Internet Appendix for
"Innovative Originality, Profitability, and Stock Returns"

David Hirshleifer Po-Hsuan Hsu Dongmei Li

In Section A, we discuss an example of a firm whose innovative originality (InnOrig) dropped. In Section B, we provide a case example to illustrate the calculation of our InnOrig measure. We then describe the motivation and construction of an alternative measure of InnOrig in Section C. In Section D, we develop a model based on limited attention and derive predictions on the unconditional and conditional return predictive power of innovative originality. In Section E, we examine whether InnOrig contains favorable information about a firm's future profitability.

A. Respironics

Respironics is a leading manufacturer of medical devices used primarily for the treatment of respiratory disorders. Figures IA1 and IA2 present the value and rank of InnOrig of Respironics in our sample period, and Figures IA3 presents its ROA. Figures IA2 and IA3 together show that when Respironics moved from the high InnOrig group (rank of 3—top tercile) to the low InnOrig group (rank of 1—bottom tercile), its profitability (ROA) moved downward accordingly. This example is consistent with the positive correlation between a firm's InnOrig and its future performance identified in the paper.

B. An example of the calculation of InnOrig

In this section, we illustrate the calculation of InnOrig for Incyte in 1996. As defined in the paper, a firm's InnOrig is the average number of unique technology classes assigned to patents cited by a firm's recently granted patents (last five years). To compute Incyte's InnOrig in 1996, we first list the 12 patents granted to Incyte from 1992 to 1996 as well as the technology classes (both primary and secondary) of the patents cited by these patents in Table IA1. We then compute the number of unique technology classes cited by each patent. Incyte's InnOrig in 1996 is the average number of unique technology classes across all these 12 patents, 3.67, as reported in the last column in Table IA2.

C. Alternative measure of innovative originality based on HHI

As discussed in the paper and previous sections, our primary measure of a patent's InnOrig is the number of unique technological classes (*N*) assigned to the patents cited by the focal patent. Our primary measure of a firm's InnOrig is the average InnOrig for patents granted over the past five years. It is based on *N*, which reflects the capability of a firm's managers and scientists to combine different technologies in an original way. These measures are motivated by the popular view of innovation as recombinant search.

In the paper, we discuss the shortcoming of using one minus the traditional Herfindahl index to measure InnOrig. Here, as a robustness check, we construct an alternative measure of a patent's InnOrig based on the reciprocal of a *generalized* Herfindahl index. The traditional Herfindahl index (HHI) is the sum of the squared shares of patents cited in each technology class and is a common measure of the distribution of the patents cited across the *N* classes. By construction, HHI ranges between 1/*N* (perfect diversification) and one (perfect concentration).[1]

If patents cited are equally distributed across the *N* classes, the reciprocal of HHI is equivalent to *N*, our primary measure of patent's InnOrig. When the *N* classes have unequal shares, the reciprocal of HHI indicates the 'equivalent' number of classes cited by the focal patent, which decreases with the dispersion in the shares for a given *N*. This property of the reciprocal of HHI rests on the fact that the shares of patents cited are squared prior to being summed, putting more weight on those classes with larger shares (henceforth, the dispersion effect).

For example, if a patent cites one patent in class A and nine patents in class B, the 'equivalent' *N* (i.e., the inverse of the traditional HHI) of this patent is 1.22. However, in terms of measuring the range of knowledge draw upon by the focal patent, the share of each class is unlikely to be linearly increasing in the *importance* of each class in making a patent more original and more complex to evaluate. For example, if class A represents a new cutting-edge technology, while class B represents a mature technology, then the number of existing patents granted in class A could be far fewer than that in class B. Even though the focal patent only cites one patent from class A, this one patent at the forefront technology could be as crucial as, or more crucial than, the nine patents cited from class B in creating the focal patent. So in this example, the inverse of the traditional HHI would underestimate the patent's knowledge scope and originality score due to the dispersion

[1] When *N* is one, HHI is one and reflects perfect concentration.

effect. This suggests modifying the HHI to be less sensitive to shares relative to the number of classes, to better reflect the range of knowledge that a patent draws from.

To address these limitations of the traditional Herfindahl index in capturing a patent's knowledge scope and originality, we use the *reciprocal* of a generalized HHI (GHHI) as an alternative measure of a patent's innovative originality:

$$GHHI = \sum_{j=1}^{N} (NS_j)^{\alpha}/N^2,$$

where N denotes the number of technology classes cited, S_j denotes the share of cited patents belonging to the j-th technology class, and $1 < \alpha < 2$. Having $\alpha < 2$ reduces convexity, and hence reduces the influence of shares relative to the number of technology classes. In the limit, GHHI becomes $1/N$ as α approaches 1 (and its reciprocal is exactly our InnOrig), and becomes the traditional HHI as α approaches 2.[2] For our robustness check, we set α to the mid-point of this range, 1.5, so that GHHI reflects both the number of technology classes and the shares.

We then compute the alternative firm InnOrig measure as the average of the reciprocals of GHHI of all patents granted to a firm over the last five years. In Table IA3, we present the returns of one-way sorted portfolios based on this alternative InnOrig measure.

D. A model of limited investor attention to innovative originality

This model considers the implications of a well-established psychological constraint, limited investor attention, on the ability of innovative originality (InnOrig) to forecast abnormal returns. The basic argument is simple. As discussed in the main text, InnOrig is a low-salience historical statistic about the firm's innovative activities; empirically, InnOrig is a positive indicator of the average future citations received by a firm's patents and its profitability. Owing to low salience, a fraction of the investors neglect the favorable information about future profitability contained in high InnOrig. In consequence, the stock price underweights this information, so high InnOrig is associated with underpricing and low InnOrig with overpricing. Hence InnOrig is a positive predictor of future abnormal stock returns.

When the prior uncertainty about the value of the stock (without any conditioning on InnOrig) is higher, heavier weight should optimally be placed on InnOrig by investors in forming posterior beliefs about value. So neglect of InnOrig causes greater mispricing. Hence the ability of InnOrig to predict returns is stronger when prior valuation uncertainty is greater. Furthermore, we show

[2] When the shares are equal, GHHI is equivalent to $1/N$ regardless of the value of α.

that the ability of InnOrig to predict returns is stronger when the fraction of attentive investors is lower, and when InnOrig is a stronger positive predictor of fundamentals.

Attention requires effort, and the amount of information available exceeds our ability to process it. So attention must be selective (see, e.g., Kahneman 1973). Evidence from the experimental laboratory indicates that limited attention affects how both individual investors and financial professionals interpret public information (see the review of Libby, Bloomfield, and Nelson 2002). This suggests that limited attention may affect the valuation of public information in securities markets.

As in recent theoretical literature on limited attention, in our model some investors condition only on subsets of publicly available information signals in valuing a stock. Some investors attend to the implications of InnOrig for the firm's future prospects, and some do not. Risk averse investors who are fully attentive to the relevant information item are willing to bear only a limited amount of risk in order to exploit mispricing. In consequence, equilibrium stock prices reflect a weighted average of the beliefs of investors who attend to different signals, with weights that depend on the relative numbers in each investor group and their risk tolerances. In equilibrium, prices underreact to InnOrig because of the subset of investors who do not incorporate this information into their expectations of future cash flows.

This model builds on a recent theoretical literature on how constraints on information processing affect investor behavior. The approach followed here is similar in spirit to that of Hirshleifer and Teoh (2003) and Hirshleifer, Lim, and Teoh (2011), who study the effects on market prices of investors neglecting relevant accounting information or strategic aspects of the disclosure and reporting environment. Here we examine the implications of limited attention for neglect of information relating to innovative activity of the firm. Other recent papers model the allocation of attentional resources (Gabaix and Laibson 2005; Peng 2005; and Hirshleifer, Lim, and Teoh 2008), how limited learning capacity affects asset price comovement (Peng and Xiong 2006) and the speed of price adjustment to fundamental shocks (Peng 2005; Peng and Xiong 2006), and how neglect of demographic information affects asset prices (DellaVigna and Pollet 2007).

D1. The economic setting

There are two types of investors: those who ignore the public information contained in innovative originality about future cash flows, and those who attend to all publicly available information. There are two dates. At date 1, innovative originality is publicly revealed. We denote the revealed level of InnOrig by y. Those investors who attend to y update their prior beliefs accordingly. All consumption occurs at date 2. At date 2, the stock's terminal cash flows S_2 are realized.

Prices are set by trading in a securities market with no private information, so that a rational individual has nothing to learn from market price. An inattentive investor who is unaware of his/her signal neglect also thinks he/she has nothing to learn from market price. We therefore assume that inattentive investors do not update their beliefs based upon market price.[3]

In principle, a discrepancy between an investor's valuation and the market price could alert the investor to his information neglect. In general, however, the same constraints on processing power and memory that make it hard to attend to some public signal also make it hard to use price or other indicators to compensate optimally for the failure to attend to it. Since in reality people face many relevant signals, they try to leverage their attention by focusing on more important information items. However, we cannot know perfectly which are more important before processing them, which makes it hard to determine how to optimally compensate for information neglect. Section AIII discusses evidence suggesting that individuals fail to compensate fully for the consequences of limited attention in making decisions. So long as some fraction of inattentive investors have imperfect self-awareness, results similar to those derived here will obtain. Furthermore, even if individuals always attend to market price and draw inferences from this about their information neglect, similar results to those derived here could be obtained so long as there is noise in market price arising from liquidity trading. In such a setting, an individual who attends to a given public signal in effect has a sort of 'private' information, so individuals who attend to the public signals will profit at the expense of liquidity traders, in the spirit of the models of Grossman and Stiglitz (1976) and Diamond and Verrecchia (1981).

In general, attending to more information is costly. Since investors have finite cognitive resources, attending to some information implies less time and resources for other activities. We assume that there are two investor groups indexed by i who attend to different information sets.

[3] As discussed in previous theoretical literature on limited attention in securities markets, observing the 'wrong' price is an event which, as perceived by the investor, is not supposed to occur in equilibrium. In the Perfect Bayesian Equilibrium concept of game theory setting the individual's posterior beliefs in such a situation equal to the prior belief can be consistent with equilibrium. Similar results would hold so long as some disagreement remains between the attentive and inattentive investors, i.e., inattentive investors do not always abandon their beliefs in favor of the information implicit in the market price.

Fully attentive investors attend to all date 1 publicly available information including InnOrig; investors with limited attention neglect InnOrig.

We assume that investors have a mean-variance utility function,

$$E^i[C_2^i] - \left(\frac{A}{2}\right) V^i(C_2^i), \tag{1}$$

where C_2^i is terminal consumption, A is the coefficient of absolute risk aversion, V denotes variance, and i superscripts denote the expectation or variance as formed by group i. Specifically, we use an 'a' superscript to denote the attentive group which conditions on y, and a 'u' superscript to denote the inattentive group, which does not condition on y. (The 'u' superscript stands for "unconditional," as the inattentive group does not condition on the signal y.)

Investors have an initial wealth endowment (i.e., claims to terminal consumption) of W, and zero shares of the risky security. There is also an exogenous per capita supply of the single risky security of x_0 (i.e., supply per member of the decision-making population). (This is without loss of generality; the same results would apply if the investors were endowed with x_0 directly.)

At date 1, each individual can buy or sell the security in exchange for 'cash' (claims to terminal consumption) at price S_1. The position in the security he/she attains is denoted x^i. Let S_2, the terminal cash flow, be the true value of the stock, which is conclusively revealed to all at date 2. For brevity, we do not include any cost of attending to the information signal (for the attentive group) in the expression for terminal consumption since, conditional upon attending, such a cost is a constant that would not affect the optimal investment positions or any other expressions in the following derivations. So the consumption of an individual in attention group i is

$$C_2^i = W + x^i(S_2 - S_1). \tag{2}$$

Thus, an individual in attention group i solves for the x^i that maximizes the objective

$$x^i\left(E^i[S_2] - S_1\right) - \left(\frac{A}{2}\right) V^i\left(x^i S_2\right). \tag{3}$$

As a preliminary building block, we verify a standard finding that in equilibrium stock prices are a weighted average of investor expectations of terminal cash flows as adjusted by a risk premium. We start by calculating optimal investment positions. Differentiating the objective with respect to x^i, equating to zero and solving yields

$$x^i = \frac{E^i[S_2] - S_1}{AV^i(S_2)}. \tag{4}$$

Letting f^i denote the fraction of investors in attention group i, the security price is determined by the market clearing condition

$$\sum_i f^i x^i = x_0. \tag{5}$$

Substituting for x^i from (4), and solving for S_1 gives

$$S_1 = \frac{\sum_i \lambda^i E^i[S_2] - Ax_0}{\sum_i \lambda^i}, \tag{6}$$

where

$$\lambda^i \equiv \frac{f^i}{V^i(S_2)}. \tag{7}$$

By normality, the λ^i's are constants independent of the values of the signal realizations used by investors to condition beliefs. Of course, with normal distributions, an investor who fails to condition on a signal will have higher variance than an investor who does condition on that signal.

This confirms that, in equilibrium, prices are a weighted average of the beliefs about terminal cash flows of different investors adjusted by a risk premium ($Ax_0 / \sum_i \lambda^i$), with weight $\lambda^i / \sum_i \lambda^i$ on each group's belief. Both attention groups influence prices significantly owing to the finite risk-bearing capacity of each group. By (7), *ceteris paribus*, λ^i is increasing in f^i. Thus, the greater the fraction of investors who are inattentive to InnOrig, the greater the weight that inattentive investors play in determining prices. Similar pricing equations are found in several behavioral models, such as Hirshleifer and Teoh (2003) and Hirshleifer, Lim, and Teoh (2011).

It also captures formally the idea, common to behavioral theories of anomalies, that arbitrage of market inefficiencies is imperfect. It can be argued that arbitrageurs such as hedge funds will profit by trading against mispricing, thereby growing in risk-bearing capacity. However, a literature in behavioral finance and accounting has argued that arbitrage is limited by market frictions, psychological effects, and dynamic considerations; see, e.g., Shleifer and Vishny (1997) and Hirshleifer (2001).

In this setting rational investors exploit a trading strategy that earns predictable abnormal returns relative to a fully rational asset pricing benchmark. Nevertheless, even though markets are perfect and there are no restrictions on either long positions or short-selling, fully attentive investors do not completely arbitrage away the mispricing generated by inattentive investors because doing so is risky.

D2. Innovative originality and return predictability

Suppose that, at date 1, fraction f^a of investors attend to InnOrig, and fraction $f^u \equiv 1 - f^a$ ignore InnOrig and remain at their prior belief. It is not essential for the main conclusions of this paper that these investors completely ignore InnOrig. They could attend to some effects but ignore the positive implications of higher InnOrig for future profits. For tractability, we assume multivariate normality of the stochastic variables. As a result, date 2 cash flows can be expressed as a linear function of y (date 1 InnOrig, a normally distributed variable with mean $\bar{y}$ and standard deviation σ_y), and a noise term δ as

$$S_2 = \beta_0 + \beta_{Sy} y + \delta, \tag{8}$$

where $cov(\delta, y) = 0$.

Consistent with the empirical evidence (see Table III in the main text and Table IA.I in the Internet Appendix), we assume that InnOrig is a positive predictor of future cash flows, i.e., $\beta_{Sy} > 0$. Therefore, high value at date 2 is associated with high InnOrig at date 1. The strength of this relation is given by the regression coefficient $\beta_{Sy} \equiv cov(S_2, y)/\sigma_y^2$, where σ_y is the standard deviation of y.

We next examine the relation between the date 1 InnOrig and expected future abnormal stock returns. For tractability, we examine price changes rather than percentage returns, as is standard in much of the literature on information in securities markets (see, e.g., Verrecchia (2001) and Peng (2005)). We begin by calculating, conditional on InnOrig, the expected future value of the stock $E[S_2|y]$. Using standard properties of conditional expected values with multivariate normal distributions,

$$E^a[S_2] \equiv E[S_2|y] = E[S_2] + \beta_{Sy}(y - \bar{y}), \tag{9}$$

where $\bar{y}$ is the prior expectation of date 1 InnOrig. So the sensitivity of $E[S_2|y]$ to the level of InnOrig, $\frac{\partial(E[S_2|y])}{\partial y}$, is simply β_{Sy}.

We then examine how the current price, $S_1(y)$, relates to the level of InnOrig. By a standard formula for normal distributions,

$$V(S_2|y) = (1 - \rho^2)V(S_2), \tag{10}$$

where ρ is the correlation between S_2 and y. Since $\beta_{Sy} = \rho\sigma_S/\sigma_y$, where σ_S is the standard deviation of S_2, our assumption that $\beta_{Sy} > 0$ implies that $\rho > 0$. In addition, we assume $\rho < 1$, i.e., the variation in future cash flows is not solely determined by the variation in InnOrig. (ρ^2 is the R-square of the regression specified by (8).)

Substituting (7), (9), and (10) in (6), and recognizing that owing to limited attention, $E^u[S_2] = E[S_2]$ and $V^u(S_2) = V(S_2)$, gives

$$S_1(y) = \frac{\frac{f^a}{(1-\rho^2)}\left[E[S_2] + \beta_{Sy}(y-\bar{y})\right] + f^u E[S_2] - Ax_0 V(S_2)}{\frac{f^a}{(1-\rho^2)} + f^u}. \tag{11}$$

Let $W^a \equiv \frac{\lambda^a}{\lambda^a + \lambda^u}$ denote the weight of attentive investors in current price as defined in Section AI. It is clear from (11) that the sensitivity of the current price to y, $\frac{\partial S_1(y)}{\partial y}$, is simply $W^a \beta_{Sy}$, which increases in W^a. The last term shows that greater $V(S_2)$ reduces current stock price and increases the risk premium. But this effect does not depend on the value of y, so it does not affect the cross-sectional prediction derived from varying y.

As discussed earlier, it is more tractable to define returns as price changes. Let R denote the price changes, $S_2 - S_1$. It follows from (9) and (11) that the expected return conditional on InnOrig ($E[R]$) is

$$E[R] \equiv E[S_2|y] - S_1(y) = \frac{\beta_{Sy}(y-\bar{y})f^u}{\frac{f^a}{(1-\rho^2)} + f^u} + \frac{Ax_0 V(S_2)}{\frac{f^a}{(1-\rho^2)} + f^u}. \tag{12}$$

We label the first term the *expected future abnormal return*, which depends on the revealed level of InnOrig (y). The second term relates to risk premium. So higher $V(S_2)$ causes a higher expected return on the stock owing to rational risk aversion, which does not depend on y. Therefore, the sensitivity of $E[R]$ to InnOrig and the interactions of this sensitivity with other characteristics (discussed later) are the same as those for the future abnormal returns. For brevity, we only show the derivations based on $E[R]$ below.

From (12), we show that the sensitivity of expected returns (and future abnormal returns) to InnOrig is

$$\frac{\partial E[R]}{\partial y} = \frac{\partial E[S_2|y]}{\partial y} - \frac{\partial S_1(y)}{\partial y} = \beta_{Sy}(1 - W^a) = \frac{\beta_{Sy} f^u (1-\rho^2)}{(1 - f^u \rho^2)}, \tag{13}$$

which is positive as long as $\beta_{Sy} > 0$ and $f^u > 0$, where the last equality in (13) follows from the definition of W^a. In other words, InnOrig is a positive predictor of future abnormal returns if

InnOrig is positively associated with future value and if a non-zero fraction of investors neglect the favorable information in InnOrig.

The sensitivity in (13) is decreasing in W^a, the influence of attentive investors' beliefs in the current price. This is intuitive since the weaker is the influence of attentive investors, the more the current price deviates from the efficient price that would prevail if all investors were attentive; and the more sensitive the abnormal return is to InnOrig. Furthermore, W^a is increasing in the fraction of attentive investors, f^a, and decreasing in uncertainty about future cash flows as reflected in $V(S_2)$ or σ_S. (Note that $V(S_2) \equiv \sigma_S^2$.) More formally, W^a can be expressed as

$$W^a = \frac{\frac{f^a}{(1-\rho^2)V(S_2)}}{\frac{f^a}{(1-\rho^2)V(S_2)} + \frac{f^u}{V(S_2)}} = \frac{f^a}{1 - f^u\rho^2}. \tag{14}$$

Hence the derivative of W^a with respect to f^a is

$$\frac{\partial W^a}{\partial f^a} = \frac{1-\rho^2}{(1-f^u\rho^2)^2} > 0, \tag{15}$$

and the derivative of W^a with respect to σ_S is

$$\frac{\partial W^a}{\partial \sigma_S} = \left[\frac{2f^a f^u \rho}{(1-f^u\rho^2)^2}\right]\left(\frac{\partial \rho}{\partial \sigma_S}\right) = \left[\frac{2f^a f^u \rho}{(1-f^u\rho^2)^2}\right]\left(\frac{-\beta_{Sy}\sigma_y}{\sigma_S^2}\right) = \frac{-2(1-f^u)f^u\rho^2}{(1-f^u\rho^2)^2\sigma_S} < 0, \tag{16}$$

where the second and third equalities follow from $\rho = \beta_{Sy}\sigma_y/\sigma_S$ and $\frac{\partial \rho}{\partial \sigma_S} = \frac{-\beta_{Sy}\sigma_y}{\sigma_S^2}$.

Intuitively, the smaller the fraction of attentive investors, the less influence they have on the current price and hence the larger mispricing owing to neglect of InnOrig. In addition, when the prior uncertainty about the value of the stock (without any conditioning on InnOrig) is higher, heavier weight should optimally be placed on InnOrig by investors in forming posterior beliefs about value. So neglect of InnOrig causes greater mispricing.

Since the sensitivity of expected returns (and future abnormal returns) to InnOrig is decreasing in W^a as shown in (13), it is increasing in the uncertainty about S_2 and the fraction of inattentive investors (or decreasing in the fraction of attentive investors). More formally, taking the derivative of (13) with respect to σ_S and substituting $\frac{\partial W^a}{\partial \sigma_S}$ derived in (16) gives

$$\frac{\partial^2(E[R])}{\partial y \partial \sigma_S} = -\beta_{Sy}\frac{\partial W^a}{\partial \sigma_S} = \frac{2f^u(1-f^u)\rho^3}{\sigma_y(1-f^u\rho^2)^2}, \tag{17}$$

which is positive as long as $f^u > 0$. Similarly, taking the derivative of (13) with respect to f^u and substituting $\frac{\partial W^a}{\partial f^a}$ derived in (15) gives

$$\frac{\partial^2(E[R])}{\partial y \partial f^u} = -\beta_{Sy}\frac{\partial W^a}{\partial f^u} = \beta_{Sy}\frac{\partial W^a}{\partial f^a} = \frac{\beta_{Sy}(1-\rho^2)}{(1-f^u\rho^2)^2}, \tag{18}$$

which is positive as long as $\beta_{Sy} > 0$.

In sum, greater prior uncertainty about the stock payoff makes the return predictive power of InnOrig stronger. This prior uncertainty will be higher in a more opaque information environment (e.g., younger firms, firms with more opaque financial reports). Furthermore, the strength of the return predictive power based upon InnOrig is increasing in f^u, the degree of inattention in the market, e.g., lower analyst following relative to the supply of information.

We also explore the interaction of this return predictive power of InnOrig with β_{Sy}, the strength of the predictive relationship between InnOrig and future cash flows. Taking the derivative of (13) with respect to β_{Sy} gives

$$\frac{\partial^2(E[R])}{\partial y \partial \beta_{Sy}} = \frac{f^u[1-3\rho^2+f^u\rho^2(1+\rho^2)]}{(1-f^u\rho^2)^2}, \tag{19}$$

which is positive if $f^u > \frac{3\rho^2-1}{\rho^4+\rho^2}$, i.e., the fraction of inattentive investors is large enough. A sufficient condition for this to hold for any non-zero f^u is $\rho^2 \leq 1/3$, i.e., the R-square from the regression of future cash flows on InnOrig is less than or equal to 1/3. We verify that this sufficient condition holds in the data in untabulated results. Therefore, $\frac{\partial^2(E[R])}{\partial y \partial \beta_{Sy}} > 0$ for all $f^u > 0$. In other words, the model predicts that the strength of the return predictive power of InnOrig increases with β_{Sy} as long as there are inattentive investors in the market.

In addition, all the above derivations apply to date 1 mispricing as well, and hence to future abnormal returns. To see this, we define mispricing in the current stock price as the difference between S_1 and S_1^*, the price that would be set if all investors were attentive ($f^a = 1$). Setting f^a to 1 and f^u to 0 in (11) gives

$$S_1^* = E[S_2|y] - (1-\rho^2)Ax_0V(S_2). \tag{20}$$

Therefore, mispricing is

$$S_1^* - S_1 = E[R] - (1-\rho^2)Ax_0V(S_2), \tag{21}$$

where the equality follows from (12). In other words, expected returns are the sum of mispricing and the risk premium. It is clear that mispricing and $E[R]$ depend on InnOrig in the same way since the second term in (21) does not vary with InnOrig. Consequently, the interactions of this sensitivity with prior valuation uncertainty, investor attention, and the strength of the predictive relation between InnOrig and future profits apply to mispricing and future abnormal returns as well.

The above analysis is summarized in the following propositions:

Proposition 1. If fraction $f^u > 0$ of investors neglect the favorable information in IO, then higher IO is associated with greater subsequent abnormal returns.

Proposition 2. The greater the valuation uncertainty as reflected in σ_S, the steeper the relation between IO and subsequent abnormal returns.

Proposition 3. The greater the fraction of inattentive investors, f^u, the steeper the relation between IO and subsequent abnormal returns.

Proposition 4. If the fraction of inattentive investors is large enough or if the correlation between IO and future cash flows is low enough, then the stronger the predictive relationship between IO and future profits as reflected in the regression coefficient, β_{Sy}, the steeper the relation between IO and subsequent abnormal returns.

D3. Do investors fully compensate for limited attention?

A key assumption of our model is that individuals with limited attention trade based upon their beliefs. As a result, limited attention affects the equilibrium price. Casual observation suggests that investors often do make trades based on beliefs that do not fully reflect publicly available information. Intuitively, ignoring an information item and failing to adjust for the fact that the item has been neglected go hand in hand, as both kinds of neglect are natural results of limited cognitive capacity. A more detailed discussion and defense of the proposition that investors neglect information yet trade and influence price is provided in Hirshleifer, Lim, and Teoh (2011) Section 4.

There is extensive evidence from both psychology and experimental markets that people both neglect signals and do not adjust for the fact that they are neglecting them, such as studies that show that the form of presentation of information affects individuals' judgments and decisions (see, e.g., the review of Libby, Bloomfield, and Nelson 2002). Experimental studies have found that different presentations of equivalent information about a firm affect the valuations and trades of investors and experienced financial analysts. Presentation effects have been documented for various forms of accounting reporting and disclosure contexts. In principle, if an investor understood that owing to limited attention certain formats were hard to process, the investor could self-debias by, for example, mentally rearranging the format of presentation. However, such self-debiasing often does not occur.

There is other evidence that limited attention affects capital markets; indeed, Daniel, Hirshleifer, and Teoh (2002) argue that limited attention may underlie a wide range of anomalous patterns in securities market trading and prices. In an experimental setting, Gillette et al. (1999) document investor misreactions to public information arrival. Perhaps the most striking indication of limited attention in public markets is that stock prices react to news that is already public information (Huberman and Regev 2001, and Ho and Michaely 1988). More broadly, Hong, Torous, and Valkanov (2007) report evidence that industry stock returns lead aggregate market returns, potentially consistent with gradual diffusion of information about fundamentals across markets. Hou and Moskowitz (2005) provide a measure of investor neglect of a stock, the lag in the relation between the return on the overall market and the stock's return. They find that stocks with long delay (which can be viewed as low-attention stocks) have stronger post-earnings announcement drift. Many short-horizon event studies confirm that stock markets react immediately to relevant news, but long-horizon event studies provide evidence suggesting that there is underreaction to various kinds of public news events (see, e.g., the review of Hirshleifer 2001). However, there has been a great deal of debate as to the appropriate methodology for testing market efficiency using long-run abnormal returns. There is also evidence suggesting that investors' and analysts' assessments are influenced by the format and salience with which public signals are presented. For example, Hand (1990) finds that the reannounced gains from debt-equity swaps in quarterly earnings announcements were significantly related to mean abnormal returns. Schrand and Walther (2000) provide evidence that managers strategically select the form of the prior-period earnings benchmark when announcing earnings. Prior period special gains were more likely to be mentioned than prior period special losses in the sample, apparently to lower the benchmark for current-period evaluation. Miller (2002) finds that firms at the end of periods of sustained earnings increases shift from long-term forecasts to short-term forecasts, thereby deferring the need to forecast adversely. Plumlee (2003) finds that analyst forecasts of effective tax rates impound the effects of complex tax-law changes less accurately than less complex changes.

E. InnOrig and level of future profitability

Following Fama and French (2000), we conduct annual cross-sectional regressions of individual firms' future profitability on InnOrig and other control variables (profitability, change in profitability, market-to-book assets, advertising expenses, capital expenditure, R&D, innovative efficiency, and industry effects). We set missing values for InnOrig, IE, advertising expenses, and R&D expenses to zero. We also control for a dummy variable that equals one for firms with no R&D expenses over the last five years and its interactions with all the other control variables. For brevity, we omit the slopes on these terms related to this dummy and the industry dummies in the tabulation of results. To reduce the influence of outliers and facilitate the interpretation, we winsorize all variables at the 1% and 99% levels and standardize all independent variables (except the dummies) to zero mean and one standard deviation.

In Table IA4, we report the relation between InnOrig and the level of profitability over each of the next five years for ROE and ROA in Panels A1 and A2, respectively. In each panel, we control for citations-based IE (CIE) on the top and patents-based IE (PIE) at the bottom. To control for persistence in profitability, we control for five lags of ROE and ROA. The *t*-statistics in parentheses are based on Newey-West standard errors adjusted for autocorrelation and heteroscedasticity. As expected, InnOrig is associated with significantly higher future profitability in each of these five years, regardless which types of IE we control for. These effects are economically substantial. For example, in the top half of Panel A1, the slopes on InnOrig are 1.86%, 2.26%, 2.14%, 1.81%, and 1.78% for each of the next five years, respectively. All of them have *t*-statistics above 5. Furthermore, these results also indicate that the positive link between InnOrig and future profitability is fairly persistent.

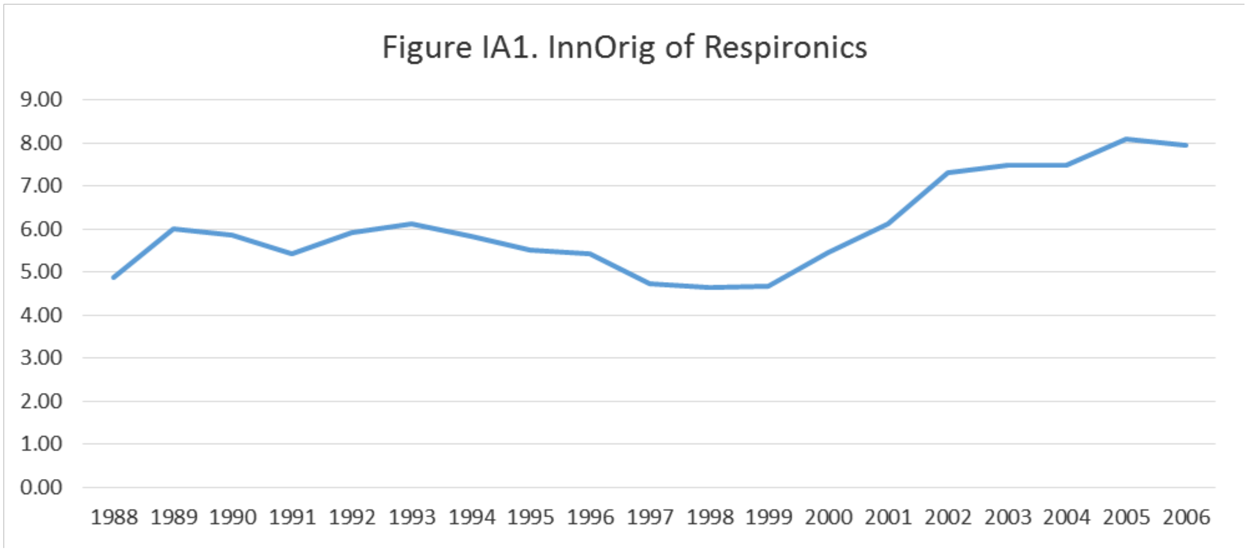
Figure IA1. InnOrig of Respironics
9.00
8.00
7.00
6.00
5.00
4.00
3.00
2.00
1.00
0.00
1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006

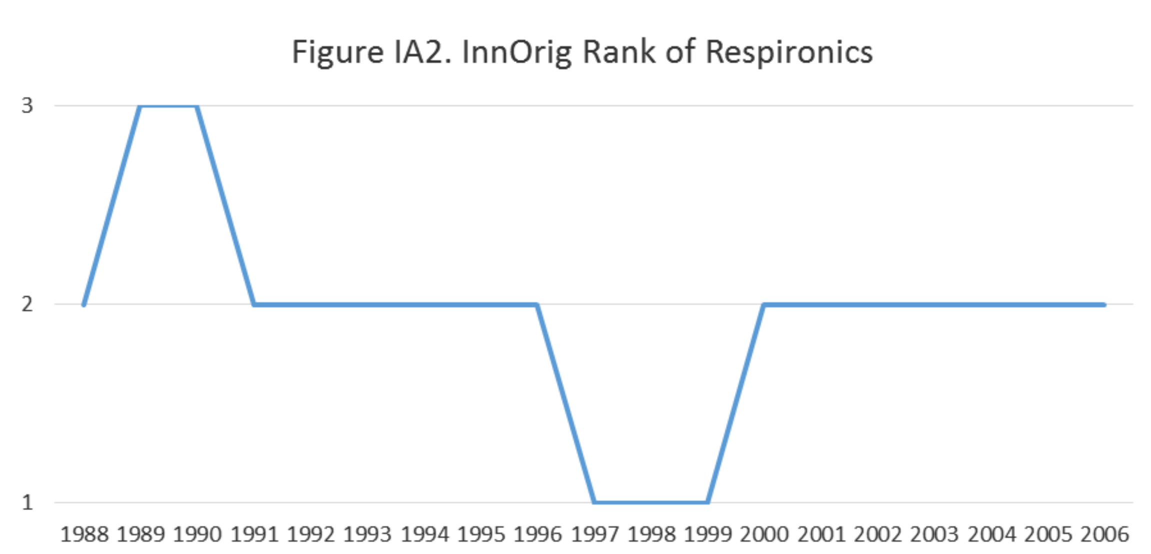
Figure IA2. InnOrig Rank of Respironics
3
2
1
1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006

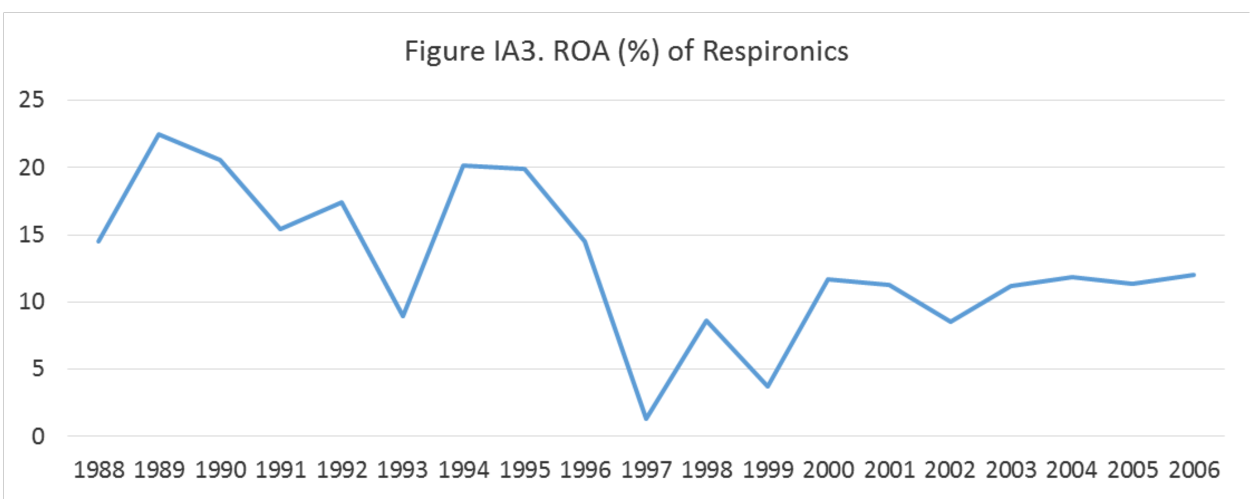
Figure IA3. ROA (%) of Respironics
25
20
15
10
5
0
1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006

Table IA1. The Technology Classes Cited by Incyte's Patents

In this table, we present the patent number for the patents granted to Incyte in 1992-1996, the grant year, the patent number of patents cited by these patents, and the primary and secondary technology classes of those cited patents.

Incyte's patent	Grant year	Cited patents	Class of cited patents (primary and secondary)
5112608	1992	4265233	424
			435
			602
			604
		4904469	424
		5006252	210
			530
5196196	1993	4265233	424
			435
			602
			604
		4904469	424
		5006252	210
			530
		5112608	424
5206017	1993	4265233	424
			435
			602
			604
		4904469	424
		5006252	210
			530
		5112608	424
5210027	1993	4705777	514
			530
			930

Incyte's patent	Grant year	Cited patents	Class of cited patents (primary and secondary)
5234912	1993	4863740	424
		5089274	424
			514
			530
5326562	1994	5006252	210
			530
		5112608	424
5334584	1994	5089274	424
			514
			530
		5171739	514
5457090	1995	5187089	435
5470825	1995	4935370	435
5476839	1995	4833092	436
			436
		4935370	435
			435
5495001	1996	4265233	424
			435
			602
			604
		4904469	424
		5006252	210
			530
		5112608	424
5532216	1996	4950600	435
		5089274	424
			514
			530
		5348942	514

Table IA2. The Unique Technology Classes Cited by Incyte's Patents

In this table, we present the patent number for the patents granted to Incyte in 1992-1996, the unique technology classes (primary and secondary) cited by these patents, the number of patents cited in each technology class, the number of unique classes cited by each patent, and Incyte's InnOrig in 1996.

Incyte's patent	Class of cited patents	Number of citations	Number of unique classes	InnOrig
5112608	210	1	6	3.667
5112608	424	2		
5112608	435	1		
5112608	530	1		
5112608	602	1		
5112608	604	1		
5196196	210	1	6	
5196196	424	3		
5196196	435	1		
5196196	530	1		
5196196	602	1		
5196196	604	1		
5206017	210	1	6	
5206017	424	3		
5206017	435	1		
5206017	530	1		
5206017	602	1		
5206017	604	1		
5210027	514	1	3	
5210027	530	1		
5210027	930	1		
5234912	424	2	3	
5234912	514	1		
5234912	530	1		
5326562	210	1	3	
5326562	424	1		
5326562	530	1		
5334584	424	1	3	
5334584	514	2		
5334584	530	1		
5457090	435	1	1	
5470825	435	1	1	
5476839	435	2	2	
5476839	436	2		
5495001	210	1	6	
5495001	424	3		
5495001	435	1		
5495001	530	1		
5495001	602	1		
5495001	604	1		
5532216	424	1	4	
5532216	435	1		
5532216	514	2		
5532216	530	1		

Table IA3. Return Predictive Power of Alternative Measure of Innovative Originality

At the end of June of year *t* from 1982 to 2007, we sort firms with non-missing alternative innovative originality (InnOrig) into three groups (Low, Middle, or High) based on the 30th and 70th percentiles of the alternative InnOrig measure in year *t* – 1 following Table IV in the paper. In addition, we assign all firms with missing InnOrig but positive five-year R&D capital into the "Low" group. The alternative InnOrig is defined as the reciprocal of GHHI, defined in Section C of the Internet Appendix. We hold these portfolios over the next twelve months (July of year *t* to June of year *t* + 1) and compute their value-weighted average monthly returns in excess of one-month Treasury bill rate (Exret). We also construct a high-minus-low (High–low) portfolio by holding a long position in the high InnOrig portfolio and a short position in the low InnOrig portfolio. We report the average industry- and characteristic-adjusted returns of all portfolios. The industry-adjusted returns (Ind-adjret) are based on the difference between individual firms' returns and the returns of firms in the same industry (based on Fama-French 48 industry classifications). Following Daniel, Grinblatt, Titman, and Wermers (DGTW 1997) and Wermers (2004), the characteristic-adjusted returns (Char-adjret) are based on the difference between individual firms' returns and the DGTW benchmark portfolio returns. We report the alphas from the regression of the time-series of portfolio excess returns on various factor models: the Fama-French three factors (the market factor–MKT, the size factor–SMB, and the value factor–HML) plus the momentum (UMD) factor (4F) model, 4F plus the investment-minus-consumption (IMC) factor (Papanikolaou 2011), the liquidity (LIQ) factor (Pastor and Stambaugh 2003), the citations- or patents-based innovative efficient-minus-inefficient (EMI1 or EMI2) factor as in Hirshleifer, Hsu, and Li (2013), the robust-minus-weak (RMW) factor and the conservative-minus-aggressive (CMA) factor as in Fama and French (2015), or the undervalued-minus-overvalued (UMO) factor of Hirshleifer and Jiang (2010). We also report the alpha from the investment-based factor model (q-factor model) of Hou, Xue, and Zhang (HXZ 2015) and the mispricing factor model of Stambaugh and Yuan (2017). All returns and alphas are value-weighted and expressed in percentage. The *t*-statistics are reported in parentheses.

A. Excess returns and adjusted returns				B. Alphas from factor models								
InnOrig	Exret	Ind-adjret	Char-adjret	4F	4F + IMC	4F + LIQ	4F + citations-based EMI	4F + patents-based EMI	4F + RMW + CMA	4F + UMO	HXZ	Mispricing
Low	0.54	-0.14	-0.06	-0.10	-0.09	-0.11	-0.05	-0.03	-0.02	-0.06	-0.03	-0.03
	(1.83)	(-1.99)	(-0.25)	(-1.30)	(-1.22)	(-1.47)	(-0.61)	(-0.43)	(-0.25)	(-0.82)	(-0.31)	(-0.34)
Middle	0.73	0.05	0.08	0.18	0.19	0.19	0.11	0.09	0.19	0.18	0.16	0.11
	(2.86)	(1.12)	(0.32)	(2.80)	(2.79)	(2.86)	(1.73)	(1.44)	(2.91)	(2.69)	(2.55)	(1.53)
High	0.76	0.06	0.15	0.18	0.17	0.17	0.15	0.13	0.13	0.15	0.09	0.12
	(2.95)	(1.20)	(0.61)	(2.06)	(2.03)	(1.88)	(1.76)	(1.46)	(1.51)	(1.75)	(1.09)	(1.44)
High-Low	0.22	0.20	0.22	0.27	0.27	0.28	0.20	0.16	0.15	0.21	0.12	0.15
	(1.75)	(2.20)	(1.86)	(2.31)	(2.23)	(2.30)	(1.69)	(1.31)	(1.24)	(1.76)	(1.00)	(1.20)

Table IA4. Innovative Originality and Future Profitability

This table reports the average slopes (in %) and their Newey-West (1987) autocorrelation-adjusted heteroscedasticity-robust t-statistics in parentheses from annual Fama and MacBeth (1973) cross-sectional regressions. In Panel A, we regress the level of profitability in year $t + k$ (k = 1 to 5) on innovative originality (*InnOrig*) and other control variables in year t from 1981 to 2006. ΔROE_t (ΔROA_t) is the change in *ROE* (*ROA*) between year t and year t – 1. *R&D* is R&D expenditure divided by assets. *Capex* is capital expenditure divided by assets. *MTB* is market-to-book assets. *Adv* is advertising expense divided by assets. We also control for five lags of ROE and ROA in year $t - k$ (k = 1 to 5). All the other control variables are defined as in Table 3. We also control for industry dummies based on the Fama and French (1997) 48 industries and five lagged ROE/ROA (slopes omitted). We set missing values for InnOrig, IE, advertising expenses, and R&D expenses to zero. In addition, we control for a dummy, which equals one for firms with no R&D investment over the past five years and 0 otherwise, and its interactions with all the other control variables. We omit the slopes on the 48 industry dummies, the slopes on the missing dummy, and its interactions with other control variables for brevity. We winsorize all variables at the 1% and 99% levels and standardize all independent variables (except the dummies) to zero mean and one standard deviation. Financial and utility firms are excluded.

Panel A1. InnOrig and level of future ROE

Dependent	$InnOrig_t$	ROE_t	ΔROE_t	ADV_t	$R\&D_t$	$Capex_t$	CIE_t	MTB_t	Intercept
ROE_{t+1}	1.86	20.26	-0.88	0.58	-6.30	-0.30	-0.33	2.71	-7.58
	(5.56)	(9.11)	(-0.47)	(1.87)	(-5.42)	(-0.63)	(-1.47)	(3.69)	(-5.22)
ROE_{t+2}	2.26	6.54	-0.35	0.59	-6.17	-0.33	-0.70	1.05	-8.30
	(6.84)	(3.58)	(-0.20)	(2.23)	(-5.62)	(-0.91)	(-2.53)	(1.20)	(-6.54)
ROE_{t+3}	2.14	6.86	-2.73	0.37	-5.62	0.06	-0.33	0.27	-8.33
	(8.11)	(3.22)	(-1.37)	(0.96)	(-6.02)	(0.16)	(-1.32)	(0.34)	(-6.55)
ROE_{t+4}	1.81	4.75	-1.48	0.60	-5.25	0.03	-0.34	0.03	-6.84
	(6.81)	(3.52)	(-1.00)	(2.46)	(-7.63)	(0.07)	(-1.42)	(0.03)	(-5.10)
ROE_{t+5}	1.78	4.88	-2.24	0.57	-4.23	0.70	-0.19	-0.15	-5.11
	(6.10)	(6.00)	(-3.85)	(1.64)	(-6.32)	(1.79)	(-1.13)	(-0.27)	(-5.19)
Dependent	$InnOrig_t$	ROE_t	ΔROE_t	ADV_t	$R\&D_t$	$Capex_t$	PIE_t	MTB_t	Intercept
ROE_{t+1}	1.84	20.21	-0.83	0.59	-6.31	-0.27	-0.38	2.71	-7.55
	(5.60)	(9.00)	(-0.44)	(1.88)	(-5.42)	(-0.58)	(-1.29)	(3.67)	(-5.19)
ROE_{t+2}	2.09	6.61	-0.41	0.62	-6.14	-0.35	-0.31	1.00	-8.27
	(6.14)	(3.64)	(-0.23)	(2.30)	(-5.64)	(-0.98)	(-1.61)	(1.15)	(-6.59)
ROE_{t+3}	2.16	6.87	-2.76	0.37	-5.64	0.09	-0.52	0.27	-8.30
	(8.43)	(3.19)	(-1.36)	(0.95)	(-5.99)	(0.27)	(-2.06)	(0.36)	(-6.55)
ROE_{t+4}	1.81	4.76	-1.48	0.60	-5.24	0.05	-0.37	0.05	-6.83
	(7.88)	(3.49)	(-1.00)	(2.39)	(-7.64)	(0.11)	(-1.60)	(0.06)	(-5.07)
ROE_{t+5}	1.77	4.86	-2.24	0.57	-4.23	0.70	-0.26	-0.15	-5.06
	(6.41)	(5.99)	(-3.90)	(1.64)	(-6.35)	(1.80)	(-1.07)	(-0.28)	(-5.12)

Panel A2. InnOrig and level of future ROA

Dependent	$InnOrig_t$	ROA_t	ΔROA_t	ADV_t	$R\&D_t$	$Capex_t$	CIE_t	MTB_t	Intercept
ROA_{t+1}	0.46	6.65	1.34	0.21	-1.30	-0.04	-0.10	1.02	-0.58
	(6.90)	(3.63)	(1.53)	(2.18)	(-4.98)	(-0.33)	(-2.73)	(4.79)	(-0.81)
ROA_{t+2}	0.51	0.99	0.71	0.25	-1.17	0.18	-0.13	0.11	-0.98
	(8.06)	(1.45)	(1.66)	(3.05)	(-4.42)	(0.99)	(-3.02)	(0.42)	(-1.28)
ROA_{t+3}	0.51	1.13	-0.10	0.15	-1.12	0.18	-0.08	-0.02	-0.92
	(10.44)	(1.99)	(-0.26)	(1.56)	(-4.26)	(1.40)	(-1.44)	(-0.12)	(-1.23)
ROA_{t+4}	0.41	0.95	-0.06	0.18	-1.05	0.09	-0.06	-0.02	-0.56
	(6.49)	(1.65)	(-0.14)	(2.65)	(-5.19)	(0.91)	(-1.12)	(-0.09)	(-0.80)
ROA_{t+5}	0.42	1.03	-0.42	0.20	-0.91	0.19	-0.04	-0.03	-0.20
	(6.36)	(5.26)	(-2.69)	(2.62)	(-4.00)	(1.55)	(-0.88)	(-0.28)	(-0.31)
Dependent	$InnOrig_t$	ROA_t	ΔROA_t	ADV_t	$R\&D_t$	$Capex_t$	PIE_t	MTB_t	Intercept
ROA_{t+1}	0.46	6.64	1.35	0.21	-1.30	-0.03	-0.15	1.02	-0.58
	(6.90)	(3.62)	(1.54)	(2.17)	(-5.00)	(-0.29)	(-2.07)	(4.74)	(-0.81)
ROA_{t+2}	0.49	1.01	0.70	0.25	-1.16	0.18	-0.10	0.10	-0.98
	(7.41)	(1.49)	(1.62)	(3.07)	(-4.44)	(1.00)	(-2.13)	(0.40)	(-1.27)
ROA_{t+3}	0.51	1.12	-0.10	0.15	-1.12	0.19	-0.12	-0.02	-0.90
	(9.89)	(1.97)	(-0.24)	(1.55)	(-4.26)	(1.48)	(-1.75)	(-0.11)	(-1.21)
ROA_{t+4}	0.40	0.95	-0.06	0.18	-1.05	0.09	-0.06	-0.01	-0.56
	(7.80)	(1.66)	(-0.13)	(2.70)	(-5.16)	(0.90)	(-0.94)	(-0.08)	(-0.79)
ROA_{t+5}	0.42	1.03	-0.42	0.20	-0.91	0.19	-0.06	-0.04	-0.19
	(6.90)	(5.19)	(-2.71)	(2.64)	(-4.01)	(1.58)	(-1.00)	(-0.29)	(-0.29)

References

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Daniel, Kent, David Hirshleifer, and Siew Hong Teoh, 2002, Investor psychology in capital markets: Evidence and policy implications, *Journal of Monetary Economics* 49, 139–209.

Fodor Stephen P., J. Leighton Read, Michael C. Pirrung, Lubert Stryer, Amy T. Lu, Dennis Solas, 1991, Light-directed, spatially addressable parallel chemical synthesis, *Science* 251, 767–773.

DellaVigna, Stefano, and Joshua M. Pollet, 2007, Demographics and industry returns, *American Economic Review* 97, 1667–1702.

Diamond, Douglas W., and Robert E. Verrecchia 1981, Information aggregation in a noisy rational expectations economy, *Journal of Financial Economics* 9, 221–235.

Fama, Eugene, and French, Kenneth, 2000, Forecasting profitability and earnings, *Journal of Business* 73, 161–175.

Gabaix, Xavier, and David Laibson, 2005, Bounded rationality and directed cognition, Working paper, Harvard University.

Gillette, Ann, Douglas Stevens, Susan Watts, and Willams Arlington, 1999, Price and volume reactions to public information releases: An experimental approach incorporating traders' subjective beliefs, *Contemporary Accounting Research* 16, 437–479.

Grossman, Sanford J., and Joseph E. Stiglitz, 1976, Information and competitive price systems, *American Economic Review* 66, 246–253.

Hand, John R. M., 1990, A test of the extended functional fixation hypothesis, *Accounting Review* 65, 740–763.

Hirshleifer, David, 2001, Investor psychology and asset pricing, *Journal of Finance* 64, 1533–1597.

Hirshleifer, David, Sonya S. Lim, and Siew Hong Teoh, 2008, Disclosure to an audience with limited attention, Working paper, Merage Business School, UC Irvine.

Hirshleifer, David, Sonya Lim, and Siew Hong Teoh, 2011, Limited investor attention and stock market misreactions to accounting information, *Review of Asset Pricing Studies* 1, 35–73.

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

Ho, Thomas S. Y., and Roni Michaely, 1988, Information quality and market efficiency, *Journal of Financial and Quantitative Analysis* 23, 53–70.

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

Hou, Kewei and Tobias J. Moskowitz, 2005, Market frictions, price delay, and the cross-section of expected returns, *Review of Financial Studies* 18, 981–1020.

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

Kahneman, Daniel, 1973, Attention and Effort (Prentice-Hall, Englewood Cliffs, New Jersey).

Libby, Robert, Robert Bloomfield, and Mark W. Nelson, 2002, Experimental research in financial accounting, *Accounting, Organizations and Society* 27, 775–810.

Miller, Gregory S., 2002, Earnings performance and discretionary disclosure, *Journal of Accounting Research* 40, 173–204.

Peng, Lin, 2005, Learning with information capacity constraints, *Journal of Financial and Quantitative Analysis* 40, 307–329.

Peng, Lin, and Wei Xiong, 2006, Investor attention, overconfidence and category learning, *Journal of Financial Economics* 80, 563–602.
Plumlee, Marlene A., 2003, The effect of information complexity on analysts' use of that information, *Accounting Review* 78, 275–296.
Schrand, Catherine M., and Beverly R. Walther, 2000, Strategic benchmarks in earnings announcements: The selective disclosure of prior-period earnings components, *Accounting Review* 75, 151–177.
Shleifer, Andrei, and Robert W. Vishny, 1997, The limits to arbitrage, *Journal of Finance* 52, 35–55.
Verrecchia, Robert, 2001, Essays on disclosure, *Journal of Accounting and Economics* 32, 97–180.
Wermers, Russ, 2004, Is money really 'smart'? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence. Working paper, University of Maryland.