Illiquidity, R&D investment, and stock returns

Shamim Ahmed* Ziwen Bu† Xiaoxia Ye‡§

July 11, 2022

*University of Liverpool Management School, University of Liverpool, Liverpool L69 7ZH, United Kingdom; Shamim.Ahmed@liverpool.ac.uk.

†Birmingham Business School, University of Birmingham, Birmingham B15 2TT, United Kingdom; Z.Bu.1@bham.ac.uk.

‡Corresponding author. University of Liverpool Management School, University of Liverpool, Liverpool L69 7ZH, United Kingdom; Xiaoxia.Ye@liverpool.ac.uk.

§We are grateful to Sanjay Chugh (Editor) and an anonymous referee for their constructive comments and suggestions, which have improved the quality of the paper considerably. We also thank Jonathan Berk, Thierry Foucault, Andrew Ellul, Albert Menkveld, Seok Young Hong, Sanjay Banerji, Manish Gupta, Weimin Liu, and Abhay Abhyankar for many useful comments and suggestions, as well as participants at several conferences and university seminars.

Abstract

We propose a dynamic model of research and development (R&D) venture, which predicts that the positive relation between the firm's R&D investment and the expected stock returns strengthens with illiquidity. Consistent with the model's prediction, empirical evidence based on cross-sectional regressions and double-sorted portfolios largely suggests a stronger and positive R&D-return relation among illiquid stocks. A further analysis shows that the important role of illiquidity in the R&D-return relation cannot be explained by factors such as financial constraints, innovation ability, and product market competition. Collectively, our results suggest that stock illiquidity is an independent driver of the R&D premium.

JEL Classification: G12; G14; O32

Keywords: Illiquidity; Research and development investment; Risk premium; Stock returns

1. Introduction

Investment in research and development (R&D) plays a critical role in improving firms' productivity, and thereby promoting a nation's long-term economic growth (Hall, 1996). Over the years, firms in the United States have invested heavily in R&D activities. For example, according to the National Science Foundation's data, privately funded industrial R&D increased to $356 billion in 2015, from $247 billion in 2009. Despite its growing importance and the fact that such investment is often more inflexible than capital investment, still less is known about the economic origins of the R&D premium documented in the literature (see Chan et al., 1990; Chambers et al., 2002; Eberhart et al., 2004; Lev and Sougiannis, 1996). Both classic and new generation asset pricing models find it difficult to rationalize positive premiums associated with R&D investments (see, for example, Daniel et al., 2020; Hou et al., 2015). In addition, though existing studies confirm a significant relation between R&D investment and future stock returns, there is far less consensus on why it exists (see Cohen et al., 2013; Gu, 2016; Li, 2011). Investigating the mechanism that potentially drives the R&D premium is important because it will enhance the financial market's ability to efficiently value investments in R&D (as in Chan et al., 2001; Cohen et al., 2013; Eberhart et al., 2004). It will also complement our understanding of financing, obsolescence, and other risk implications for new or ongoing R&D ventures (as in Berk et al., 2004; Carpenter and Petersen, 2002; Chambers et al., 2002; Gu, 2016; Li, 2011).

This paper contributes to such exploratory investigation by developing a continuous-time partial equilibrium model of an R&D venture, which offers a substantive insight into the R&D-return relation from the perspective of stock illiquidity due to the informational friction channel. For example, the information asymmetries, between the investors and the R&D manager; the uninformed and informed traders, on the likelihood of R&D success. The channel emerges naturally in the context of R&D because of the highly heterogeneous complexity in assessing R&D success (see, among others, Balachandra and Friar, 1997; Hannon et al., 2015; Jo and Park, 2019; Pammolli et al., 2011). Because of the relative uniqueness of the R&D process, the extent of information asymmetry associated with an R&D venture is also larger than that associated with tangible (for example, property, plant, and equipment) and financial investments (Aboody and Lev, 2000). Surprisingly, to the best of our knowledge, this channel approximated by the level of stock illiquidity has not been studied in conjuncture with the significantly positive R&D-return relation documented in previous literature. This is our main focus in the paper.

Building on the dynamic model of Berk et al. (2004) for a single multistage R&D venture, we bring the informational friction about the likelihood of R&D success into the valuation framework through the parameter uncertainty effect similar to that in Pástor and Veronesi (2003). We show that information friction lowers the value of a firm's R&D project and therefore increases its expected excess returns in equilibrium. Based on the market structure literature (see, among others, Copeland and Galai, 1983; Glosten and Milgrom, 1985; Stoll, 2000), our model further links the informational friction to the stock illiquidity of firms. Specifically, a firm's R&D value is a function of the probability of R&D success (R&D success intensity). The true value of the success intensity parameter is unknown to the market makers and uninformed investors who have to infer the value from limited information available to them. When information used for inferring the success intensity is sparse (for example, new technology with few prior studies), the true value of the parameter is more likely to be far from the estimate, which translates into higher losses for the market makers trading with informed traders. Therefore, the higher the uncertainty, the higher the bid-ask spreads set by the market makers in attempts to avoid losses, i.e., higher illiquidity. The parameter uncertainty is positively linked to the risk premium of the R&D project and has an amplifying effect on the exposure to the systematic risk. We also show that this amplifying effect is even stronger when there exists uncertainty aversion (see, for example, Camerer and Weber, 1992). In sum, the intuition is that the R&D-return relation is more significant when there is more uncertainty about the project success intensity, and the stock illiquidity is a measure of this uncertainty. To formalize this intuition and to motivate analyses, we develop a new testable hypothesis, based on our model, that the positive relation between a firm's R&D investment and expected stock returns strengthens with the level of illiquidity.

We test our hypothesis both by Fama and MacBeth (1973) cross-sectional regressions and by independently double sorting stocks on illiquidity and R&D intensity. The cross-sectional regressions, which also control for the potential effects of size, book-to-market equity, short-term reversal, momentum, asset growth, profitability, and idiosyncratic volatility, confirm that the positive R&D-return relation is much stronger among illiquid stocks. For example, when R&D-to-market (i.e., R&D expenditure scaled by market value of equity) and bid-ask spread are used, respectively, as measures of a firm's R&D intensity and stock illiquidity, the average coefficient estimates on R&D-to-market are 5.43 (t-statistic = 5.28) and 0.97 (t-statistic = 1.16), respectively, for high and low illiquid stocks. Importantly, the average spread between the estimated slope coefficients on R&D-to-market for high and low illiquid stocks is 4.46, which is statistically significant with a t-statistic of 4.55. The results remain qualitatively similar, when we conduct regression analyses using: the all-but-microcaps sample, which excludes stocks with a market value of equity below the 20th percentile of the NYSE market capitalization distribution; industry dummies as additional control variables; subperiods; stock returns adjusted for delistings, following Shumway (1997); samples partitioned into high-tech firms and low-tech firms; alternative versions of the bid-ask spread; other R&D intensity measures such as R&D expenditure scaled by total assets, R&D expenditure scaled by sales, R&D expenditure scaled by capital expenditure, and R&D capital scaled by total assets; and the Amihud (2002) illiquidity measure. In a separate cross-sectional regression analysis, we also create interaction terms involving illiquidity dummy variables. After controlling for firm-level attributes, such as size, book-to-market equity, short-term reversal, momentum, asset growth, profitability, and idiosyncratic volatility, the results remain robust and support our theoretical model's prediction of a stronger R&D-return relation among illiquid stocks. For example, we find that, all else being equal, a one-standard deviation increase in R&D-to-market ratio is associated with 56 basis points higher future returns per month for high illiquid stocks relative to all other stocks.

The model's prediction is also supported by the results from independent double-sorted portfolio analyses. Similar to the results from the Fama and MacBeth (1973) cross-sectional regressions, the positive R&D-return relation strengthens with the level of stock illiquidity. For example, among stocks with high illiquidity, the high-minus-low R&D-to-market portfolio delivers a value-weighted average monthly return of 1.57% (t-statistic = 4.79). The value-weighted characteristic- and industry-adjusted returns of 1.18% and 1.50% per month on this hedge portfolio are statistically significant, with t-statistics of 4.39 and 5.09, respectively. The R&D premium among illiquid stocks persists even after adjusting for risk using empirical factor pricing models. The value-weighted average monthly abnormal returns on the hedge portfolio relative to the Fama and French (1993) three-factor model, the Fama and French (1993) and Carhart (1997) four-factor model, the Fama and French (2015) five-factor model, and the Hou et al. (2015) q-factor model are 1.59%, 1.36%, 1.47%, and 1.53%, with t-statistics of 4.85, 3.50, 4.65, and 3.80, respectively. In contrast, among stocks with low illiquidity, the high-minus-low R&D-to-market portfolio generates a substantially lower monthly return of 0.08% (t-statistic = 0.51). The value-weighted characteristic-adjusted return and the value-weighted industry-adjusted return on the portfolio are also negligible at 0.03% (t-statistic = 0.21) and 0.09% (t-statistic = 0.69) per month, respectively. Furthermore, the average monthly abnormal returns on the portfolio relative to the four asset pricing models are much lower and statistically indistinguishable from zero at conventional levels. Specifically, the three-factor, the four-factor, the five-factor, and the q-factor model alphas are, respectively, −0.17% (t-statistic = −1.07), −0.07% (t-statistic = −0.46), −0.12% (t-statistic = −0.72), and 0.02% (t-statistic = 0.10). The results, supporting our hypothesis, remain robust to using: the all-but-microcaps sample; alternative measures of R&D intensity and stock illiquidity; double-sorted quartile or quintile portfolios; and portfolios from independent triple sorts on size, illiquidity, and R&D intensity.

Our paper complements and extends a growing body of literature on the relation between R&D investment and stock returns. The theoretical foundation of our model is closely related to the dynamic model by Berk et al. (2004), who show how systematic and technical (idiosyncratic) risks jointly determine the value of a multistage R&D venture. Our paper is also closely related to Copeland and Galai (1983), Glosten and Milgrom (1985), and Stoll (2000), in that, we link the informational friction to the firm's market illiquidity, which independently explains the R&D-return relation in the cross-section. We theoretically illustrate and empirically show that the positive R&D-return relation increases with the level of stock illiquidity.

Some studies demonstrate that financial constraints (Li, 2011), innovation ability (Cohen et al., 2013), and product market competition (Gu, 2016) affect the R&D premium in the cross-section of stocks. In fact, both Li (2011) and Gu (2016) develop real option models, which, respectively, predict that the positive R&D-return relation increases with the financial constraints and the product market competition. Our model differs from theirs in that its mechanism is based on a firm's stock illiquidity emanating from uncertainty faced by the market makers and uninformed investors. Moreover, using triple-sorting portfolio and cross-sectional regression approaches, and proxies for financial constraints (Hadlock and Pierce, 2010; Kaplan and Zingales, 1997), innovation ability (Cohen et al., 2013), and product market competition (Hou and Robinson, 2006; Peress, 2010), we show that the important role of stock illiquidity in the R&D-return relation cannot be explained by these factors. Our paper is also related more generally to earlier studies that document significant relation between stock returns and R&D expenditure (Chan et al., 1990), capital (Lev and Sougiannis, 1996), intensity (Chan et al., 2001), asset (Chambers et al., 2002), growth (Eberhart et al., 2004), information quality (Huang et al., 2022), innovative efficiency (Hirshleifer et al., 2013), and innovative originality (Hirshleifer et al., 2018).

2. The Model

The firm operates in continuous time and has a single R&D project, which requires passing through a sequence of K discrete stages of development successfully. Once the R&D venture is completed, the firm will generate a stream of stochastic cash flows. Building on Berk et al. (2004), the model incorporates investors' uncertainty about the likelihood of R&D success into the valuation framework. We also analytically link the informational friction to the level of stock illiquidity based on a simplified version of the model by Glosten and Milgrom (1985).

2.1 Valuation without uncertainty about R&D success intensity

To begin with, we follow Berk et al. (2004, Section 4) to value an R&D firm for the case of no uncertainty about the success intensity. The firm value at time t depends on the number of successfully completed stages, n, and the future cash flow, c_t, which is modeled as

$$\frac{dc_t}{c_t} = \mu dt + \sigma dw_t, \tag{1}$$

where w_t is a standard Brownian motion. Under a partial equilibrium setting, the pricing kernel, Λ_t, is exogenously given by the process

$$\frac{d\Lambda_t}{\Lambda_t} = -r_f dt + \eta dh_t, \tag{2}$$

where r_f is the constant risk-free interest rate and h_t is a standard Brownian motion. We denote the market price of risk as $\zeta = \sigma\eta\rho$, where ρ is the correlation between the two Brownian motion processes w_t and h_t. Therefore, under the risk-neutral measure, c_t follows

$$\frac{dc_t}{c_t} = \mu^{\mathbb{Q}} dt + \sigma dw_t^{\mathbb{Q}}, \tag{3}$$

where $w_t^{\mathbb{Q}}$ is a standard Brownian motion under the risk-neutral measure $\mathbb{Q}$, and $\mu^{\mathbb{Q}} = \mu - \zeta$. Completion of the R&D project involves passing through a sequence of K discrete stages of development successfully. The R&D expenditure for each stage is given by $a + bc_t$.[1] The success intensity, λ, is a constant. We denote the obsolescence risk-adjusted discount rate as $r = r_f + \phi$, where ϕ is the obsolescence intensity. Given these basic settings, by Proposition 7 in Berk et al. (2004), the value of the R&D project, $V(\lambda, c)$, is

$$V(\lambda, c) = \begin{cases} \sum_{i=n}^{K} F_n^i c^\gamma (\log c)^{i-n} + S_n c + G_n & c \geqslant c_n^* \\ O_n c^\beta & c < c_n^*, \end{cases} \tag{4}$$

where

$$G_n = -a \sum_{i=1}^{K-n} \frac{\lambda^{i-1}}{(\lambda + r)^i}, \tag{5}$$

$$S_n = -b \sum_{i=1}^{K-n} \frac{\lambda^{i-1}}{(\lambda + r - \mu^{\mathbb{Q}})^i} + \frac{1}{r - \mu^{\mathbb{Q}}} \left(\frac{\lambda}{\lambda + r - \mu^{\mathbb{Q}}} \right)^{K-n}, \tag{6}$$

and F_n^i is defined through the following recursive system

$$F_n^i = \begin{cases} 0 & i = K \\ \frac{2\lambda}{(i-n)\psi} F_{n+1}^i + \frac{\sigma^2(i-n+1)}{\psi} F_n^{i+1} & n < i < K \\ \frac{2\sigma^2 (c_n^*)^{-\gamma}}{\psi+\theta} \left[(1-\beta) S_n c_n^* - \beta G_n \right] - & i = n \\ \sum_{j=n+1}^{K-1} F_n^j \left[(\log c_n^*)^{j-n} - \frac{2\sigma^2(j-n)}{\varphi+\theta} (\log c_n^*)^{j-n-1} \right] & \end{cases}; \tag{7}$$

[1] We follow Berk et al. (2004)'s setting for the cost of investment allowing both fixed (a) and variable (bc_t) components. Including a variable component allows for increases (decreases) in scale during the development process that are an involuntary consequence of increased (decreased) anticipated cash flows. All the later numerical results are robust to setting $b = 0$, which is a special case of the general model we present here.

$$O_n = \left[S_n c_n^* + G_n + \sum_{i=n}^{K-1} F_n^i (c_n^*)^\gamma (\log c_n^*)^{i-n} \right] (c_n^*)^{-\beta}; \tag{8}$$

and c_n^* solves the following equation:

$$(\beta - 1)(1 - \gamma) S_n c_n^* - \beta\gamma G_n = \frac{(c_n^*)^\gamma}{2\sigma^2} \sum_{i=n+1}^{K-1} \left[(i-n)(\psi - \theta) F_n^i - 4\lambda F_{n+1}^i \right] (\log c_n^*)^{i-n-1}; \tag{9}$$

and

$$\beta = \frac{(\sigma^2 - 2\mu^{\mathbb{Q}}) + \theta}{\sigma^2}, \qquad \theta = \sqrt{8r\sigma^2 + (\sigma^2 - 2\mu^{\mathbb{Q}})^2},$$
$$\psi = \sqrt{8(r+\lambda)\sigma^2 + (\sigma^2 - 2\mu^{\mathbb{Q}})^2}, \qquad \gamma = \frac{(\sigma^2 - 2\mu^{\mathbb{Q}}) - \psi}{2\sigma^2}.$$

The risk premium, $rp(\lambda, c)$, is given by

$$\frac{\partial V(\lambda, c)/\partial c}{V(\lambda, c)} \zeta c = \begin{cases} \zeta \dfrac{S_n c + c^\gamma \sum_{i=n}^{N-1} \left[\gamma (\log c)^{i-n} + (i-n)(\log c)^{i-n-1} \right] F_n^i}{S_n c + G_n + c^\gamma \sum_{i=n}^{N-1} (\log c)^{i-n} F_n^i} & c \geqslant c_n^* \\ \zeta\beta & c < c_n^*. \end{cases} \tag{10}$$

From Berk et al. (2004, Table 2), we know that $V(\lambda, c)$ increases with c; and from Gu (2016, Fig. 1) and Li (2011, Figure 4), we also know that $rp(\lambda, c)$ increases with the R&D expenditure. This positive R&D premium in Berk et al. (2004)'s model is also verified in Figure IA1 of the Internet Appendix. It is interesting to examine the relation between $V(\lambda, c)$ and λ. Clearly, $V(\lambda, c)$ is an increasing function of λ as the higher is the R&D success intensity the more valuable is the project. This result is summarized in the following corollary.

Corollary 1.

$$\frac{\partial V(\lambda, c)}{\partial \lambda} > 0. \tag{11}$$

Due to the implicit nature in solving F_n^i and c_n^* (see equations (7) and (9)), analytical results of the relation are unavailable. We therefore resort to numerical procedures for illustration. Based on a set of parameters similar to Berk et al. (2004, Table 1), we find that $V(\lambda, c)$ increases with λ and is a concave function of λ. These observations are shown in Figure 1. The concavity implies that increasing dispersion in λ decreases the expected value of $V(\lambda, c)$, which has important implications when we introduce uncertainty about λ into the model.

2.2 Investors' uncertainty about R&D success intensity

In what follows, the investors do not know the success intensity for sure, but the firm's manager has perfect knowledge of the success intensity. The investors are aware of the fact that the firm's manager knows the success intensity; therefore, they agree that conditional on a value of the success intensity the firm's R&D value follows equation (4). The true value of the success intensity is unknown to the market makers and uninformed investors who have to infer the value from limited information available to them. For example, they have to estimate the success intensity from past relevant information available to them. We model the investors' belief of the success intensity as the Cox proportional hazard function due to Cox (1972), which is the most widely used functional form for modeling the relationship of covariates to a survival outcome. The Cox proportional hazard function also has wide-ranging applications in the finance literature. For example, Duffie et al. (2007) employ this functional form to model corporate defaults. To emphasize the firm-specific and time-varying nature of the estimation, we add subscripts of j and t to the variables in this subsection. Specifically, the investors' belief of the success intensity of firm j at time t is written as a simplified proportional-hazards form:

$$\lambda(x_{j,t}, u) = \exp(u x_{j,t}), \tag{12}$$

where $x_{j,t}$ is a firm-specific and time-varying variate determining the R&D success intensity via u, an unknown and project type-specific parameter. $x_{j,t}$ can be considered as the key indicator of R&D success that aggregates various factors affecting R&D production, for example, team social capital, knowledge sharing, team-efficacy, etc. (see Jo and Park, 2019).

The investors have to infer u from past information available to them. The information includes past R&D success or failure records of the same project type,

$$\mathbf{N}(m_{j,t}, T_{j,t}) = \{N_{i,s} : i = 1, \ldots, m_{j,t}; t - T_{j,t} \leqslant s \leqslant t\}, \tag{13}$$

where $N_{i,s}$ is the counting process with intensity $\lambda(x_{i,s}, u)$, which counts the number of stages that the R&D has been successful up to time s; and the past observations of the time-varying variate,

$$\mathbf{X}(m_{j,t}, T_{j,t}) = \{x_{i,s} : i = 1, \ldots, m_{j,t}; t - T_{j,t} \leqslant s \leqslant t\}. \tag{14}$$

Here, $m_{j,t}$ is the number of similar projects and $T_{j,t}$ is the number of years of the data that are available to the investors and relevant for assessing the R&D success intensity of firm j at time t. Intuitively, the smaller $m_{j,t}$ and $T_{j,t}$, the less information available for the R&D assessment at time t for firm j; therefore, the harder to pin down the success intensity accurately. This intuition is formalized in the following proposition.

Proposition 1. *At time t for firm j, given the information of $\mathbf{N}(m_{j,t}, T_{j,t})$ and $\mathbf{X}(m_{j,t}, T_{j,t})$, the investors' belief of u has a mean being the maximum likelihood estimate, $\hat{u}_{j,t}$, which solves*

$$0 = \sum_{i=1}^{m_{j,t}} \int_{t-T_{j,t}}^{t} \left[x_{i,s} - \overline{x}\left(\hat{u}_{j,t}, s\right)\right] dN_{i,s}, \ \textit{where } \overline{x}\left(\hat{u}_{j,t}, s\right) = \sum_{i=1}^{m_{j,t}} \frac{e^{\hat{u}_{j,t} x_{i,s}} x_{i,s}}{\sum_{i=1}^{m_{j,t}} e^{\hat{u}_{j,t} x_{i,s}}}; \tag{15}$$

and a variance, $\sigma^2_{\hat{u}_{j,t}}$, given by

$$\sigma^2_{\hat{u}_{j,t}} = \left[\sum_{i=1}^{m_{j,t}} \int_{t-T_{j,t}}^{t} W(\hat{u}_{j,t}, s) dN_{i,s}\right]^{-1}, \ \textit{where}\ W(\hat{u}_{j,t}, s) = \sum_{i=1}^{m_{j,t}} \frac{e^{\hat{u}_{j,t} x_{i,s}} \left[x_{i,s} - \overline{x}\left(\hat{u}_{j,t}, s\right)\right]^2}{\sum_{i=1}^{m_{j,t}} e^{\hat{u}_{j,t} x_{i,s}}}, \tag{16}$$

where $\hat{u}_{j,t}$ is a consistent estimate of u and $\sigma_{\hat{u}_{j,t}}$ is a decreasing function of $m_{j,t}$ and $T_{j,t}$.

Proof. Equations (15) and (16) are direct results from Section 3.1 of Therneau and Grambsch (2000). Since $W(\hat{u}_{j,t}, s)$ is nonnegative for all s, $\sigma_{\hat{u}_{j,t}}$ is a decreasing function of $m_{j,t}$ and $T_{j,t}$. ■

At any time t, the investors' beliefs about u are summarized by a probability density function, $f(\hat{u}_{j,t}, \sigma^2_{\hat{u}_{j,t}})$, and investors update their beliefs about u over time as the available information changes. It is worth emphasizing that although we assume that the true parameter, u, is project type-specific and not time-varying, the firm-specific and limited information available and relevant for the estimation makes u's estimate, $\hat{u}_{j,t}$, firm-specific and time-varying. Therefore, the uncertainty is typically idiosyncratic and does not go away with t goes large since $m_{j,t}$ and $T_{j,t}$ are constantly changing and limited information availability. To illustrate the effects of $T_{j,t}$ and $m_{j,t}$ on the uncertainty about R&D success intensity, we assume $x_{j,t}$ to be independent and identically distributed and uniformly distributed within zero and one, simulate $\mathbf{N}$ and $\mathbf{X}$ up to $m_{j,t} = 20$ and $T_{j,t} = 6$, and estimate $\sigma_{\hat{u}_{j,t}}$ using part of the simulated sample. The estimates of $\sigma_{\hat{u}_{j,t}}$ are plotted in Figure 2. It can be seen that the uncertainty measured by $\sigma_{\hat{u}_{j,t}}$ clearly increases with the reducing sample size. For example, the uncertainty in a sample with two years ($T_{j,t} = 2$) of five similar projects' ($m_{j,t} = 5$) data is about five times of that in a sample with six years ($T_{j,t} = 6$) of 20 similar projects' ($m_{j,t} = 20$) data.

2.3 Valuation effect of investors' uncertainty

Since in what follows all variables are for firm j at time t, for notational convenience, we drop the subscripts of j and t. Given a value of u, the investors agree on the functional form of $V(\lambda(u), c)$, i.e., equation (4) that links the value of the R&D project to u as they believe the firm's manager has perfect information about the success intensity. However, the true value of u is unknown to the investors. Therefore, by the law of iterated expectations, the R&D project's value under uncertainty is the average of $V(\lambda(\bar{u}), c)$ for all possible values of $\bar{u}$, weighted by the estimated probabilities assigned to each $\bar{u}$.[2] This is similar to the notion in Pástor and Veronesi (2003) for valuing stocks whose average future profitability is unknown. Given Proposition 1, we have the R&D project's value under the investors' uncertainty about the R&D success intensity as

$$\widehat{V}(\hat{u}, \sigma_{\hat{u}}, c) = \int_{-\infty}^{\infty} V(\lambda(\bar{u}), c) f_{(\hat{u}, \sigma^2_{\hat{u}})}(\bar{u}) d\bar{u}, \tag{17}$$

[2] This can also be interpreted as a certainty-equivalence aggregating different investors' risk assessments. We thank an anonymous referee for pointing this out.

where $f_{(\hat{u},\sigma_{\hat{u}}^2)}$ is a probability density function with mean $\hat{u}$ and variance $\sigma_{\hat{u}}^2$. Here, the source of uncertainty is assumed to be idiosyncratic and therefore the probability density function remains identical under both physical and risk-neutral measures. Given the concavity in Figure 1, it is straightforward to see by Jensen's inequality that, all else being equal, $\widehat{V}(\hat{u}, \sigma_{\hat{u}}, c)$ decreases with $\sigma_{\hat{u}}$. We summarize this in Corollary 2 below and illustrate in Figure 3. With investors' uncertainty about the R&D success intensity, the project is valued less relative to the without uncertainty scenario, and therefore earns a higher risk premium. In other words, the uncertainty about the R&D success intensity makes the project more sensitive to systematic risk and have higher expected returns in equilibrium.[3]

Corollary 2.

$$\frac{\partial \widehat{V}(\hat{u}, \sigma_{\hat{u}}, c)}{\partial \sigma_{\hat{u}}} < 0. \tag{18}$$

2.4 Illiquidity effect of uncertainty

In addition to the effects on the risk premium, the uncertainty about the R&D success intensity also affects the level of stock liquidity of the R&D firms through an informational channel. Copeland and Galai (1983) and Glosten and Milgrom (1985) show theoretically that informed trading due to information asymmetry increases the bid-ask spreads in equilibrium as the market makers adjust the spreads to protect themselves from losses. Stoll (2000) shows empirically that the informational friction component exists in the observed bid-ask spreads. As the market makers also have knowledge of $\sigma_{\hat{u}}$, they know that when $\sigma_{\hat{u}}$ is large the true value of u is more likely to be far from the estimate, $\hat{u}$. This means that the insiders' information becomes better (i.e., finer) with higher uncertainty about the R&D success intensity measured by $\sigma_{\hat{u}}$, which in turn increases the bid-ask spreads (Glosten and Milgrom, 1985, Proposition 5).

We follow Glosten and Milgrom (1985, Section 3) and develop a parsimonious model of bid-ask spread that explicitly links stock illiquidity to $\sigma_{\hat{u}}$. For simplicity, we assume under the belief of uninformed traders and market makers, u only takes two values, $\hat{u} - \sigma_{\hat{u}}$ and $\hat{u} + \sigma_{\hat{u}}$, with equal probability. The mean and standard deviation of this distribution match those in Proposition 1. Therefore, the expected value of the R&D firm is

$$\widehat{V} = 0.5[V(\lambda(\hat{u} - \sigma_{\hat{u}}), c) + V(\lambda(\hat{u} + \sigma_{\hat{u}}), c)]. \tag{19}$$

[3] In a recent work, Huang et al. (2022) empirically verify a negative relation between expected stock returns and R&D information quality (RDIQ), conjecturing an underpricing mechanism similar to our theoretical results here. The RDIQ is somewhat related to our $\sigma_{\hat{u}}$; therefore, to certain extent their empirical results verify our theoretical prediction. However, the RDIQ is measured using only R-squares from time-series regressions of firm's sales growth on its realized R&D capital. Lack of cross-sectional analysis (see Section 2.2 for a cross-sectional analysis) makes RDIQ a less relevant measure than stock illiquidity as far as the informational friction and the uncertainty are concerned, which are the focus of our study.

Consistent with Glosten and Milgrom (1985), we further assume that the market makers have zero expected profit in equilibrium due to competition and the probability of an uninformed buyer (seller) accepting the ask (bid) price is 50%.[4] Denote the proportion of informed traders in the trading population as α. Then, the bid and ask prices can be solved from equations of zero expected profit for the market makers. Specifically, the expected profit from the ask-side given the ask price, A, is

$$0.5(1-\alpha)(A-\widehat{V})-0.5\alpha[V(\lambda(\hat{u}+\sigma_{\hat{u}}),c)-A]=0, \tag{20}$$

where the first term on the left-hand-side is the expected income from trading with uninformed buyers, while the second term is the expected loss from trading with informed buyers. Similarly, the expected profit from the bid-side given the bid price, B, is

$$0.5(1-\alpha)(\widehat{V}-B)-0.5\alpha[B-V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)]=0. \tag{21}$$

It can be shown that the relative bid-ask spread, defined as the ratio of the difference between the ask and bid prices to the value of the R&D firm, $(A-B)/\widehat{V}$, increases with $\sigma_{\hat{u}}$. This result is formalized in the following proposition and illustrated in Figure 4.

Proposition 2. *Given $\hat{u}$, $\sigma_{\hat{u}}$, c, and α, the ask price, A, and the bid price, B, are given by*

$$A=\alpha V(\lambda(\hat{u}+\sigma_{\hat{u}}),c)+(1-\alpha)\widehat{V}, \tag{22}$$
$$B=\alpha V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)+(1-\alpha)\widehat{V}, \tag{23}$$

respectively. The relative bid-ask spread is

$$\textit{bid-ask spread}=\frac{A-B}{\widehat{V}}=\alpha\frac{V(\lambda(\hat{u}+\sigma_{\hat{u}}),c)-V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)}{\widehat{V}}, \tag{24}$$

and increases with $\sigma_{\hat{u}}$, that is,

$$\frac{\partial \textit{bid-ask spread}}{\partial\sigma_{\hat{u}}}>0. \tag{25}$$

Proof. See Appendix A. ■

To include spreads due to factors unrelated to uncertainty, for example, real friction cost of providing liquidity, we add a fixed spread κ to equation (24), that is,

$$\text{bid-ask spread}=\kappa+\alpha\frac{V(\lambda(\hat{u}+\sigma_{\hat{u}}),c)-V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)}{\widehat{V}}. \tag{26}$$

[4] This is precisely the case of perfectly inelastic liquidity supply and demand in Glosten and Milgrom (1985, Section 3).

Given reasonable parameter values, the model produces realistic bid-ask spreads comparable to the empirical counterpart. Since both the bid-ask spread and the firm's R&D risk premium are related to $\sigma_{\hat{u}}$, we can establish their theoretical relations for the later empirical testing.

2.5 Implications for the R&D-return relation

Given equation (19), the firm's risk premium with uncertainty about λ can be derived as

$$\widehat{rp}(\hat{u}, \sigma_{\hat{u}}, c) = \frac{\partial \widehat{V}(\hat{u}, \sigma_{\hat{u}}, c)/\partial c}{\widehat{V}(\hat{u}, \sigma_{\hat{u}}, c)} \zeta c. \tag{27}$$

As mentioned earlier, $rp(\lambda, c)$ increases with the R&D expenditure. With uncertainty, this is still true, that is, $\widehat{rp}(\hat{u}, \sigma_{\hat{u}}, c)$ increases with the R&D expenditure as well. To quantify the R&D-return relation, we first define an R&D intensity measure, RDME, as

$$\text{RDME} = \frac{a + bc}{\widehat{V}}, \tag{28}$$

where "ME" in RDME stands for firm's market value of equity, which equals $\widehat{V}$ in our case as the firm considered here is unlevered. With high and low RDME, the relation between the R&D premium and RDME is measured as a slope coefficient:

$$\text{slope}(\sigma_{\hat{u}}) = \frac{\widehat{rp}_{\text{RDME}_{\text{high}}} - \widehat{rp}_{\text{RDME}_{\text{low}}}}{\text{RDME}_{\text{high}} - \text{RDME}_{\text{low}}}. \tag{29}$$

In the previous subsection, we mentioned that the uncertainty about the R&D success intensity makes the project more sensitive to systematic risk and have higher expected returns in equilibrium. We now show that the uncertainty, $\sigma_{\hat{u}}$, positively correlated with the bid-ask spread($\sigma_{\hat{u}}$) has an amplifying effect on the relation between firm's R&D risk premium and RDME.[5] That is, slope($\sigma_{\hat{u}}$) is an increasing function of bid-ask spread ($\sigma_{\hat{u}}$) via $\sigma_{\hat{u}}$. The results are shown in Figure IA1 of the Internet Appendix (see there for a discussion of the results).

So far, we have assumed that uninformed investors are not uncertainty averse (see Camerer and Weber, 1992, for a review). In our context, it is natural to relax this assumption and allow the uninformed investors to have uncertainty aversion. We extend our analysis to incorporate uncertainty aversion using the "Source Function" approach of Abdellaoui et al. (2011) in the Internet Appendix. The intuition is that the uncertainty averse investors assign low weights for good outcomes, which enhance risk aversion or pessimism. Due to the results in Corollaries 1 and 2, the uncertainty aversion on the uncertainty about the success intensity will further lowers the value of an R&D project due to higher weights on lower success intensity, therefore, intensify

[5] To avoid confusion, the empirical counterpart of RDME in subsequent sections is denoted in *italic*.

the uncertainty effect we show here. Indeed, in the Internet Appendix, we show that the level of stock illiquidity has even stronger amplifying effects on the R&D-return relation when uncertainty aversion is allowed. Given these insights, we develop the following testable hypothesis.

Hypothesis: The positive R&D-return relation strengthens with firm's stock illiquidity.

3. Data and summary statistics

3.1 Data and measures of R&D intensity and illiquidity

We obtain data on daily stock prices (ask and bid) and returns, monthly stock returns, shares outstanding, and trading volume from the Center for Research in Security Prices (CRSP) and a host of accounting information from the Compustat Annual Industrial Files. Our preliminary sample includes all NYSE-, AMEX-, and NASDAQ-listed ordinary common stocks (CRSP share code of 10 or 11). From this preliminary sample we exclude financial firms, which have four-digit standard industrial classification codes between 6000 and 6999, and firms with a nonpositive book value of equity. Our sample period starts in January 1975 and ends in December 2016, since the accounting treatment of R&D expenses reporting was standardized in 1975. Table IA1 in the Internet Appendix reports summary statistics for this preliminary sample.

To perform all of the empirical analyses, we construct two working samples, namely, the full sample and the all-but-microcaps sample, from the intersection of the Compustat database and the preliminary sample considered in Table IA1. Specifically, our full sample consists of all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year and a valid illiquidity estimate in a given month are obtainable. In the all-but-microcaps sample, we exclude stocks with a market value of equity below the 20th percentile of the NYSE market capitalization distribution and retain the remaining NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks with both nonmissing R&D intensity and valid illiquidity estimates. Microcap firms tend to experience a higher level of stock illiquidity and to have a higher R&D intensity. Using the all-but-microcaps sample therefore helps mitigate the possible undue influence of microcaps on the results obtained for the full sample. In line with Fama and French (1992), we further employ all accounting variables at the end of June of calendar year t by using accounting information available for the fiscal year ending in the calendar year $t-1$ from the Compustat database. This ensures that the accounting information is already incorporated into firms' stock returns.

Consistent with the literature (see, among others, Chan et al., 2001; Gu, 2016; Li, 2011), we mainly use two alternative measures of a firm's R&D intensity. The first is R&D expenditure scaled by market value of equity, denoted $RDME$, while the second is R&D expenditure scaled by total assets, denoted $RDAT$. More specifically, $RDME$ is computed at the end of June of year t as the ratio of the R&D expenditure (Compustat item XRD) at the end of the fiscal year ending in the calendar year $t-1$ to the market value of equity at the end of June of year t. On the other hand, $RDAT$ is computed in June of year t as the ratio of R&D expenditure to total assets (Compustat item AT) at the end of the fiscal year ending in the calendar year $t-1$. Previous studies (see Chan et al., 2001; Li, 2011) also document the existence of positive R&D premium using other R&D intensity measures such as R&D expenditure scaled by sales, R&D expenditure scaled by capital expenditure, and R&D capital scaled by total assets. Table IA32 in the Internet Appendix shows that our findings are robust to these alternative measures.

The stock illiquidity measure that we primarily adopt in this paper is the average bid-ask spread (denoted BAS), computed using one month's worth of daily data with a minimum of 15 valid daily observations. As an alternative to BAS, we further adopt the Amihud (2002) illiquidity measure, denoted $ILLIQ$.[6] Both the bid-ask spread and the Amihud (2002) illiquidity measure are extensively used in the asset pricing literature to compute stock illiquidity (see, among others, Amihud and Mendelson, 1989; Amihud et al., 2015; Han and Lesmond, 2011).

To control for firm attributes that may be related to future stock returns, we construct a number of additional variables. These include past stock returns, book-to-market equity, firm size, asset growth, profitability, and idiosyncratic volatility. In line with Fama and French (1992, 1993), we compute the book-to-market (BM) ratio at the end of June of year t as the ratio of the book value of equity at the end of the fiscal year ending in the calendar year $t-1$ to the market value of equity at the end of December of the calendar year $t-1$. The market value of equity (ME) is calculated as absolute price per share times number of equity shares outstanding at the end of June of each year t. Asset growth (AG) is computed in June of each year t as the annual growth rate of total assets. Profitability (denoted OP) is the operating profitability computed in June of each year t as total revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense, all scaled by book value of equity at the end of the fiscal year ending in the calendar year $t-1$. Idiosyncratic volatility ($IVOL$) is the computed at the end of each month τ as the standard deviation of the residuals from regressing daily stock returns (in excess of the risk-free rate) on the Fama and French (1993) three factor returns over the month τ. We require a minimum of 15 valid daily return observations to estimate $IVOL$.

We also compute the average abnormal returns on portfolios (described in Sections 4.3 and 5) relative to several prominent factor models. These are the Fama and French (1993) three-factor (FF3) model, the Fama and French (1993) and Carhart (1997) four-factor (FFC) model, the Fama and French (2015) five-factor (FF5) model, and the Hou et al. (2015) q-factor (HXZ) model.

[6] See Appendix B for details on BAS and $ILLIQ$.

For this purpose, time-series data on the pricing factors (i.e., market, size, value, momentum, profitability, and investment) of the FF3, FFC, and FF5 models, and monthly risk-free returns are retrieved from the Internet Data Library maintained by Kenneth French. The time-series data on the HXZ model factors are sourced from Lu Zhang's website.[7]

3.2 Summary statistics

Our empirical analyses begin by presenting summary statistics for the full sample.[8] Panel A of Table 1 reports the time-series averages of the cross-sectional summary statistics for our chosen R&D intensity and illiquidity measures. As can be seen, in the average month, the mean value of BAS is 0.06 and the median value of BAS is 0.04. The average cross-sectional standard deviation of BAS is 0.06. In the average month, the maximum value of 0.67 for BAS is more than 11 standard deviations above the mean. Analyzing the R&D intensity measure $RDME$, we find that, in the average year, the mean and median values are 0.07 and 0.03, respectively. The average cross-sectional standard deviation of $RDME$ is 0.14. The maximum value of $RDME$ in the average year of 2.88 is more than 20 standard deviations above the mean. There are 2037 stocks with nonmissing values of both BAS and $RDME$ in the average month.

We report equal-weighted average values of characteristics, including BAS, ME, BM, $R_{-1,0}$, $R_{-12,-2}$, AG, OP, and $IVOL$ for portfolios sorted on R&D intensity in Panel B of Table 1. At the end of June of each year t, stocks are sorted into three portfolios using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of $RDME$. We keep the stocks in the assigned portfolios for the subsequent 12 months. This procedure is repeated at the end of June of each year t. It is noticeable that stocks in the High $RDME$ portfolio tend to be the most illiquid, followed by stocks in the Low $RDME$ portfolio. This suggests a nonlinear relation between R&D intensity and stock illiquidity. The BM ratios also display a nonlinear relation with R&D intensity, and the pattern is similar to that for illiquidity. In contrast, stocks in the High $RDME$ portfolio tend to be the smallest, with an average market capitalization of $662.93 million. Past one-month return increases monotonically with $RDME$, whereas AG and OP decrease monotonically. Furthermore, past one-year return displays a nearly decreasing pattern, while $IVOL$ exhibits a nearly increasing pattern.

Panel C of Table 1 reports the value-weighted and equal-weighted average monthly excess returns for the portfolios sorted on R&D intensity. The idea behind this empirical exercise is to investigate whether conventional univariate sorting on R&D intensity alone yields a pattern in average excess returns. We group stocks into three portfolios using the breakpoints set to the bottom 30%, middle 40%, and top 30% of the ranked values of $RDME$ at the end of June of

[7] See https://sites.google.com/site/theqfactormodel/?pli=1.

[8] Table IA2 in the Internet Appendix provides summary statistics for the all-but-microcaps sample.

each year t. Analyzing the portfolios, we observe that the value-weighted average monthly excess portfolio returns increase monotonically, from 0.55% for the Low $RDME$ portfolio to 1.03% for the High $RDME$ portfolio. A similar monotonically increasing pattern in excess return arises for the equal-weighted portfolios. The value-weighted (equal-weighted) average return of 0.48% (0.93%) per month for the high-minus-low $RDME$ portfolio is statistically significant, with a corresponding Newey and West (1987)-adjusted t-statistic of 2.61 (4.88).

In Panel D of Table 1, we repeat the univariate portfolio analyses of Panel C but with stocks sorted into quintile portfolios. Both the value-weighted and equal-weighted average excess returns on R&D intensity-sorted portfolios increase nearly monotonically with $RDME$ (the exception is portfolio 2). Moreover, the high-minus-low $RDME$ portfolio generates an equal-weighted average return of 1.04% per month, which is both economically large and statistically significant (t-statistic = 4.47). Although the value-weighted average monthly return of 0.31% for the high-minus-low $RDME$ portfolio is statistically indistinguishable from zero at conventional levels (t-statistic = 1.41), the portfolio generates an average return of 2.22% per month (t-statistic = 3.99) among illiquid stocks (see Table IA22 in the Internet Appendix).

4. Empirical results[9]

4.1 Cross-sectional regressions

At the end of each month τ, stocks are sorted into three illiquidity groups using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of BAS in month τ. Then, in each month from July of year t to June of year t+1, we run the Fama and MacBeth (1973) cross-sectional regressions in the following form for the Low (or, equivalently, liquid) and High (or, equivalently, illiquid) groups separately:

$$\begin{aligned} R_{\tau+1} = {} & \delta_{0,\tau} + \delta_{1,\tau} R_{-1,0_\tau} + \delta_{2,\tau} R_{-12,-2_\tau} + \delta_{3,\tau} \ln(BM)_\tau + \delta_{4,\tau} \ln(ME)_\tau \\ & + \delta_{5,\tau} AG_\tau + \delta_{6,\tau} OP_\tau + \delta_{7,\tau} IVOL_\tau + \delta_{8,\tau} RDME_\tau + \epsilon_{\tau+1}, \end{aligned} \tag{30}$$

where R is the monthly return on an individual stock, $R_{-1,0}$ is the prior one-month return, $R_{-12,-2}$ is the prior one-year return skipping the last month, $\ln(BM)$ is the natural logarithm of book-to-market ratio, and $\ln(ME)$ is the natural logarithm of market value of equity. To combat the possible undue influence of outlier observations on the results, we winsorize all independent variables at the 1% and 99% levels on a monthly basis prior to running the cross-

[9] We report empirical results based on the $\{RDME, BAS\}$ combination of R&D intensity and illiquidity measures throughout this paper. The results obtained by using alternative R&D intensity and illiquidity measure combinations, such as $\{RDME, ILLIQ\}$ and $\{RDAT, BAS\}$, are qualitatively similar to those presented in this paper. To conserve space, these additional results are reported in the Internet Appendix Tables IA15 to IA20.

sectional regressions. Moreover, we compute the time-series averages of the differences in slope and intercept coefficient estimates between the High and Low illiquidity groups.

Table 2 reports the time-series average slope and intercept coefficient estimates from the cross-sectional regressions. After controlling for the effects of size, book-to-market, short-term reversal, momentum, asset growth, profitability, and idiosyncratic volatility, the average coefficient estimate on R&D intensity appears to be much larger for illiquid stocks (High group) than for liquid stocks (Low group), both in magnitude and in statistical significance. Specifically, in Panel A, which utilizes the full sample to construct the three illiquidity groups, we see that the estimated average coefficient on $RDME$ for stocks in the High illiquidity group is 5.43 with a t-statistic of 5.28. In contrast, the estimated average coefficient of 0.97 on $RDME$ for stocks in the Low illiquidity group is statistically indistinguishable from zero (t-statistic = 1.16). More importantly, the average spread between the estimated slope coefficients on $RDME$ for stocks in the High and Low illiquidity groups is 4.46, with a corresponding t-statistic of 4.55. We observe qualitatively similar results in Panel B, which utilizes the all-but-microcaps sample. These results also suggest that the presence of microcap stocks in the full sample are unlikely to drive the positive and significant R&D-return relation observed among illiquid stocks in Panel A.[10]

Our empirical results remain very robust to an extended sample period that begins in January 1963 and ends in December 2016.[11] The estimated average slope spreads on $RDME$ between the High and Low illiquidity groups are 3.47 (t-statistic = 4.20) and 4.30 (t-statistic = 2.74) for the full sample and the all-but-microcaps sample, respectively. We present these robustness results in the Internet Appendix Table IA4. When the monthly stock returns during the period from January 1975 through December 2016 are adjusted following Shumway (1997) for delistings, the results reported in Table IA5 of the Internet Appendix are very similar to those in Table 2. We also conduct cross-sectional regression analyses for subperiods. Our motivation for this empirical investigation is the reduction in the NYSE and AMEX tick size from \$1/16th to \$1/100th in January 2001. The NASDAQ completed its implementation of decimalization shortly after, in April 2001. It is possible that the results in Table 2 are mainly driven by observations from the predecimalization period, which is largely characterized by a higher level of stock illiquidity (see, for example, Chordia et al., 2008, 2011). We therefore skip observations for 2001 to avoid confounding effects (if any) and split the sample period into two subperiods: 1975–2000 and 2002–2016. Tables IA6 and IA7 in the Internet Appendix provide strong empirical evidence that a significantly positive R&D-return relation among illiquid stocks holds for both subperiods.

[10] Similar empirical results are obtained when we include industry dummies as additional controls in the Fama and MacBeth (1973) cross-sectional regressions; these are reported in the Internet Appendix Table IA3.

[11] We follow Gu (2016), who examines the positive R&D-return relation in competitive and noncompetitive industries by using R&D expenditure data from 1963 to 2013. But most studies (see, for example, Cohen et al., 2013; Li, 2011) use data from or after 1975 due to accounting standardization of R&D expenses reporting.

Furthermore, we investigate whether our empirical findings differ across high-tech and low-tech industries. The idea behind this examination follows from Chan et al. (1990), who show that high-tech firms experience higher abnormal returns on average than low-tech firms following an announcement of an increase in R&D spending. To rule out the possible impact of industry types on our empirical results, we proceed by partitioning both the full sample and the all-but-microcaps sample into high-tech firms and low-tech firms, based on the Fama and French 12-industry group classification system. Analogous to Table 2, we then conduct the monthly Fama and MacBeth (1973) cross-sectional regression analyses using each of these subsamples. Table IA8 in the Internet Appendix shows evidence of a significantly positive R&D-return relation among illiquid stocks in both high-tech and low-tech industries. This suggests that the findings in Table 2 are unlikely to be driven by firms in a particular industry.

4.2 Cross-sectional regressions with illiquidity dummies

To be consistent with the cross-sectional analyses in Section 4.1, two illiquidity dummy variables are created by sorting stocks into three groups based on the breakpoints set to the 30th and 70th percentiles of the ranked values of BAS in month τ. Then, in each month from July of year t to June of year $t+1$, we estimate the following cross-sectional regressions:

$$R_{\tau+1} = \delta_{0,\tau} + \delta'_{\tau}\mathcal{C}_{\tau} + \epsilon_{\tau+1}, \tag{31}$$

$$R_{\tau+1} = \delta_{0,\tau} + \delta'_{\tau}\mathcal{C}_{\tau} + \delta_{9,\tau}RDME_{\tau} * BAS_{L_\tau} + \delta_{10,\tau}BAS_{L_\tau} + \epsilon_{\tau+1}, \tag{32}$$

$$R_{\tau+1} = \delta_{0,\tau} + \delta'_{\tau}\mathcal{C}_{\tau} + \delta_{9,\tau}RDME_{\tau} * BAS_{H_\tau} + \delta_{10,\tau}BAS_{H_\tau} + \epsilon_{\tau+1}, \tag{33}$$

$$\begin{aligned} R_{\tau+1} = \delta_{0,\tau} &+ \delta'_{\tau}\mathcal{C}_{\tau} + \delta_{9,\tau}RDME_{\tau} * BAS_{H_\tau} + \delta_{10,\tau}RDME_{\tau} * BAS_{L_\tau} \\ &+ \delta_{11,\tau}BAS_{H_\tau} + \delta_{12,\tau}BAS_{L_\tau} + \epsilon_{\tau+1}, \end{aligned} \tag{34}$$

where $\mathcal{C}_{\tau}$ is a vector of explanatory variables including $R_{-1,0}$, $R_{-12,-2}$, $\ln(BM)$, $\ln(ME)$, AG, OP, $IVOL$, and $RDME$; BAS_H (BAS_L) is a dummy variable that is equal to one for stocks in the High (Low) illiquidity group in a given month τ and is zero otherwise. We also winsorize all independent variables (except the dummies) at the 1% and 99% levels on a month-by-month basis prior to running the regressions.

The time-series averages of the cross-sectional regression estimates in Table 3 confirm that the positive R&D-return relation strengthens with the level of stock illiquidity. As can be seen from the baseline regressions (equation (31)) in Panel A, using the full sample, R&D intensity has a significantly positive relation with future stock returns. The average coefficient estimate on $RDME$ is 4.21, with a corresponding t-statistic of 5.19. When considering the regressions specified by equation (32), we see that the average coefficient estimate on the interaction variable

$RDME * BAS_L$ is -4.38 (t-statistic $= -5.18$). The sum of the average coefficient estimates on $RDME$ and $RDME * BAS_L$ is 0.39 (t-statistic $= 0.44$). This suggests that R&D intensity has no significantly positive effect on future returns for stocks in the Low illiquidity group.

Turning now to the cross-sectional regressions specified by equation (33), we find that the estimated average coefficient on the variable of interest, $RDME * BAS_H$, is 3.98, which is both economically substantial and statistically significant (t-statistic $= 4.45$). To get an idea of the magnitude of this result, we multiply the average coefficient estimate by the standard deviation of $RDME$ in the average month. All else being equal, a one-standard deviation increase in $RDME$ (=0.14) is associated with a future return per month for stocks in the High illiquidity group that is 56 basis points higher than that for all stocks in the other two groups. The sum of the average coefficients on $RDME$ and $RDME * BAS_H$, which is the slope estimate of $RDME$ for stocks in the High illiquidity group, turns out to be 6.10 and statistically significant (t-statistic $= 5.84$). We find, for the cross-sectional regressions in equation (34), that the average coefficient estimates on the interaction variables $RDME * BAS_L$ and $RDME * BAS_H$ are, respectively, -2.55 (t-statistic $= -3.23$) and 3.30 (t-statistic $= 3.49$). Panel B of Table 3 reports the results based on the all-but-microcaps sample, which are qualitatively similar to those obtained using the full sample. Taken together, the results in Panels A and B show that, after controlling for firm-specific characteristics, such as size, book-to-market equity, short-term reversal, momentum, asset growth, profitability, and idiosyncratic volatility, the power of R&D intensity to predict future returns is stronger in the cross-section of illiquid stocks.

4.3 Portfolio-level analysis

At the end of each month τ, we sort stocks into three groups based on the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of illiquidity measured by BAS in month τ. Independently, at the end of June of each year t, we also sort stocks into three groups based on the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of $RDME$ computed at the end of June of year t. The stocks are held in the assigned R&D groups for the subsequent 12 months. The monthly intersections of the illiquidity and R&D intensity groups result in nine portfolios.[12] We compute the value-weighted average excess returns for portfolios based on the full sample, while the equal-weighted average excess returns for portfolios are computed using the all-but-microcaps sample. To ensure that our results from both the full sample and the all-but-microcaps sample are robust to firm characteristics and industry effects, we further compute characteristic-and industry-adjusted returns of portfolios. Following the exact procedure in Daniel et al. (1997),

[12] We also perform 4×4 and 5×5 independent-sort portfolio analyses. Internet Appendix Tables IA21 and IA22 show that the results are much stronger than those reported in this paper.

characteristic-adjusted returns are computed as the difference between individual stocks' returns and 125 ($5 \times 5 \times 5$) size/book-to-market/momentum benchmark portfolio returns. The industry-adjusted returns are calculated as the difference between individual stocks' returns and the returns of stocks in the same industry based on the Fama and French 12 industry classifications.

In Table 4, we report average monthly returns in excess of one-month Treasury bill rate (excess returns), characteristic-adjusted returns, industry-adjusted returns, and alphas (i.e., abnormal returns relative to the asset pricing models) on R&D intensity portfolios in the Low illiquidity and the High illiquidity groups. The results show that the positive R&D-return relation exists only among illiquid stocks. In Panel A, which utilizes stocks in the full sample, we find that the value-weighted average monthly return of 1.57% for the high-minus-low $RDME$ portfolio in the High illiquidity group is statistically significant (t-statistic = 4.79). The value-weighted characteristic-adjusted (industry-adjusted) return of this hedge portfolio turns out to be 1.18% (1.50%) per month, with a corresponding t-statistic of 4.39 (5.09). Moreover, the value-weighted average excess returns, the characteristic-adjusted returns, and the industry-adjusted returns of $RDME$ portfolios increase monotonically when moving from the Low R&D-intensity portfolio, $RDME_L$, to the High R&D-intensity portfolio, $RDME_H$. A monotonically increasing pattern can also be seen for the monthly average abnormal returns relative to the factor pricing models, which once again shows the existence of a positive R&D-return relation among illiquid stocks. The FF3, FFC, FF5, and HXZ model alphas on the high-minus-low $RDME$ portfolio are quite substantial and statistically significant; they are 1.59% (t-statistic = 4.85), 1.36% (t-statistic = 3.50), 1.47% (t-statistic = 4.65), and 1.53% (t-statistic = 3.80), respectively.

However, a completely different picture emerges for R&D-intensity portfolios in the Low illiquidity group. For example, the value-weighted average return is only 0.08% per month (t-statistic = 0.51) for the high-minus-low $RDME$ portfolio. The value-weighted characteristic-adjusted return and the value-weighted industry-adjusted return on the portfolio are also economically very small at 0.03% (t-statistic = 0.21) and 0.09% (t-statistic = 0.69) per month, respectively. Neither the average excess returns nor the average abnormal returns on portfolios displays a monotonically increasing pattern similar to that of the High illiquidity group. For the high-minus-low $RDME$ portfolio, the value-weighted average abnormal returns relative to the asset pricing models are all statistically insignificant. In particular, the FF3, FFC, FF5, and HXZ model alphas for this portfolio are, respectively, -0.17% (t-statistic = -1.07), -0.07% (t-statistic = -0.46), -0.12% (t-statistic = -0.72), and 0.02% (t-statistic = 0.10) per month.

The results based on the all-but-microcaps sample in Panel B are consistent with those based on the full sample in Panel A. To further rule out the possibility that extremely small stocks alone drive the above results, we conduct independent triple sorts on size, illiquidity, and R&D

intensity using the full sample. The results reported in Tables IA23 and IA24 of the Internet Appendix suggest that the positive R&D-return relation among illiquid stocks is unlikely to be driven by extremely small stocks. Taken together, the results from the portfolio analyses show evidence of an economically large and statistically significant R&D premium only among illiquid stocks. Moreover, the R&D premium among illiquid stocks persists even after adjusting for risk using all four premier asset pricing models. These results strongly support our hypothesis and are consistent with earlier findings in Sections 4.1 and 4.2 from cross-sectional regressions.

5. Tests of alternative mechanisms

5.1 Financial constraints

First, we examine the effect of financial constraints on our finding that the return predictive power of R&D intensity strengthens with the level of illiquidity. Li (2011) shows that the positive R&D-return relation exists only among financially constrained firms. In this subsection, we show empirically that financial constraints cannot explain the important role that illiquidity plays in the relation between R&D intensity and future stock returns. To begin with, we divide both the full sample and the all-but-microcaps sample into financially constrained and financially unconstrained subsamples based on the median value of the distribution of the KZ index developed in Kaplan and Zingales (1997).[13] More precisely, the financially constrained subsample includes stocks above the median, whereas the financially unconstrained subsample includes stocks below the median. For each of these subsamples, we then continue independent double sorts, on illiquidity and R&D intensity, similar to the procedure used in Table 4.

Table 5 shows the results for financially constrained and unconstrained subsamples. We report the value-weighted (equal-weighted) average excess returns and abnormal returns for portfolios based on the full sample (all-but-microcaps sample). As can be seen in Panels A and B, controlling for financial constraints does not weaken the return predictive power of R&D intensity among illiquid stocks. The significantly positive return and abnormal returns on the high-minus-low R&D-intensity portfolio prevail only in the High illiquidity group, in both the financially constrained and the financially unconstrained subsamples. For example, when analyzed using the financially constrained subsample of the full sample, the value-weighted average monthly return of 1.36% on the high-minus-low $RDME$ portfolio in the High illiquidity group is statistically significant (t-statistic = 3.00). Moreover, the value-weighted characteristic-adjusted

[13] For robustness check, we also use the SA index and show, in the Internet Appendix Table IA25, that the empirical results are qualitatively similar to those reported in this subsection. By construction, the value of the KZ index/SA index is higher for a more financially constrained firm. The details of the computation of the KZ index are presented in Kaplan and Zingales (1997) and Lamont et al. (2001). The SA index, which is a combination of asset size and firm age, is computed following Hadlock and Pierce (2010).

return and the value-weighted industry-adjusted return of 1.35% and 1.34% per month on this hedge portfolio are statistically significant, with t-statistics of 3.00 and 3.25, respectively. After adjusting for risk, the R&D premium still persists; the monthly alphas relative to the FF3, FFC, FF5, and HXZ models are, respectively, 1.62% (t-statistic = 3.90), 1.30% (t-statistic = 2.63), 1.87% (t-statistic = 4.60), and 1.74% (t-statistic = 3.28). Furthermore, the value-weighted average excess returns and abnormal returns on portfolios increase monotonically with R&D intensity. However, the value-weighted average return, the characteristic-adjusted return, and the industry-adjusted return on the high-minus-low $RDME$ portfolio in the Low illiquidity group are only 0.16% (t-statistic = 0.55), 0.15% (t-statistic = 0.58), and 0.17% (t-statistic = 0.76) per month, respectively. The corresponding monthly alphas relative to the FF3, FFC, FF5, and HXZ models are, respectively, −0.06% (t-statistic = −0.24), 0.01% (t-statistic = 0.03), 0.19% (t-statistic = 0.64), and 0.37% (t-statistic = 1.20). The absence of a significantly positive R&D premium in the Low illiquidity group suggests that financial constraints do not affect the relation between R&D intensity and future returns once the level of stock illiquidity is controlled for.

An examination of the financially unconstrained subsample of the full sample reveals that the value-weighted average monthly return, the characteristic-adjusted return, and the industry-adjusted return on the high-minus-low $RDME$ portfolio formed in the High illiquidity group are, respectively, 1.14%, 1.12%, and 1.08%. Each of these average returns is also found to be statistically significant with t-statistics of 3.08, 3.30, and 2.98, respectively. The average abnormal returns on this hedge portfolio relative to the FF3, FFC, FF5, and HXZ models are, respectively, 1.01% (t-statistic = 2.67), 1.09% (t-statistic = 2.77), 1.20% (t-statistic = 3.19), and 1.50% (t-statistic = 3.54) per month. In addition, a monotonically increasing pattern prevails for the value-weighted average excess returns and abnormal returns on portfolios when moving from the Low R&D-intensity portfolio, $RDME_L$, to the High R&D-intensity portfolio, $RDME_H$. In sharp contrast, we observe that the value-weighted average return, the characteristic-adjusted return, and the industry-adjusted return on the high-minus-low $RDME$ portfolio formed in the Low illiquidity group are only 0.12% (t-statistic = 0.60), 0.13% (t-statistic = 0.65), and 0.12% (t-statistic = 0.79) per month, respectively. The monthly average alphas estimates for this spread portfolio relative to the FF3, FFC, FF5, and HXZ models are also very small, both in magnitude and in statistical significance; they are, respectively, −0.08% (t-statistic = −0.44), 0.08% (t-statistic = 0.45), 0.07% (t-statistic = 0.38), and 0.23% (t-statistic = 1.13).

When we analyze the results based on the all-but-microcaps sample, we see a pattern very similar to that of the full sample. It is worth highlighting that, in both the full sample and the all-but-microcaps sample, the positive R&D-return relation among illiquid stocks appears to be much stronger for the financially constrained subsample – an empirical finding similar to that reported in Li (2011). Collectively, the results in Table 5 show that financial constraints cannot explain the important role that stock-level illiquidity plays in the positive R&D-return relation.

5.2 Innovation ability

We next investigate the effect of a firm's innovation ability on our empirical finding. Cohen et al. (2013) show that R&D intensity predicts future stock returns only when firms have a high ability to translate R&D activities into real sales growth. Specifically, GoodR&D firms (i.e., firms with high R&D intensity and high innovation ability) earn substantially higher future stock returns than BadR&D firms (i.e., firms with high R&D intensity and low innovation ability). In this subsection, we show that our finding is unaffected after controlling for innovation ability. Consistent with Cohen et al. (2013), we compute a firm's ability to innovate as the average of the slope coefficient estimates from the rolling time-series regressions of firm-level sales growth on the past five years of R&D investments (i.e., R&D expenditure scaled by sales).[14] Both the full sample and the all-but-microcaps sample are divided separately into low innovation ability and high innovation ability subsamples. Analogous to the preceding section, we then repeat independent double sorts, on stock illiquidity and R&D intensity, for each of these subsamples.

Table 6 demonstrates that innovation ability does not weaken our preceding finding that the positive relation between R&D intensity and future stock returns strengthens with the level of illiquidity. Both in the low innovation ability and the high innovation ability subsamples, the high-minus-low R&D-intensity portfolio generates a positive and significant return only in the High illiquidity group. For example, when the low innovation ability subsample constructed from the full sample is examined in Panel A of Table 6, we find that the high-minus-low $RDME$ portfolio in the High illiquidity group generates a value-weighted average monthly return of 0.68% (t-statistic = 2.00). The value-weighted characteristic- and industry-adjusted returns of this hedge portfolio are 0.57% and 0.74% per month, with t-statistics of 1.98 and 2.51, respectively. The corresponding monthly alpha estimates relative to the FF3, FFC, FF5, and HXZ models are 0.62% (t-statistic = 1.99), 0.55% (t-statistic = 1.97), 0.74% (t-statistic = 2.01), and 0.73% (t-statistic = 2.03), respectively. On the other hand, the high-minus-low $RDME$ portfolio in the Low illiquidity group earns only 0.18% per month (t-statistic = 0.86) in value-weighted average return, −0.13% per month (t-statistic = −0.68) in value-weighted characteristic-adjusted return, and 0.10% per month (t-statistic = 0.44) in value-weighted industry-adjusted return. The value-weighted average monthly abnormal returns of this hedge portfolio relative to the FF3, FFC, FF5, and HXZ models are, respectively, 0.05% (t-statistic = 0.20), 0.06% (t-statistic = 0.30),

[14] Cohen et al. (2013) use eight years of past data as a back window for each firm-level regression. The authors show that their results are insensitive to estimation windows of 6 to 10 years. We use a back window of six years of past data to run these regressions. This allows us to have more observations estimated for a firm's innovation ability. The details of the estimation procedure can be found in Cohen et al. (2013).

0.13% (t-statistic = 0.56), and 0.08% (t-statistic = 0.36).

When the high innovation ability subsample constructed from the full sample is examined in Panel B, we observe that both the value-weighted average excess returns and abnormal returns on portfolios increase monotonically with R&D intensity in the High illiquidity group. The high-minus-low $RDME$ portfolio delivers a value-weighted average return of 0.90% per month, with a corresponding t-statistic of 2.63. This hedge portfolio also earns 0.77% per month (t-statistic = 2.38) in value-weighted characteristic-adjusted return and 0.75% per month (t-statistic = 2.19) in value-weighted industry-adjusted return. Furthermore, the average abnormal returns on the portfolio relative to the factor pricing models are large and statistically significant. But in the Low illiquidity group, the abnormal returns do not show a pattern similar to that in the High illiquidity group. In addition, the average return and abnormal returns on the high-minus-low R&D-intensity portfolio are much smaller and are all statistically insignificant. For example, the value-weighted average return, the characteristic-adjusted return, and the industry-adjusted return on the high-minus-low $RDME$ portfolio are 0.20% (t-statistic = 0.75), −0.22% (t-statistic = −0.85), and 0.17% (t-statistic = 0.75) per month, respectively. These results also indicate that a firm's ability to innovate is unlikely to affect the significantly positive relation between R&D intensity and future returns once its stock-level illiquidity is controlled for.

A pattern very similar to that of the full sample emerges when examining the low innovation ability and the high innovation ability subsamples constructed from the all-but-microcaps sample. To sum up, the results in Table 6 indicate that the positive R&D-return relation among illiquid stocks is stronger in the high innovation ability subsample, which is consistent with Cohen et al. (2013). More importantly, illiquidity independently drives a significant portion of the positive relation between R&D intensity and future stock returns.

5.3 Product market competition

Finally, we explore the effect of product market competition on our finding. Gu (2016) has recently shown that the positive relation between R&D intensity and future stock returns exists only for firms in competitive industries. In this subsection, we show empirically that the strengthening of the cross-sectional return predictive power of R&D intensity with the level of stock illiquidity remains unaffected after controlling for product market competition. As in Gu (2016), we follow Hou and Robinson (2006) to measure product market competition using the Herfindahl-Hirschman Index.[15] We divide both the full sample and the all-but-microcaps sample

[15] The Herfindahl-Hirschman Index is defined as $\text{HHI}_{l,t} = \sum_{j=1}^{o_l} \vartheta_{j,l,t}^2$, where $\vartheta_{j,l,t}$ is the market share of net sales (Compustat item SALE) of firm j in industry l in year t, o_l is the number of firms in industry l in year t, and $\text{HHI}_{l,t}$ is the Herfindahl-Hirschman Index of industry l in year t. Small (large) values of HHI imply that the product market is shared by many competing firms (a few large firms). We compute HHI in every year and use the average value over the previous three years. This ensures that potential data errors do not have undue influence on our measure of product market competition. More details can be found in Hou and Robinson (2006).

separately into low competition and high competition subsamples based on the median value of the distribution of HHI. Using each of these subsamples, we then conduct independent double sorts, on stock illiquidity and firm's R&D intensity, similar to the procedure in Table 4.

Table 7 reports the results for the low competition (i.e., above the median) and the high competition (i.e., below the median) subsamples. We see that after controlling for product market competition, our main finding holds regardless of the sample composition. In Panel A, where the low competition subsample of the full sample is examined, the value-weighted average monthly return, the characteristic-adjusted return, and the industry-adjusted return on the high-minus-low $RDME$ portfolio formed in the High illiquidity group turn out to be 1.68%, 1.41%, and 1.51%, with t-statistics of 3.90, 3.51, and 3.70, respectively. The R&D premium persists after adjusting for risk using the asset pricing models. Specifically, this hedge portfolio generates monthly average abnormal returns of 1.68% (t-statistic = 4.02) relative to the FF3 model, 1.47% (t-statistic = 2.89) relative to the FFC model, 1.73% (t-statistic = 4.16) relative to the FF5 model, and 1.85% (t-statistic = 3.46) relative to the HXZ model. Moreover, the value-weighted average excess returns and abnormal returns on portfolios increase monotonically with R&D intensity. By contrast, the high-minus-low $RDME$ portfolio in the Low illiquidity group generates a value-weighted average monthly return of only −0.01% (t-statistic = −0.04). The value-weighted characteristic-adjusted return and the value-weighted industry-adjusted return on the portfolio are also very small at 0.00% (t-statistic = −0.01) and 0.10% (t-statistic = 0.69) per month, respectively. The corresponding monthly alpha estimates relative to the FF3, FFC, FF5, and HXZ models are all negative and statistically insignificant at conventional levels.

Examining the high competition subsample of the full sample, it can be seen that the high-minus-low $RDME$ portfolio in the High illiquidity group generates a value-weighted average monthly return of 1.86%, which is economically large and statistically significant (t-statistic = 4.55). The value-weighted characteristic- and industry-adjusted returns on the portfolio are also substantial and statistically significant; they are, respectively, 1.75% (t-statistic = 5.05) and 1.73% (t-statistic = 4.77) per month. The average abnormal returns on the portfolio relative to the FF3, FFC, FF5, and HXZ models are, respectively, 1.78% (t-statistic = 4.42), 1.74% (t-statistic = 4.10), 1.79% (t-statistic = 3.71), and 1.87% (t-statistic = 3.51) per month. The high-minus-low $RDME$ portfolio in the Low illiquidity group also earns 0.68% per month (t-statistic = 2.92) in value-weighted average return, 0.61% per month (t-statistic = 2.95) in value-weighted characteristic-adjusted return, and 0.59% per month (t-statistic = 2.92) in value-weighted industry-adjusted return. The monthly average abnormal returns on this hedge portfolio relative to the FF3, FFC, FF5, and HXZ models are 0.52% (t-statistic = 2.27), 0.52%

For robustness check, we also use the price-cost margin (as in Peress, 2010) and show, in the Internet Appendix Table IA26, that the empirical results are qualitatively similar to those reported in this subsection.

(t-statistic = 1.91), 0.36% (t-statistic = 1.39), and 0.41% (t-statistic = 1.35), respectively. It is important to mention that although an R&D premium exists in the Low illiquidity group, it is roughly three times smaller than that in the High illiquidity group. The value-weighted average abnormal returns on the high-minus-low $RDME$ portfolio are also about three times smaller than those in the High illiquidity group. These results, therefore, corroborate our preceding finding that the positive R&D-return relation is much stronger among illiquid stocks.

We see a pattern similar to that of the full sample when we focus on the results based on the all-but-microcaps sample. To summarize, the results in Table 7 reveal a more pronounced power of R&D intensity to predict stock returns in the high competition subsample, which is in line with Gu (2016). Moreover, product market competition cannot explain the important role that stock illiquidity plays in the positive relation between R&D intensity and future stock returns.[16]

6. Conclusion

This paper examines the effect of illiquidity on the relation between R&D investment and expected equity returns. To do so, we propose a dynamic model of R&D venture, which incorporates uncertainty about the likelihood of project success into the valuation framework. Uncertainty, which emerges naturally in the context of R&D due to the highly heterogeneous complexity in assessing R&D venture's success, lowers the value of an R&D project. Thus, its prevalence increases expected excess returns in equilibrium. Our model also links information friction to stock illiquidity. In particular, a firm's R&D value is a function of the probability of project success intensity. The true value of the R&D success intensity parameter is unknown to the market makers and uninformed investors who have to decipher the value from limited information available to them. In fact, the true value of the parameter is more likely to be far from the estimate when information used for inferring the success intensity is sparse. This translates into higher losses for the market makers trading with informed traders. Thus, the higher the uncertainty, the higher the bid-ask spreads set by the market makers in attempts to protect themselves from possible losses. The parameter uncertainty is positively related to the firm's R&D risk premium and has an amplifying effect on the exposure to the systematic risk. Taken together, the model predicts that the positive relation between a firm's R&D investment and expected stock returns strengthens with the level of illiquidity.

Consistent with our theoretical model's prediction, we find robust empirical evidence that the positive R&D-return relation exists among illiquid stocks. A further analysis indicates that the important role of illiquidity in the relation between R&D intensity and future stock returns cannot be explained by factors such as financial constraints, innovation ability, and product

[16] We provide a battery of additional robustness checks in the Internet Appendix.

market competition, which have been shown to affect the R&D premium in prior studies. All these results suggest that stock illiquidity is an independent driver of the positive R&D-return relation. This paper thereby complements and extends the extant literature by providing an understanding of the R&D investment premium from the perspective of a firm's stock illiquidity.

Appendix A.

Proof of Proposition 2.

Equations (22) and (23) are direct results from equations (20) and (21), respectively. To prove equation (25), we have

$$\begin{aligned}
&\frac{\partial \text{bid-ask spread}}{\partial \sigma_{\hat{u}}} \\
&= \alpha \frac{\left[\frac{\partial V(\lambda(\hat{u}+\sigma_{\hat{u}}),c)}{\partial \sigma_{\hat{u}}} - \frac{\partial V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)}{\partial \sigma_{\hat{u}}}\right]\widehat{V} - [V(\lambda(\hat{u}+\sigma_{\hat{u}}),c) - V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)]\frac{\partial \widehat{V}}{\partial \sigma_{\hat{u}}}}{\widehat{V}^2} \\
&= \alpha \frac{\left[\frac{\partial V(\lambda,c)}{\partial \lambda}|_{\lambda=\lambda(\hat{u}+\sigma_{\hat{u}})} + \frac{\partial V(\lambda,c)}{\partial \lambda}|_{\lambda=\lambda(\hat{u}-\sigma_{\hat{u}})}\right]\widehat{V} - [V(\lambda(\hat{u}+\sigma_{\hat{u}}),c) - V(\lambda(\hat{u}-\sigma_{\hat{u}}),c)]\frac{\partial \widehat{V}}{\partial \sigma_{\hat{u}}}}{\widehat{V}^2} \\
&> 0,
\end{aligned}$$

where the last line is by Corollaries 1 and 2. This completes the proof of Proposition 2. ■

Appendix B. Illiquidity measures

BAS is calculated at the end of each month τ as

$$BAS_{j,\tau} = \frac{1}{D}\sum_{d=1}^{D}\frac{Ask_{j,d} - Bid_{j,d}}{Mid_{j,d}},$$

where $Ask_{j,d}$ and $Bid_{j,d}$ are, respectively, the ask and bid prices of stock j on day d from the CRSP daily database, $Mid_{j,d}$ is the average of $Ask_{j,d}$ and $Bid_{j,d}$, and D is the number of days in the estimation period. A higher value of BAS implies a higher level of illiquidity (or, equivalently, a lower level of liquidity). For robustness checks of our empirical findings, we also estimate two more versions of BAS. The first is computed at the end of each month τ using one year's worth of daily data. In a slightly different manner, the second is calculated at the end of June of each year t using one year's worth of daily data. We require a minimum of 200 valid daily observations in both versions. The value of BAS is updated on a monthly basis in the former, while on a yearly basis in the latter version. The results based on these alternative versions of BAS are qualitatively similar to those reported in this paper and are provided in the Internet Appendix Tables IA9 to IA14.

The Amihud (2002) illiquidity measure is defined as

$$ILLIQ_j = \frac{1}{D}\sum_{d=1}^{D}\frac{|R_{j,d}|}{DVOL_{j,d}},$$

where $R_{j,d}$ is the return of stock j on day d, $DVOL_{j,d}$ is the dollar trading volume of stock j on day d, and D is the number of days in the estimation period. A higher value of $ILLIQ$ indicates a higher level of stock illiquidity. We follow Gao and Ritter (2010) to adjust trading volume of NASDAQ-listed stocks and compute $ILLIQ$ at the end of each month τ using one month's worth of daily data with at least 15 positive-volume days.

Appendix C. Supplementary data

Supplementary results related to this article can be found in the Internet Appendix.

References

Abdellaoui, M., Baillon, A., Placido, L., Wakker, P. P., 2011. The rich domain of uncertainty: Source functions and their experimental implementation. American Economic Review 101, 695–723.

Aboody, D., Lev, B., 2000. Information asymmetry, R&D, and insider gains. Journal of Finance 55, 2747–2766.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Amihud, Y., Hameed, A., Kang, W., Zhang, H., 2015. The illiquidity premium: International evidence. Journal of Financial Economics 117, 350–368.

Amihud, Y., Mendelson, H., 1989. The effects of beta, bid-ask spread, residual risk, and size on stock returns. Journal of Finance 44, 479–486.

Balachandra, R., Friar, J. H., 1997. Factors for success in R&D projects and new product innovation: A contextual framework. IEEE Transactions on Engineering Management 44, 276–287.

Berk, J. B., Green, R. C., Naik, V., 2004. Valuation and return dynamics of new ventures. Review of Financial Studies 17, 1–35.

Camerer, C., Weber, M., 1992. Recent developments in modeling preferences: Uncertainty and ambiguity. Journal of Risk and Uncertainty 5, 325–370.

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

Carpenter, R. E., Petersen, B. C., 2002. Capital market imperfections, high-tech investment, and new equity financing. Economic Journal 112, 54–72.

Chambers, D., Jennings, R., Thompson, R. B., 2002. Excess returns to R&D-intensive firms. Review of Accounting Studies 7, 133–158.

Chan, L. K. C., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. Journal of Finance 56, 2431–2456.

Chan, S. H., Martin, J. D., Kensinger, J. W., 1990. Corporate research and development expenditures and share value. Journal of Financial Economics 26, 255–276.

Chordia, T., Roll, R., Subrahmanyam, A., 2008. Liquidity and market efficiency. Journal of Financial Economics 87, 249–268.

Chordia, T., Roll, R., Subrahmanyam, A., 2011. Recent trends in trading activity and market quality. Journal of Financial Economics 101, 243–263.

Cohen, L., Diether, K., Malloy, C., 2013. Misvaluing innovation. Review of Financial Studies 26, 635–666.

Copeland, T. E., Galai, D., 1983. Information effects on the bid-ask spread. Journal of Finance 38, 1457–1469.

Cox, D. R., 1972. Regression models and life-tables. Journal of the Royal Statistical Society: Series B (Methodological) 34, 187–202.

Daniel, K., Grinblatt, M., Titman, S., Wermers, R., 1997. Measuring mutual fund performance with characteristic-based benchmarks. Journal of Finance 52, 1035–1058.

Daniel, K., Hirshleifer, D., Sun, L., 2020. Short and long horizon behavioral factors. Review of Financial Studies 33, 1673–1736.

Duffie, D., Saita, L., Wang, K., 2007. Multi-period corporate default prediction with stochastic covariates. Journal of Financial Economics 83, 635–665.

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

Fama, E. F., French, K. R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Fama, E. F., French, K. R., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3–56.

Fama, E. F., French, K. R., 1997. Industry costs of equity. Journal of Financial Economics 43, 153–193.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Gao, X., Ritter, J. R., 2010. The marketing of seasoned equity offerings. Journal of Financial Economics 97, 33–52.

Glosten, L. R., Milgrom, P. R., 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics 14, 71–100.

Gu, L., 2016. Product market competition, R&D investment, and stock returns. Journal of Financial Economics 119, 441–455.

Hadlock, C. J., Pierce, J. R., 2010. New evidence on measuring financial constraints: Moving beyond the KZ index. Review of Financial Studies 23, 1909–1940.

Hall, B. H., 1996. The private and social returns to research and development. In: Smith, B. L. R., Barfield, C. E. (Eds.), Technology, R&D, and the Economy. Brookings Institution Press, Washington, D.C., pp. 140–183.

Han, Y., Lesmond, D., 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. Review of Financial Studies 24, 1590–1629.

Hannon, E., Smits, S., Weig, F., 2015. Brightening the black box of R&D. McKinsey Quarterly 2, 1–11.

Hirshleifer, D., Hsu, P.-H., Li, D., 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107, 632–654.

Hirshleifer, D., Hsu, P.-H., Li, D., 2018. Innovative originality, profitability, and stock returns. Review of Financial Studies 31, 2553–2605.

Hou, K., Robinson, D. T., 2006. Industry concentration and average stock returns. Journal of Finance 61, 1927–1956.

Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. Review of Financial Studies 28, 650–705.

Huang, T., Li, J., Wu, F., Zhu, N., 2022. R&D information quality and stock returns. Journal of Financial Markets 57, 100599.

Jo, D., Park, J., 2019. Understanding the success factors of R&D organization. In: Lee, R. (Ed.), Software Engineering Research, Management and Applications. Springer International Publishing, Cham, pp. 75–90.

Kaplan, S. N., Zingales, L., 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? Quarterly Journal of Economics 112, 169–215.

Lamont, O., Polk, C., Saá-Requejo, J., 2001. Financial constraints and stock returns. Review of Financial Studies 14, 529–554.

Lev, B., Sougiannis, T., 1996. The capitalization, amortization, and value-relevance of R&D. Journal of Accounting and Economics 21, 107–138.

Li, D., 2011. Financial constraints, R&D investment, and stock returns. Review of Financial Studies 24, 2974–3007.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Pammolli, F., Magazzini, L., Riccaboni, M., 2011. The productivity crisis in pharmaceutical R&D. Nature Reviews Drug Discovery 10, 428–438.

Pástor, Ľ., Veronesi, P., 2003. Stock valuation and learning about profitability. Journal of Finance 58, 1749–1789.

Peress, J., 2010. Product market competition, insider trading, and stock market efficiency. Journal of Finance 65, 1–43.

Shumway, T., 1997. The delisting bias in CRSP data. Journal of Finance 52, 327–340.

Stoll, H. R., 2000. Friction. Journal of Finance 55, 1479–1514.

Therneau, T. M., Grambsch, P. M., 2000. Modeling Survival Data: Extending the Cox Model. Springer-Verlag, New York.

Table 1
Summary statistics: Full sample

The sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a nonmissing research and development (R&D) intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. Panel A reports the time-series averages of the cross-sectional mean, standard deviation (SD), minimum, fifth percentile (5%), 25th percentile (25%), median, 75th percentile (75%), 95th percentile (95%), and maximum values of the distribution of each R&D intensity measure and illiquidity measure. BAS is the average bid-ask spread, which is computed using one month's worth of daily data and requires a minimum of 15 valid daily observations. $RDME$ is computed at the end of June of each year t using R&D expenditure at the end of the fiscal year ending in the calendar year $t-1$ scaled by ME at the end of June of year t. ME is the market value of equity (in million \$) computed as the number of shares outstanding times the absolute price of one share at the end of June of each year t. $\mathcal{M}$ is the number of firms in an average month. At the end of June of each year t, stocks are sorted into three portfolios using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of $RDME$ computed at the end of June of year t and the portfolios are rebalanced annually. Panel B, for each $RDME$-sorted portfolio, reports the time-series averages of the cross-sectional means for characteristics of constituent firms. BM is the book-to-market ratio computed in June of each year t as the ratio of the book value of equity at the end of the fiscal year ending in the calendar year $t-1$ to the market value of equity at the end of December of the calendar year $t-1$. $R_{-1,0}$ is the prior one-month return and $R_{-12,-2}$ is the prior year's return skipping the last month. AG is the asset growth computed in June of year t as the annual growth rate of total assets. OP is the operating profitability computed in June of year t as total revenues minus cost of goods sold, minus selling, general, and administrative expenses, minus interest expense, all divided by book value of equity at the end of the fiscal year ending in the calendar year $t-1$. $IVOL$ is the idiosyncratic volatility computed at the end of each month τ as the standard deviation of the residuals from regressing daily stock returns (in excess of the risk-free rate) on the Fama and French (1993) three factor returns over the month τ (with a minimum of 15 valid daily observations). Panel C reports both the value-weighted (VW) and equal-weighted (EW) average monthly excess returns (in %) of the portfolios sorted on $RDME$. Panel D reports both VW and EW average monthly excess returns (in %) of quintile portfolios sorted on $RDME$. H−L denotes the high-minus-low portfolio, that is, long stocks in the High portfolio and short stocks in the Low portfolio. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). The sample period is from January 1975 to December 2016.

Panel A: Summary statistics of R&D intensity and illiquidity measures

	Mean	SD	Minimum	5%	25%	Median	75%	95%	Maximum	$\mathcal{M}$
BAS	0.06	0.06	0.00	0.02	0.03	0.04	0.07	0.17	0.67	2037
$RDME$	0.07	0.14	0.00	0.00	0.01	0.03	0.07	0.25	2.88	2037

Panel B: Characteristics of portfolios sorted on $RDME$

Portfolio	BAS	ME	BM	$R_{-1,0}$	$R_{-12,-2}$	AG	OP	$IVOL$
Low	0.061	2804.972	0.808	0.012	0.174	0.231	0.201	0.028
Medium	0.051	2979.094	0.624	0.013	0.177	0.214	0.117	0.027
High	0.072	662.931	0.928	0.019	0.103	0.092	−0.153	0.037

Panel C: Portfolio excess returns

		Low	Medium	High	H−L
$RDME$	VW	0.55	0.67	1.03	0.48
		(2.47)	(3.17)	(3.68)	(2.61)
	EW	0.72	0.83	1.65	0.93
		(2.74)	(2.92)	(4.51)	(4.88)

Panel D: Excess returns of quintile portfolios

		Low	2	3	4	High	H−L
$RDME$	VW	0.70	0.53	0.71	0.83	1.01	0.31
		(2.91)	(2.48)	(3.17)	(3.31)	(3.37)	(1.41)
	EW	0.82	0.55	0.87	1.12	1.86	1.04
		(3.09)	(2.03)	(3.01)	(3.52)	(4.78)	(4.47)

Table 2
Cross-sectional regressions

The table reports the time-series average slope and intercept coefficient estimates from the Fama and MacBeth (1973) cross-sectional regressions. At the end of each month τ, stocks are sorted into three groups using the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of illiquidity measured by the average bid-ask spread in month τ. Then, in each month from July of year t to June of year t+1 for the Low (High) illiquidity group formed in the previous month, we run a cross-sectional regression of monthly returns on lagged variables including prior one-month return ($R_{-1,0}$), prior year's return skipping the last month ($R_{-12,-2}$, i.e., return from month $\tau-12$ to month $\tau-2$), the natural logarithm of market value of equity ($\ln(ME)$), the natural logarithm of book-to-market ratio ($\ln(BM)$), asset growth (AG), profitability (OP), idiosyncratic volatility ($IVOL$), and research and development (R&D) intensity measured by $RDME$ (i.e., R&D expenditure scaled by ME). In Panel A, the full sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. In Panel B, the all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates are used to compute the breakpoints for illiquidity. High−Low is the time-series average of the difference in slope or intercept coefficient estimates between the High and Low illiquidity groups. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). All independent variables are winsorized at the 1% and 99% levels on a monthly basis prior to running the cross-sectional regressions. The sample period is from January 1975 to December 2016. See also notes to Table 1.

	Intercept	$R_{-1,0}$	$R_{-12,-2}$	$\ln(BM)$	$\ln(ME)$	AG	OP	$IVOL$	$RDME$
Panel A: Full sample									
Low	0.72	−5.04	0.59	0.08	−0.05	−0.28	0.46	0.02	0.97
	(1.90)	(−8.45)	(3.10)	(1.02)	(−1.69)	(−1.59)	(1.63)	(0.43)	(1.16)
High	2.72	−7.53	−0.45	0.09	−0.44	−1.21	0.35	−0.07	5.43
	(5.08)	(−11.03)	(−1.33)	(0.87)	(−5.70)	(−6.81)	(1.71)	(−1.63)	(5.28)
High−Low	2.00	−2.49	−1.04	0.00	−0.39	−0.93	−0.10	−0.10	4.46
	(4.60)	(−3.45)	(−4.07)	(0.04)	(−5.04)	(−3.85)	(−0.30)	(−1.43)	(4.55)
Panel B: All-but-microcaps sample									
Low	1.00	−6.86	0.26	0.04	−0.07	−0.12	0.66	0.23	−0.08
	(1.97)	(−9.43)	(1.08)	(0.37)	(−1.88)	(−0.49)	(2.06)	(3.04)	(−0.06)
High	3.18	−2.97	0.21	−0.03	−0.28	−0.54	0.40	−0.38	6.40
	(4.53)	(−4.62)	(0.65)	(−0.23)	(−3.11)	(−3.14)	(1.18)	(−4.69)	(3.43)
High−Low	2.18	3.90	−0.05	−0.07	−0.20	−0.42	−0.26	−0.61	6.47
	(3.42)	(4.66)	(−0.19)	(−0.59)	(−2.63)	(−1.44)	(−0.59)	(−5.31)	(3.64)

Table 3
Cross-sectional regressions with illiquidity dummies

The table reports the time-series averages of the Fama and MacBeth (1973) cross-sectional regression coefficient estimates. At the end of each month τ, two dummy variables are created by sorting stocks into three groups based on the breakpoints for the bottom 30% (Low), middle 40% (Medium), and top 30% (High) of the ranked values of illiquidity measured by the average bid-ask spread (BAS) in month τ. Then, in each month from July of year t to June of year t+1, we run cross-sectional regressions of monthly returns on lagged variables including prior one-month return ($R_{-1,0}$), prior year's return skipping the last month ($R_{-12,-2}$, i.e., return from month $\tau-12$ to month $\tau-2$), the natural logarithm of market value of equity ($\ln(ME)$), the natural logarithm of book-to-market ratio ($\ln(BM)$), asset growth (AG), profitability (OP), idiosyncratic volatility ($IVOL$), research and development (R&D) intensity measured by $RDME$ (i.e., R&D expenditure scaled by ME), and illiquidity dummy and interaction variables. BAS_H (BAS_L) is a dummy variable equal to one for stocks in the High (Low) illiquidity group in a given month and is zero otherwise. In Panel A, the full sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. In Panel B, the all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates are used to compute the breakpoints for illiquidity. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). All independent variables (except the dummies) are winsorized at the 1% and 99% levels on a monthly basis prior to running the cross-sectional regressions. The sample period is from January 1975 to December 2016. See also notes to Table 1.

Intercept	$R_{-1,0}$	$R_{-12,-2}$	$\ln(BM)$	$\ln(ME)$	AG	OP	$IVOL$	$RDME$	$RDME*BAS_H$	$RDME*BAS_L$	BAS_H	BAS_L
Panel A: Full sample												
1.31	−5.98	0.29	0.15	−0.08	−0.87	0.41	−0.05	4.21				
(3.34)	(−11.81)	(1.30)	(1.91)	(−2.47)	(−7.97)	(2.85)	(−1.28)	(5.19)				
1.33	−5.95	0.30	0.15	−0.10	−0.84	0.42	−0.05	4.77		−4.38		0.19
(3.27)	(−11.72)	(1.39)	(2.00)	(−2.82)	(−7.69)	(2.95)	(−1.53)	(5.65)		(−5.18)		(1.68)
1.48	−6.08	0.30	0.14	−0.10	−0.83	0.40	−0.02	2.12	3.98		−0.52	
(3.87)	(−12.16)	(1.41)	(1.92)	(−3.15)	(−7.81)	(2.87)	(−0.52)	(2.85)	(4.45)		(−3.61)	
1.49	−6.04	0.31	0.14	−0.11	−0.81	0.41	−0.02	2.81	3.30	−2.55	−0.49	0.07
(3.74)	(−12.04)	(1.49)	(1.96)	(−3.32)	(−7.61)	(2.95)	(−0.65)	(3.61)	(3.49)	(−3.23)	(−3.62)	(0.69)
Panel B: All-but-microcaps sample												
1.76	−4.15	0.43	0.06	−0.14	−0.40	0.35	−0.17	3.60				
(3.65)	(−7.43)	(1.71)	(0.63)	(−3.39)	(−3.27)	(1.43)	(−3.01)	(2.55)				
2.02	−4.13	0.43	0.08	−0.16	−0.39	0.38	−0.23	4.19		−4.58		−0.05
(4.12)	(−7.45)	(1.69)	(0.79)	(−3.72)	(−3.25)	(1.57)	(−4.00)	(2.64)		(−3.08)		(−0.52)
1.82	−4.12	0.45	0.07	−0.14	−0.38	0.37	−0.18	1.60	4.57		−0.13	
(3.84)	(−7.55)	(1.85)	(0.73)	(−3.47)	(−3.18)	(1.55)	(−3.29)	(1.97)	(3.33)		(−0.97)	
2.06	−4.10	0.44	0.08	−0.15	−0.38	0.39	−0.22	2.24	3.85	−2.63	−0.14	−0.12
(4.27)	(−7.54)	(1.82)	(0.87)	(−3.73)	(−3.20)	(1.67)	(−3.95)	(2.01)	(2.65)	(−1.75)	(−1.01)	(−1.30)

Table 4
Double sorts on illiquidity and R&D intensity

The table reports monthly returns of portfolios sorted on illiquidity and research and development (R&D) intensity. Illiquidity is measured by the average bid-ask spread (BAS). R&D intensity is measured by $RDME$, which is R&D expenditure scaled by market value of equity. At the end of each month τ, stocks are sorted into three groups using the breakpoints for the bottom 30% (Low (L)), middle 40% (Medium (M)), and top 30% (High (H)) of the ranked values of BAS in month τ. Independently, at the end of June of each year t, stocks are sorted into three groups using the breakpoints for the bottom 30% (L), middle 40% (M), and top 30% (H) of the ranked values of $RDME$ computed at the end of June of year t and the stocks are held in the assigned groups for the subsequent 12 months. Nine portfolios are formed every month from the intersections of the illiquidity and R&D intensity groups. The full sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. The all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and we use the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates to compute the breakpoints for illiquidity and R&D intensity separately. Panel A (Panel B) reports the value-weighted (equal-weighted) average monthly returns (in %) on portfolios. Excess return is the portfolio return in excess of the one-month Treasury bill rate. Characteristic-adjusted (Char-adj) returns are computed by adjusting returns using 125 (5 × 5 × 5) size/book-to-market/momentum benchmark portfolios (as in Daniel et al., 1997). Industry-adjusted (Ind-adj) returns are computed by adjusting returns using the Fama and French 12 industry portfolios (as in Fama and French, 1997). The alphas (in %) are estimated from the time-series regressions of portfolio excess returns on various factor models including the Fama and French (1993) three-factor (FF3) model, the Fama and French (1993) and Carhart (1997) four-factor (FFC) model, the Fama and French (2015) five-factor (FF5) model, and the Hou et al. (2015) q-factor (HXZ) model. H−L denotes the high-minus-low portfolio. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). The sample period is from January 1975 to December 2016. See also notes to Table 1.

	Low illiquidity				High illiquidity			
	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L
Panel A: Full sample								
Excess return	0.73	0.70	0.81	0.08	−0.31	0.44	1.26	1.57
	(3.84)	(3.89)	(3.32)	(0.51)	(−0.58)	(0.89)	(2.46)	(4.79)
Char-adj return	−0.36	−0.28	−0.33	0.03	−0.11	−0.08	1.07	1.18
	(−4.38)	(−4.14)	(−3.65)	(0.21)	(−0.51)	(−0.45)	(3.88)	(4.39)
Ind-adj return	−0.32	−0.34	−0.23	0.09	−1.30	−0.58	0.20	1.50
	(−4.44)	(−5.22)	(−2.39)	(0.69)	(−3.35)	(−1.78)	(0.53)	(5.09)
FF3 α	0.25	0.19	0.08	−0.17	−1.49	−0.64	0.10	1.59
	(2.87)	(2.97)	(0.54)	(−1.07)	(−5.28)	(−2.65)	(0.38)	(4.85)
FFC α	0.21	0.22	0.14	−0.07	−0.74	−0.09	0.62	1.36
	(2.34)	(3.18)	(0.91)	(−0.46)	(−2.25)	(−0.36)	(2.12)	(3.50)
FF5 α	0.12	0.02	0.00	−0.12	−0.72	0.07	0.75	1.47
	(1.35)	(0.39)	(0.03)	(−0.72)	(−2.73)	(0.33)	(2.80)	(4.65)
HXZ α	0.12	0.06	0.14	0.02	−0.25	0.51	1.28	1.53
	(1.19)	(0.88)	(0.85)	(0.10)	(−0.78)	(1.92)	(3.69)	(3.80)
Panel B: All-but-microcaps sample								
Excess return	0.88	0.84	0.97	0.09	0.15	0.71	1.10	0.95
	(4.31)	(4.21)	(4.27)	(1.00)	(0.62)	(1.78)	(2.55)	(3.39)
Char-adj return	−0.26	−0.26	−0.18	0.08	−0.98	−0.50	−0.19	0.79
	(−2.86)	(−3.65)	(−2.11)	(0.93)	(−6.69)	(−3.81)	(−0.97)	(3.43)
Ind-adj return	−0.20	−0.21	−0.10	0.10	−0.80	−0.39	−0.05	0.75
	(−2.20)	(−2.80)	(−0.89)	(1.35)	(−3.36)	(−1.75)	(−0.21)	(3.82)
FF3 α	0.25	0.18	0.24	−0.01	−0.81	−0.24	0.04	0.85
	(2.14)	(2.08)	(2.14)	(−0.16)	(−4.76)	(−1.39)	(0.20)	(3.38)
FFC α	0.24	0.20	0.28	0.04	−0.47	0.15	0.50	0.97
	(2.04)	(2.40)	(2.45)	(0.36)	(−2.74)	(0.82)	(2.32)	(3.56)
FF5 α	0.00	−0.04	0.06	0.06	−0.38	0.29	0.60	0.98
	(−0.00)	(−0.51)	(0.62)	(0.63)	(−2.12)	(1.61)	(2.84)	(4.23)
HXZ α	−0.01	0.00	0.13	0.14	−0.20	0.57	1.00	1.20
	(−0.05)	(0.04)	(1.13)	(1.37)	(−0.99)	(2.46)	(3.64)	(4.28)

Table 5
Double sorts on illiquidity and R&D intensity after controlling for financial constraints

The table reports monthly returns of portfolios sorted on illiquidity and research and development (R&D) intensity for both financially constrained and financially unconstrained subsamples. Financial constraint is measured by the KZ index developed in Kaplan and Zingales (1997). Illiquidity is measured by the average bid-ask spread. R&D intensity is measured by $RDME$, which is R&D expenditure scaled by market value of equity. The full sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. The all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and we use the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates for portfolio analysis. At the end of June of each year t, we first assign stocks in the full sample to the financially constrained (unconstrained) subsample based on the median value of the distribution of the KZ index (i.e., financially constrained (unconstrained) subsample includes stocks above (below) the median). Then for each subsample, we conduct independent double sorts on illiquidity and R&D intensity as in Table 4. A similar procedure is followed for stocks in the all-but-microcaps sample. We report the value-weighted (equal-weighted) average monthly returns (in %) of portfolios based on the full sample (all-but-microcaps sample). Excess return is the portfolio return in excess of the one-month Treasury bill rate. Characteristic-adjusted (Char-adj) returns are computed by adjusting returns using 125 ($5 \times 5 \times 5$) size/book-to-market/momentum benchmark portfolios (as in Daniel et al., 1997). Industry-adjusted (Ind-adj) returns are computed by adjusting returns using the Fama and French 12 industry portfolios (as in Fama and French, 1997). The alphas (in %) are estimated from the time-series regressions of portfolio excess returns on various factor models including the Fama and French (1993) three-factor (FF3) model, the Fama and French (1993) and Carhart (1997) four-factor (FFC) model, the Fama and French (2015) five-factor (FF5) model, and the Hou et al. (2015) q-factor (HXZ) model. H−L denotes the high-minus-low portfolio. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). The sample period is from January 1975 to December 2016. See also notes to Tables 1 and 4.

	Full sample								All-but-microcaps sample							
	Low illiquidity				High illiquidity				Low illiquidity				High illiquidity			
	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L
Panel A: Financially constrained subsample																
Excess return	0.81	0.55	0.97	0.16	0.02	0.39	1.38	1.36	1.03	0.72	0.86	−0.17	0.46	0.54	1.26	0.80
	(3.57)	(2.57)	(2.90)	(0.55)	(0.03)	(0.72)	(2.29)	(3.00)	(4.05)	(2.98)	(3.29)	(−1.06)	(1.03)	(0.75)	(2.38)	(2.64)
Char-adj return	−0.36	−0.49	−0.21	0.15	−0.97	−0.71	0.38	1.35	−0.16	−0.59	−0.22	−0.06	−0.79	−0.71	−0.02	0.77
	(−2.69)	(−4.12)	(−0.99)	(0.58)	(−2.63)	(−2.77)	(1.16)	(3.00)	(−1.11)	(−4.39)	(−1.18)	(−0.28)	(−2.85)	(−2.52)	(−0.09)	(2.15)
Ind-adj return	−0.25	−0.48	−0.08	0.17	−1.02	−0.68	0.32	1.34	−0.17	−0.51	−0.27	−0.10	−0.58	−0.51	0.19	0.77
	(−2.31)	(−3.89)	(−0.41)	(0.76)	(−2.08)	(−1.75)	(0.71)	(3.25)	(−1.32)	(−4.46)	(−1.50)	(−0.50)	(−1.64)	(−1.75)	(0.28)	(2.06)
FF3 α	0.20	−0.14	0.14	−0.06	−1.55	−0.92	0.07	1.62	0.36	0.03	0.09	−0.27	−0.95	−0.73	−0.13	0.82
	(1.37)	(−1.00)	(0.54)	(−0.24)	(−4.78)	(−2.84)	(0.20)	(3.90)	(2.28)	(0.19)	(0.55)	(−1.54)	(−3.28)	(−3.17)	(−0.50)	(2.08)
FFC α	0.13	−0.08	0.14	0.01	−0.62	−0.40	0.68	1.30	0.32	0.03	0.11	−0.21	−0.31	−0.24	0.44	0.75
	(0.89)	(−0.61)	(0.55)	(0.03)	(−1.44)	(−1.25)	(1.85)	(2.63)	(2.01)	(0.18)	(0.70)	(−1.19)	(−1.00)	(−1.44)	(1.24)	(2.07)
FF5 α	−0.07	−0.28	0.12	0.19	−1.25	−0.33	0.62	1.87	0.08	−0.21	−0.15	−0.22	−0.81	−0.28	0.12	0.93
	(−0.46)	(−1.87)	(0.45)	(0.64)	(−3.46)	(−1.07)	(1.83)	(4.60)	(0.54)	(−1.52)	(−0.98)	(−1.21)	(−2.47)	(−1.08)	(0.47)	(2.68)
HXZ α	−0.13	−0.11	0.24	0.37	−0.53	0.14	1.21	1.74	0.04	−0.14	−0.09	−0.13	−0.41	0.04	0.52	0.93
	(−0.85)	(−0.71)	(0.85)	(1.20)	(−1.10)	(0.46)	(3.14)	(3.28)	(0.27)	(−0.95)	(−0.56)	(−0.75)	(−0.89)	(0.15)	(1.80)	(2.02)
Panel B: Financially unconstrained subsample																
Excess return	0.70	0.72	0.82	0.12	0.08	0.61	1.22	1.14	0.81	0.87	0.95	0.14	0.39	0.78	1.07	0.68
	(3.76)	(3.73)	(3.35)	(0.60)	(0.18)	(1.28)	(2.63)	(3.08)	(4.25)	(4.42)	(4.33)	(1.47)	(1.06)	(2.03)	(2.71)	(3.09)
Char-adj return	−0.24	−0.21	−0.11	0.13	−1.00	−0.46	0.12	1.12	−0.23	−0.24	−0.34	−0.11	−0.76	−0.25	−0.05	0.71
	(−2.76)	(−2.64)	(−0.58)	(0.65)	(−4.54)	(−1.93)	(0.48)	(3.30)	(−2.42)	(−2.87)	(−3.02)	(−0.84)	(−4.90)	(−1.26)	(−0.26)	(3.10)
Ind-adj return	−0.33	−0.30	−0.21	0.12	−0.88	−0.39	0.20	1.08	−0.35	−0.36	−0.28	0.07	−0.78	−0.19	−0.04	0.74
	(−4.41)	(−4.18)	(−1.55)	(0.79)	(−2.94)	(−1.18)	(0.57)	(2.98)	(−3.83)	(−4.67)	(−2.91)	(0.60)	(−3.43)	(−0.79)	(−0.19)	(2.80)
FF3 α	0.25	0.21	0.17	−0.08	−0.79	−0.37	0.22	1.01	0.24	0.22	0.27	0.03	−0.49	−0.07	0.13	0.62
	(2.75)	(2.44)	(0.98)	(−0.44)	(−2.97)	(−1.29)	(0.76)	(2.67)	(2.17)	(2.59)	(2.30)	(0.36)	(−3.08)	(−0.36)	(0.64)	(2.83)
FFC α	0.18	0.23	0.26	0.08	−0.45	0.16	0.64	1.09	0.23	0.23	0.32	0.09	−0.18	0.26	0.51	0.69
	(2.06)	(2.32)	(1.54)	(0.45)	(−1.64)	(0.57)	(2.06)	(2.77)	(2.19)	(2.89)	(2.67)	(1.02)	(−1.15)	(1.39)	(2.36)	(3.15)
FF5 α	0.05	−0.03	0.12	0.07	−0.32	0.29	0.88	1.20	−0.04	0.01	0.08	0.12	−0.06	0.42	0.72	0.78
	(0.53)	(−0.32)	(0.71)	(0.38)	(−1.22)	(1.09)	(2.96)	(3.19)	(−0.44)	(0.17)	(0.78)	(1.30)	(−0.35)	(2.31)	(3.45)	(3.85)
HXZ α	0.01	−0.05	0.24	0.23	−0.16	0.70	1.34	1.50	−0.04	0.05	0.15	0.19	0.09	0.67	0.96	0.87
	(0.07)	(−0.46)	(1.30)	(1.13)	(−0.52)	(2.36)	(3.61)	(3.54)	(−0.38)	(0.56)	(1.30)	(1.61)	(0.43)	(3.03)	(3.95)	(4.19)

Table 6
Double sorts on illiquidity and R&D intensity after controlling for innovation ability

The table reports monthly returns of portfolios sorted on illiquidity and research and development (R&D) intensity for both low innovation ability and high innovation ability subsamples. Following Cohen et al. (2013), innovation ability is computed as the average of the slope coefficient estimates from the rolling time-series regressions of firm-level sales growth on the past five years of R&D investments. Illiquidity is measured by the average bid-ask spread. R&D intensity is measured by $RDME$, which is R&D expenditure scaled by market value of equity. The full sample includes all NYSE-, AMEX-, and NASDAQ-listed non-financial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. The all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and we use the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates for portfolio analysis. At the end of June of each year t, we first assign stocks in the full sample to the low (high) innovation ability subsample based on the median value of the distribution of the innovation ability (i.e., low (high) innovation ability subsample includes stocks below (above) the median). Then for each subsample, we conduct independent double sorts on illiquidity and R&D intensity as in Table 4. A similar procedure is followed for stocks in the all-but-microcaps sample. We report the value-weighted (equal-weighted) average monthly returns (in %) of portfolios based on the full sample (all-but-microcaps sample). Excess return is the portfolio return in excess of the one-month Treasury bill rate. Characteristic-adjusted (Char-adj) returns are computed by adjusting returns using 125 (5 × 5 × 5) size/book-to-market/momentum benchmark portfolios (as in Daniel et al., 1997). Industry-adjusted (Ind-adj) returns are computed by adjusting returns using the Fama and French 12 industry portfolios (as in Fama and French, 1997). The alphas (in %) are estimated from the time-series regressions of portfolio excess returns on various factor models including the Fama and French (1993) three-factor (FF3) model, the Fama and French (1993) and Carhart (1997) four-factor (FFC) model, the Fama and French (2015) five-factor (FF5) model, and the Hou et al. (2015) q-factor (HXZ) model. H−L denotes the high-minus-low portfolio. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). The sample period is from January 1975 to December 2016. See also notes to Tables 1 and 4.

	Full sample								All-but-microcaps sample							
	Low illiquidity				High illiquidity				Low illiquidity				High illiquidity			
	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L
Panel A: Low innovation ability subsample																
Excess return	0.61	0.57	0.79	0.18	0.34	0.91	1.02	0.68	0.95	0.86	0.98	0.03	0.74	1.02	1.47	0.73
	(2.86)	(2.10)	(3.09)	(0.86)	(0.77)	(2.37)	(2.20)	(2.00)	(4.62)	(3.73)	(3.90)	(0.23)	(2.13)	(2.83)	(3.25)	(3.13)
Char-adj return	−0.11	−0.29	−0.24	−0.13	−0.65	−0.27	−0.08	0.57	−0.17	−0.26	−0.25	−0.08	−0.31	−0.16	0.24	0.55
	(−0.81)	(−1.97)	(−1.36)	(−0.68)	(−2.82)	(−1.27)	(−0.31)	(1.98)	(−1.46)	(−2.44)	(−1.94)	(−0.52)	(−2.14)	(−1.40)	(1.50)	(2.66)
Ind-adj return	−0.33	−0.38	−0.23	0.10	−0.78	−0.11	−0.04	0.74	−0.14	−0.27	−0.22	−0.08	−0.35	−0.09	0.28	0.63
	(−2.64)	(−2.81)	(−1.25)	(0.44)	(−3.11)	(−0.54)	(−0.15)	(2.51)	(−1.31)	(−2.37)	(−1.43)	(−0.56)	(−1.82)	(−0.43)	(1.03)	(3.12)
FF3 α	0.13	−0.04	0.18	0.05	−0.58	−0.01	0.04	0.62	0.37	0.16	0.30	−0.07	−0.21	0.06	0.36	0.57
	(1.01)	(−0.24)	(0.85)	(0.20)	(−2.11)	(−0.02)	(0.14)	(1.99)	(2.90)	(1.45)	(1.91)	(−0.47)	(−1.45)	(0.50)	(1.77)	(2.72)
FFC α	0.10	−0.10	0.16	0.06	−0.23	0.20	0.32	0.55	0.37	0.19	0.34	−0.03	0.04	0.27	0.70	0.66
	(0.71)	(−0.55)	(0.80)	(0.30)	(−0.80)	(0.72)	(1.20)	(1.97)	(3.05)	(1.82)	(2.12)	(−0.22)	(0.28)	(2.07)	(3.16)	(3.13)
FF5 α	−0.12	−0.35	0.01	0.13	−0.35	0.23	0.39	0.74	0.08	−0.07	0.17	0.09	−0.03	0.22	0.65	0.68
	(−0.94)	(−1.91)	(0.05)	(0.56)	(−1.18)	(0.89)	(1.38)	(2.01)	(0.75)	(−0.58)	(1.09)	(0.67)	(−0.20)	(1.59)	(2.95)	(3.39)
HXZ α	−0.07	−0.41	0.01	0.08	−0.10	0.48	0.63	0.73	0.13	0.02	0.20	0.07	0.13	0.41	1.03	0.90
	(−0.55)	(−2.24)	(0.05)	(0.36)	(−0.33)	(1.83)	(2.10)	(2.03)	(0.94)	(0.12)	(1.19)	(0.55)	(0.71)	(2.86)	(3.90)	(4.16)
Panel B: High innovation ability subsample																
Excess return	0.81	0.92	1.01	0.20	0.44	1.21	1.34	0.90	0.67	0.87	0.90	0.23	0.53	0.78	1.37	0.84
	(4.26)	(3.39)	(3.16)	(0.75)	(1.05)	(3.14)	(3.31)	(2.63)	(3.27)	(4.23)	(3.81)	(1.52)	(1.67)	(2.22)	(3.34)	(3.15)
Char-adj return	−0.02	0.00	−0.24	−0.22	−0.72	−0.11	0.05	0.77	−0.20	−0.09	−0.18	0.02	−0.43	−0.33	0.13	0.56
	(−0.17)	(−0.03)	(−1.02)	(−0.85)	(−3.33)	(−0.60)	(0.21)	(2.38)	(−1.45)	(−0.69)	(−1.19)	(0.15)	(−2.66)	(−2.46)	(0.72)	(2.23)
Ind-adj return	−0.29	−0.16	−0.12	0.17	−0.59	0.05	0.16	0.75	−0.34	−0.21	−0.24	0.10	−0.40	−0.25	0.27	0.67
	(−2.28)	(−1.17)	(−0.64)	(0.75)	(−2.52)	(0.26)	(0.60)	(2.19)	(−2.69)	(−1.62)	(−1.48)	(0.64)	(−1.92)	(−1.37)	(0.93)	(2.28)
FF3 α	0.31	0.27	0.34	0.03	−0.46	0.30	0.42	0.88	0.15	0.27	0.29	0.14	−0.30	−0.09	0.49	0.79
	(2.60)	(1.43)	(1.47)	(0.13)	(−1.77)	(1.34)	(1.62)	(2.65)	(1.01)	(2.28)	(1.62)	(0.94)	(−1.81)	(−0.49)	(2.32)	(3.16)
FFC α	0.22	0.33	0.34	0.12	−0.23	0.64	0.77	1.00	0.12	0.29	0.31	0.19	−0.08	0.18	0.77	0.85
	(1.90)	(1.70)	(1.56)	(0.52)	(−0.93)	(2.62)	(2.62)	(2.69)	(0.84)	(2.62)	(1.77)	(1.26)	(−0.50)	(0.98)	(3.67)	(3.34)
FF5 α	0.15	0.00	0.21	0.06	−0.26	0.55	0.79	1.05	−0.14	0.04	0.03	0.17	−0.22	0.11	0.77	0.99
	(1.18)	(−0.02)	(1.08)	(0.28)	(−0.95)	(2.21)	(2.75)	(2.73)	(−1.05)	(0.36)	(0.15)	(1.05)	(−1.11)	(0.54)	(3.87)	(3.64)
HXZ α	0.11	0.01	0.22	0.11	−0.10	0.81	1.19	1.29	−0.14	0.05	0.14	0.28	−0.06	0.31	1.03	1.09
	(0.84)	(0.05)	(1.11)	(0.50)	(−0.34)	(2.93)	(3.56)	(3.09)	(−0.88)	(0.41)	(0.70)	(1.57)	(−0.32)	(1.52)	(4.31)	(3.59)

Table 7

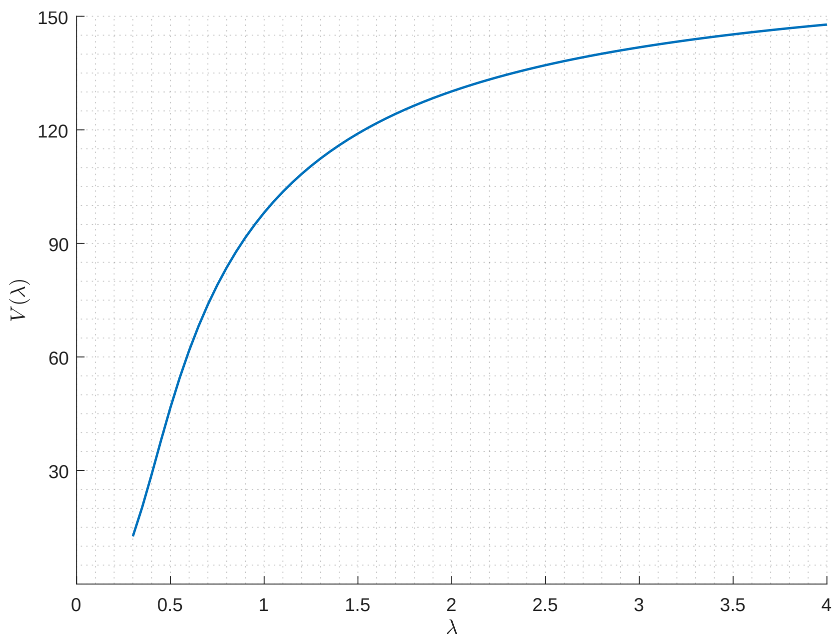
Double sorts on illiquidity and R&D intensity after controlling for product market competition

The table reports monthly returns of portfolios sorted on illiquidity and research and development (R&D) intensity for both low competition and high competition subsamples. Product market competition is measured by the Herfindahl-Hirschman Index (HHI) as in Hou and Robinson (2006). Illiquidity is measured by the average bid-ask spread. R&D intensity is measured by $RDME$, which is R&D expenditure scaled by market value of equity. The full sample includes all NYSE-, AMEX-, and NASDAQ-listed nonfinancial ordinary common stocks for which both a valid (i.e., nonmissing) R&D intensity estimate in a given year t and a valid illiquidity estimate in a given month τ are available. The all-but-microcaps sample excludes stocks with an end-of-June market value of equity below the 20th percentile of the NYSE market capitalization distribution and we use the remaining stocks with both nonmissing R&D intensity and valid illiquidity estimates for portfolio analysis. At the end of June of each year t, we first assign stocks in the full sample to the low (high) competition subsample based on the median value of the distribution of HHI (i.e., low (high) competition subsample includes stocks above (below) the median). Then for each subsample, we conduct independent double sorts on illiquidity and R&D intensity as in Table 4. A similar procedure is followed for stocks in the all-but-microcaps sample. We report the value-weighted (equal-weighted) average monthly returns (in %) of portfolios based on the full sample (all-but-microcaps sample). Excess return is the portfolio return in excess of the one-month Treasury bill rate. Characteristic-adjusted (Char-adj) returns are computed by adjusting returns using 125 (5 × 5 × 5) size/book-to-market/momentum benchmark portfolios (as in Daniel et al., 1997). Industry-adjusted (Ind-adj) returns are computed by adjusting returns using the Fama and French 12 industry portfolios (as in Fama and French, 1997). The alphas (in %) are estimated from the time-series regressions of portfolio excess returns on various factor models including the Fama and French (1993) three-factor (FF3) model, the Fama and French (1993) and Carhart (1997) four-factor (FFC) model, the Fama and French (2015) five-factor (FF5) model, and the Hou et al. (2015) q-factor (HXZ) model. H−L denotes the high-minus-low portfolio. Numbers in parentheses are t-statistics adjusted following Newey and West (1987). The sample period is from January 1975 to December 2016. See also notes to Tables 1 and 4.

	Full sample							
	Low illiquidity				High illiquidity			
	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L
Panel A: Low competition subsample								
Excess return	0.72	0.64	0.71	−0.01	−0.49	0.43	1.19	1.68
	(3.74)	(3.20)	(2.90)	(−0.04)	(−0.87)	(0.84)	(2.19)	(3.90)
Char-adj return	−0.27	−0.34	−0.27	0.00	−1.31	−0.57	0.10	1.41
	(−2.82)	(−3.18)	(−2.06)	(−0.01)	(−4.32)	(−2.38)	(0.34)	(3.51)
Ind-adj return	−0.41	−0.36	−0.31	0.10	−1.42	−0.54	0.09	1.51
	(−4.78)	(−4.26)	(−3.25)	(0.69)	(−3.29)	(−1.52)	(0.24)	(3.70)
FF3 α	0.25	0.09	−0.06	−0.31	−1.79	−0.68	−0.11	1.68
	(2.21)	(0.82)	(−0.40)	(−1.61)	(−5.77)	(−2.37)	(−0.36)	(4.02)
FFC α	0.16	0.11	0.01	−0.15	−1.06	−0.20	0.41	1.47
	(1.30)	(1.04)	(0.08)	(−0.83)	(−2.76)	(−0.72)	(1.27)	(2.89)
FF5 α	0.06	−0.20	−0.20	−0.26	−1.25	−0.11	0.48	1.73
	(0.47)	(−2.06)	(−1.44)	(−1.32)	(−4.17)	(−0.39)	(1.56)	(4.16)
HXZ α	0.02	−0.15	−0.06	−0.08	−0.83	0.40	1.02	1.85
	(0.19)	(−1.27)	(−0.37)	(−0.41)	(−2.13)	(1.45)	(2.80)	(3.46)
Panel B: High competition subsample								
Excess return	0.64	0.95	1.32	0.68	−0.26	0.44	1.60	1.86
	(2.88)	(4.73)	(4.75)	(2.92)	(−0.44)	(0.83)	(2.73)	(4.55)
Char-adj return	−0.40	−0.07	0.21	0.61	−1.25	−0.54	0.50	1.75
	(−3.84)	(−0.72)	(1.18)	(2.95)	(−4.22)	(−2.12)	(1.75)	(5.05)
Ind-adj return	−0.37	−0.22	0.22	0.59	−1.27	−0.68	0.46	1.73
	(−3.00)	(−2.82)	(1.13)	(2.92)	(−3.15)	(−1.75)	(1.13)	(4.77)
FF3 α	0.17	0.45	0.69	0.52	−1.30	−0.58	0.48	1.78
	(1.28)	(4.82)	(3.78)	(2.27)	(−3.68)	(−1.97)	(1.52)	(4.42)
FFC α	0.25	0.49	0.77	0.52	−0.66	0.01	1.08	1.74
	(1.42)	(5.01)	(4.08)	(1.91)	(−1.68)	(0.03)	(3.20)	(4.10)
FF5 α	0.26	0.34	0.62	0.36	−0.28	0.24	1.51	1.79
	(1.62)	(3.18)	(3.19)	(1.39)	(−0.80)	(0.73)	(3.84)	(3.71)
HXZ α	0.36	0.35	0.77	0.41	0.01	0.69	1.88	1.87
	(1.59)	(3.30)	(3.81)	(1.35)	(0.22)	(1.79)	(4.39)	(3.51)

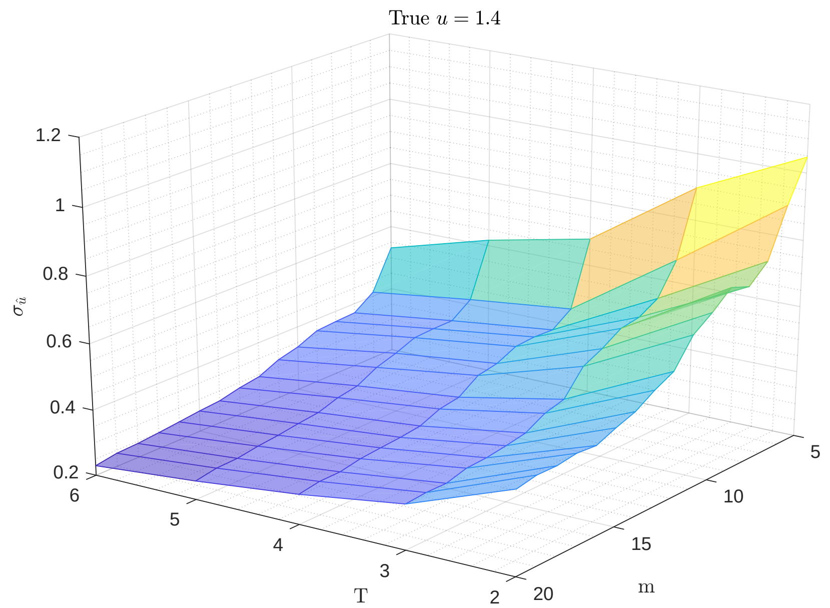
	All-but-microcaps sample							
	Low illiquidity				High illiquidity			
	$RDME_L$	$RDME_M$	$RDME_H$	H−L	$RDME_L$	$RDME_M$	$RDME_H$	H−L
Panel A: Low competition subsample								
Excess return	0.72	0.80	0.71	−0.01	0.74	0.76	1.36	0.60
	(3.03)	(3.56)	(2.63)	(−0.05)	(1.62)	(1.64)	(2.74)	(2.12)
Char-adj return	−0.21	−0.15	−0.27	−0.06	−0.38	−0.38	0.09	0.47
	(−1.54)	(−1.23)	(−2.52)	(−0.40)	(−1.91)	(−2.64)	(0.46)	(2.07)
Ind-adj return	−0.25	−0.22	−0.30	−0.05	−0.21	−0.20	0.34	0.55
	(−2.12)	(−1.94)	(−2.46)	(−0.55)	(−0.67)	(−1.01)	(1.16)	(2.56)
FF3 α	0.18	0.24	0.02	−0.16	−0.37	−0.30	0.21	0.58
	(1.13)	(1.74)	(0.14)	(−1.07)	(−1.66)	(−1.88)	(0.99)	(2.17)
FFC α	0.16	0.26	0.10	−0.06	0.08	0.09	0.66	0.58
	(1.02)	(2.05)	(0.74)	(−0.47)	(0.30)	(0.28)	(3.07)	(2.02)
FF5 α	−0.10	−0.04	−0.17	−0.07	−0.43	−0.04	0.42	0.85
	(−0.73)	(−0.34)	(−1.30)	(−0.48)	(−1.86)	(−0.24)	(1.81)	(3.56)
HXZ α	−0.11	0.01	−0.09	0.02	−0.21	0.17	0.76	0.97
	(−0.64)	(0.07)	(−0.58)	(0.13)	(−0.58)	(0.93)	(2.97)	(2.96)
Panel B: High competition subsample								
Excess return	0.81	0.80	1.01	0.20	0.36	0.95	1.29	0.93
	(3.16)	(3.68)	(4.11)	(0.85)	(0.85)	(2.16)	(2.59)	(3.08)
Char-adj return	−0.14	−0.01	0.15	0.29	−0.75	−0.19	0.19	0.94
	(−0.78)	(−0.08)	(0.88)	(1.20)	(−4.93)	(−1.33)	(0.90)	(3.37)
Ind-adj return	−0.11	−0.12	−0.01	0.10	−0.58	−0.12	0.21	0.79
	(−0.53)	(−0.82)	(−0.07)	(0.44)	(−2.55)	(−0.52)	(0.71)	(3.35)
FF3 α	0.31	0.28	0.51	0.20	−0.61	0.10	0.32	0.93
	(1.59)	(2.14)	(2.77)	(0.83)	(−3.59)	(0.57)	(1.40)	(3.02)
FFC α	0.36	0.27	0.44	0.08	−0.35	0.38	0.66	1.01
	(1.98)	(2.10)	(2.32)	(0.34)	(−2.00)	(1.99)	(2.90)	(3.43)
FF5 α	0.17	0.12	0.35	0.18	−0.42	0.54	0.68	1.10
	(0.89)	(1.00)	(1.83)	(0.64)	(−2.46)	(3.01)	(3.06)	(3.90)
HXZ α	0.18	0.09	0.35	0.17	−0.28	0.72	1.02	1.30
	(0.85)	(0.67)	(1.87)	(0.57)	(−1.46)	(3.07)	(3.78)	(4.02)

Figure 1
R&D valuation and success intensity



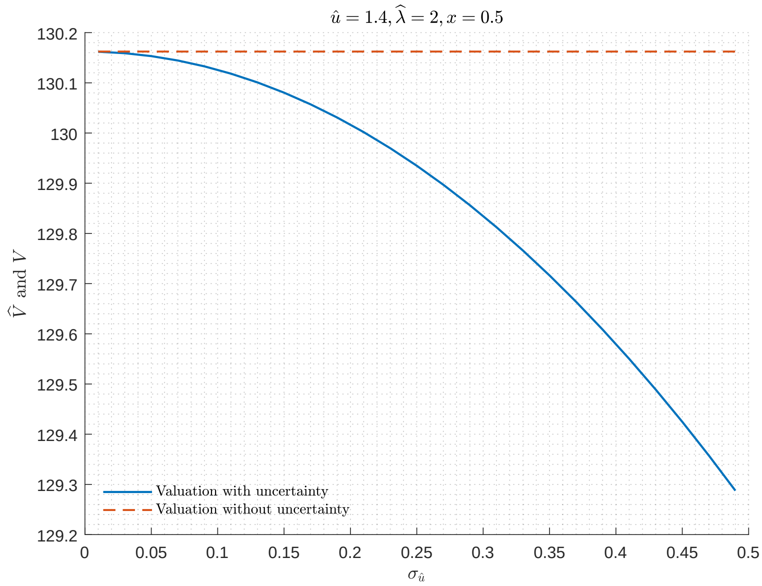
The figure plots $V(\lambda, c)$ as a function as λ. The parameters are set as following: $K = 5$, $n = 2$, $c = 10$, $\sigma = 0.2$, $\mu = 0.1$, $r_f = 0.03$, $\zeta = 0.08$, $\phi = 0.05$, $a = 10$, and $b = 0.6$.

Figure 2
The effects of varying T and m on $\sigma_{\hat{u}}$



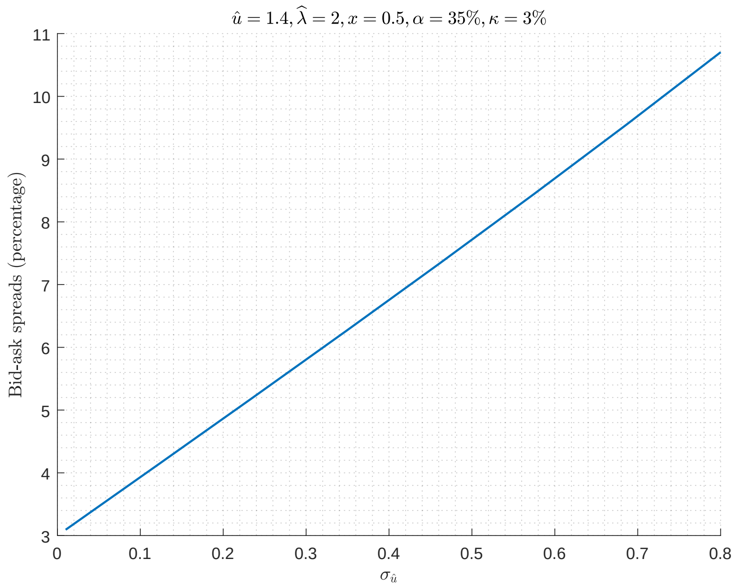
The figure plots $\sigma_{\hat{u}}$ estimated using varying sizes of simulated $\mathbf{N}$ and $\mathbf{X}$ up to $m = 20$ and $T = 6$. The true parameter $u = 1.4$.

Figure 3
Uncertainty and R&D valuation



The figure plots $\widehat{V}(\hat{u}, \sigma_{\hat{u}}, c)$ as a function of $\sigma_{\hat{u}}$. The parameters are set as following: $K = 5$, $n = 2$, $c = 10$, $\sigma = 0.2$, $\mu = 0.1$, $r_f = 0.03$, $\zeta = 0.08$, $\phi = 0.05$, $a = 10$, and $b = 0.6$. Additional parameter values are shown on the top of the plot.

Figure 4
Uncertainty and bid-ask spreads



The figure plots relative bid-ask spread as a function of $\sigma_{\hat{u}}$. The parameters are set as following: $K = 5$, $n = 2$, $c = 10$, $\sigma = 0.2$, $\mu = 0.1$, $r_f = 0.03$, $\zeta = 0.08$, $\phi = 0.05$, $a = 10$, $b = 0.64$. Additional parameter values are shown on the top of the plot.