Panel A describes the means, standard deviations, time series *t*-statistics, and the ex post Sharpe ratios of the monthly returns of EMI, Fama-French three factors (MKT, SMB, and HML), two investment-related factors (INV and ROA), the momentum factor (MOM), and the mispricing factor (UMO). The average return of EMI is 0.41% per month, which is lower than the average returns of MKT (0.66%), ROA (0.82%), MOM (0.80%), and UMO (0.90%); however, it is higher than the average

[23] The EMI factor return series is the same as the size-adjusted returns for the high-minus-low IE portfolios discussed in Section 4.2.

returns of SMB (0.06%), HML (0.39%), and INV (0.22%).

Furthermore, the standard deviation of EMI is 1.84%, which is considerably lower than those of all the other factors except INV (1.79%). This finding indicates that investing based upon innovative efficiency is even more attractive than its substantial returns would suggest. Indeed EMI offers an ex post Sharpe ratio of 0.22, which is higher than that of all the other factors except UMO (0.28).[24]

Panel B reports the correlation between different factor returns, and shows that EMI is distinct from other familiar factors. EMI has a correlation of –0.10 with MKT, –0.20 with SMB, and 0.13 with HML, all of which are small in magnitude. Moreover, EMI has a correlation of 0.08 with INV, 0.21 with ROA, –0.02 with MOM, and 0.21 with UMO. The fairly high correlation between EMI and ROA suggests a link between IE and profitability. The high correlation between EMI and UMO, on the other hand, suggests a link between IE and systematic mispricing.

These findings suggest that investors may be able to do substantially better than the market portfolio, or the three Fama-French factors in optimal combination, by further including the EMI factor in their portfolios. Panel C describes the maximum ex post Sharpe ratios achievable by combining the various factors to form the tangency portfolio, which is, according to the mean-variance portfolio theory, the optimal portfolio of risky assets to select when a risk-free asset is available.

The first row shows that the monthly ex post Sharpe ratio of MKT is 0.15. The second row shows that when SMB is available as well, it receives negative weighting in the optimal portfolio (–12%), but that the maximum achievable Sharpe ratio remains the same. The third row shows that when HML is also available, it is weighted extremely heavily (52%), and almost doubles the Sharpe ratio, bringing it to 0.29. The fourth row shows that EMI substantially increases the Sharpe ratio of

[24] Ex post Sharpe ratio estimates are upward biased (MacKinlay 1995). However, adjusting for the bias would not change the qualitative nature of our conclusions. Moreover, we find that EMI offers an ex post Sharpe ratio of 0.19 after dropping the year 2000, in which EMI reaches its highest annualized return.

the tangency portfolio, to 0.38. Moreover, the weight on EMI (38%) is much higher than that of any of the other three factors. The reason that EMI dominates in the tangency portfolio is that it combines three good features: a substantial average return, a very low standard deviation, and a very low (in some cases negative) correlation with the other three Fama-French factors.

The improvement in the maximum Sharpe ratio from inclusion of EMI is a challenge to rational risk premia as an explanation for IE-based return predictability. Previous research on the equity premium puzzle (Mehra and Prescott 1985) already indicates that the high Sharpe ratio of the stock market presents a difficult challenge for rational asset pricing theory. Furthermore, EMI retains its substantial role in the tangency portfolio, with weights ranging from 22% to 33% when combined with INV, ROA, MOM, or UMO. In unreported results, we find that even when all the eight factors are included, the weight on EMI is still 20%, which is substantially higher than that of any of the other factors except MKT and UMO (26% and 33%, respectively). These findings suggest that EMI captures risk or mispricing effects above and beyond those captured by the above well-known factors.

7. Conclusion

We find that firms that are more efficient in innovation have superior future operating performance and stock returns. The relation between innovative efficiency and operating performance is robust to measures of operating performance (ROA or cash flows) and control variables such as R&D intensity, R&D growth, lagged operating performance, and change in adjusted patent citations scaled by average total assets. The positive relation of innovation efficiency (IE) with subsequent stock returns is robust to alternative IE measures, and to controlling for other firm characteristics known to predict returns, such as size, book-to-market, momentum, R&D

intensity, R&D growth, change in adjusted patent citations scaled by average total assets, investment intensity, ROA, asset growth, net stock issuance, and institutional ownership. Empirical factor pricing models, such as the CAPM, the Carhart model, and the investment-based three-factor model, do not fully explain this relation. Adding the mispricing factor UMO to these models improves the models' explanatory power substantially, but substantial and significant abnormal return performance remains. These findings suggest that mispricing plays an important role in explaining the IE-return relation.

Further analyses show that proxies for investor inattention and valuation uncertainty are associated with stronger ability of IE to predict returns. These findings provide further support for psychological bias or constraints contributing to the IE-return relation. The high Sharpe ratio of the EMI (Efficient Minus Inefficient) factor also suggests this relation is not entirely explained by rational pricing. Finally, regardless of the source of the effect, the heavy weight of the EMI factor in the tangency portfolio suggests that innovative efficiency captures pricing effects above and beyond those captured by the other well-known factors.

As a policy matter, if capital markets fail to reward firms that are more efficient at innovation, there will be potential misallocation of resources in which firms that are highly effective at innovation are undercapitalized relative to firms that are less effective at innovating. Our findings suggest that market participants should direct greater attention to innovative efficiency, and that business students should be taught that innovative efficiency can be a useful input for firm valuation.

REFERENCES

Andrade, Sandro C., Jiangze Bian, and Timothy R. Burch, 2010, Does information dissemination mitigate bubbles? The role of analyst coverage in China, working paper, University of Miami.

Balakrishnan, Karthik, Eli Bartov, and Lucile Faurel, 2010, Post loss/profit announcement drift, *Journal of Accounting and Economics* 50, 20-41.

Barberis, Nicholas, and Andrei Shleifer, 2003, Style investing, *Journal of Financial Economics* 68, 161-199.

Berk, Jonathan B., 1995, A critique of size related anomalies, *Review of Financial Studies* 8, 275-286.

Berk, Jonathan, Richard. C. Green, and Vasant Naik, 2004, Valuation and return dynamics of new ventures, *Review of Financial Studies* 17, 1-35.

Bernard, Victor L., and Jacob K. Thomas, 1989, Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research* 27, 1-36.

Brennan, Michael J., Narasimhan Jegadeesh, and Bhaskaran Swaminathan, 1993, Investment analysis and the adjustment of stock prices to common information, *Review of Financial Studies* 6, 799-824.

Bushee, Brian J., and Gregory S. Miller, 2007, Investor relations, firm visibility, and investor following, working paper, University of Pennsylvania and University of Michigan.

Chambers, Anne E., and Stephen H. Penman, 1984, Timeliness of reporting and the stock price reaction to earnings announcement, *Journal of Accounting Research* 22, 21-47.

Chambers, Dennis, Ross Jennings, and Robert B. Thompson II, 2002, Excess returns to R&D-intensive firms, *Review of Accounting Studies* 7, 133-158.

Chan, Louis K. C., Josef Lakonishok, and Theodore Sougiannis, 2001, The stock market valuation of research and development expenditures, *Journal of Finance* 56, 2431-2456.

Chen, Long, Robert Novy-Marx, and Lu Zhang, 2010, An alternative three-factor model, working paper, Washington University in St. Louis, University of Rochester, and Ohio State University.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57-82.

Cochrane, John H., 1991, Production-based asset pricing and the link between stock returns and economic fluctuations, *Journal of Finance* 46, 209-237.

Cochrane, John H., 1996, A cross-sectional test of an investment-based asset pricing model, *Journal of Political Economy* 104, 572-620.

Cohen, Lauren, and Andrea Frazzini, 2008, Economic links and predictable returns, *Journal of Finance* 63, 1977-2011.

Cooper, Michael J., Huseyin Gulen, and Michael J. Schill, 2008, Asset growth and the cross-section of stock returns, *Journal of Finance* 63, 1609-1651.

Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam, 2001, Overconfidence, arbitrage, and equilibrium asset pricing, *Journal of Finance* 56, 921-965.

Daniel, Kent, and Sheridan Titman, 2006, Market reactions to tangible and intangible information, *Journal of Finance* 61, 1605-1643.

DellaVigna, Stefano, and Joshua Pollet, 2009, Investor inattention and Friday earnings announcements, *Journal of Finance* 64, 709-749.

Dierickx, Ingemar, and Karel Cool, 1989, Asset stock accumulation and sustainability of competitive advantage, *Management Science* 35, 1504-1513.

Eberhart, Allan C., William F. Maxwell, and Akhtar R. Siddique, 2004, An examination of long-term abnormal stock returns and operating performance following R&D increases, *Journal of Finance* 59, 623-650.

Eberhart, Allan C., William F. Maxwell, and Akhtar R. Siddique, 2008, A reexamination of the tradeoff between the future benefit and riskiness of R&D increases, *Journal of Accounting Research* 46, 27-52.

Einhorn, Hillel J., 1980, Overconfidence in judgment, *New Directions for Methodology of Social and Behavioral Science* 4, 1-16.

Fama, Eugene, and Kenneth French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3-56.

Fama, Eugene, and Kenneth French, 1996, Multifactor explanations of asset pricing anomalies, *Journal of Finance* 51, 55-84.

Fama, Eugene, and Kenneth French, 1997, Industry costs of equity, *Journal of Financial Economics* 43, 153-193.

Fama, Eugene, and Kenneth French, 2008, Dissecting anomalies, *Journal of Finance* 63, 1653-1678.

Fama, Eugene, and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy* 71, 607-636.

Fang, Lily, and Joel Peress, 2009, Media coverage and the cross-section of stock returns, *Journal of Finance* 64, 2023-2052.

Garleanu, Nicolae, Leonid Kogan, and Stavros Panageas, 2011, The demographics of innovation and asset returns, working paper, University of California at Berkeley, MIT, and University of Chicago.

Garleanu, Nicolae, Stavros Panageas, and Jianfeng Yu, 2011, Technological growth and asset pricing, working paper, University of California at Berkeley, University of Chicago, and University of Minnesota.

Greenwood, Jeremy, and Boyan Jovanovic, 1999, The information-technology revolution and the stock market, *American Economic Review, Papers and Proceedings* 89, 116-122.

Griliches, Zvi, 1990, Patent statistics as economic indicators: A survey, *Journal of Economic Literature* 28, 1661-1707.

Gu, Feng, 2005, Innovation, future earnings, and market efficiency, *Journal of Accounting Auditing and Finance* 20, 385-418.

Hall, Bronwyn H., 2005, Exploring the patent explosion, *Journal of Technology Transfer* 30, 35-48.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, 2001, The NBER patent citation data file: Lessons, insights and methodological tools, NBER Working Paper 8498.

Hall, Bronwyn H., and Rosemarie Ham Ziedonis, 2001, The patent paradox revisited: An empirical study of patenting in the US semiconductor industry, 1979-1995, *RAND Journal of Economics* 32, 101-128.

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, 2005, Market value and patent citations, *RAND Journal of Economics* 36, 16-38.

Harhoff, Dietmar, Francis Narin, F.M. Scherer, and Katrin Vopel, 1999, Citation frequency and the value of patented inventions, *Review of Economics and Statistics* 81, 511-515.

Hirshleifer, David, and Danling Jiang, 2010, A financing-based misvaluation factor and the cross section of expected returns, *Review of Financial Studies* 23, 3401-3436.

Hirshleifer, David, Sonya Lim, and Siew Hong Teoh, 2009, Driven to distraction: Extraneous events and underreaction to earnings news, *Journal of Finance* 63, 2287-2323.

Hirshleifer, David, and Siew Hong Teoh, 2003, Limited attention, information disclosure, and financial reporting, *Journal of Accounting and Economics* 36, 337-386.

Hong, Harrison, and Jeremy C. Stein, 1999, A unified theory of underreaction, momentum trading, and overreaction in asset markets, *Journal of Finance* 54, 2143-2184.

Hong, Harrison, Terence Lim, and Jeremy C. Stein, 2000, Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies, *Journal of Finance* 55, 265-295.

Hong, Harrison, Walter Torous, and Rossen Valkanov, 2007, Do industries lead stock markets? *Journal of Financial Economics* 83, 367-396.

Hou, Kewei, 2007, Industry information diffusion and the lead-lag effect in stock returns, *Review of Financial Studies* 20, 1113-1138.

Hou, Kewei, Lin Peng, and Wei Xiong, 2009, A tale of two anomalies: The implication of investor attention for price and earnings momentum, working paper, Ohio State University, City University of New York, and Princeton University.

Hsu, Po-Hsuan, 2009, Technological innovations and aggregate risk premiums, *Journal of Financial Economics* 94, 264-279.

Huberman, Gur, and Tomer Regev, 2001, Contagious speculation and a cure for cancer, *Journal of Finance* 56, 387-396.

Irvine, Paul, 2003, The incremental impact of analyst initiation of coverage, *Journal of Corporate Finance* 9, 431-451.

Kahneman, Daniel, and Dan Lovallo, 1993, Timid choices and bold forecasts: A cognitive perspective on risk taking, *Management Science* 39, 17-31.

Klibanoff, Peter, Owen Lamont, and Thierry A. Wizman, 1998, Investor reaction to salient news in closed-end country funds, *Journal of Finance* 53, 673-699.

Koester, Allison, Russell Lundholm, and Mark Soliman, 2010, Attracting attention in a limited attention world: An exploration of the forces behind positive extreme earnings surprises, working paper, University of Washington and University of British Columbia.

Kothari, S.P., Ted E. Laguerre, and Andrew J. Leone, 2002, Capitalization versus expensing: evidence on the uncertainty of future earnings from capital expenditures versus R&D outlays. *Review of Accounting Studies* 7, 355-382.

Kumar, Alok, 2009, Hard-to-value stocks, behavioral biases, and informed trading, *Journal of Financial and Quantitative Analysis* 44, 1375-1401.

Lanjouw, Jean, and Mark Schankerman, 2004, Patent quality and research productivity: Measuring innovation with multiple indicators, *Economic Journal* 114, 441-465.

LaPlante, Michele, and Chris J. Muscarella, 1997, Do institutions receive comparable execution in the NYSE and Nasdaq markets? A transaction study of block trades, *Journal of Financial Economics* 45, 97-134.

Lerner, Joshua, 1994, The importance of patent scope: An empirical analysis, *RAND Journal of Economics* 25, 319-333.

Lev, Baruch, Bharat Sarath, and Theodore Sougiannis, 2005, R&D reporting biases and their consequences, *Contemporary Accounting Research* 22, 977-1026.

Lev, Baruch, and Theodore Sougiannis, 1996, The capitalization, amortization, and value-relevance of R&D, *Journal of Accounting and Economics* 21, 107-138.

Lev, Baruch, 2001, *Intangibles: Management, Measurement and Reporting*. Washington, D.C.: Brookings Institution Press.

Li, Dongmei, 2010, Financial constraints, R&D investment, and stock returns, forthcoming, *Review of Financial Studies.*

Liu, Laura X., Toni Whited, and Lu Zhang, 2009, Investment-based expected stock returns, *Journal of Political Economy* 117, 1105-1139.

Lovallo, Dan, and Daniel Kahneman, 2003, Delusions of success – How optimism undermines executives' decision, *Harvard Business Review* 81, 56-63.

Lyandres, Evgeny, Le Sun, and Lu Zhang, 2008, The new issues puzzle: Testing the investment-based explanation. *Review of Financial Studies* 21, 2825-2855.

MacKinlay, A. Craig, 1995, Multifactor models do not explain deviations from the CAPM, *Journal of Financial Economics* 38, 3-28.

Mashruwala, Christina, Shivaram Rajgopal, and Terry Shevlin, 2006, Why is the accrual anomaly not arbitraged away? The role of idiosyncratic risk and transaction costs, *Journal of Accounting and Economics* 42, 3-33.

Matolcsy, Zoltan P., and Anne Wyatt, 2008, The association between technological conditions and the market value of equity, *Accounting Review* 83, 479-518.

Mehra, Rajnish, and Edward C. Prescott, 1985, The equity premium: A puzzle, *Journal of Monetary Economics* 15, 145-161.

Merton, Robert C, 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Pandit, Shail, Charles E. Wasley, and Tzachi Zach, 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.

Pastor, Lubos, and Pietro Veronesi, 2009, Technological revolutions and stock prices, *American Economic Review* 99, 1451-1483.

Peng, Lin, and Wei Xiong, 2006, Investor attention and time-varying comovements, *Journal of Financial Economics* 80, 563-602.

Polk, Christopher, and Paola Sapienza, 2009, The stock market and corporate investment: A test of catering theory, *Review of Financial Studies* 22, 187-217.

Pontiff, Jeffrey, and Artemiza Woodgate, 2008, Share issuance and cross-sectional returns, *Journal of Finance* 63, 921-945.

Song, Hyunjin, and Norbert Schwarz, 2010, If it's easy to read, it's easy to do, pretty, good, and true: Fluency effects on judgment, choice, and processing style, *The Psychologist* 23, 108-111.

Trajtenberg, Manuel, 1990, A penny for your quotes: Patent citations and the value of innovations, *RAND Journal of Economics* 21, 172-187.

Teoh, Siew Hong, Yong Yang, and Yinglei Zhang, 2009, R-square and market efficiency, working paper, University of California, Irvine and Chinese University of Hong Kong.

Table 1 Innovative efficiency measures of selected industries

This table reports the pooled mean, standard deviation (Stdev), 25th percentile, median, and 75th percentile of three innovative efficiency (IE) measures for selected two-digit SIC industries. The three IE measures in year t are computed as patents granted to a firm in year t scaled by R&D capital in year $t-2$ (Patents/RDC), patents granted to a firm in year t scaled by R&D expenses in year $t-2$ (Patents/RD), and adjusted patent citations received in year t by patents granted to a firm in years $t-1$ to $t-5$ scaled by the sum of R&D expenses in years $t-3$ to $t-7$ (Citations/RD). R&D capital in year $t-2$ is computed as $R\&D_{i,t-2} + 0.8*R\&D_{i,t-3} + 0.6*R\&D_{i,t-4} + 0.4*R\&D_{i,t-5} + 0.2*R\&D_{i,t-6}$. The numerator of Citations/RD for firm i in year t is computed as the sum of $C_{i,n(t-j)}$, i.e., the sum of adjusted number of citations received in year t by patent n_{t-j} ($n_{t-j} = 1 \ldots N_{t-j}$) issued to firm i in year $t-j$ (j = 1, 2, 3, 4, 5). The adjusted number of citations in year t by patent n_{t-j}, $C_{i,n(t-j)}$, is the ratio of number of citations received by patent n_{t-j} to the mean number of citations received by patents of the same subcategory/grant year group cited in year t. N_{t-j} is the total number of patents issued to firm i in year $t-j$ that are cited in year t. The denominator of Citations/RD for firm i in year t is computed as $R\&D_{i,t-3} + R\&D_{i,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}$. More details are provided in Section 2.1. We construct these three IE measures for each year from 1981 to 2006.

Industries	Mean	Stdev	25%	Median	75%
A. Patents/RDC					
Biotech and pharmaceuticals (28)	0.49	5.80	0.00	0.07	0.23
Computers and machinery (35)	0.55	6.48	0.00	0.07	0.29
Electrical and electronics (36)	0.56	4.34	0.00	0.09	0.33
Transportation equipment (37)	2.11	40.78	0.02	0.11	0.37
Medical and scientific instruments (38)	0.61	3.87	0.00	0.11	0.37
Other	1.31	38.69	0.00	0.00	0.18
B. Patents/RD					
Biotech and pharmaceuticals (28)	0.86	14.33	0.00	0.18	0.56
Computers and machinery (35)	0.91	4.05	0.00	0.18	0.69
Electrical and electronics (36)	0.96	3.83	0.00	0.23	0.79
Transportation equipment (37)	1.64	9.52	0.05	0.28	0.89
Medical and scientific instruments (38)	1.06	3.80	0.00	0.27	0.89
Other	0.96	8.94	0.00	0.00	0.48
C. Citations/RD					
Biotech and pharmaceuticals (28)	2.57	33.67	0.00	0.26	1.07
Computers and machinery (35)	2.49	13.62	0.00	0.29	1.16
Electrical and electronics (36)	2.62	14.86	0.00	0.32	1.33
Transportation equipment (37)	10.66	123.17	0.05	0.35	1.41
Medical and scientific instruments (38)	4.76	57.14	0.00	0.36	1.65
Other	3.42	49.07	0.00	0.02	0.75

Table 2 Summary statistics

At the end of February of year *t* we sort firms into three groups (low, middle, and high) based on the 33th and 66th percentiles of our primary IE measure, Patents/RDC, in year *t* – 1. Patents/RDC is defined in Table 1. Panel A reports the annual average number of firms and pooled median characteristics of the three IE groups. Market equity is CRSP price per share times the number of shares outstanding at the end of February of year *t*. Future citations is the number of citations received from the grant year till 2006. Patents/RD and Citations/RD are defined in Table 1. RDS is R&D expenses divided by sales. RDG is the growth in annual R&D expenses. ΔAPC is change in adjusted patent citations (i.e., Citations defined in Table 1) scaled by total assets averaged over years *t* and *t* – 1. ROA is income before extraordinary items divided by lagged total assets. Book-to-market (BTM) is the ratio of book equity of fiscal year ending in year *t* – 1 to market equity at the end of year *t* – 1. Book equity is the Compustat book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit, minus the book value of preferred stock. Depending on availability, we use redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. Net stock issues (NS) is the change in the natural log of the split-adjusted shares outstanding. The split-adjusted shares outstanding is Compustat shares outstanding times the Compustat adjustment factor. Asset growth (AG) is the change in total assets divided by lagged total assets. CapEx/Assets is capital expenditure divided by lagged total assets. Momentum is the prior six-month returns (with one-month gap between the holding period and the current month). Institutional ownership (IO) denotes the fraction of firm shares outstanding owned by institutional investors. Panel B reports the Pearson correlation coefficients (below-diagonal) and Spearman rank correlation coefficients (above-diagonal) between selected variables. *p*-values are reported in parentheses.

A. Descriptive statistics			
	Patents/RDC		
	Low	Middle	High
Number of firms	574	274	424
Market equity ($million)	42.05	630.59	208.54
% of total market equity	13.4%	23.7%	17.5%
Patents/RDC	0.00	0.08	0.41
Future citations/RDC	0.00	0.88	5.57
Future citations excluding self-citations/RDC	0.00	0.77	4.86
Future self-citations/RDC	0.00	0.03	0.25
Patents/RD	0.00	0.20	0.97
Citations/RD	0.00	0.30	1.14
R&D expenses ($million)	2.00	28.50	7.37
RDS (%)	5.37	6.93	3.98
RDG	0.09	0.08	0.10
ΔAPC (%)	0.00	0.00	0.33
ROA (%)	2.55	4.75	4.93
ROA in year after portfolio formation (%)	2.15	4.56	4.64
BTM	0.55	0.48	0.50
NS	0.01	0.01	0.01
AG	0.06	0.06	0.07
CapEx/Assets	0.04	0.05	0.06
Momentum	–0.06	0.02	0.00
IO	0.19	0.54	0.41

Table 2 (continued)

B. Correlation coefficients (Pearson correlations are shown below the diagonal with Spearman rank correlations above)

	Patents/ RDC	Patents/ RD	Citations/ RD	Future citations/ RDC	RDS	RDG	ΔAPC	ROA	Size	BTM
Patents/RDC	-	0.99 (0.00)	0.63 (0.00)	0.92 (0.00)	–0.06 (0.00)	0.04 (0.00)	0.16 (0.00)	0.10 (0.00)	0.31 (0.00)	–0.03 (0.00)
Patents/RD	0.58 (0.00)	-	0.62 (0.00)	0.91 (0.00)	–0.09 (0.00)	0.03 (0.00)	0.15 (0.00)	0.11 (0.00)	0.30 (0.00)	–0.01 (0.04)
Citations/RD	0.09 (0.00)	0.20 (0.00)	-	0.61 (0.00)	–0.02 (0.00)	0.04 (0.00)	0.29 (0.00)	0.10 (0.00)	0.31 (0.00)	–0.05 (0.00)
Future citations/RDC	0.83 (0.00)	0.53 (0.00)	0.14 (0.00)	-	–0.01 (0.04)	0.04 (0.00)	0.17 (0.00)	0.11 (0.00)	0.31 (0.00)	–0.04 (0.00)
RDS	0.00 (0.84)	0.00 (0.62)	0.00 (0.76)	0.00 (0.86)	-	0.14 (0.00)	0.09 (0.00)	–0.28 (0.00)	0.03 (0.00)	–0.30 (0.00)
RDG	0.01 (0.00)	0.01 (0.04)	0.01 (0.07)	0.01 (0.02)	0.00 (0.67)	-	0.02 (0.00)	0.16 (0.00)	0.11 (0.00)	–0.15 (0.00)
ΔAPC	0.01 (0.01)	0.07 (0.00)	0.12 (0.00)	0.02 (0.00)	0.00 (0.30)	–0.00 (0.78)	-	–0.01 (0.15)	0.05 (0.00)	–0.04 (0.00)
ROA	0.00 (0.41)	0.01 (0.21)	0.00 (0.38)	0.00 (0.96)	–0.01 (0.00)	–0.02 (0.00)	–0.03 (0.00)	-	0.34 (0.00)	–0.17 (0.00)
Size	0.00 (0.41)	–0.01 (0.03)	–0.01 (0.09)	0.00 (0.41)	0.00 (0.54)	–0.00 (0.54)	0.00 (0.49)	0.04 (0.00)	-	–0.30 (0.00)
BTM	0.02 (0.00)	0.01 (0.04)	0.00 (0.96)	0.01 (0.02)	–0.01 (0.10)	–0.01 (0.01)	–0.01 (0.00)	0.00 (0.33)	–0.04 (0.00)	-

Table 3 Innovative efficiency and operating performance

This table reports the average slopes (in %) and their time series *t*-statistics in parentheses from annual Fama and MacBeth (1973) cross-sectional regressions of individual stocks' operating performance in year $t + 1$ on innovative efficiency (IE) and other control variables in year *t*. We measure operating performance by ROA (income before extraordinary items divided by lagged total assets) and CF (income before extraordinary items plus depreciation divided by lagged total assets) in Panels A and B, respectively. ln(1+IE) is the natural log of one plus each of the three IE measures (defined in Table 1). ln(1+RDS) is the natural log of one plus R&D expenses divided by sales. R&D growth (RDG) is the growth in annual R&D expenses. ΔAPC is change in adjusted patent citations scaled by average total assets (defined in Table 2). All regressions include industry dummies for Fama and French's (1997) 48 industries. We winsorize all variables at the 1% and 99% levels and standardize all independent variables to zero mean and one standard deviation.

A. Dependent variable = ROA next year			
Independent variables	Patents/RDC	Patents/RD	Citations/RD
ln(1+IE)	0.23	0.32	0.41
	(3.55)	(3.95)	(5.45)
ln(1+RDS)	–1.01	–1.01	–1.00
	(–3.29)	(–3.27)	(–3.25)
RDG	0.05	0.02	0.02
	(0.31)	(0.13)	(0.10)
ΔAPC	–0.68	–0.70	–0.78
	(–2.83)	(–2.91)	(–3.19)
ROA	24.11	24.10	24.14
	(13.33)	(13.34)	(13.34)
B. Dependent variable = CF next year			
Independent variables	Patents/RDC	Patents/RD	Citations/RD
ln(1+IE)	0.27	0.35	0.44
	(4.20)	(4.49)	(5.93)
ln(1+RDS)	–1.10	–1.09	–1.08
	(–4.34)	(–4.31)	(–4.30)
RDG	0.00	–0.03	–0.04
	(0.00)	(–0.22)	(–0.26)
ΔAPC	–0.71	–0.72	–0.80
	(–3.17)	(–3.23)	(–3.51)
CF	23.46	23.46	23.49
	(15.49)	(15.49)	(15.50)

Table 4 Fama-MacBeth regressions of stock returns on innovative efficiency and other variables

This table reports the average slopes (in %) and their time series *t*-statistics in parentheses from monthly Fama and MacBeth (1973) cross-sectional regressions of individual stocks' excess returns from July of year *t* to June of year *t* + 1 on ln(1+IE) and other control variables in fiscal year ending in year *t* – 1 except size and momentum. The details of IE are defined in Table 1. ln(1+RDS) is the natural log of one plus R&D expenses divided by sales. R&D growth (RDG) is the growth in annual R&D expenses, and ΔAPC is change in adjusted patent citations scaled by average total assets (defined in Table 2). ln(Size) is the natural log of market equity at the end of June of year *t*. ln(BTM) is the natural log of book-to-market equity ratio. Book equity is the Compustat book value of stockholders' equity, plus balance sheet deferred taxes and investment tax credit, minus the book value of preferred stock. Depending on availability, we use redemption, liquidation, or par value (in that order) to estimate the book value of preferred stock. Market equity (ME) is CRSP price per share times the number of shares outstanding at the end of year *t* – 1. Momentum is the prior six-month returns (with one-month gap between the holding period and the current month). CapEx/Assets is capital expenditure divided by lagged total assets. ROA is income before extraordinary items divided by lagged total assets. Asset growth (AG) is the change in total assets divided by lagged total assets. Net stock issues (NS) is the change in the natural log of the split-adjusted shares outstanding. The split-adjusted shares outstanding is Compustat shares outstanding times the Compustat adjustment factor. Institutional ownership (IO) denotes the fraction of firm shares outstanding owned by institutional investors. To control for industry effects, we also include industry dummies based on the 48 industry classification defined in Fama and French (1997) in all regressions. All independent variables are normalized to have zero mean and one standard deviation after winsorization at the 1% and 99% levels. The return data are from July of 1982 to June of 2008.

Independent variables	Patents/RDC	Patents/RD	Citations/RD
ln(1+IE)	0.06	0.10	0.09
	(2.91)	(3.13)	(4.32)
ln(1+RDS)	0.04	0.00	0.04
	(0.77)	(0.04)	(0.81)
RDG	0.22	0.11	0.22
	(0.67)	(0.27)	(0.67)
ΔAPC	–0.11	–0.10	–0.14
	(–1.97)	(–1.59)	(–2.47)
ln(Size)	–0.33	–0.30	–0.34
	(–3.21)	(–2.91)	(–3.32)
ln(BTM)	0.51	0.46	0.51
	(6.54)	(5.55)	(6.54)
Momentum	–0.12	–0.11	–0.12
	(–1.08)	(–0.97)	(–1.07)
CapEx/Asset	0.01	0.03	0.01
	(0.11)	(0.54)	(0.11)
ROA	0.29	0.34	0.29
	(2.08)	(2.40)	(2.08)
AG	–0.28	–0.29	–0.28
	(–4.84)	(–4.53)	(–4.80)
NS	–0.08	–0.07	–0.08
	(–1.22)	(–1.07)	(–1.20)
IO	0.13	0.12	0.13
	(2.44)	(2.14)	(2.35)

Table 5 Innovative efficiency portfolio returns and standard risk factor models

At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint at the end of February of year *t* and three IE groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of IE in year t – 1. We measure IE by Patents/RDC, Patents/RD, and Citations/RD defined in Table 1 in Panels A, B, and C, respectively. We hold these portfolios over the next twelve months and compute value-weighted monthly returns of these size-IE portfolios (S/H, B/H, S/M, B/M, S/L, and B/L). We then calculate monthly size-adjusted returns of the low, middle, and high IE portfolios as (S/L+B/L)/2, (S/M+B/M)/2, and (S/H+B/H)/2, respectively. We also form a high-minus-low IE portfolio (H–L) with size-adjusted return as (S/H+B/H)/2–(S/L+B/L)/2. This table reports the monthly average size-adjusted excess returns to these portfolios and the intercepts (α, in percentage) and risk factor loadings from regressing portfolio excess returns on factor returns. Heteroscedasticity-robust *t*-statistics are reported in parentheses. Excess return is the difference between size-adjusted portfolio returns and the one-month Treasury bill rate. MKT, SMB, and HML are the market, size, and book-to-market factors of Fama and French (1993). MOM is the momentum factor of Carhart (1997).

Innovative efficiency	Excess return (%)	CAPM		Carhart four-factor model				
		α	MKT	α	MKT	SMB	HML	MOM
A. Patents/RDC								
Low	0.42	–0.40	1.23	–0.23	1.08	0.57	–0.18	–0.05
t	(1.25)	(–2.92)	(33.48)	(–2.56)	(43.45)	(17.24)	(–3.91)	(–1.41)
Middle	0.79	0.03	1.16	0.22	1.02	0.44	–0.20	–0.06
t	(2.58)	(0.22)	(43.60)	(2.51)	(41.40)	(8.36)	(–4.17)	(–1.91)
High	0.82	0.03	1.19	0.20	1.06	0.47	–0.17	–0.05
t	(2.61)	(0.28)	(42.66)	(2.31)	(41.15)	(8.64)	(–4.52)	(–1.47)
H–L	0.41	0.44	–0.04	0.42	–0.02	–0.10	0.02	–0.00
t	(3.91)	(4.14)	(–1.64)	(3.91)	(–0.83)	(–1.93)	(0.35)	(–0.03)
B. Patents/RD								
Low	0.39	–0.43	1.24	–0.23	1.08	0.58	–0.22	–0.05
t	(1.17)	(–3.02)	(32.07)	(–2.60)	(40.67)	(16.35)	(–4.38)	(–1.35)
Middle	0.80	0.01	1.20	0.23	1.04	0.44	–0.27	–0.05
t	(2.51)	(0.05)	(44.50)	(2.58)	(40.79)	(7.86)	(–5.74)	(–1.44)
High	0.81	0.05	1.15	0.18	1.04	0.46	–0.10	–0.06
t	(2.68)	(0.43)	(42.41)	(2.18)	(41.32)	(9.49)	(–2.96)	(–1.94)
H–L	0.42	0.48	–0.09	0.42	–0.03	–0.12	0.11	–0.01
t	(3.83)	(4.34)	(–2.92)	(3.76)	(–1.11)	(–2.33)	(2.19)	(–0.31)
C. Citations/RD								
Low	0.52	–0.27	1.19	–0.14	1.06	0.54	–0.15	–0.02
t	(1.63)	(–2.07)	(37.04)	(–1.59)	(40.89)	(15.43)	(–3.40)	(–0.59)
Middle	0.76	0.02	1.11	0.12	1.02	0.40	–0.06	–0.05
t	(2.61)	(0.21)	(43.07)	(1.40)	(48.57)	(12.42)	(–1.51)	(–1.66)
High	0.80	0.00	1.20	0.23	1.04	0.45	–0.24	–0.07
t	(2.52)	(0.04)	(41.70)	(2.95)	(40.86)	(8.08)	(–6.54)	(–2.46)
H–L	0.28	0.28	0.01	0.38	–0.02	–0.09	–0.09	–0.05
t	(2.71)	(2.53)	(0.33)	(3.28)	(–0.50)	(–1.71)	(–1.64)	(–1.59)

Table 6 Innovative efficiency portfolio returns and the Carhart model with UMO

At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint at the end of February of year *t* and three IE groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of IE in year *t* – 1. We measure IE by Patents/RDC, Patents/RD, and Citations/RD defined in Table 1 in Panels A, B, and C, respectively. We hold these portfolios over the next twelve months and compute value-weighted monthly returns of these size-IE portfolios (S/H, B/H, S/M, B/M, S/L, and B/L). We then calculate monthly size-adjusted returns of the low, middle, and high IE portfolios as (S/L+B/L)/2, (S/M+B/M)/2, and (S/H+B/H)/2, respectively. We also form a high-minus-low IE portfolio (H–L) with size-adjusted return as (S/H+B/H)/2–(S/L+B/L)/2. This table reports the monthly average size-adjusted excess returns to these portfolios and the intercepts (α, in percentage) and factor loadings from regressing portfolio excess returns on factor returns. Heteroscedasticity-robust *t*-statistics are reported in parentheses. Excess return is the difference between size-adjusted portfolio returns and the one-month Treasury bill rate. MKT, SMB, and HML are the market, size, and book-to-market factors of Fama and French (1993). MOM is the momentum factor of Carhart (1997). UMO (Undervalued Minus Overvalued) is the mispricing factor of Hirshleifer and Jiang (2010).

Innovative efficiency	Excess return (%)	Carhart four-factor model plus UMO					
		α	MKT	SMB	HML	MOM	UMO
A. Patents/RDC							
Low	0.42	–0.14	1.05	0.57	–0.11	–0.01	–0.14
t	(1.25)	(–1.52)	(39.69)	(16.74)	(–2.29)	(–0.24)	(–2.90)
Middle	0.79	0.15	1.04	0.44	–0.25	–0.08	0.10
t	(2.58)	(1.66)	(38.29)	(8.38)	(–4.63)	(–2.85)	(1.62)
High	0.82	0.16	1.07	0.47	–0.19	–0.06	0.05
t	(2.61)	(1.85)	(40.54)	(8.53)	(–3.86)	(–2.09)	(0.96)
H–L	0.41	0.30	0.02	–0.09	–0.08	–0.05	0.19
t	(3.91)	(2.89)	(0.83)	(–1.88)	(–1.47)	(–1.54)	(3.44)
B. Patents/RD							
Low	0.39	–0.14	1.04	0.57	–0.13	–0.00	–0.15
t	(1.17)	(–1.44)	(37.79)	(15.89)	(–2.70)	(–0.09)	(–3.00)
Middle	0.80	0.19	1.05	0.45	–0.30	–0.06	0.06
t	(2.51)	(1.97)	(37.90)	(7.77)	(–5.32)	(–2.10)	(1.01)
High	0.81	0.13	1.06	0.46	–0.15	–0.08	0.08
t	(2.68)	(1.53)	(40.67)	(9.49)	(–3.18)	(–2.80)	(1.51)
H–L	0.42	0.27	0.02	–0.11	–0.01	–0.07	0.23
t	(3.83)	(2.51)	(0.78)	(–2.29)	(–0.25)	(–1.99)	(4.41)
C. Citations/RD							
Low	0.52	–0.09	1.04	0.53	–0.10	0.01	–0.08
t	(1.63)	(–0.96)	(38.73)	(15.05)	(–2.02)	(0.29)	(–1.77)
Middle	0.76	0.03	1.06	0.41	–0.13	–0.09	0.13
t	(2.61)	(0.37)	(45.76)	(12.61)	(–3.12)	(–2.93)	(2.89)
High	0.80	0.23	1.04	0.45	–0.24	–0.07	0.00
t	(2.52)	(2.76)	(40.75)	(7.82)	(–4.49)	(–2.58)	(0.01)
H–L	0.28	0.32	0.01	–0.08	–0.14	–0.08	0.09
t	(2.71)	(2.69)	(0.19)	(–1.67)	(–1.98)	(–2.17)	(1.27)

Table 7 Innovative efficiency portfolio returns and the investment-based model with and without UMO

At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint at the end of February of year *t* and three IE groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of IE in year *t* – 1. We measure IE by Patents/RDC, Patents/RD, and Citations/RD defined in Table 1 in Panels A, B, and C, respectively. We hold these portfolios over the next twelve months and compute value-weighted monthly returns of these size-IE portfolios (S/H, B/H, S/M, B/M, S/L, and B/L). We then calculate monthly size-adjusted returns of the low, middle, and high IE portfolios as (S/L+B/L)/2, (S/M+B/M)/2, and (S/H+B/H)/2, respectively. We also form a high-minus-low IE portfolio (H–L) with size-adjusted return as (S/H+B/H)/2–(S/L+B/L)/2. This table reports the monthly average size-adjusted excess returns to these portfolios and the intercepts (α, in percentage) and factor loadings from regressing portfolio excess returns on factor returns. Heteroscedasticity-robust *t*-statistics are reported in parentheses. Excess return is the difference between size-adjusted portfolio returns and the one-month Treasury bill rate. MKT is the market factor of Fama and French (1993). INV and ROA are the investment and profitability factors from Chen, Novy-Marx, and Zhang (2010). UMO (Undervalued Minus Overvalued) is the mispricing factor of Hirshleifer and Jiang (2010).

Innovative efficiency	Excess return (%)	Investment-based three factors				Investment-based three factors plus UMO				
		α	MKT	INV	ROA	α	MKT	INV	ROA	UMO
A. Patents/RDC										
Low	0.42	0.00	1.08	–0.20	–0.32	0.05	1.06	–0.15	–0.30	–0.07
t	(1.25)	(0.00)	(28.96)	(–2.62)	(–6.74)	(0.39)	(26.53)	(–1.49)	(–4.62)	(–0.81)
Middle	0.79	0.29	1.06	–0.19	–0.19	0.29	1.06	–0.19	–0.19	–0.00
t	(2.58)	(1.89)	(30.47)	(–2.38)	(–3.21)	(2.34)	(31.15)	(–1.52)	(–2.27)	(–0.02)
High	0.82	0.34	1.08	–0.15	–0.24	0.32	1.08	–0.16	–0.24	0.02
t	(2.61)	(2.18)	(29.98)	(–1.85)	(–4.22)	(2.65)	(31.36)	(–1.24)	(–2.93)	(0.18)
H–L	0.41	0.33	–0.01	0.05	0.08	0.27	0.02	–0.01	0.05	0.09
t	(3.91)	(3.11)	(–0.21)	(0.71)	(2.01)	(2.55)	(0.71)	(–0.08)	(1.18)	(1.53)
B. Patents/RD										
Low	0.39	–0.01	1.08	–0.20	–0.33	0.06	1.06	–0.14	–0.31	–0.09
t	(1.17)	(–0.06)	(29.76)	(–2.58)	(–7.01)	(0.44)	(27.71)	(–1.34)	(–4.67)	(–1.03)
Middle	0.80	0.31	1.08	–0.23	–0.22	0.34	1.07	–0.21	–0.21	–0.03
t	(2.51)	(1.94)	(30.54)	(–2.69)	(–3.59)	(2.64)	(31.79)	(–1.52)	(–2.34)	(–0.23)
High	0.81	0.30	1.06	–0.11	–0.20	0.28	1.07	–0.12	–0.21	0.03
t	(2.68)	(2.07)	(29.71)	(–1.37)	(–3.89)	(2.33)	(30.44)	(–1.06)	(–2.83)	(0.29)
H–L	0.42	0.31	–0.02	0.09	0.13	0.22	0.01	0.02	0.10	0.12
t	(3.83)	(2.90)	(–0.87)	(1.36)	(3.32)	(2.01)	(0.38)	(0.23)	(2.27)	(2.32)
C. Citations/RD										
Low	0.52	0.07	1.07	–0.14	–0.27	0.09	1.06	–0.11	–0.26	–0.04
t	(1.63)	(0.47)	(30.06)	(–1.81)	(–5.94)	(0.70)	(27.56)	(–1.08)	(–4.04)	(–0.39)
Middle	0.76	0.18	1.05	–0.06	–0.13	0.15	1.06	–0.08	–0.13	0.03
t	(2.61)	(1.39)	(31.67)	(–0.80)	(–2.92)	(1.34)	(31.51)	(–0.79)	(–2.39)	(0.38)
High	0.80	0.36	1.07	–0.24	–0.26	0.38	1.06	–0.22	–0.25	–0.03
t	(2.52)	(2.46)	(32.49)	(–3.21)	(–4.84)	(3.25)	(32.16)	(–1.79)	(–3.12)	(–0.26)
H–L	0.28	0.30	0.00	–0.10	0.01	0.29	0.00	–0.11	0.01	0.01
t	(2.71)	(2.70)	(0.01)	(–1.54)	(0.28)	(2.37)	(0.06)	(–1.27)	(0.21)	(0.09)

Table 8 Fama-MacBeth regressions: subsample analysis

This table reports the average slopes (in %) of ln(1+IE) and their time series *t*-statistics in parentheses from monthly Fama and MacBeth (1973) cross-sectional regressions of individual stocks' excess returns from July of year *t* to June of year *t* + 1 on ln(1+IE) and other control variables in fiscal year ending in year *t* – 1 except size and momentum in subsamples split by the medians of firm size, analyst coverage, firm age, turnover, and idiosyncratic volatility (IVOL). We measure IE by Patents/RDC, Patents/RD, and Citations/RD as defined in Table 1. Firm size is the market equity at the end of February of year *t*. Analyst coverage (AC) is the average monthly number of stock analyst reports on earnings estimates in year *t* – 1. Firm age denotes the number of years listed on Compustat with non-missing price data at the end of year *t* – 1. Turnover is the average monthly turnover over the prior twelve months at the end of February of year *t*, and the monthly turnover is the number of shares traded during a month divided by the number of shares outstanding at the end of the month. IVOL is the standard deviation of the residuals from regressing daily stock excess returns on market excess returns over a maximum of 250 days ending on December 31 of year *t* – 1. All regressions include the following control variables defined in Table 4: ln(1+RDS), RDG, ΔAPC, ln(Size), ln(BTM), momentum, CapEx/Assets, ROA, asset growth, net stock issues, and institutional ownership. We also include industry dummies based on the 48 industry classification defined in Fama and French (1997) in all regressions. All independent variables are normalized to zero mean and one standard deviation after winsorization at the 1% and 99% levels. The return data are from July of 1982 to June of 2008.

	Patents/RDC	Patents/RD	Citations/RD
Panel A: Subsamples split by investor attention proxies			
Small size	0.08	0.14	0.11
	(3.24)	(3.45)	(4.59)
Big size	0.04	0.03	0.04
	(1.32)	(0.84)	(1.49)
Low AC	0.10	0.16	0.13
	(2.61)	(2.80)	(3.45)
High AC	0.05	0.05	0.05
	(1.77)	(1.46)	(1.70)
Panel B: Subsamples split by valuation uncertainty proxies			
Young age	0.11	0.23	0.12
	(3.54)	(3.76)	(3.99)
Old age	0.00	0.01	0.05
	(0.12)	(0.16)	(2.00)
High turnover	0.10	0.13	0.10
	(3.17)	(3.18)	(3.53)
Low turnover	0.01	0.04	0.09
	(0.34)	(0.81)	(2.67)
High IVOL	0.08	0.13	0.15
	(2.28)	(2.43)	(4.24)
Low IVOL	0.05	0.09	0.05
	(2.65)	(2.64)	(2.26)

Table 9 Size/innovative efficiency portfolio returns and standard risk factor models with or without UMO

At the end of February of year t from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint and three innovative efficiency (IE) groups (low "L", middle "M", or high "H") based on the 33[th] and 66[th] percentiles of Patents/RDC as defined in Table 1. Size is the market equity at the end of February of year t, and IE is measured in year t – 1. The intersection of these portfolios forms six size-IE portfolios (S/L, S/M, S/H, B/L, B/M, and B/H). We also form a zero-investment portfolio (H–L) that goes long on the high IE portfolio and short on the low IE portfolio within each size group. We hold the portfolios over the next 12 months and rebalance them each year. This table reports the monthly average value-weighted excess returns (in percentage) to these portfolios and the intercepts (α, in percentage) and factor loadings from regressing portfolio excess returns on factor returns. Heteroscedasticity-robust t-statistics are reported in parentheses. Excess return is the difference between portfolio returns and the one-month Treasury bill rate. MKT, SMB, and HML are the market, size, and book-to-market factors of Fama and French (1993). MOM is the momentum factor of Carhart (1997). UMO (Undervalued Minus Overvalued) is the mispricing factor of Hirshleifer and Jiang (2010).

Size/ IE	Excess return	CAPM		Carhart four-factor model					Carhart four-factor model plus UMO					
		α	MKT	α	MKT	SMB	HML	MOM	A	MKT	SMB	HML	MOM	UMO
S/L	0.38	–0.52	1.35	-0.28	1.12	1.07	-0.20	-0.08	-0.22	1.09	1.07	-0.14	-0.05	-0.11
t	(0.94)	(–2.36)	(26.18)	(-2.89)	(40.96)	(18.61)	(-3.87)	(-2.30)	(-1.99)	(36.51)	(18.14)	(-2.49)	(-1.44)	(-1.52)
S/M	0.91	0.03	1.32	0.25	1.10	1.09	-0.15	-0.10	0.18	1.13	1.09	-0.21	-0.13	0.11
t	(2.23)	(0.12)	(25.00)	(1.75)	(26.33)	(10.41)	(-1.87)	(-1.71)	(1.20)	(26.22)	(10.28)	(-2.07)	(-2.47)	(1.07)
S/H	0.93	0.07	1.29	0.18	1.12	1.01	-0.05	-0.04	0.13	1.14	1.01	-0.09	-0.06	0.08
t	(2.44)	(0.34)	(27.05)	(1.44)	(28.73)	(12.73)	(-0.94)	(-0.87)	(0.97)	(29.37)	(12.58)	(-1.26)	(-1.53)	(1.04)
S/H–L	0.55	0.59	–0.07	0.46	0.00	-0.06	0.15	0.04	0.34	0.04	-0.06	0.04	-0.01	0.19
t	(4.70)	(5.00)	(–2.29)	(3.76)	(-0.07)	(-1.24)	(3.09)	(1.12)	(2.87)	(1.35)	(-1.15)	(0.72)	(-0.25)	(2.68)
B/L	0.45	–0.28	1.11	-0.17	1.04	0.07	-0.17	-0.01	-0.07	1.00	0.07	-0.08	0.04	-0.16
t	(1.52)	(–2.19)	(28.75)	(-1.22)	(23.40)	(1.22)	(-2.54)	(-0.20)	(-0.46)	(20.50)	(1.08)	(-1.15)	(0.71)	(-2.38)
B/M	0.68	0.02	0.99	0.18	0.94	-0.22	-0.24	-0.02	0.13	0.96	-0.21	-0.29	-0.04	0.08
t	(2.67)	(0.27)	(48.07)	(2.46)	(50.74)	(-7.72)	(-6.90)	(-0.88)	(1.63)	(45.09)	(-7.23)	(-7.54)	(-1.62)	(1.92)
B/H	0.72	–0.00	1.09	0.21	1.00	-0.07	-0.28	-0.05	0.20	1.00	-0.07	-0.29	-0.06	0.02
t	(2.53)	(–0.04)	(37.07)	(2.11)	(33.61)	(-1.39)	(-6.10)	(-1.81)	(1.88)	(31.08)	(-1.35)	(-5.08)	(-1.91)	(0.39)
B/H–L	0.27	0.28	–0.02	0.38	-0.04	-0.14	-0.11	-0.04	0.27	0.00	-0.13	-0.21	-0.09	0.18
t	(1.74)	(1.79)	(–0.53)	(2.25)	(-0.93)	(-1.85)	(-1.38)	(-0.96)	(1.56)	(0.05)	(-1.78)	(-2.21)	(-1.78)	(2.14)

Table 10 Size/innovative efficiency portfolio returns and the investment-based model with and without UMO

At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint and three innovative efficiency (IE) groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of Patents/RDC as defined in Table 1. Size is the market equity at the end of February of year *t*, and IE is measured in year $t-1$. The intersection of these portfolios forms six size-IE portfolios (S/L, S/M, S/H, B/L, B/M, and B/H). We also form a zero-investment portfolio (H–L) that is long the high innovative efficiency portfolio and short the low innovative efficiency portfolio within each size group. We hold the portfolios over the next 12 months and rebalance them each year. This table reports the monthly average value-weighted excess percent returns (in percentage) to these portfolios and the intercepts (α, in percentage) and factor loadings from regressing portfolio excess returns on factor returns. Heteroscedasticity-robust *t*-statistics are reported in parentheses. Excess return is the difference between portfolio returns and the one-month Treasury bill rate. MKT is the market factor of Fama and French (1993). INV and ROA are the investment and profitability factors from Chen, Novy-Marx, and Zhang (2010). UMO (Undervalued Minus Overvalued) is the mispricing factor of Hirshleifer and Jiang (2010).

		Investment-based three-factor model				Investment-based three-factor model plus UMO				
Size/IE	Excess return	α	MKT	INV	ROA	α	MKT	INV	ROA	UMO
S/L	0.38	0.07	1.13	-0.26	-0.47	0.13	1.11	-0.21	-0.45	-0.08
t	(0.94)	(0.30)	(19.39)	(-1.95)	(-6.12)	(0.60)	(17.77)	(-1.11)	(-4.01)	(-0.49)
S/M	0.91	0.55	1.12	-0.37	-0.38	0.51	1.14	-0.40	-0.40	0.06
t	(2.23)	(1.73)	(15.32)	(-2.20)	(-3.04)	(2.10)	(16.70)	(-1.49)	(-2.23)	(0.25)
S/H	0.93	0.50	1.13	-0.12	-0.36	0.45	1.15	-0.17	-0.38	0.07
t	(2.44)	(1.81)	(17.56)	(-0.89)	(-3.62)	(2.06)	(18.85)	(-0.76)	(-2.66)	(0.36)
S/H–L	0.55	0.43	0.00	0.13	0.12	0.32	0.04	0.04	0.08	0.15
t	(4.70)	(3.35)	(-0.10)	(1.78)	(2.72)	(2.66)	(1.20)	(0.46)	(1.52)	(1.93)
B/L	0.45	-0.07	1.03	-0.14	-0.16	-0.02	1.01	-0.10	-0.14	-0.07
t	(1.52)	(-0.59)	(26.76)	(-1.83)	(-3.79)	(-0.17)	(22.13)	(-1.06)	(-3.02)	(-0.83)
B/M	0.68	0.02	0.99	-0.02	0.01	0.07	0.97	0.02	0.02	-0.07
t	(2.67)	(0.25)	(42.77)	(-0.41)	(0.20)	(0.77)	(37.64)	(0.31)	(0.78)	(-1.31)
B/H	0.72	0.17	1.02	-0.17	-0.11	0.20	1.02	-0.15	-0.11	-0.03
t	(2.53)	(1.67)	(35.93)	(-2.67)	(-3.68)	(1.76)	(31.21)	(-1.87)	(-2.75)	(-0.51)
B/H–L	0.27	0.24	-0.01	-0.04	0.04	0.22	0.00	-0.06	0.03	0.03
t	(1.74)	(1.58)	(-0.21)	(-0.36)	(0.82)	(1.28)	(0.02)	(-0.48)	(0.58)	(0.40)

Table 11 Summary statistics of monthly factor returns and Sharpe Ratio (SR)

At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint and three innovative efficiency (IE) groups (low "L", middle "M", or high "H") based on the 33^{th} and 66^{th} percentiles of our primary IE measure, Patents/RDC (defined in Table 1), in year *t* – 1. Size is the market equity at the end of February of year *t*. The intersection of these portfolios forms six size-IE portfolios (S/L, S/M, S/H, B/L, B/M, and B/H). Value-weighted monthly returns on these six portfolios are computed from March of year *t* to February of year *t* + 1. The innovative efficiency factor—EMI (Efficient Minus Inefficient)—is (S/H+B/H)/2–(S/L+B/L)/2. MKT is the return on the value-weighted NYSE/AMEX/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. INV and ROA are the investment and profitability factors from Chen, Novy-Marx, and Zhang (2010). UMO (Undervalued Minus Overvalued) is the mispricing factor of Hirshleifer and Jiang (2010). Panel A reports the mean, standard deviation (Stdev), *t*-statistics, and ex post Sharpe ratio (SR) for these factors. Panel B reports the Pearson correlation coefficients among these factors. Panel C reports the monthly Sharpe ratios of ex post tangency portfolios based on investing in subsets of these factor-mimicking portfolios. Portfolio weights are determined by $\Omega^{-1}r$, normalized to sum to one. Ω is the sample covariance matrix and r is the column vector of average excess returns of the factor-mimicking portfolios. All returns and standard deviations are in percentage.

Panel A: Summary statistics of factor-mimicking portfolios

	EMI	MKT	SMB	HML	INV	ROA	MOM	UMO
Mean	0.41	0.66	0.06	0.39	0.22	0.82	0.80	0.90
Stdev	1.84	4.29	3.25	3.06	1.79	4.61	4.22	3.26
t-stat	3.91	2.73	0.32	2.27	2.15	3.16	3.34	4.87
Ex post SR	0.22	0.15	0.02	0.13	0.12	0.18	0.19	0.28

Panel B: Correlation matrix of factor-mimicking portfolios

	EMI	MKT	SMB	HML	INV	ROA	MOM	UMO
EMI	1.00							
MKT	–0.10	1.00						
SMB	–0.20	0.20	1.00					
HML	0.13	–0.49	–0.42	1.00				
INV	0.08	–0.31	–0.16	0.43	1.00			
ROA	0.21	–0.37	–0.50	0.47	0.14	1.00		
MOM	–0.02	–0.09	0.11	–0.08	0.14	0.20	1.00	
UMO	0.21	–0.62	–0.28	0.66	0.50	0.57	0.34	1.00

Panel C: Ex post tangency portfolio

Portfolio weights								Tangency portfolio		
MKT	SMB	HML	EMI	INV	ROA	MOM	UMO	Mean	Stdev	Ex post SR
1								0.66	4.29	0.15
1.12	–0.12							0.73	4.75	0.15
0.34	0.14	0.52						0.44	1.50	0.29
0.21	0.12	0.29	0.38					0.41	1.10	0.38
0.19	0.10	0.22	0.33	0.16				0.39	0.99	0.39
0.20	0.17	0.20	0.29		0.14			0.45	1.06	0.43
0.19	0.08	0.26	0.32			0.14		0.48	1.08	0.44
0.27	0.09	0.02	0.22				0.40	0.64	1.20	0.53

Figure 1 EMI (Efficient Minus Inefficient) and market factor returns over time

This figure plots the return (on a per annum basis) for the EMI (Efficient Minus Inefficient) factor and the market factor from 1982 to 2008. Rm–Rf (or MKT) is the return on the value-weighted NYSE/AMEX/NASDAQ portfolio minus the one-month Treasury bill rate. At the end of February of year *t* from 1982 to 2007, we sort firms independently into two size groups (small "S" or big "B") based on the NYSE median size breakpoint and three innovative efficiency (IE) groups (low "L", middle "M", or high "H") based on the 33th and 66th percentiles of our primary IE measure, Patents/RDC (defined in Table 1), in year t – 1. Size is the market equity at the end of February of year *t*. The intersection of these portfolios forms six size-IE portfolios (S/L, S/M, S/H, B/L, B/M, and B/H). Value-weighted monthly excess returns on these six double-sorted portfolios are computed from March of year *t* to February of year *t* + 1. The innovative efficiency factor—EMI (Efficient Minus Inefficient)—is (S/H+B/H)/2–(S/L+B/L)/2.