Innovation under Ambiguity and Risk

Gabriela Coiculescu[*], Yehuda Izhakian[†] and S. Abraham Ravid[‡§]

July 29, 2019

Abstract

We explore the implications of ambiguity (Knightian uncertainty) and risk for innovation decisions through the lens of real options. Our hypotheses are supported by a real options model, and are based on a new risk- and outcome-independent measure of ambiguity. We expect ambiguity to decrease innovation investment, whereas risk should increase innovation investment. The latter prediction is also consistent with prior work. Empirically, we find a consistently significant negative effect of ambiguity on R&D investment, as well as on patents and citations. We also find a significant positive effect of risk on R&D, but the effect of risk on patents and citations is negative and significant, which suggests that in the face of higher risk firms may wait and delay patenting. The effect of ambiguity is more important for high tech firms, which invest heavily in research and in patenting, consistent with our intuition.

Keywords: Ambiguity measurement, Ambiguity aversion, Risk aversion, Innovation, Patents

JEL Classification: C65, D81, D83, G13, G22, G30, G31, G39, O30, O32

[*]Sy Syms School of Business, Yeshiva University; gabriela.coiculescu@yu.edu

[†]Zicklin School of Business, Baruch College; yud@stern.nyu.edu

[‡]Sy Syms School of Business, Yeshiva University; ravid@yu.edu

[§]We are indebted to Yakov Amihud, Slava Fos, Elad Green, Shu Han, Deepak Hegde, James Ohlson, Joaquin Poblete, Alejandro Rivera, Eduardo Schwartz, Alex Vedrashko, Tom Åstebro, as well as participants at the 2018 Midwest Finance Association conference, the 2018 Northwestern (Searle Center) Conference on Innovation Economics, and seminar participants at HEC Paris, Lund University, New York University, Simon Fraser University, Hong Kong Poly, and Yeshiva University, for helpful discussions and comments.

Electronic copy available at: https://ssrn.com/abstract=3428896

1 Introduction and Relation to the Literature

A large and growing body of literature investigates the determinants of innovation decisions, including industry competition (Aghion et al., 2005), institutional ownership (Aghion et al., 2013) and organizational structure (Lerner et al., 2011, Seru, 2014, Bernstein, 2015). Special attention has been paid to *risk* as a driver of innovative activity. However, *risk*, the uncertainty of outcomes, usually measured by the variance of equity returns, assumes a unique known distribution of future outcomes. In reality, it may be very difficult (and, perhaps, impossible) to predict a distribution of future outcomes for a new innovative product such as a new drug (Krieger et al., 2017). Therefore, the concept of *ambiguity*—the uncertainty of probabilities—seems a natural lens through which managers may also assess future prospects. Our paper introduces a new fully developed concept of ambiguity which is theoretically sound and empirically testable, to investigate how each type of uncertainty affects the innovative decision, and which type of uncertainty may be more salient for innovating firms.

Several early studies analyzed investment decisions as real options (Brennan and Schwartz, 1985, McDonald and Siegel, 1986). Building on this concept, Schwartz (2004) and Kraft et al. (2018) view R&D and patent decisions as real options, implying a positive effect of risk on R&D. Bloom (2007, 2014) shows, however, that risk can negatively affect R&D investment, due to adjustment costs and an increase in the value of the option to wait.

As in Schwartz (2004), Bloom (2014) and Kraft et al. (2018), we consider patent and R&D decisions as real options. However, while in these studies the values of real options are subject to risk only, a different strand of literature shows that option values are significantly affected by ambiguity (Izhakian and Yermack, 2017, Augustin and Izhakian, 2019). Our testable hypotheses combine the insights from these two approaches, and are supported by a one-period stylized model of an optimal investment decision through the lens of real options in the presence of ambiguity (presented in the Appendix).

Experimental studies show that decision makers tend to be ambiguity averse in the sense discussed in this paper. Ellsberg (1961) and Halevy (2007) show that, while making decisions, decision makers prefer alternatives involving clear probabilities (risk, the *known unknowns*) over alternatives involving vague probabilities (ambiguity, the *unknown unknowns*), even if normative theories (Von-Neumann and Morgenstern, 1944, Savage, 1954) imply indifference. This phenomenon of am-

Electronic copy available at: https://ssrn.com/abstract=3428896

biguity aversion has been shown to be economically relevant and to persist in experimental market settings and among business owners and managers.[1]

The effect of ambiguity on investment decisions is very different than the effect of risk. Any investor will invest more as risk goes up, since higher risk increases the upside potential and the value of the option. However, intuitively, an ambiguity-averse investor overweights the likelihood of bad outcomes and underweights the likelihood of good outcomes (e.g., Tversky and Kahneman, 1992, Izhakian, 2017).[2] Thus, higher ambiguity reduces the perceived attractiveness of investment opportunities (i.e., a lower perceived expected payoff), which negatively affects investment decisions.

While ambiguity may affect any investment decision, in practice, we expect our ideas to have more bite in cases where there is high uncertainty regarding the future prospects of the investment in question, rather than say, in renovations or expansions of existing product lines. For example, it would be difficult to view an investment in refurbishing an office building as a real option. However, an investment in a new lab, for example, may create a real option to commercially license a new drug, and thus is closer to our hypotheses. The former would appear under capital expenditures (CAPEX) and the latter would appear under R&D in firms' accounting statements. This distinction is consistent with the accounting treatment of R&D as expenses and the requirement, on the other hand, to depreciate CAPEX investments. At the same time, following Kumar and Li (2016), we take into account the possibility that R&D expenditures might actually understate long-run innovative capacity investments. Therefore, while in our main tests we measure innovation investments using R&D expenses, in a robustness test we use the sum of R&D and CAPEX, and we find that the results continue to hold.

We find a significant negative effect of ambiguity on measures of innovation in various stages of product development including R&D investment and patents. These findings are consistent with our stylized model and are also in line with Herron and Izhakian (2017), who show that ambiguity matters to firm payout policies. Our study is also related to the literature studying the effect of risk on corporate real investments which comes up with a mixed verdict. One strand in this literature concludes that risk increases corporate investments. This literature goes back a few decades. Early work (e.g., Hartman, 1972, Abel, 1983) suggests that since the marginal value

[1]See, for example, Mangelsdorff and Weber (1994), Viscusi and Chesson (1999), Abdellaoui et al. (2005), Du and Budescu (2005), Maffioletti and Santoni (2005), Wakker et al. (2007).

[2]Behavior consistent with this way of thinking was found in several experimental studies (e.g., Wu and Gonzalez, 1999, Abdellaoui and Kemel, 2013, Crockett et al., 2019).

Electronic copy available at: https://ssrn.com/abstract=3428896

product of capital is a convex function of the risk faced by the firm, greater risk raises the marginal valuation of one additional unit of capital, thereby increasing investment. The other strand in this literature concludes that risk decreases corporate investments due to the irreversibility effect: delaying the decision to invest in order to wait for new information (e.g., Bernanke, 1983, Pindyck, 1988). In other words, the opportunity cost associated with an irreversible investment increases in risk. However, according to Caballero (1991) and Abel and Eberly (1994), even in the presence of irreversibility, risk has a non-negative effect on investment if the firm operates in a competitive market (see also Abel and Eberly, 1994). Bloom (2007, 2014) follows these ideas and shows that in a dynamic framework risk may lead firms to react slowly or to reduce investment. We indeed document a significant negative effect of risk on patents and citations, which is consistent with Bloom's (2007) framework. However, in our sample, we find a significant positive effect of risk on R&D decisions, which is consistent with the real options perspective.[3]

Other empirical treatments of the effect of risk on innovation investment include Bernstein et al. (2017) who suggest that macroeconomic risk, measured by negative housing shocks, reduces employees' interest in risky and exploratory projects. Krieger et al. (2017) investigate the tradeoff between conservative and riskier investments in drug development.

The remainder of this paper is organized as follows. Section 2 presents a discussion of ambiguity and develops the hypotheses. Section 3 describes the sample selection and data construction, including the estimates of the ambiguity and risk variables that are central to our investigation. Section 4 presents the empirical methodology, and 5 reports the results. Section 6 concludes.

2 Ambiguity

2.1 Decision theoretic framework

There have been several earlier theoretical studies of ambiguity, and our work extends and generalizes their thinking. A path-breaking set of papers provides an initial axiomatization for decision making in the presence of ambiguity (Gilboa and Schmeidler, 1989, Schmeidler, 1989). However,

[3]In a recent contribution, related, but distinct from ours, Kumar and Li (2018) document a positive association between idiosyncratic volatility and the response rate of subsequent innovation-related investment (either R&D or the sum of R&D and CAPEX). In Kumar and Li (2018), the response rate of innovation-related investment is defined as the *absolute* percentage change in innovation investment. This result is interpreted in light of a feedback model in which idiosyncratic volatility proxies for investors' private information regarding the prospects of the firm's innovation projects.

Electronic copy available at: https://ssrn.com/abstract=3428896

these papers do not separate ambiguity from attitude towards ambiguity. Later studies develop a theory where ambiguity is separated from preferences for ambiguity but, in these studies, preferences are outcome-dependent (e.g., Klibanoff et al., 2005, Nau, 2006, Chew and Sagi, 2008).

To derive a risk-independent measure of ambiguity, preferences for ambiguity must be outcome independent, so that the measure itself is outcome-independent. To illustrate, consider an innovation investment whose payoff is determined by a flip of an unbalanced coin, and for which the manager does not know the odds of heads or tails. The payoff of the innovation investment is \$1,000,000 in case of heads and \$0 in case of tails. Suppose that prior to flipping the coin, the payoff in case of heads is suddenly changed to \$2,000,000. Since this change in payoff provides no new information about the probabilities involved, the manager has no reason to change the assessed probabilities or the perceived degree of ambiguity. In other words, ambiguity is outcome-independent up to a state space partition, since it applies exclusively to probabilities. However, the risk does increase in this example, since it is outcome-dependent.[4]

Izhakian's (2017) expected utility with uncertain probabilities (EUUP) framework introduces outcome-independent preferences for ambiguity, which allow us to distinguish the concepts of risk and ambiguity, and to specify distinct preferences with respect to both. Importantly, EUUP allows us to measure ambiguity independently of risk and of the attitude toward risk (Izhakian, 2018). Under EUUP, a decision maker acts as if she solves a two-stage decision-making problem. In the first stage, she forms a representation of perceived probabilities for each relevant event, based on her perceived ambiguity and her attitude toward this ambiguity. In the second stage, she considers the expected utility associated with a set of possible outcomes, where the expectation is taken with respect to her perceived probabilities. The main idea of EUUP is that in the presence of ambiguity (i.e., when probabilities are uncertain), preferences for ambiguity are applied exclusively to probabilities (outcome-independence) such that aversion to ambiguity is defined as aversion to mean-preserving spreads in probabilities. As such, the Rothschild and Stiglitz (1970) approach, typically applied to outcomes when examining risk, is applied to probabilities when examining ambiguity. In this framework, an ambiguity-averse decision maker overweights the probabilities of bad outcomes and underweights the probabilities of good outcomes. In particular, the higher the ambiguity or the aversion to ambiguity, the lower the perceived probabilities of good outcomes

[4] A similar example is presented in Izhakian (2018).

Electronic copy available at: https://ssrn.com/abstract=3428896

and the higher the perceived probabilities of bad outcomes. As a result, when ambiguity rises, the perceived expected utility computed with perceived probabilities falls. We formally describe this decision theory framework in the Appendix.

Based on EUUP, Izhakian (2018) shows that the degree of ambiguity can be measured by the volatility of probabilities—just as the degree of risk has been measured by the volatility of outcomes. This measure accounts for the variance of all moments of the outcome distribution, and can be utilized in empirical investigations.[5]

2.2 Real options view

Innovation investments (R&D or patent) can be viewed as real options. Consider, for example, a decision to invest in a new drug or a new technology. The firm will make an initial investment in innovation only if the value of the option created is positive, given the "exercise price" (i.e., the eventual outlay for production). R&D provides the foundation to develop a new enterprise. However, the firm can also decide to shelve the drug at a later stage if more information suggests that the likelihoods of unfavorable outcome are high. Patents create real options as well. For example, an article in the trade publication Tomorrow's Pharmacist (Torjesen, 5/12/2015) states an open secret in the industry: "Pharmaceutical companies will patent any molecule that shows promise early in the development process."[6] In general, the drug development process, with the various phases of FDA approval, can be viewed as a sequence of real options.

It is well known that the value of a (real) option increases in risk. In contrast, this is not the case for the effect of ambiguity. When valuing a real option using EUUP or any other frameworks of decision making under ambiguity, decision makers act as if they overweight the probabilities of bad outcomes (out of the money) and underweight the probabilities of good outcomes (in the money). Thus, ambiguity reduces the perceived value of the option. As we show in the appendix, a real option, valued by an ambiguity-averse decision maker, declines in value as ambiguity increases and increases in value as risk increases. In particular, the higher the ambiguity, the lower is the perceived probability of a positive payoff, which the decision maker uses to form the expected value of the (real) option. For a similar reason, employees tend to exercise their options early when the

[5]The EUUP measure of ambiguity is employed in several empirical studies using equity market data (e.g., Izhakian and Yermack, 2017, Brenner and Izhakian, 2018, Augustin and Izhakian, 2019).

[6]https://www.pharmaceutical-journal.com/publications/tomorrows-pharmacist/drug-development-the-journey-of-a-medicine-from-lab-to-shelf/20068196.article?firstPass=false.

Electronic copy available at: https://ssrn.com/abstract=3428896

expected ambiguity increases (Izhakian and Yermack, 2017) and CDS spreads decrease in ambiguity (Augustin and Izhakian, 2019).

Assuming no conflicts of interest, managers act to maximize the value of the firm. Thus, higher perceived risk encourages innovative investments. In contrast, higher perceived ambiguity suppresses innovative investments. Next, we provide a simple illustrative numerical example that shows how our intuition works. A more developed theoretical model is provided in the Appendix.

2.3 Binomial example

Consider a one-period binomial real option for a project which requires an eventual investment of \$100. Suppose that the payout of the project may either be $H = \$120$ or $L = \$80$. The firm can buy the option to invest and then decide whether or not to invest the required amount when the state of the world materializes. In the case of the high payoff (i.e., $H = \$120$) the option pays the difference between the investment I (which is \$100) and the project's value (i.e., $H - I = \$120 - \$100$). If the low case materializes, the firm will not pursue the investment. For simplicity, assume that the risk-free rate is zero.

Suppose that the manger is risk neutral.[7] When the probabilities of both the bad and the good outcomes are known to be 50% (no ambiguity is present), the variance of the probabilities is 0. Therefore, the value of the option (in terms of expected utility) is $C = 0.5 \times (120 - 100) = 10$. If the variance of the payoff of the project increases, such that the outcomes in the good and bad states are respectively 130 or 70 (i.e., a higher but mean-preserving spread), then the value of the option increases to $C = 0.5 \times (130 - 100) = 15$. Thus, an increase in risk is associated with a higher value of the option. This is naturally less pronounced for risk-averse decision makers or different but going in the same direction for Black and Scholes type models.[8]

To examine the impact of ambiguity, assume instead that the future payoffs remain the same, 80 or 120, but the probabilities of these future payoffs occurring are ambiguous. The distributions of payoffs can be either $(0.4, 0.6)$ or $(0.6, 0.4)$. The manager, who does not have any information

[7] The EUUP framework allows different combinations of risk attitudes and ambiguity attitudes. Typically, we expect decision makers to be both risk-averse and ambiguity-averse. However, in order to focus on ambiguity, the current example is a simplification in which we have a risk-neutral but ambiguity-averse investor. While options in general can be valued using a framework similar to Black and Scholes, there is no market for the real options we consider in our setting, so arbitrage based option pricing may not be possible.

[8] Consider, for example, a risk-averse decision maker with the utility function $\mathrm{U}(c) = \sqrt{c}$. In this case, the value (in terms of expected utility) of the option on the less risky asset is $C = 0.5 \times \sqrt{120 - 100} = 2.24$, while the value of the option on the more risky asset is $C = 0.5 \times \sqrt{130 - 100} = 2.74$.

Electronic copy available at: https://ssrn.com/abstract=3428896

regarding the precision of these probability estimates, acts as if she assigns an equal weight to each state probability distribution. Thus, the expected probability of the good state is $\mathrm{E}[\varphi(H)] = 0.5 \times 0.4 + 0.5 \times 0.6 = 0.5$ and its variance is $\mathrm{Var}[\varphi(H)] = 0.5 \times (0.4 - 0.5)^2 + 0.5 \times (0.6 - 0.5)^2 = 0.01$. The same values apply for the bad state. This implies that the degree of ambiguity (expected variance of the probabilities, see Appendix) is $\mho^2 = 0.5 \times 0.01 + 0.5 \times 0.01 = 0.01$.

In EUUP, an ambiguity-averse decision maker forms perceived probabilities by certainty equivalent probabilities and uses them to assess her expected utility. A certainty equivalent probability is the unique certain probability value that the decision maker is willing to accept in exchange for the uncertain probability of a given event. This is analogous to the certainty equivalent outcome (based on risk). By Equation (11) in Appendix A, the perceived probability of the preferable payoff is $\mathrm{E}[\varphi(H)] \times \left(1 + \frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)} \mathrm{Var}[\varphi(H)]\right)$, where $\varphi(\cdot)$ is the marginal probability (probability mass function), and $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)}$ is the coefficient of (constant absolute) ambiguity aversion.

Assume first an ambiguity-neutral decision maker. The preference for ambiguity of this decision maker is characterized by a linear function $\Upsilon(\cdot)$, implying that perceived probabilities are equal to the expected probabilities. Accordingly, the value of the option (in terms of expected utility) remains the same and equal to $C = 0.5 \times (120 - 100) = 10$.

Now assume instead an ambiguity-averse decision maker with a constant absolute ambiguity aversion $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)} = \eta = 2$. Due to aversion to ambiguity, this decision maker does not form perceived probabilities through a linear compounding of probabilities, but aggregates probabilities in a non-linear way as described above. As a result, the value of the option (in terms of expected utility) becomes $C = 0.5 \times (1 - 2 \times 0.01) \times (120 - 100) = 9.8$.[9] For a decision maker with higher aversion to ambiguity, say $\eta = 4$, the value of the option (in terms of expected utility) drops even further to $C = 0.5 \times (1 - 4 \times 0.01) \times (120 - 100) = 9.6$. Thus, an increase in aversion to ambiguity *decreases* the option value. In the data, we naturally cannot observe either aversion to ambiguity or to risk. However, we can compute the degree of ambiguity.

Assume now that the ambiguity of the payoff of the project increases. For example, if future payoffs are distributed either $(0.3, 0.7)$ or $(0.7, 0.3)$ with equal likelihood (a mean-preserving spread in probabilities), then the expected probability of the good (and the bad) state remains unchanged: $\mathrm{E}[\varphi(H)] = 0.5 \times 0.3 + 0.5 \times 0.7 = 0.5$, but the variance of its probabilities increases to $\mathrm{Var}[\varphi(H)] =$

[9]When the decision maker is risk-averse with the utility function $\mathrm{U}(c) = \sqrt{c}$, the value (in terms of expected utility) is $C = 0.5 \times (1 - 2 \times 0.01) \times \sqrt{120 - 100} = 2.19$.

Electronic copy available at: https://ssrn.com/abstract=3428896

$0.5 \times (0.3-0.5)^2 + 0.5 \times (0.7-0.5)^2 = 0.04$, implying a degree of ambiguity of $\mho^2[X] = 0.04$. Assuming a coefficient of ambiguity aversion $\eta = 2$, the value of the option then drops to $C = 0.5 \times (1 - 2 \times 0.04) \times (120 - 100) = 9.2$.

This simple example illustrates our main predictions based on the real options view. An increase in risk (variance of outcomes) increases the value of the real option, thus increasing the investment in innovation. In contrast, an increase in ambiguity decreases option value, leading to a lower investment in innovation. Since risk has been investigated extensively in prior studies, we propose a hypothesis based on our simple real options model, and a competing hypothesis based on Bloom (2007, 2014). To our knowledge, there is no competing hypothesis regarding ambiguity.

2.4 Hypotheses

We propose two competing hypotheses for the effect of risk on innovation.

Hypothesis 1a *Investments in innovation are higher for higher degrees of firm (project) risk.*

Hypothesis 1b *Investments in innovation are lower for higher degree of firm (project) risk.*

Hypothesis 1a coincides with Schwartz (2004) and Kraft et al. (2018), and follows directly from the stylized model presented in the Appendix and illustrated by the binomial example in Section 2.3. This hypothesis also coincides with earlier corporate investments literature (e.g., Hartman, 1972, Abel, 1983). Hypothesis 1b is motivated by Bloom (2007, 2014), who argues that when R&D is below the optimum, firms may want to raise R&D, but higher risk induces a pause in R&D investment ("delay effect"). This hypothesis also coincides with the idea that, as risk increases, the option to wait increases in value (e.g., Bernanke, 1983, Pindyck, 1988).

Higher ambiguity always implies lower perceived probabilities of the good states (in which the innovative investment bears fruit), and therefore a lower value of the (real) option. A lower value of the real option results in less investment in R&D or less patenting.

Hypothesis 2 *Investments in innovation are lower for higher degrees of firm (project) ambiguity.*

Below we test these hypotheses on R&D and patent data.

Electronic copy available at: https://ssrn.com/abstract=3428896

3 Data

The primary data sources for the analysis are the intraday trade and quote (TAQ) data for the estimation of the degrees of ambiguity and risk; the patent database of Kogan et al. (2017) for historical information on patents; and Compustat for accounting data. In robustness tests, we also use institutional ownership data from the Thompson Reuters 13F database[10], as well as the Bushee (1998) classification of institutional owners.

3.1 Sample construction

In order to construct our sample, we start with all firm-quarters with strictly positive sales and assets in the Compustat Fundamentals Quarterly files for fiscal years 1993-2016. We start our sample in 1993, since the TAQ data, which we use to compute our ambiguity and risk measures, is available only from 1993. We organize the data by *calendar quarter-year*. For example, the first quarter of 2000 includes all firm-quarters with fiscal quarter ending in February, March or April 2000. We augment this dataset with the entire history of patents for Compustat firms using the Kogan et al. (2017) patent dataset.[11]

Next, we attempt to identify firm reorganizations that are not accompanied by a change in the Compustat firm identifier (*gvkey*). Specifically, following Bloom et al. (2013), whenever we observe extremely large jumps (greater than 200% or lower than -67%) in annual sales, employment, or assets, we treat the firm as a new entity and assign it a new identifier (*new gvkey*), even if the Compustat *gvkey* remains the same. This approach is more general than including a full set of *gvkey* fixed effects, because it allows the fixed effect to change over time, when the firm undergoes major changes.

As in most other papers on patents, our measure of the patenting process is patent applications. However, patent applications are observed only conditional on the patent being eventually granted. Since our patent data (Kogan et al., 2017) ends in 2010, we are missing patents applied for in the later years of our sample period, but granted after 2010. To reduce this truncation bias (Dass et al., 2017), and following Dong et al. (2017), we drop the last two years of the patent data, ending the

[10]Following Ben-David et al. (2018), after June 2013, we calculate institutional ownership using the 13F data parsed directly from the SEC EDGAR filings system, and available on WRDS.

[11]The Compustat Fundamentals Annual file starts in 1950. The Compustat Fundamentals Quarterly file starts in 1962, but the coverage is sparse until 1982. For each firm, we use the entire patent history from the Kogan et al. (2017) dataset, starting in 1925, the first year of CRSP data.

Electronic copy available at: https://ssrn.com/abstract=3428896

patent sample in 2008. Furthermore, patent citations are subject to truncation, because we only observe citations made by patents granted by 2010. To correct for this bias, we scale the citation count for each patent by the average number of citations received by all patents in the same 3-digit USPTO technology class and filed in the same year. This is the so-called fixed-effects approach from Hall et al. (2001)

We use the historical SIC code from CRSP to identify industries; when the historical SIC code is missing in CRSP, we use the historical SIC code from Compustat. When both are missing, we use the SIC code of the largest business or operating segment from the Compustat Segment Files. We exclude utilities (SIC codes 4900-4999), financials (SIC codes 6000-6999), public service, international affairs firms and non-operating establishments (SIC codes 9000-9999). For R&D, the sample period is 1993-2016.

Since our paper analyzes the effect of ambiguity on both R&D and patenting, we present empirical results for three different samples: the sample of firms with at least one quarter of positive R&D expenditures (*R&D Sample*), the sample of firms with at least one patent application (*Patent Sample*), and the sample of firms with at least one citation (*Citation Sample*), conditional on non-missing data for all variables of interest during the sample period (1993-2016 for R&D and 1993-2008 for patents and citations). This is common for papers on innovation, given that if we include the universe of all firms, most of them have neither any patents nor positive R&D. For all samples, we require firms to have available data for at least four quarters for all variables of interest. In addition, for the *Patent Sample* and the *Citation Sample*, we require firms to have at least four years (16 quarters) of patent data before the first quarter in the sample (the pre-sample period). For firms that enter Compustat after 1993, we use the first four years of data as the pre-sample period, and we include the following years in the sample.[12]

In some specifications, in order to eliminate microstructure effects that might affect our measures of ambiguity and risk, we exclude penny stocks, very small firms and very young firms. Penny stocks are stocks with a price less than $5 at the end of the quarter. Very small firms are firms with a market capitalization less than $10 million at the end of the quarter. Very young firms are firms with less than 5 years in Compustat.

There are 105,037 firm-quarters for 4,053 different firms in the *R&D Sample*, 54,093 firm-

[12]See the discussion in Section 4 of the Blundell et al. (1999) pre-sample mean scaling fixed effect estimator.

Electronic copy available at: https://ssrn.com/abstract=3428896

quarters for 2,108 different firms in the *Patent Sample*, and 51,392 firm-quarters for 1,967 different firms in the *Citation Sample*. The *R&D* and *Patent* samples do not overlap completely: for quarters ending on or before December 31, 2008 (the period when the two samples potentially overlap) only 43,090 firm-quarters are in both samples, while 29,900 firm-quarters are only in the *R&D Sample*, and 11,003 firm-quarters are only in the *Patent Sample*.[13]

3.2 Measures of innovation

The hypotheses derived by our stylized model of real options under ambiguity and risk apply in principle to "investment projects." However, the main source of ambiguity and risk is the innovative activity of the firm, rather than say, routine maintenance. As discussed above, the accounting treatment of R&D (expensing) seems to recognize that R&D buys you an option, rather than a known asset that needs to be depreciated, as is the case for CAPEX. Kumar and Li (2016) point out that part of the capital expenditures of innovative firms may in fact reflect investments in innovative capacity, such as the construction of a research facility or purchasing patents. Hence, R&D expenditures might actually understate the actual investment in innovation for these firms. To address this concern, we measure innovation by both R&D and the sum of R&D and CAPEX, as well as patents and citations.

Our measures of innovation intensity are defined as follows. RD_ASSETS_{t+1} is defined as research and development expenses in quarter $t+1$, scaled by total assets at the beginning of the quarter. It is possible that the firm adjusts its R&D with a lag. Thus, to reflect a potential delayed response of R&D to ambiguity and risk, we also analyze the R&D intensity one year ahead, $RD\_ASSETS_{t+1\ldots t+4}$, which is defined as total research and development expenditures in the four quarters $t+1\ldots t+4$, scaled by total assets at the beginning of quarter $t+1$. For robustness, we use two alternative measures of investment in innovation: $RD\_CAPEX\_ASSETS_{t+1}$ is the sum of R&D and CAPEX, scaled by total assets at the beginning of quarter, and $RD\_ADJ\_ASSETS_{t+1}$ is R&D scaled by assets at the beginning of the quarter, adjusted to include capitalized R&D. To eliminate the effect of outliers, we drop firm-quarters with RD_ASSETS, RD_CAPEX_ASSETS or RD_ADJ_ASSETS above the 99$^{\text{th}}$ percentile.

[13]The fact that 20.34% (11,003 out of 54,093) of firm-quarters in the *Patent Sample* do not have positive R&D expenditures during the sample period is consistent with Koh and Reeb (2015), who find that a significant number of firms with missing R&D in Compustat actually file and receive patents.

Electronic copy available at: https://ssrn.com/abstract=3428896

To measure innovation intensities, we also consider patents and citations, up to three years (12 quarters) ahead. $PATENTS_{t+1}$ is the number of patents applied for during the quarter, conditional on being granted by 2010. To reduce the bias caused by the application-grant lag, following Dong et al. (2017), we end the sample for patents and citations regressions in 2008. We follow numerous innovation papers, including recent contributions (e.g., He and Tian, 2013, Dong et al., 2017), and use citation counts as a proxy for the quality of the firm's patents (i.e., citations-weighted patents, Trajtenberg, 1990). $CITATIONS_{t+1}$ is the number of citations received by 2010 by the patents that the firm filed during quarter $t+1$, excluding self-cites, and corrected for citation truncation using the fixed-effects approach described by Hall et al. (2001). Namely, the raw number of citations, excluding self-cites, is scaled by the average number of citations received by all patents in the same 3-digit USPTO technology class filed in the same year.

3.3 Estimating ambiguity

Our goal is to analyze the effect of ambiguity on innovation, and ideally, we would like to estimate the ambiguity associated with the firm's innovative projects. In practice, we can only observe stock returns. Therefore, our empirical measure is the ambiguity extracted from a company's equity. Intuitively, ambiguity represents the uncertainty in future outcome *probabilities*, as opposed to risk, which measures the uncertainty in future *outcomes*. As leverage may affect the measure of ambiguity estimated from equity data, we compute unlevered intraday returns using the book value of total debt and the market value of equity estimated at every five-minute interval.[14]

Utilizing the EUUP framework, the degree of ambiguity can be measured by the volatility of uncertain *probabilities*, just as the degree of risk can be measured by the volatility of uncertain *outcomes*. Formally, the measure of ambiguity is defined as:

$$\mho^2[X] \equiv \int \mathrm{E}[\varphi(x)] \, \mathrm{Var}[\varphi(x)] \, dx, \tag{1}$$

where $\varphi(\cdot)$ is an uncertain probability density function; and the expectation $\mathrm{E}[\cdot]$ and the variance $\mathrm{Var}[\cdot]$ are taken with respect to the second-order probability measure ξ on a set $\mathcal{P}$ of probability measures (Izhakian, 2018).[15] Equation (1) represents a probability-weighted average of the variances of probabilities. The measure of ambiguity, defined in Equation (1), is distinct from aversion

[14] The correlation between the ambiguity measure computed using unlevered returns and the one computed using (levered) stock returns is almost 0.99, so unlevering the returns does not drive the results.

[15] In a finite state space, $\mho^2[X] \equiv \sum_j \mathrm{E}[\varphi(x_j)] \, \mathrm{Var}[\varphi(x_j)]$, where $\varphi(\cdot)$ is an uncertain probability mass function.

Electronic copy available at: https://ssrn.com/abstract=3428896

to ambiguity. The former, which is a matter of beliefs (or information), is estimated from the data, while the latter, which is a matter of subjective attitudes, is endogenously determined by the empirical estimations.

We follow recent literature and estimate the empirical degree of firm-level ambiguity using intraday stock data from the *TAQ* database (e.g., Izhakian and Yermack, 2017, Augustin and Izhakian, 2019). We compute the degree of ambiguity for each stock each month and use its trailing three-month moving average.

As investors share the same information set, all have an identical set of priors over the intraday return distribution. Each prior in the set is represented by the observed daily intraday returns on the firm's equity, and the number of priors in the set depends on the number of trading days in the month. The set of priors thus consists of 18–22 realized distributions over a month. For practical implementations, we discretize return distributions into n bins $B_j = (r_j, r_{j-1}]$ of equal size, such that each distribution is represented as a histogram. The height of the bar of a particular bin is computed as the fraction of daily intraday returns observed in that bin, and thus represents the probability of the returns in that bin. Equipped with these 18–22 daily return histograms, we compute the expected probability of being in a particular bin across the daily return distributions, $\mathrm{E}\left[\mathrm{P}\left(B_j\right)\right]$, as well as the variance of these probabilities, $\mathrm{Var}\left[\mathrm{P}\left(B_j\right)\right]$. To this end, an equal likelihood is assigned to each histogram.[16] Using these values, the monthly degree of ambiguity of firm i is then computed as follows:

$$\mho^2\left[r_i\right] \equiv \frac{1}{\sqrt{w\left(1-w\right)}}\sum_{j=1}^{n}\mathrm{E}\left[\mathrm{P}_i\left[B_j\right]\right]\mathrm{Var}\left[\mathrm{P}_i\left[B_j\right]\right]. \tag{2}$$

To minimize the impact of bin size on the scale of ambiguity, we apply a variation of Sheppard's correction and scale the probability weighted-average variance of probabilities to the size of the bins by $\frac{1}{\sqrt{w(1-w)}}$, where $w = r_{i,j} - r_{i,j-1}$.

In our implementation, we sample five-minute stock returns from 9:30 to 16:00 to eliminates micro-structure effects (Andersen et al., 2001, Ait-Sahalia et al., 2005, M.Bandi and R.Russell, 2006, Y.Liu et al., 2015). Thus, we obtain daily histograms of up to 78 intraday returns. If we observe

[16] This is consistent with the *principle of insufficient reason*, which states that given n possibilities that are indistinguishable except for their names, each possibility should be assigned a probability equal to $\frac{1}{n}$ (Bernoulli, 1713, de Laplace, 1814). It is also consistent with the idea of the simplest non-informative prior in Bayesian probability (Bayes et al., 1763), which assigns equal probabilities to all possibilities; and the principle of maximum entropy (Jaynes, 1957), which states that the probability distribution which best describes the current state of knowledge is the one with the largest entropy.

Electronic copy available at: https://ssrn.com/abstract=3428896

no trade in a specific time interval for a given stock, we compute returns based on the volume-weighted average of the nearest trading prices. We ignore returns between closing and next-day opening prices to eliminate the impact of overnight price changes and dividend distributions. We drop all days with fewer than 10 different five-minute returns, then we drop months with fewer than 10 intraday return distributions. In addition, we drop extreme returns ($\pm 5\%$ log returns over five minutes), as many such returns are due to improper orders that are often later canceled by the stock exchange.

For the bin formation, we divide the range of daily returns into 162 intervals. We form a grid of 160 bins, from -40% to $+40\%$, each of width 0.5%, in addition to the left and right tails, defined as $(\infty, -40\%]$ and $(+40\%, +\infty)$, respectively. We compute the mean and the variance of probabilities for each interval, assigning equal likelihood to each distribution (i.e., all histograms are equally likely).[17] Some bins may not be populated with return realizations. Therefore, we assume a normal return distribution and use its moments to extrapolate return probabilities. That is, $\mathrm{P}_i[B_j] = \left[\Phi(r_j;\mu_i,\sigma_i) - \Phi(r_{j-1};\mu_i,\sigma_i)\right]$, where $\Phi(\cdot)$ denotes the cumulative normal probability distribution, characterized by its mean μ_i and the variance σ_i^2 of the returns. As in French et al. (1987), we apply the Scholes and Williams (1977) adjustment for non-synchronous trading to estimate variance of returns.[18] This adjustment further eliminates any micro-structure effects caused by bid-ask bounce, although our use of five-minute returns minimizes micro-structure effects.[19] Finally, $AMBIGUITY_{i,t}$, our measure of ambiguity of firm i in quarter t, is the average of the monthly ambiguity $\mho^2[r_i]$ over all months during quarter t.[20]

[17] The assignment of equal likelihoods is equivalent to assuming that the daily ratios $\frac{\mu}{\sigma}$ are Student's-t distributed. When $\frac{\mu}{\sigma}$ is Student's t-distributed, cumulative probabilities are uniformly distributed (e.g., Proposition 1.27, page 21 Kendall and Stuart, 2010).

[18] Scholes and Williams (1977) suggest adjusting the volatility of returns for non-synchronous trading as $\sigma_t^2 = \frac{1}{N_t}\sum_{\ell=1}^{N_t}\left(r_{t,\ell} - \mathrm{E}[r_{t,\ell}]\right)^2 + 2\frac{1}{N_t-1}\sum_{\ell=2}^{N_t}\left(r_{t,\ell} - \mathrm{E}[r_{t,\ell}]\right)\left(r_{t,\ell-1} - \mathrm{E}[r_{t,\ell-1}]\right)$.

[19] In a battery of robustness tests, Brenner and Izhakian (2018) and Augustin and Izhakian (2019) rule out the concern that $\mho^2$ may capture other well-known uncertainty factors including skewness, kurtosis, variance of variance, variance of mean, downside risk, mixed data sampling measure of forecasted volatility (MIDAS), investors' sentiment, and jumps, among several others. Their tests also rule out the concern that the empirical implementation is sensitive to the selection of bin size and the data frequency.

[20] We also considered an alternative frequency for our estimates of risk and ambiguity. Using daily data stock data from CRSP, we extract the intraday return distribution from open, close, high and low price quotes using Garman and Klass (1980). This method allows us to use all stocks, not only those included in the TAQ database, but we lose much of the information that intraday volatility allows us to include in the measure described in the text. Thus, it is a more crude measure of ambiguity. When we ran our regressions on daily data, the results were qualitatively similar but somewhat less significant.

Electronic copy available at: https://ssrn.com/abstract=3428896

3.4 Estimating risk

Along with ambiguity, risk serves as an important explanatory variable in our analysis. For consistency, we compute risk using the same five-minute returns that we use to compute ambiguity. For each individual firm i on each day, we compute the variance of five-minute intraday returns, applying the Scholes and Williams (1977) correction for non-synchronous trading and a correction for heteroscedasticity.[21] Each month, we estimate risk as the mean of the daily variance estimates. In our analysis, as with ambiguity, we use the quarterly mean of monthly risk estimates. $RISK_{i,t}$.

3.5 Control variables

We control for variables that are known in the literature to be correlated with innovation. Our firm-level controls include: log sales (LN_SALES);[22] Tobin's Q (Q); log ratio of physical capital per employee (LN_K_L); cash-flow ($CASH\_FLOW$); leverage ($LEVERAGE$); log firm age (LN_AGE); log of one plus R&D capital ($LN\_RD\_CAPITAL$); a dummy for Nasdaq listing ($NASDAQ$), and a control variable for missing R&D expenditures in Compustat (MISSING_RD).[23] All variables are described in detail in Appendix B.

We drop firm-quarters with $AMBIGUITY$, $RISK$ or $CASH\_FLOW$ below the 1$^{\text{st}}$ percentile and above the 99$^{\text{st}}$ percentile over the entire sample period. We also drop firm-quarters with K_L above the 99$^{\text{st}}$ sample percentile. Following Lanjouw and Schankerman (2004), Aghion et al. (2013) and others, we winsorize Q by setting it equal to 0.10 for values below 0.10 and to 20 for values above 20. All balance sheet and income statement variables are deflated using the quarterly GDP deflator from St Louis Fed (2009=100).

3.6 Summary statistics

Table 1 presents descriptive statistics for the *R&D Sample* (Panel A), and the *Patent Sample* (Panel B). In the *R&D Sample*, the median firm has sales of $75.842 million per quarter, while in the *Patent Sample*, the median firm is much larger, with sales of $179.731 million per quarter. The median firm age, approximated by the number of quarters the firm is listed in Compustat, is 13.5

[21] See, for example, French et al. (1987).

[22] Recall, that we keep only firm-quarter observation with strictly positive sales and assets.

[23] We control for missing R&D following Koh and Reeb (2015), who find that a significant number of firms with missing R&D in Compustat actually file and receive patents.

Electronic copy available at: https://ssrn.com/abstract=3428896

years (54 quarters) in the *R&D Sample*, and 18 years (72 quarters) in the *Patent Sample*. Overall, these differences suggest that R&D investment and patenting may take place at different stages in the firms' life cycle.[24]

In the *R&D Sample*, the median (mean) RD_ASSETS is 1.4% (2.1%) per quarter, and 5.2% (8%) per year.[25] We also note that 23.9% of the firm-quarters in the *R&D Sample* have missing R&D in the Compustat Fundamentals Quarterly file. For these innovative firms the median (mean) $CAPEX\_ASSETS$ ratio is only 0.8% (1.2%) per quarter, and 3.3% (4.7%) per year.[26] Conditional on filing at least one patent across all sample quarters (*Patent Sample*, Panel B), the median (mean) firm files 2 (24.96) patents per year and receives 1.19 (26.31) citations.[27] This indicates that the distribution of the number of patents and citations is heavily skewed, as previously documented in the literature.

The median (mean) $AMBIGUITY$ in the *R&D Sample* (Panel A) is 0.013 (0.02). Given that $AMBIGUITY$ measures the expected variance of probabilities, this implies that the median expected standard deviation of probabilities is $\sqrt{0.013} = 11.4\%$. The median $RISK$ of 0.002 per day corresponds to an annualized stock return volatility of approximately $\sqrt{250 \times 0.002} = 70.71\%$ (or, equivalently, $\sqrt{20 \times 0.002} = 20\%$ per month).[28] The medians for $AMBIGUITY$ and $RISK$ are quite stable in the *R&D Sample* (Panel A) and the *Patent Sample* (Panel B).

We then split the sample in two sub-samples: high-tech and non high-tech industries. Following Brown et al. (2009), we classify the following seven three-digit SIC code industries as high-tech industries: drugs (SIC 283), office and computing equipment (SIC 357), communications equipment (SIC 366), electronic components (SIC 367), scientific instruments (SIC 382), medical instruments

[24]Part of the difference in firm age between the *R&D Sample* and the *Patent Sample* comes from the fact that, for firms that enter Compustat (broadly speaking, IPO firms) during the sample period, we use the first four years of data to construct pre-sample means of the dependent count variables ($PATENTS$ and $CITATIONS$), effectively removing these years from the actual sample.

[25]For variables calculated over the four quarters $t+1 \ldots t+4$, we require the firm to be in the sample in all four quarters. For this reason the mean and median for annual variables are not necessarily exactly four times larger than for the corresponding quarterly variables.

[26]In untabulated analysis, we also calculated statistics for the sample of firms requiring at least one quarter of positive $CAPEX$, instead of one quarter of positive R&D. In that sample, the median (mean) $CAPEX\_ASSETS$ is 0.9% (1.4%) per quarter and 3.8% (5.7%) per year, while the median RD_ASSETS, R&D divided by total assets, is 0% (1.2%) per quarter and 0.4% (4.5%) per year.

[27]Recall that the number of citations for each patent is scaled by the average number of citations received by all patents in the same technology class filed in the same year, which corresponds to the fixed-effects approach in (Hall et al., 2001).

[28]When we exclude firms with a stock price lower than \$5, market capitalization less than \$10 million and fewer than 5 years in Compustat, the median $RISK$ in the *R&D Sample* falls to 0.001, which corresponds to an annualized stock return volatility of approximately $\sqrt{250 \times 0.001} = 50\%$ (or, equivalently, $\sqrt{20 \times 0.001} = 14.14\%$ per month).

Electronic copy available at: https://ssrn.com/abstract=3428896

(SIC 384), and software (SIC 737). This classification to high-tech and non high-tech industries splits the *R&D Sample* approximately in half: there are 55,184 firm-quarters (2,460 distinct firms) in the high-tech sample, and 49,060 firm-quarters (1,738 distinct firms) in the non high-tech sample. As expected, the R&D intensity is larger in high-tech industries. This is also reflected in the ratio of R&D to total (R&D plus CAPEX) investment: the median (mean) RD_RATIO is 77.9% (68.2%) in high-tech industries, and only 18.6% (31.4%) in non high-tech industries. High-tech firms are in general smaller, younger, have less leverage, less tangible capital and more intangible capital than non high-tech firms. In addition, high-tech firms appear to have higher risk and lower ambiguity than non high-tech firms.[29]

Table 2 presents averages of within-firm Pearson correlation coefficients for the explanatory variables for all firms in the *R&D Sample* (Panels A and B) and in the *Patent Sample* (Panel C). Once we remove penny stocks, very small and very young firms, the correlation between $AMBIGUITY$ and $RISK$ decreases from 0.025 in Panel A and becomes negative, -0.280 in Panel B. The correlations are very similar in the *R&D Sample* and the *Patent Sample*, as well as for high-tech and non high-tech firms within each sample. For illustrative purposes, Panel C presents correlations for high-tech firms in the *Patent sample*.

4 Empirical methodology

In our empirical exploration, we utilize two main models. First, to analyze the effect of ambiguity on innovation input, we estimate the following model using OLS:

$$RD\_ASSETS_{i,t+1} = \alpha + \beta_1 AMBIGUITY_{i,t} + \beta_2 RISK_{i,t} + \Gamma' X_{i,t} + \mu_i + \nu_t + \epsilon_{i,t}, \tag{3}$$

where i stands for the firm and t for the quarter; $X_{i,t}$ is a vector of control variables; μ_i denotes firm fixed effects; and ν_t denotes quarter-year fixed effects. Quarter-year fixed effects absorb any time effects that are constant across all firms, including seasonality effects. Standard errors are clustered by firm. The coefficient estimates for this model are presented in Tables 3, 4 and 5.

Second, as common in the patent literature, we estimate the following count model:

$$\mathbb{E}\left[OUT_{i,k,t+n} | X_{i,t}, \chi_i, \xi_k, \nu_t\right] = \exp\left[\alpha + \beta_1 AMBIGUITY_{i,t} + \beta_2 RISK_{i,t} + \Gamma' X_{i,t} + \chi_i + \xi_k + \nu_t\right], \tag{4}$$

[29] Untabulated tests for differences in means and medians for all variables between high-tech and non high-tech firms are significant at the 1% level, except for the patent variables in the *R&D sample*.

Electronic copy available at: https://ssrn.com/abstract=3428896

where $\mathbb{E}[\cdot]$ stands for expected value; $OUT_{i,k,t+n}$ is innovation output—either $PATENTS_{i,k,t+n}$ or $CITATIONS_{i,k,t+n}$—for firm i, in industry k in quarter $t+n$ ($n = 1 \ldots 12$); $X_{i,t}$ is a vector of the firm's control variables; χ_i denotes Blundell et al. (1999) pre-sample firm fixed effects; ξ_k denotes industry (3-digit SIC code) fixed effects; and ν_t denotes quarter-year fixed effects. We also estimate the same equation for the total number of patents or citations over each of the following three years ($OUT_{i,k,t+1 \ldots t+4}$, $OUT_{i,k,t+5 \ldots t+8}$ and $OUT_{i,k,t+9 \ldots t+12}$). Standard errors are clustered by firm. Equation (4) is estimated using both a Poisson and a Negative Binomial model. Estimation results are presented in Tables 6 and 7.

In the count models for $PATENTS$ and $CITATIONS$, we follow the recent innovation literature (e.g., Aghion et al., 2013, Bloom et al., 2013), and control for unobserved, time-invariant, firm-level heterogeneity using the pre-sample mean scaling fixed effect estimator of Blundell et al. (1999). This approach exploits the history of patent data for each firm and uses the log of pre-sample averages of the count dependent variable as a proxy for unobserved heterogeneity. We calculate pre-sample means of the dependent count variables ($PREPATENTS$ and $PRECITATIONS$) starting in the first quarter when the firm (*permco*) appears in the CRSP *dsenames* dataset. We require firms to have at least four years of pre-sample data (16 quarters) in order to calculate pre-sample averages of the dependent variables. For firms that enter Compustat after 1993 (the first year of patent data included in our regression sample), we use the first 16 quarters of data to calculate pre-sample averages, and we include the following quarters in the sample.[30] In addition to including the log of $PREPATENTS$ ($PRECITATIONS$), the count models for $PATENTS$ ($CITATIONS$) include an indicator variable for whether the firm had any patents (citations) in the pre-sample period.[31]

[30] Bloom et al. (2013) use a similar approach, also requiring 4 years of pre-sample data in their dataset covering the 1981-2001 period.

[31] In addition to the pre-sample mean scaling fixed effect, the models in Tables 6 and 7 include three-digit SIC code fixed effects and year fixed effects. We use three-digit SIC codes instead of four-digit SIC codes because our sample includes both NYSE/AMEX and Nasdaq firms, and, according to WRDS documentation, CRSP provides only the three-digit SIC code for Nasdaq firms.

Electronic copy available at: https://ssrn.com/abstract=3428896

5 Empirical findings

5.1 R&D investment

Table 3 presents coefficient estimates for OLS regressions for forward one-quarter and one-year R&D investment, as a function of $AMBIGUITY$, controlling for $RISK$ and other explanatory variables. We find that ambiguity has a negative and significant effect on R&D, while risk has a positive and significant effect, both one quarter ahead (Column (1)) and one year ahead (Column (2)). These findings are in line with the predication of our real options model and support Hypotheses 2 and 1a. Panel B of Table 3 shows that these results are robust to excluding penny stocks (stocks with price less than \$5 at the end of the previous quarter), very small firms (firms with market capitalization less than \$10 million at the and of the previous quarter), and very young firms (firms with less than 5 years in Compustat).

The effect of ambiguity is driven mainly by high-tech firms: the coefficient estimates on $AMBIGUITY$ in this sub-sample are larger and significant at the 1% level, both forward one quarter and one year R&D (Columns (3) and (4)). For non high-tech firms, the effect of ambiguity is not significant for forward one quarter R&D (Column (5)), and it is only marginally significant forward one year R&D (Column (6)). The coefficients of all control variables have the expected signs: R&D is higher in small firms, in firms with high growth opportunities (high Q), low tangibility (LN_K_L), low cash-flows and low leverage. The effect of age is positive and significant in 3 Panel A, but it becomes insignificant when we exclude firms with less than 5 years in Compustat (Panel B).

In terms of economic magnitude, the coefficient estimates in Table 3 imply that a one standard deviation increase in $AMBIGUITY$ across all firms (0.021) decreases the R&D intensity one quarter ahead (RD_ASSETS_{t+1}) by $-0.016 \times 0.021 = -0.00034$, which represents approximately 1.4% of the empirical standard deviation of the dependent variable (0.024). In the high-tech sub-sample, the economic effect is larger: a one standard deviation increase in $AMBIGUITY$ (0.02) decreases the R&D intensity one quarter ahead by $-0.037 \times 0.02 = -0.00149$, which represents approximately 3% of the empirical standard deviation of the dependent variable in that sub-sample (0.025). Similarly, a one standard deviation increase in $RISK$ increases the R&D intensity across all firms by approximately 6.7% of the empirical standard deviation for all firms (Column (1)), and by approximately 10.8% of the empirical standard deviation for high-tech firms (Column (3)).

Electronic copy available at: https://ssrn.com/abstract=3428896

Thus, the economic effect of ambiguity is lower, but comparable to that of risk. Furthermore, we see that ambiguity matters to high tech, high growth firms, which by definition are engaged in new and hard to predict lines of business. Panel B of Table 3 also shows that the effect of ambiguity is robust to eliminating penny stocks, very small and very young firms, so it is unlikely to be driven by microstructure effects.[32]

Table 4 presents results for different splits of the sub-sample of high-tech firms. To split this sub-sample, we first calculate the average sales, age and leverage for each firm over all quarters in the sample. Then we define small firms to be those with average sales below the sample median and large firms to be those with average sales above the sample median. Similarly, we define young/old firms and low-leverage/high-leverage firms. The coefficient of $AMBIGUITY$ is significant for small and large firms, young and older firms, and low-leverage and high-leverage firms. Moreover, it is larger in absolute value in the sub-samples of small firms (Column (1)), young firms (Column 5), and low-leverage firms (Column 9). The interaction effect between $AMBIGUITY$, $RISK$ and various measures is sometimes significant indicating that the effects may be more important for specific types of firms.

The finding that the effect of ambiguity is significantly stronger for low-leverage firms, together with the fact that R&D in general is higher in low-leverage firms, suggests that $AMBIGUITY$ matters for high growth firms (i.e., firms that engage in higher levels of R&D). In addition, in low-leverage firms, the ambiguity estimated from stock market data is closer conceptually to the ambiguity associated with the firm's assets, since unlevering is not as important. Intuitively, ambiguity matters more for smaller firms, which have less of a track record and the prospects of their outcomes may be more difficult to be established. The fact that the effect of ambiguity is stronger for low-leverage firms makes us confident that our findings in the rest of the paper are not driven by measurement errors associated with unlevering stock returns.

Table 5 reports additional three robustness tests for our R&D results. First, Panel A shows that our findings are robust to controlling for institutional ownership. Bushee (1998) finds that, while total institutional ownership decreases the probability that firms cut R&D in order to reverse an earnings decline, ownership by transient institutional investors (i.e., investors with diversified

[32]Recall that to eliminate microstructure effects, we compute ambiguity using five-minute returns and not higher frequency returns. We also apply the Scholes and Williams (1977) correction, which further eliminates possible microstructure effects.

Electronic copy available at: https://ssrn.com/abstract=3428896

portfolios and high turnover) has the opposite effect. In that latter case, transient institutional ownership encourages myopic investment behavior. On the other hand, Aghion et al. (2013) find that institutional ownership increases innovation, as measured by citation-weighted patents. Moreover, using Bushee's (1989) classification of institutional owners, Aghion et al. (2013) find a positive effect of institutional ownership on citation-weighted patents for both dedicated and transient institutional ownership, and no effect for quasi-indexer institutional ownership. Panel A of Table 5 shows that our R&D findings are unaffected when we augment the regression to include institutional ownership variables. Moreover, the institutional ownership variables themselves are not significant. In untabulated analysis, we find that our results are also robust to controlling for total institutional ownership instead of including dedicated and transient institutional ownership separately in the regression and that total institutional ownership is itself not significant.

Second, Panel B of Table 5 shows that our findings are robust to measuring innovation investment by the sum of R&D and CAPEX. As discussed above, a significant share of capital investment of R&D-active firms is in fact investment in innovative capacity which are not included in R&D expenditures, but are instead included in capital expenditures.[33] Thus, in this robustness test, the dependent variable is the sum of R&D and CAPEX, scaled by assets (RD_CAPEX_ASSETS). Panel B of Table 5 shows that the results are similar to the results in Table 3, when the dependent variable is RD_ASSETS, especially for high-tech firms, where the concern that CAPEX might include innovation investments is greater.[34]

Third, Panel C of Table 5 shows that our findings are robust to measuring innovation investment by R&D adjusted to the book value of total assets including capitalized R&D. As shown in Table 1, high-tech firms have both smaller size (measured with either assets or sales) and larger stocks of capitalized R&D expenditures than firms in traditional industries. We follow Chan et al. (2001), Lev et al. (2005) and Chambers et al. (2002), and adjust the book value of total assets to include capitalized R&D ($RD\_CAPITAL$). Accordingly, the dependent variable in Panel C, of Table 5 is

[33] At the same time, not all investments in innovative capacity would be included in capital expenditures. For example, the purchase of inventories would be reflected as an increase in total assets, but are not included in capital expenditures (Kumar and Li, 2016).

[34] We also regress CAPEX, both one quarter ahead and one year ahead, scaled by assets at the beginning of quarter $t+1$ ($CAPEX\_ASSETS_{t+1}$ and $CAPEX\_ASSETS_{t+1...t+4}$) on $AMBIGUITY$, $RISK$ and the same control variables as for the R&D regressions. When we require firms to have at least one quarter with positive R&D, without requiring positive CAPEX, which is the sample used in Tables 3, 4 and 5, we do not find a significant effect of $AMBIGUITY$ or $RISK$ on CAPEX either one year ahead or one quarter ahead when we exclude penny stocks, very small firms or very young firms. This suggests that R&D-active firms are different than the other firms, supporting our real options of innovative (vs. maintenance) investment.

Electronic copy available at: https://ssrn.com/abstract=3428896

R&D scaled by adjusted assets (RD_ADJ_AT), where adjusted assets include the book value on the balance sheet plus capitalized R&D.[35] This adjustment, however does not affect our results. Finally, in untabulated analysis, we find that the results are also robust to excluding observations with $AMBIGUITY$ and $RISK$ calculated during recession quarters (2001q2-2001q4 and 2008q1-2009q2).

In summary, our findings for R&D investment are broadly consistent with the real options view, supporting Hypotheses 2 and 1a. Namely, our findings show that investments in R&D decrease with ambiguity and increase with risk. This is particularly true for the firms that fit the model best: high-tech firms.

5.2 Patents and citations

We turn now to examine the effect of ambiguity and risk on innovation outputs: patents and citations. Tables 6 and 7 present results for Poisson and Negative Binomial regressions for $PATENTS$ (Panel A) and $CITATIONS$ (Panel B), restricting the sample to high-tech firms.[36] We further exclude penny stocks, very small firms and very young firms. All regression tests include three-digit SIC code fixed effects, Blundell et al. (1999) pre-sample firm fixed effects, and quarter-year fixed effects. We run the regression tests separately for each quarter $t+1,\ldots,t+12$, but for brevity we report findings only for quarter $t+1$, as well for the combined quarters $t+1\ldots t+4$ (Year 1), $t+5\ldots t+8$ (Year 2) and $t+9\ldots t+12$ (Year 3). Importantly, the findings for each of years 1, 2 and 3 are not driven by individual quarters within that year. Table 6 shows that the coefficient estimates of both $AMBIGUITY$ and $RISK$ are negative. The negative effect of risk on patents and citations is significant for both variables in Poisson regressions, but ambiguity is no longer significant in Negative Binomial regressions.

The negative effect of ambiguity on patenting activity is in line with the predictions implied by the real options concept, supporting Hypothesis 2. We note that, in any setup, an ambiguity averse manager should invest less as ambiguity increases. Risk is significant and negative throughout, which is not consistent with the real options concept, stated in Hypothesis 1a. However, the

[35]These regressions include the same control variables Table 3, Panel A, with one difference: as we adjust total assets to include capitalized R&D, we apply the same adjustment to the denominator in Tobin's Q.

[36]When we pool high-tech and non high-tech firms together, the results are qualitatively similar, but the significance is lower. We do not find a significant effect of ambiguity on patents and citation in non high-tech industries, so the effect in the overall sample is driven by firms in high-tech industries. Therefore, we restrict our analysis to these industries going forward.

Electronic copy available at: https://ssrn.com/abstract=3428896

negative effect of risk on patenting activity is consistent with the ideas in Bloom (2007, 2014). It may be that the R&D decision is better modeled by a real option whereas patenting decisions conform better to the dynamic set up in Bloom (2007), and with the idea that the delaying patenting is more valuable when risk is relatively high.

The bottom part of each panel in Table 6 presents marginal effects, which are defined as differences between the predicted number of counts (patents or citations) at the 90$^{\text{th}}$ and the 10$^{\text{th}}$ percentiles for both $AMBIGUITY$ and $RISK$. In the Poisson model, the predicted number of patents three years ahead at the 10$^{\text{th}}$ percentile of $AMBIGUITY$ in the *Patent Sample* is 6.153, while the predicted number of patents at the 90$^{\text{th}}$ percentile of $AMBIGUITY$ is 5.210. The marginal effect of ambiguity is thus to decrease the predicted number of patents by about 0.943, and this effect is statistically significant at the 5% level (Table 6, Panel A). Similarly, the predicted number of citations received for patents filed three years ahead is 8.438 at the 10$^{\text{th}}$ percentile of $AMBIGUITY$, but only 7.242 at the 90$^{\text{th}}$ percentile of $AMBIGUITY$.[37] The marginal effect is -1.196 citations, and is statistically significant at the 10% level (Table 6, Panel B). These marginal effects are economically important, given that the median (mean) high-tech firm in the *Patent Sample* files 3 (30.478) patent applications during a sample year and receives 2.217 (32.882) citations for these patents, as reported in Panel B of Table 1.

In the Poisson model, the marginal effect of increasing $RISK$ from the 10$^{\text{th}}$ to the 90$^{\text{th}}$ percentile is to decrease the predicted number of patents three years ahead by 4.092 (Table 6, Panel A), and the predicted number of citations three years ahead by 5.962 (Table 6, Panel B). Our findings thus lend support to Hypothesis 1b, based on Bloom (2007), suggesting that firms may delay, and hence decrease, investment in innovation in the face of increased risk.

Next we perform two sets of robustness tests for our patent results. First, Table 7 restricts the sample to high-tech firms with average number of patents above the sample median. For these patent-intensive firms, the median (mean) number of patents one year ahead is 13 (63.3), compared to 3 (30.478) for all high-tech patenting firms, as reported in Panel B of Table 1. The effect of ambiguity on both patents and citations is stronger for patent-intensive high-tech firms (Table 7) than for high-tech firms in general (Table 6). In Table 7, $AMBIGUITY$ is significant in both

[37] Recall that the citation count for each patent is scaled by the average number of citations received by all patents in the same technology field filed in the same year, which corresponds to the fixed-effect approach for dealing with citation truncation discussed in Hall et al. (2001).

Electronic copy available at: https://ssrn.com/abstract=3428896

the Poisson and the Negative Binomial regressions. Moreover, in the negative binomial regression tests of Table 7, the effect of $AMBIGUITY$ completely subsumes the effect of $RISK$, which is no longer significant.[38]

Second, Table 8 replicates the analysis from Table 6, augmenting the set of controls to include institutional ownership variables. We find that our patent results are robust to controlling for institutional ownership. Moreover, in Poisson regressions, we find a positive and significant effect of dedicated institutional ownership on both patents and citations. The effect of transient institutional ownership is insignificant in both Poisson and Negative Binomial models. This finding is partly in line with Aghion et al. (2013), who find a positive and significant effect of both dedicated and transient institutional ownership on citations. Our patent findings are also robust to controlling for total institutional ownership instead of including dedicated and transient institutional ownership separately in the regression. Total institutional ownership is itself not significant.

Overall, Tables 6, 7 and 8 show that both ambiguity and risk have a negative and significant effect on patents and citations up to three years into the future. The findings for ambiguity are in line with the prediction from our stylized real options model, while the findings for risk suggest instead that when faced with increased risk firms may in fact decrease investments in innovation, consistent with Bloom (2007). The effect of ambiguity is particularly important for high tech firms engaged in research and patenting, as we expect.

In untabulated analysis, we also find that both the R&D and the patent results are robust to controlling for the disagreement among analyst forecasts, measured by the standard deviation of the price forecasts and scaled by the average stock price or by the average forecast. The disagreement of analyst forecasts itself has a positive and significant effect on R&D, patents and citations, when it is scaled by the average stock price. On the contrary, when it is scaled by the average forecast, the effect on R&D is only marginally significant, and the effect on patents and citations is insignificant.

[38]One concern with the models presented in Table 7 is the relatively high correlation between $AMBIGUITY$ and $RISK$ in the sub-sample of patent-intensive high-tech firms. In this sub-sample, the correlation between $AMBIGUITY$ and $RISK$ is -0.472, compared to -0.264 for the sample used in Table 6. In general, the correlation between $AMBIGUITY$ and $RISK$ is negative and larger in absolute value in sub-samples of large firms. Izhakian et al. (2018) show that a possible reason for this is that larger firms typically do not have many organic growth opportunities, and the opportunities that do exist are likely to be typified by ambiguous prospects. Smaller and younger firms tend to have organic growth opportunities (expansion of existing activities) whose characteristics are similar to those of the firms assets in place. For robustness, we run the regression tests in Tables 6 and 7 including only $AMBIGUITY$, without $RISK$. The coefficient on $AMBIGUITY$ is always negative and significant at levels similar to those reported in the tables. This indicates that the correlation between $AMBIGUITY$ and $RISK$ does not drive the results.

Electronic copy available at: https://ssrn.com/abstract=3428896

6 Conclusion

A number of recent papers document the impact of various factors on innovation. One of the most important questions in innovation research is the effect of uncertainty on investment in R&D and in patents, which by definition are both paths into the unknown. We analyze two different types of uncertainty—ambiguity and risk—which ex-ante may lead to very different firm decisions. We focus on the distinction between ambiguity and risk as drivers of innovation.

To support our hypotheses, we present a stylized model, which shows that firms should increase investment in innovative projects as risk increases, but decrease investment as ambiguity, defined as the expected variance of probabilities, goes up. We contrast these predictions with the ideas in Bloom (2007, 2014), who suggests that risk may deter innovation. Empirically, we find broad support for the proposition that firms facing high ambiguity decrease both R&D and patents, as predicted by our model. This is particularly true for high-tech, high-growth firms, which are the types of firms expected to be concerned about ambiguity in addition to risk. However, we observe two different effects of risk on innovation. Riskier firms indeed invest more in R&D, but riskier firms also file fewer patents and receive fewer citations, which is consistent with the idea that in uncertain times delay and the option to wait are more valuable.

Our findings may generalize. The fact that ambiguity and risk can have similar, but also opposite effects, may help explain various phenomena such as ostensible under-investment or over-investment.

Electronic copy available at: https://ssrn.com/abstract=3428896

References

Abdellaoui, M. and E. Kemel (2013) "Eliciting prospect theory when consequences are measured in time units: "Time is not money," *Management Science*, Vol. 60, No. 7, pp. 1844–1859.

Abdellaoui, M., F. Vossmann, and M. Weber (2005) "Choice-Based Elicitation and Decomposition of Decision Weights for Gains and Losses Under Uncertainty," *Management Science*, Vol. 51, No. 9, pp. 1384–1399.

Abel, A. B. (1983) "Optimal investment under uncertainty," *The American Economic Review*, Vol. 73, No. 1, p. 228.

Abel, A. B. and J. C. Eberly (1994) "A Unified Model of Investment Under Uncertainty," *The American Economic Review*, Vol. 84, No. 5, pp. 1369–1384.

Aghion, P., N. Bloom, R. Blundell, R. Griffith, and P. Howitt (2005) "Competition and Innovation: an Inverted-U Relationship," *The Quarterly Journal of Economics*, Vol. 120, No. 2, p. 701728.

Aghion, P., J. Van Reenen, and L. Zingales (2013) "Innovation and institutional ownership," *The American Economic Review*, Vol. 103, No. 1, pp. 277–304.

Ait-Sahalia, Y., P. A. Mykland, and L. Zhang (2005) "How Often to Sample a Continuous-Time Process in the Presence of Market Microstructure Noise," *Review of Financial Studies*, Vol. 18, No. 2, pp. 351–416.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001) "The distribution of realized stock return volatility," *Journal of Financial Economics*, Vol. 61, No. 1, pp. 43–76.

Augustin, P. and Y. Izhakian (2019) "Ambiguity, Volatility and Credit Risk," *The Review of Financial Studies*, Forthcoming.

Bayes, T., R. Price, and J. Canton (1763) *An essay towards solving a problem in the doctrine of chances*: C. Davis, Printer to the Royal Society of London London, U. K.

Ben-David, D., F. Franzoni, R. Moussawi, and J. Sedunov (2018) "The Granular Nature of Large Institutional Investors," Swiss Finance Institute, Research Paper Series, No. 15-67.

Bernanke, B. S. (1983) "Irreversibility, uncertainty, and cyclical investment," *The Quarterly Journal of Economics*, Vol. 98, No. 1, pp. 85–106.

Bernoulli, J. (1713) *Ars Conjectandi (The Art of Conjecturing)*.

Bernstein, S. (2015) "Does Going Public Affect Innovation?" *Journal of Finance*, Vol. 70, No. 4, p. 13651403.

Bernstein, S., T. McQuade, and R. R. Townsend (2017) "Does Economic Insecurity Affect Employee Innovation?" *Stanford University Working Paper*.

Bloom, N. (2007) "Uncertainty and the Dynamics of R&D," *American Economic Review*, Vol. 97, No. 2, pp. 250–255.

——— (2014) "Fluctuations in uncertainty," *Journal of Economic Perspectives*, Vol. 28, No. 2, pp. 153–76.

Bloom, N., M. Schankerman, and J. Van Reenen (2013) "Technology spillovers and product market rivalry," *Econometrica*, Vol. 81, No. 4, p. 13471393.

Blundell, R., R. Griffith, and J. Van Reenen (1999) "Market share, market value and innovation in a panel of British manufacturing firms," *Review of Economic Studies*, Vol. 66, No. 3, pp. 529–554.

Brennan, M. J. and E. S. Schwartz (1985) "Evaluating Natural Resouce Investments," *Journal of Business*, Vol. 58, No. 2, pp. 135–157.

Brenner, M. and Y. Izhakian (2018) "Asset Prices and Ambiguity: Empirical Evidence," *Journal of Financial Economics*, Vol. 130, pp. 503–531.

Electronic copy available at: https://ssrn.com/abstract=3428896

Brown, J. R., S. M. Fazzari, and B. C. Petersen (2009) "Financing Innovation and Growth: Cash Flow, External Equity, and the 1990s R&D Boom," *Journal of Finance*, Vol. 64, pp. 151–185.

Bushee, B. (1998) "The Influence of Institutional Investors on Myopic R&D Investment Behavior," *Accounting Review*, Vol. 73, No. 3, pp. 305–333.

Caballero, R. J. (1991) "On the sign of the investment-uncertainty relationship," *The American Economic Review*, Vol. 81, No. 1, pp. 279–288.

Chambers, D., R. Jennings, and R. B. Thompson (2002) "Excess Returns to R&D-Intensive Firms," *Review of Accounting Studies*, Vol. 7, pp. 133–158.

Chan, L. K. C., L. Josef, and T. Sougiannis (2001) "The Stock Market Valuation of Research and Development Expenditures," *The Journal of Finance*, Vol. 56, No. 6, pp. 2431–2456.

Chew, S. H. and J. S. Sagi (2008) "Small Worlds: Modeling Attitudes Toward Sources of Uncertainty," *Journal of Economic Theory*, Vol. 139, No. 1, pp. 1–24.

Crockett, S., Y. Izhakian, and J. Jamison (2019) "Ellsberg's Hidden paradox," *SSRN eLibrary*, 3423534.

Dass, N., V. Nanda, and S. C. Xiao (2017) "Truncation bias in patent data: Implications for recent research on innovation," *Jurnal of Corporate Finance*, Vol. 44, pp. 353–374.

Dong, M., D. A. Hirshleifer, and S. H. Teoh (2017) "Stock Market Overvaluation, Moon Shots, and Corporate Innovation," *SSRN eLibrary*, 2878418.

Du, N. and D. V. Budescu (2005) "The Effects of Imprecise Probabilities and Outcomes in Evaluating Investment Options," *Management Science*, Vol. 51, No. 12, pp. 1791–1803.

Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, Vol. 75, No. 4, pp. 643–669.

French, K. R., G. W. Schwert, and R. F. Stambaugh (1987) "Expected stock returns and volatility," *Journal of Financial Economics*, Vol. 19, No. 1, pp. 3–29.

Garman, M. B. and M. J. Klass (1980) "On the estimation of security price volatilities from historical data," *Journal of business*, pp. 67–78.

Gilboa, I. and D. Schmeidler (1989) "Maxmin Expected Utility with Non-Unique Prior," *Journal of Mathematical Economics*, Vol. 18, No. 2, pp. 141–153.

Halevy, Y. (2007) "Ellsberg Revisited: An Experimental Study," *Econometrica*, Vol. 75, No. 2, pp. 503–536.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001) "The NBER patent citation data file: Lessons, insights and methodological tools," *NBER Working Paper 8498*.

Hartman, R. (1972) "The effects of price and cost uncertainty on investment," *Journal of economic theory*, Vol. 5, No. 2, pp. 258–266.

He, J. and X. Tian (2013) "The dark side of analyst coverage: The case of innovation," *Journal of Financial Economics*, Vol. 109, pp. 856–878.

Herron, R. and Y. Izhakian (2017) "Ambiguity, Risk, and Payout Policy," *SSRN eLibrary*, 2980600.

Izhakian, Y. (2017) "Expected Utility with Uncertain Probabilities Theory," *Journal of Mathematical Economics*, Vol. 69, pp. 91–103.

——— (2018) "A Theoretical Foundation of Ambiguity Measurement," *SSRN eLibrary*, 1332973.

Izhakian, Y. and D. Yermack (2017) "Risk, Ambiguity, and the Exercise of Employee Stock Options," *Journal of Financial Economics*, Vol. 124, pp. 65–85.

Electronic copy available at: https://ssrn.com/abstract=3428896

Izhakian, Y., D. Yermack, and J. Zender (2018) "Ambiguity and the Trade Off Theory of Capital Structure," *SSRN eLibrary*, 2873248.

Jaynes, E. T. (1957) "Information theory and statistical mechanics," *Physical review*, Vol. 106, No. 4, p. 620.

Kendall, M. and A. Stuart (2010) "The Advanced Theory of Statistics. Vol. 1: Distribution Theory," *London: Griffin, 2010, 6th ed.*, Vol. 1.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005) "A Smooth Model of Decision Making under Ambiguity," *Econometrica*, Vol. 73, No. 6, pp. 1849–1892.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017) "Technological Innovation, Resource Allocation, and Growth," *Quarterly Journal of Economics*, Vol. 132, p. 665712.

Koh, P.-S. and D. M. Reeb (2015) "Missing R&D," *Journal of Accounting and Economics*, Vol. 60, No. 1, pp. 73–94.

Kraft, H., E. Schwartz, and F. Weiss (2018) "Growth options and firm valuation," *European Financial Management*, Vol. 24, No. 2, pp. 209–238.

Krieger, J., D. Li, and D. Papanikolaou (2017) "Developing novel drugs," *Harvard Business School Working Paper*.

Kumar, P. and D. Li (2016) "Capital investment, innovative capacity, and stock returns," *Journal of Finance*, Vol. 71, No. 5, pp. 2059–2094.

——— (2018) "Innovation, Stock Market Feedback and Good Idiosyncratic Volatility," *SSRN eLibrary*, 3160243.

Lanjouw, J. and M. Schankerman (2004) "Patent quality and research productivity: Measuring innovation with multiple indicators," *Economic Journal*, Vol. 114, No. 495, pp. 441–465.

de Laplace, P. S. (1814) *Théorie analytique des probabilitiés*.

Lerner, J., M. Sorensen, and P. Strömberg (2011) "Private Equity and Long-Run Investment: The Case of Innovation," *Journal of Finance*, Vol. 66, No. 2, p. 445477.

Lev, B., B. Sarath, and T. Sougiannis (2005) "R&D Reporting Biases and Their Consequences," *Contemporary Accounting Research*, Vol. 22, No. 4, pp. 977–1026.

Maffioletti, A. and M. Santoni (2005) "Do Trade Union Leaders Violate Subjective Expected Utility? Some Insights From Experimental Data," *Theory and Decision*, Vol. 59, No. 3, pp. 207–253.

Mangelsdorff, L. and M. Weber (1994) "Testing Choquet Expected Utility," *Journal of Economic Behavior and Organization*, Vol. 25, No. 3, pp. 437–457.

M.Bandi, F. and J. R.Russell (2006) "Separating Microstructure Noise from Volatility," *Journal of Financial Economics*, Vol. 79, No. 3, pp. 655–692.

McDonald, R. and D. Siegel (1986) "The Value of Waiting to Invest," *Quarterly Journal of Economics*, Vol. 101, No. 4, pp. 707–727.

Nau, R. F. (2006) "Uncertainty Aversion with Second-Order Utilities and Probabilities," *Management Science*, Vol. 52, pp. 136–145.

Pindyck, R. S. (1988) "Irreversible Investment, Capacity Choice, and the Value of the Firm," *American Economic Review,*, Vol. 78, pp. 969–985.

Rothschild, M. and J. E. Stiglitz (1970) "Increasing Risk: I. A Definition," *Journal of Economic Theory*, Vol. 2, No. 3, pp. 225–243.

Electronic copy available at: https://ssrn.com/abstract=3428896

Savage, L. J. (1954) *The Foundations of Statistics*, New York, USA: Wiley.

Schmeidler, D. (1989) "Subjective Probability and Expected Utility without Additivity," *Econometrica*, Vol. 57, No. 3, pp. 571–587.

Scholes, M. and J. Williams (1977) "Estimating betas from nonsynchronous data," *Journal of Financial Economics*, Vol. 5, No. 3, pp. 309–327.

Schwartz, E. S. (2004) "Patents and R&D as real options," *Economic Notes*, Vol. 33, No. 1, pp. 23–54.

Seru, A. (2014) "Firm boundaries matter: Evidence from conglomerates and R&D activity," *Journal of Financial Economics*, Vol. 111, No. 2, pp. 381–405.

Trajtenberg, M. (1990) "A Penny for Your Quotes: Patent Citations and the Value of Innovations," *The RAND Journal of Economics*, Vol. 21, pp. 172–187.

Tversky, A. and D. Kahneman (1992) "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, Vol. 5, No. 4, pp. 297–323.

Viscusi, W. K. and H. Chesson (1999) "Hopes and Fears: the Conflicting Effects of Risk Ambiguity," *Theory and Decision*, Vol. 47, No. 2, pp. 157–184.

Von-Neumann, J. and O. Morgenstern (1944) *Theory of Games and Economic Behavior*: Princeton University Press.

Wakker, P. P., D. R. M. Timmermans, and I. Machielse (2007) "The Effects of Statistical Information on Risk and Ambiguity Attitudes, and on Rational Insurance Decisions," *Management Science*, Vol. 53, No. 11, pp. 1770–1784.

Wu, G. and R. Gonzalez (1999) "Nonlinear decision weights in choice under uncertainty," *Management science*, Vol. 45, No. 1, pp. 74–85.

Y.Liu, L., A. J. Patton, and K. Sheppard (2015) "Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes," *Journal of Econometrics*, Vol. 187, No. 1, pp. 293–311.

Electronic copy available at: https://ssrn.com/abstract=3428896

A Theoretical Appendix

A.1 A Decision Theoretic Model of Ambiguity

This discussion is based on Izhakian (2017). To formally define the uncertain payoff X in EUUP framework, let $(\mathcal{S}, \mathcal{E}, \mathrm{P})$ be a probability space, where $\mathcal{S}$ is a state space, $\mathcal{E}$ is a σ-algebra of subsets of the state space (a set of events), $\mathrm{P} \in \mathcal{P}$ is a probability measure, and the set of probability measures $\mathcal{P}$ is convex. An algebra Π of measurable subsets of $\mathcal{P}$ is equipped with a probability measure, denoted ξ. The uncertain outcome is then given by the uncertain variable, $X : \mathcal{S} \to \mathbb{R}$. Denote by $\varphi(x)$ the (uncertain) marginal probability (probability density or mass function) associated with the (uncertain) cumulative probability $\mathrm{P} \in \mathcal{P}$ of outcome x. The expected marginal and cumulative probability of outcome x, taken with respect to the second-order probability measure ξ, are then defined respectively by

$$\mathrm{E}[\varphi(x)] \equiv \int_{\mathcal{P}} \varphi(x)\, d\xi \qquad \text{and} \qquad \mathrm{E}[\mathrm{P}(x)] \equiv \int_{\mathcal{P}} \mathrm{P}(x)\, d\xi, \tag{5}$$

and the variance of the marginal probability of outcome x is defined by

$$\mathrm{Var}[\varphi(x)] \equiv \int_{\mathcal{P}} \big(\varphi(x) - \mathrm{E}[\varphi(x)]\big)^2 d\xi. \tag{6}$$

With these definitions in place, the expected outcome and the variance of outcomes are computed using the expected probabilities. That is,

$$\mathbb{E}[X] \equiv \int \mathrm{E}[\varphi(x)]\, x dx \qquad \text{and} \qquad \mathbb{V}\mathrm{ar}[X] \equiv \int \mathrm{E}[\varphi(x)] \big(x - \mathbb{E}[x]\big)^2 dx. \tag{7}$$

Notice that double-struck capital font designates expectation or variance of outcomes with respect to expected probabilities, while regular straight font designates expectation or variance of probabilities with respect to second-order probabilities.

Managers have distinct preferences for ambiguity and risk. As usual, preferences for risk are modeled by a bounded, strictly-increasing and twice-differentiable utility function $\mathrm{U} : \mathbb{R}_+ \to \mathbb{R}$. Risk aversion takes the form of a concave $\mathrm{U}(\cdot)$, risk loving takes the form of a convex $\mathrm{U}(\cdot)$, and risk neutrality takes the form of a linear $\mathrm{U}(\cdot)$. As investors are sensitive to ambiguity, they do not compound the set of priors $\mathcal{P}$ and the prior ξ over $\mathcal{P}$ in a linear way (compounded lotteries), but instead they aggregate these probabilities in a non-linear way that reflects their attitude toward ambiguity. Preferences for ambiguity are defined by preferences over mean-preserving spreads in probabilities and modeled by a strictly-increasing and twice-differentiable function over probabili-

Electronic copy available at: https://ssrn.com/abstract=3428896

ties, $\Upsilon : \mathbb{R}_+ \to \mathbb{R}$, called the *outlook function*. Similar to risk, ambiguity aversion takes the form of a concave $\Upsilon(\cdot)$, ambiguity loving takes the form of a convex $\Upsilon(\cdot)$, and ambiguity neutrality takes the form of a linear $\Upsilon(\cdot)$. In EUUP, ambiguity aversion is exhibited when an investor prefers the expectation of an uncertain probability of each payoff over the uncertain probability itself.[39] Note that in EUUP preferences for ambiguity are outcome independent. That is, preferences for ambiguity apply exclusively to probability of events, independently of the outcomes associated to these events. Like Tversky and Kahneman's (1992) cumulative prospect theory, EUUP assumes that investors have a reference point, relative to which returns are classified as either unfavorable (loss) or favorable (gain). Accordingly, we normalize U to $\mathrm{U}(k) = 0$, where k is the investors' reference point.

In the EUUP framework, the manager assesses the expected utility of a risk and ambiguous payoff by[40]

$$\begin{aligned} \mathrm{W}(X) \approx & \int_{x \leq k} \mathrm{U}(x) \underbrace{\mathrm{E}[\varphi(x)] \left(1 - \frac{\Upsilon''(1-\mathrm{E}[\mathrm{P}(x)])}{\Upsilon'(1-\mathrm{E}[\mathrm{P}(x)])} \mathrm{Var}[\varphi(x)]\right)}_{\text{Perceived Probability of Unfavorable Outcome}} dx + \\ & \int_{x \geq k} \mathrm{U}(x) \underbrace{\mathrm{E}[\varphi(x)] \left(1 + \frac{\Upsilon''(1-\mathrm{E}[\mathrm{P}(x)])}{\Upsilon'(1-\mathrm{E}[\mathrm{P}(x)])} \mathrm{Var}[\varphi(x)]\right)}_{\text{Perceived Probability of Favorable Outcome}} dx. \end{aligned} \tag{8}$$

Notice that when there is no ambiguity ($\mathcal{P}$ is a singleton) Equation (8) collapses to the conventional expected utility $\mathrm{W}(X) = \int \mathrm{U}(x)\varphi(x)\,dx$. When managers are ambiguity neutral (i.e., $\Upsilon(\cdot)$ is linear and, therefore, $\frac{\Upsilon''}{\Upsilon'} = 0$), they compound probabilities linearly and Equation (8) collapses to the conventional expected utility $\mathrm{W}(X) = \int \mathrm{U}(x)\,\mathrm{E}[\varphi(x)]\,dx$, assessed using expected probabilities. In contrast, when managers are ambiguity averse (i.e., $\Upsilon(\cdot)$ is concave), they do not aggregate probabilities linearly and the intensity of aversion to ambiguity affects the perceived probabilities. In this case, managers overweight the probabilities of the unfavorable outcomes and underweight the probabilities of the favorable outcomes. Conceptually, the perceived probability of a given

[39]Recall that risk aversion is exhibited when a manager prefers the expected outcome of the uncertain outcome over the uncertain outcome itself.

[40]This functional representation is obtained by taking the Taylor expansion of EUUP representation EUUP, proposed by Izhakian (2017). The remainder of this approximation is of order $o\left(\int \mathrm{E}\left[|\varphi(x) - \mathrm{E}[\varphi(x)]|^3\right] x dx\right)$ as $\int |\varphi(x) - \mathrm{E}[\varphi(x)]|\,dx \to 0$, meaning that the accuracy of the approximation is equivalent to the accuracy of the cubic approximation, $o\left(\mathbb{E}\left[|x - \mathbb{E}[x]|^3\right]\right)$, in which the fourth and higher absolute central moments of outcomes are of strictly smaller order than the third absolute central moment as $|x - \mathbb{E}[x]| \to 0$, and are therefore negligible.

Electronic copy available at: https://ssrn.com/abstract=3428896

outcome is the unique certain probability value that the manager is willing to accept in exchange for its uncertain probability (a certainty-equivalent probability).

The notion of mean-preserving spreads in probabilities in Equation (8) can be used to derive a measure of ambiguity (Izhakian, 2018, Theorem 6). This measure, defined as the expected variance of probabilities, is formally given by

$$\mho^2[X] \equiv \int \mathrm{E}[\varphi(x)] \operatorname{Var}[\varphi(x)]\, dx. \tag{9}$$

The measure $\mho^2$ (mho^2) can be used either in a continuous state space with infinitely many outcomes or in a discrete state space with finitely many outcomes.

To observe the distinct impact of ambiguity and ambiguity aversion on the value of an investment opportunity, consider a binomial asset with either low payoff (L) or high payoff (H). Suppose that the reference point k satisfies $L \leq k \leq \mathbb{E}[X] < H$.[41] By Equation (8), the value of this asset in terms of expected utility is

$$\begin{aligned} \mathrm{W}(X) = \; & \mathrm{U}(L)\,\mathrm{E}[\varphi(L)] \left(1 - \frac{\Upsilon''(1-\mathrm{E}[\mathrm{P}(H)])}{\Upsilon'(1-\mathrm{E}[\mathrm{P}(H)])} \operatorname{Var}[\varphi(L)]\right) + \\ & \mathrm{U}(H)\,\mathrm{E}[\varphi(H)] \left(1 + \frac{\Upsilon''(\mathrm{E}[\mathrm{P}(H)])}{\Upsilon'(\mathrm{E}[\mathrm{P}(H)])} \operatorname{Var}[\varphi(H)]\right). \end{aligned} \tag{10}$$

Expected utility in this functional representation is assessed using the manager's perceived probabilities. Ambiguity and aversion to ambiguity are modeled in Equation (10) through the manager's marginal perceived probabilities. Consider the high payoff, H. The expression

$$\mathrm{Q}(H) = \mathrm{E}[\varphi(H)] \left(1 + \frac{\Upsilon''(\mathrm{E}[\mathrm{P}(H)])}{\Upsilon'(\mathrm{E}[\mathrm{P}(H)])} \operatorname{Var}[\varphi(H)]\right) \tag{11}$$

is the marginal perceived probability of this outcome occurring.[42] This marginal perceived probability is a function of the degree of ambiguity, measured by $\operatorname{Var}[\varphi(H)]$, and the investor's attitude toward ambiguity, captured by $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)}$. For an ambiguity-averse manager with $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)} > 0$, a higher aversion to ambiguity or a higher degree of ambiguity results in lower marginal perceived probabilities of good states and higher marginal perceived probabilities of bad states. This in turn implies a lower perceived expected utility.

Note again that, in this example, if there is no ambiguity, Equation (10) collapses to the conventional expected utility with the value $\mathrm{W}(X) = \mathrm{U}(L)\,\varphi(L) + \mathrm{U}(H)\,\varphi(H)$. If managers

[41] We assume that the expected outcome is greater than the reference point; otherwise, a rational decision maker would not consider the investment opportunity.

[42] Note that, since every $\mathrm{P} \in \mathcal{P}$ is additive, $1 - \mathrm{E}[\mathrm{P}(L)] = \mathrm{E}[\mathrm{P}(H)]$. In this case, the variance of the probabilities of L is equal to the variance of the probabilities of its complementary event H, so that $\operatorname{Var}[\varphi(L)] = \operatorname{Var}[\varphi(H)]$.

Electronic copy available at: https://ssrn.com/abstract=3428896

are ambiguity neutral, they compound probabilities linearly and Equation (10) collapses to the conventional expected utility with the value $\mathrm{W}(X) = \mathrm{U}(L)\,\mathrm{E}[\varphi(L)] + \mathrm{U}(H)\,\mathrm{E}[\varphi(H)]$, assessed using expected probabilities.

A.2 A Stylized model

To support our hypotheses about the effect of ambiguity and risk on investment decisions, we employ the EUUP framework described above to develop a stylized static real options model. R&D investment or a patent filing can be considered a real option affected by ambiguity and risk. Suppose that I is the present value of the costs of developing the product, and V is the present value of the expected cash flows from this development. The payoff X from owning a product can then be written as:

$$X = \begin{cases} V - I, & \text{if} \quad V \geq I; \\ 0, & \text{if} \quad V < I. \end{cases}$$

Thus, by Schwartz (2004), Brennan and Schwartz (1985) and McDonald and Siegel (1986), the project can be viewed as a call option, where the payoff of the product is the underlying asset.

Assume a one period model with a zero risk-free rate. Suppose that the cost of developing a product using the technology is the reference point, i.e., $k = I$, satisfying $0 \leq k$. By Equation (8), the value of this (call) option is

$$C = \int_I^\infty \mathrm{E}[\varphi(x)] \left(1 + \frac{\Upsilon''(\mathrm{E}[\mathrm{P}(x)])}{\Upsilon'(\mathrm{E}[\mathrm{P}(x)])} \mathrm{Var}[\varphi(x)]\right) x dx. \tag{12}$$

When there is no ambiguity, Equation (12) collapses to the conventional expected utility case; therefore $C = \displaystyle\int_I^\infty \varphi(x)\, x dx$. When investors are ambiguity neutral, since they compound probabilities linearly, Equation (12) again collapses to the conventional expected utility, assessed using expected probabilities; therefore, $C = \displaystyle\int_I^\infty \mathrm{E}[\varphi(x)]\, x dx$.

As in Rothschild and Stiglitz (1970), an underlying security is said to become riskier if its new payoffs can be written as a mean-preserving spread of the old payoffs. Accordingly, we assume neither that risk is measured by the variance of payoffs, nor that that returns are normally distributed or that the utility is quadratic. Equation (12) suggests that the option value is increasing in the risk of the payoff of the project, since the option payoff function is convex in the state outcomes. To see this more clearly, consider a possible payoff x of the project. Assume that the risk of the

Electronic copy available at: https://ssrn.com/abstract=3428896

project increases, such that this specific payoff is now $x + \Delta$ or $x - \Delta$, with equal probabilities, i.e. $x \pm \Delta$ is a mean-preserving spread of x. Since the reference point satisfies $k = I$, the value of the call option is positively affected by the magnitude of Δ: when $\Delta \leq x - I$ the option value is unaffected, and when $\Delta > x - I$, then $\frac{1}{2}(x - I + \Delta) \geq x - I$. Thus, by Equation (12), the value of the option increases in the risk of the project.[43]

In addition to the effect of risk, a higher ambiguity implies lower perceived probabilities of the good states—a successful R&D, or valuable patent, and therefore a lower value of the option. To see this, in Equation (12), consider for example a decision maker with constant absolute ambiguity-aversion. In this case $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)} = \eta$, where η is the coefficient of absolute ambiguity aversion. Since aversion to ambiguity implies a positive $-\frac{\Upsilon''(\cdot)}{\Upsilon'(\cdot)}$, a higher ambiguity, measured by $\mho^2[X]$ (which in this case is equal to the weighted sum of $\mathrm{Var}[\varphi(x)]$), implies lower perceived probabilities (Equation (11)) and therefore a lower value of the option. A lower value of the real option implies lower investment in R&D or less patenting activity.

[43]Note that, assuming normally distributed returns, a quadratic utility function or an exponential utility function (all imply a mean-variance-ambiguity preference), risk can be measured by the variance of returns, computed using expected probabilities (Izhakian and Yermack, 2017).

Electronic copy available at: https://ssrn.com/abstract=3428896

B Variable Definitions

Variable	Definition
AGE	Number of quarters in Compustat.
$AMBIGUITY$	The ambiguity measure is defined in detail in Section 3.3.
$ASSETS$	Compustat item *atq*.
ADJ_ASSETS	Assets adjusted for capitalized R&D. Compustat item atq + $RD\_CAPITAL$
$CAPEX$	Compustat item *capexy*, adjusted for fiscal year accumulation.
$CAPEX\_ASSETS$	The ratio of CAPEX to assets at the beginning of the quarter (Compustat item *atq*).
$CASH\_FLOW$	Cash-flow. Calculated as (Income Before Extraordinary Items + Depreciation and Amortization) / Assets at the beginning of the quarter. (*ibq* + *dpq*) / lagged *atq*.
$CITATIONS$	The number of citations received by all patents applied for in a given quarter, excluding self-cites. The number of citations for each patent is scaled by the average number of citations received by all patents in the same 3-digit USPTO technology class filed in the same year (Hall et al., 2001).
K_L	The ratio of physical capital per employee. Compustat item *ppentq* divided by the number of empoyees. We estimate the number of employees at the end of each quarter by linear interpolation using the values at the beginning and at the end of the fiscal year from the Compustat Fundamentals Annual file (Compustat item *emp*). When the number of employees (*emp*) is missing either at the beginning or at the end of the fiscal year, we assign the value from the other year end point, if available, to all quarters during the year.
$INSTOWN$	Insitutional ownership, from the Thompson Reuters 13F database. Following Ben-David et al. (2018), after June 2013, we calculate institutional ownership using the 13F data parsed directly from the SEC EDGAR filings system, and available on WRDS.
$INSTOWN\_DED$	Dedicated institutional ownership, ie ownership by institutions with concentrated portfolio holdings and low turnover, according to the Bushee (1998) classification.
$INSTOWN\_QIX$	Quasi-indexer institutional ownership, ie ownership by institutions with diversified portfolios and low turnover, according to the Bushee (1998) classification.
$INSTOWN\_TRA$	Transient institutional ownership, ie ownership by institutions with diversified portfolios and high turnover, according to the Bushee (1998) classification.
$LEVERAGE$	$(dlttq + dlcq)/atq$
LN_AGE	$\ln(1 + AGE)$
LN_ASSETS	$\ln(ASSETS)$
LN_K_L	$\ln(1 + K\_L)$
LN_MCAP	$\ln(MCAP)$
$LN\_PRECITATIONS$	$\ln(1 + PRECITATIONS)$

Electronic copy available at: https://ssrn.com/abstract=3428896

$LN\_PREPATENTS$	$\ln(1 + PREPATENTS)$
$LN\_RD\_CAPITAL$	$\ln(1 + RD\_CAPITAL)$
LN_SALES	$\ln(SALES)$
$MCAP$	Market capitalization. Compustat item $prccq \times cshoq$.
$NASDAQ$	Indicator variable equal to 1 if the stock is traded on Nasdaq at the end of the quarter, and 0 otherwise.
$PATENTS$	The number of patents applied for during the quarter.
$PRECITATIONS$	The quarterly average of the number of citations received for patents applied for during the presample period. (See the definition of $PREPATENTS$.)
$PRECITATIONS > 0$	An indicator variable equal to 1 if $PRECITATIONS > 0$, and 0 otherwise.
$PREPATENTS$	The quarterly average of the number of patents applied for during the presample period (Blundell et al., 1999). We use the history of patent data for each firm (*permco*) in the Kogan et al. (2017) dataset to calculate $PREPATENTS$. For firms that enter Compustat after 1993 (the first year in our sample), we use the first four years of data as the presample period, and we start the sample with the fifth year in Compustat.
$PREPATENTS > 0$	An indicator variable equal to 1 if $PREPATENTS > 0$, and 0 otherwise.
Q	Tobin's Q. Calculated as (Market value of equity - Book value of equity - Deferred taxes + Assets) / Assets. ($cshoq \times prccq - ceqq - txdbq$ (replaced with zero when missing) $+ atq)/atq$. (In Table 5, the denominator is $atq + RD\_CAPITAL$.)
RD	R&D expenditures (Compustat item *xrdq*, replaced with zero when missing).
RD_ADJ_ASSETS	The ratio of RD to adjusted assets at the beginning of the quarter (ADJ_ASSETS).
RD_ASSETS	The ratio of RD to assets at the beginning of the quarter (Compustat item *atq*).
RD_CAPEX_ASSETS	The ratio of total investment ($RD+CAPEX$) to assets at the beginning of the quarter (Compustat item *atq*).
$RD\_CAPITAL$	Capitalized R&D expenditures. Following Lev et al. (2005), Chan et al. (2001) and Chambers et al. (2002), we capitalize the R&D expenditure in the last five years, using a depreciation rate of 20% per year, or 5% per quarter: $RD\_CAPITAL_t = \sum_{k=0}^{15} RD_{t-k} \times (1 - k \times 0.05)$
$MISSING\_RD$	An indicator variable equal to 1 if *xrdq* is missing in Compustat, and 0 otherwise.
RD_RATIO	$RD/(RD + CAPEX)$
$RISK$	The risk measure is defined in detail in Section 3.4.
$SALES$	Compustat item *saleq*.

Electronic copy available at: https://ssrn.com/abstract=3428896

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 1: Descriptive Statistics

This table presents descriptive statistics for the variables used in the analysis. The sample period is 1993-2016 in Panel A and 1993-2008 in Panel B. In Panel A, the sample consists of all firms with at least four quarters of data for all variables of interest and at least one quarter of positive R&D expenditures in Compustat during the sample period (*R&D Sample*). In Panel B, the sample consists of all firms with at least four quarters of data for all variables of interest, four years in the presample period and at least one patent application filed during the sample period (*Patent Sample*). Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B.

Panel A: R&D Sample

	All Firms						High-Tech Firms						Non High-Tech Firms					
	N	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**	**N**	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**	**N**	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**
RD_AT_{t+1}	105,037	0.021	0.024	0.000	0.014	0.032	55,184	0.030	0.025	0.012	0.025	0.042	49,060	0.011	0.018	0.000	0.003	0.014
$CAPEX\_AT_{t+1}$	104,120	0.012	0.013	0.004	0.008	0.015	54,669	0.011	0.013	0.003	0.007	0.013	48,663	0.013	0.013	0.005	0.009	0.016
RD_RATIO_{t+1}	103,912	0.507	0.366	0.028	0.600	0.843	54,590	0.682	0.290	0.553	0.779	0.901	48,538	0.314	0.342	0.000	0.186	0.617
$RD\_AT_{t+1...t+4}$	95,369	0.080	0.085	0.015	0.052	0.121	48,881	0.117	0.088	0.051	0.100	0.164	45,069	0.040	0.060	0.006	0.018	0.047
$CAPEX\_AT_{t+1...t+4}$	95,659	0.047	0.047	0.018	0.033	0.061	49,449	0.044	0.048	0.015	0.029	0.055	44,791	0.052	0.046	0.022	0.038	0.066
$RD\_RATIO_{t+1...t+4}$	93,753	0.532	0.312	0.264	0.579	0.816	47,963	0.692	0.246	0.566	0.766	0.880	44,398	0.363	0.283	0.113	0.324	0.573
$PATENTS_{t+1}$	72,990	4.668	23.677	0.000	0.000	2.000	39,370	4.503	26.237	0.000	0.000	1.000	33,031	4.900	20.326	0.000	0.000	2.000
$PATENTS_{t+1...t+4}$	65,636	20.123	97.231	0.000	1.000	7.000	34,865	19.642	108.836	0.000	1.000	7.000	29,651	20.685	80.992	0.000	1.000	8.000
$CITATIONS_{t+1}$	72,990	5.023	24.111	0.000	0.000	1.431	39,370	5.011	26.604	0.000	0.000	1.250	33,031	5.069	20.831	0.000	0.000	1.689
$CITATIONS_{t+1...t+4}$	65,636	21.559	97.844	0.000	0.356	8.593	34,865	21.740	109.332	0.000	0.344	8.175	29,651	21.315	81.177	0.000	0.467	9.213
$AMBIGUITY_t$	105,037	0.020	0.021	0.006	0.013	0.025	55,184	0.018	0.020	0.005	0.011	0.023	49,060	0.022	0.021	0.007	0.015	0.029
$RISK_t$	105,037	0.007	0.010	0.001	0.002	0.009	55,184	0.008	0.011	0.001	0.003	0.012	49,060	0.005	0.009	0.000	0.001	0.006
$SALES_t$	105,037	505.631	1842.571	19.094	75.842	324.996	55,184	223.052	832.795	12.617	37.909	136.851	49,060	828.415	2507.670	44.821	181.504	626.448
$ASSETS_t$	105,037	2152.577	8622.008	97.026	344.492	1395.651	55,184	1133.060	3974.478	68.011	189.485	726.221	49,060	3317.605	11777.210	187.701	682.348	2472.996
$MCAP_t$	105,037	2612.481	8427.358	140.771	478.761	1718.985	55,184	2148.203	7968.735	108.642	336.260	1254.447	49,060	3149.180	8931.115	203.311	697.360	2357.544
$RD\_CAPITAL_t$	105,037	188.222	743.536	9.281	34.879	111.286	55,184	208.547	725.722	13.465	42.670	127.998	49,060	167.009	766.412	5.565	26.830	94.918
Q_t	105,037	2.355	1.898	1.303	1.777	2.696	55,184	2.677	2.165	1.410	2.033	3.157	49,060	1.990	1.458	1.227	1.582	2.214
K_L_t	105,037	68.672	127.610	20.684	37.528	71.119	55,184	49.985	67.455	16.779	30.805	57.486	49,060	90.109	169.893	26.967	46.037	89.984
$CASH\_FLOW_t$	105,037	0.008	0.049	0.001	0.020	0.034	55,184	0.002	0.056	-0.013	0.018	0.035	49,060	0.015	0.039	0.010	0.022	0.034
$LEVERAGE_t$	105,037	0.168	0.187	0.002	0.116	0.274	55,184	0.119	0.170	0.000	0.038	0.189	49,060	0.223	0.191	0.055	0.205	0.332
AGE_{t+1}	105,037	74.566	60.696	28.000	54.000	105.000	55,184	58.134	46.447	24.000	45.000	77.000	49,060	93.224	69.047	34.000	72.000	151.000
$INSTOWN_t$	105,037	0.513	0.303	0.248	0.554	0.777	55,184	0.482	0.305	0.209	0.491	0.756	49,060	0.548	0.297	0.319	0.609	0.796
$INSTOWN\_DED_t$	105,037	0.060	0.087	0.000	0.018	0.090	55,184	0.055	0.084	0.000	0.012	0.082	49,060	0.064	0.091	0.000	0.025	0.097
$INSTOWN\_TRA_t$	105,037	0.133	0.116	0.040	0.108	0.197	55,184	0.136	0.122	0.036	0.109	0.204	49,060	0.129	0.109	0.044	0.108	0.189
$INSTOWN\_QIX_t$	105,037	0.293	0.219	0.099	0.264	0.462	55,184	0.266	0.216	0.080	0.216	0.421	49,060	0.325	0.219	0.139	0.315	0.496
$NASDAQ_t$	105,037	0.609	0.488	0.000	1.000	1.000	55,184	0.788	0.409	1.000	1.000	1.000	49,060	0.408	0.492	0.000	0.000	1.000
$MISSING\_RD_{t+1}$	105,037	0.239	0.427	0.000	0.000	0.000	55,184	0.078	0.268	0.000	0.000	0.000	49,060	0.417	0.493	0.000	0.000	1.000

Panel B: Patent Sample

	All Firms						High-Tech Firms						Non High-Tech Firms					
	N	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**	**N**	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**	**N**	**Mean**	**St. Dev.**	**p25**	**p50**	**p75**
RD_AT_{t+1}	54,093	0.015	0.021	0.000	0.006	0.025	21,253	0.028	0.023	0.012	0.024	0.039	30,642	0.006	0.014	0.000	0.000	0.007
$CAPEX\_AT_{t+1}$	53,580	0.014	0.014	0.005	0.010	0.018	21,013	0.012	0.014	0.004	0.008	0.015	30,389	0.015	0.014	0.006	0.011	0.019
RD_RATIO_{t+1}	53,457	0.373	0.368	0.000	0.332	0.734	20,988	0.636	0.300	0.484	0.725	0.869	30,293	0.185	0.290	0.000	0.000	0.363
$RD\_AT_{t+1...t+4}$	51,404	0.059	0.075	0.000	0.029	0.092	19,626	0.111	0.080	0.053	0.097	0.154	29,399	0.024	0.045	0.000	0.007	0.029
$CAPEX\_AT_{t+1...t+4}$	51,198	0.056	0.051	0.023	0.042	0.072	19,697	0.051	0.052	0.019	0.035	0.064	29,139	0.061	0.051	0.028	0.047	0.078
$RD\_RATIO_{t+1...t+4}$	50,554	0.399	0.332	0.000	0.386	0.705	19,211	0.650	0.255	0.505	0.715	0.849	29,004	0.228	0.263	0.000	0.129	0.403
$PATENTS_{t+1}$	54,093	5.904	27.125	0.000	0.000	2.000	21,253	7.085	34.044	0.000	0.000	3.000	30,642	5.143	21.000	0.000	0.000	2.000
$PATENTS_{t+1...t+4}$	49,614	24.961	110.045	0.000	2.000	10.000	19,124	30.478	140.099	0.000	3.000	13.000	28,178	21.178	82.112	0.000	1.000	8.000
$CITATIONS_{t+1}$	54,093	6.240	27.492	0.000	0.000	2.218	21,253	7.664	34.018	0.000	0.000	3.222	30,642	5.292	21.439	0.000	0.000	1.741
$CITATIONS_{t+1...t+4}$	49,614	26.310	110.504	0.000	1.194	10.988	19,124	32.882	139.040	0.000	2.217	14.873	28,178	21.709	82.135	0.000	0.718	8.711
$AMBIGUITY_t$	54,093	0.020	0.020	0.007	0.014	0.026	21,253	0.017	0.018	0.005	0.011	0.021	30,642	0.022	0.020	0.008	0.016	0.029
$RISK_t$	54,093	0.006	0.009	0.001	0.002	0.006	21,253	0.007	0.010	0.001	0.002	0.010	30,642	0.004	0.008	0.000	0.001	0.004
$SALES_t$	54,093	876.940	2672.365	45.593	179.731	619.041	21,253	346.563	1171.853	18.903	62.273	230.131	30,642	1271.717	3344.699	111.508	332.100	1059.372
$ASSETS_t$	54,093	3642.106	13095.580	184.742	683.221	2436.244	21,253	1566.752	5027.217	89.648	291.206	1085.435	30,642	5186.988	16661.460	374.570	1130.558	3610.753
$MCAP_t$	54,093	4282.845	13187.280	227.109	788.487	2800.276	21,253	3356.786	11620.960	148.209	490.609	1826.737	30,642	5027.167	14366.080	321.845	1054.188	3567.651
$RD\_CAPITAL_t$	54,093	232.856	981.724	1.067	28.630	117.353	21,253	301.903	1003.115	16.271	57.202	182.172	30,642	186.903	975.056	0.000	11.724	78.972
Q_t	54,093	2.190	1.654	1.283	1.698	2.485	21,253	2.671	2.075	1.461	2.067	3.150	30,642	1.844	1.147	1.211	1.534	2.084
K_L_t	54,093	84.345	164.961	25.261	42.954	82.457	21,253	53.753	57.870	21.468	36.418	63.541	30,642	108.068	210.385	29.150	49.615	103.969
$CASH\_FLOW_t$	54,093	0.019	0.038	0.012	0.024	0.037	21,253	0.013	0.048	0.004	0.023	0.040	30,642	0.023	0.029	0.015	0.025	0.036
$LEVERAGE_t$	54,093	0.197	0.180	0.027	0.175	0.308	21,253	0.124	0.162	0.000	0.058	0.206	30,642	0.247	0.174	0.121	0.238	0.348
AGE_{t+1}	54,093	92.473	59.813	41.000	72.000	141.000	21,253	66.346	43.566	35.000	52.000	84.000	30,642	111.121	62.673	50.000	110.000	170.000
$INSTOWN_t$	54,093	0.554	0.268	0.371	0.603	0.768	21,253	0.520	0.286	0.288	0.556	0.766	30,642	0.581	0.249	0.440	0.627	0.768
$INSTOWN\_DED_t$	54,093	0.081	0.097	0.000	0.046	0.133	21,253	0.071	0.094	0.000	0.027	0.117	30,642	0.088	0.099	0.000	0.060	0.144
$INSTOWN\_TRA_t$	54,093	0.138	0.118	0.049	0.111	0.199	21,253	0.142	0.126	0.042	0.112	0.211	30,642	0.134	0.110	0.053	0.110	0.190
$INSTOWN\_QIX_t$	54,093	0.319	0.205	0.162	0.297	0.454	21,253	0.290	0.212	0.116	0.254	0.427	30,642	0.342	0.197	0.203	0.324	0.472
$NASDAQ_t$	54,093	0.468	0.499	0.000	0.000	1.000	21,253	0.739	0.439	0.000	1.000	1.000	30,642	0.287	0.452	0.000	0.000	1.000

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 2: Correlations

This table presents averages of within-firm Pearson correlation coefficients for the variables used in the analysis. The sample period is 1993-2016 in Panel A and 1993-2008 in Panel B. In Panel A, the sample consists of all firms with at least four quarters of data for all variables of interest and at least one quarter of positive R&D expenditures in Compustat during the sample period (*R&D Sample*). In Panel B, the sample consists of all firms in the *R&D Sample*, excluding penny stocks, very small firms and very young firms. *Penny stocks* are stocks with a price less than $5 at the end of quarter t. *Very small firms* are firms with a market capitalization less than $10 million at the end of quarter t. *Very young firms* are firms with less than 5 years in Compustat. In Panel C, the sample consists of all firms in high-tech industries (3-digit SIC codes 283, 357, 366, 367, 382, 384, or 737) with at least four quarters of data for all variables of interest, four years in the presample period and at least one patent application filed during the sample period, excluding penny stocks, very small firms and very young firms. In Panel D, the sample is the same as in Panel C, but restricted to firms above the sample median in terms of the average number of patents applied for during the sample period. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B.

Panel A: R&D Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) $AMBIGUITY_t$	1.000											
(2) $RISK_t$	0.025	1.000										
(3) LN_SALES_t	0.034	-0.270	1.000									
(4) Q_t	-0.052	-0.245	0.018	1.000								
(5) $LN\_K\_L_t$	-0.005	0.021	0.042	-0.135	1.000							
(6) $CASH\_FLOW_t$	-0.003	-0.086	0.259	0.199	-0.099	1.000						
(7) $LEVERAGE_t$	0.012	0.123	-0.006	-0.090	0.110	-0.159	1.000					
(8) LN_AGE_{t+1}	0.097	-0.256	0.304	-0.168	-0.002	-0.095	0.014	1.000				
(9) $LN\_RD\_CAPITAL_t$	0.068	-0.143	0.297	-0.185	0.144	-0.139	0.042	0.507	1.000			
(10) $INSTOWN_t$	0.004	-0.285	0.206	0.107	-0.007	0.034	-0.092	0.241	0.144	1.000		
(11) $NASDAQ_t$	-0.045	0.232	-0.144	0.033	-0.033	0.064	-0.103	-0.114	-0.112	-0.039	1.000	
(12) $MISSING\_RD_{t+1}$	-0.012	0.058	-0.063	0.027	-0.014	-0.009	0.013	-0.109	0.012	-0.041	-0.019	1.000

Panel B: R&D Sample - Stocks with Price $\geq$ \$5, Market Cap $\geq$ \$10m and Age $\geq$ 5 Years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(1) $AMBIGUITY_t$	1.000											
(2) $RISK_t$	-0.280	1.000										
(3) LN_SALES_t	0.113	-0.270	1.000									
(4) Q_t	0.021	-0.181	0.003	1.000								
(5) $LN\_K\_L_t$	-0.020	0.018	0.042	-0.108	1.000							
(6) $CASH\_FLOW_t$	0.013	-0.051	0.249	0.222	-0.088	1.000						
(7) $LEVERAGE_t$	-0.025	0.096	0.016	-0.094	0.106	-0.162	1.000					
(8) LN_AGE_{t+1}	0.172	-0.362	0.394	-0.100	-0.006	-0.067	-0.010	1.000				
(9) $LN\_RD\_CAPITAL_t$	0.091	-0.161	0.342	-0.130	0.118	-0.120	0.045	0.439	1.000			
(10) $INSTOWN_t$	0.077	-0.268	0.193	0.077	-0.016	0.010	-0.084	0.316	0.136	1.000		
(11) $NASDAQ_t$	-0.067	0.237	-0.157	0.045	-0.043	0.072	-0.103	-0.089	-0.136	-0.062	1.000	
(12) $MISSING\_RD_{t+1}$	-0.038	0.067	-0.065	0.021	0.001	-0.020	0.019	-0.133	0.083	-0.053	-0.009	1.000

Panel C: Patent Sample - High-Tech Firms, Stocks with Price $\geq$ \$5, Market Cap $\geq$ \$10m and Age $\geq$ 5 Years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) $AMBIGUITY_t$	1.000										
(2) $RISK_t$	-0.264	1.000									
(3) LN_SALES_t	0.067	-0.272	1.000								
(4) Q_t	-0.056	-0.135	-0.001	1.000							
(5) $LN\_K\_L_t$	-0.064	0.053	0.007	-0.110	1.000						
(6) $CASH\_FLOW_t$	0.010	-0.070	0.232	0.261	-0.100	1.000					
(7) $LEVERAGE_t$	-0.011	0.148	-0.069	-0.091	0.137	-0.151	1.000				
(8) LN_AGE_{t+1}	0.155	-0.400	0.443	-0.138	-0.051	-0.077	-0.105	1.000			
(9) $LN\_RD\_CAPITAL_t$	0.049	-0.152	0.412	-0.171	0.147	-0.142	0.024	0.460	1.000		
(10) $INSTOWN_t$	0.091	-0.325	0.208	0.110	-0.055	0.058	-0.152	0.325	0.116	1.000	
(11) $NASDAQ_t$	-0.177	0.314	-0.261	-0.026	-0.131	0.060	0.031	-0.175	-0.338	-0.042	1.000

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 3: Determinants of R&D Investment

The table presents OLS regression coefficients for R&D investment. The dependent variable is RD_ASSETS_{t+1}. The sample period is 1993-2016. In Panel A, the sample consists of all firms with at least four quarters of data for all variables of interest and at least one quarter of positive R&D expenditures in Compustat during the sample period (*R&D Sample*). In Panel B, the sample consists of all firms in the *R&D Sample*, excluding penny stocks, very small firms and very young firms. *Penny stocks* are stocks with a price less than $5 at the end of quarter t. *Very small firms* are firms with a market capitalization less than $10 million at the end of quarter t. *Very young firms* are firms with less than 5 years in Compustat. All regressions in Panel B include the following control variables: LN_SALES_t, Q_t, $LN\_K\_L_t$, $CASH\_FLOW_t$, $LEVERAGE_t$, LN_AGE_t, $LN\_RD\_CAPITAL_t$, $NASDAQ_t$ and $MISSING\_RD_{t+1}$. In columns (1), (3) and (5), $MISSING\_RD_{t+1}$ is an indicator variable equal to 1 if the firm has missing R&D expenditures in Compustat in quarter $t+1$. In columns (2), (4) and (6), $MISSING\_RD_{t+1}$ is the number of quarters with missing R&D in Compustat in the period $t+1 \ldots t+4$. All regressions include firm (*new gvkey*) fixed effects and quarter-year fixed effects. Standard errors are clustered by firm. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

Panel A: R&D Sample

	All Firms		**High-Tech**		**Non High-Tech**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	One Quarter	**One Year**	**One Quarter**	**One Year**	**One Quarter**	**One Year**
	$t+1$	$t+1 \ldots t+4$	$t+1$	$t+1 \ldots t+4$	$t+1$	$t+1 \ldots t+4$
$AMBIGUITY_t$	-0.016***	-0.071***	-0.037***	-0.139***	-0.000	-0.021*
	(0.004)	(0.014)	(0.007)	(0.025)	(0.004)	(0.012)
$RISK_t$	0.161***	0.609***	0.246***	0.874***	0.048***	0.234***
	(0.016)	(0.060)	(0.024)	(0.093)	(0.017)	(0.065)
LN_SALES_t	-0.002***	-0.012***	-0.004***	-0.017***	-0.001***	-0.007***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
Q_t	0.001***	0.006***	0.001***	0.006***	0.001***	0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
$LN\_K\_L_t$	-0.002***	-0.009***	-0.003***	-0.012***	-0.001**	-0.004***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.002)
$CASH\_FLOW_t$	-0.033***	-0.043***	-0.033***	-0.037***	-0.023***	-0.033**
	(0.003)	(0.009)	(0.003)	(0.012)	(0.004)	(0.014)
$LEVERAGE_t$	-0.004***	-0.020***	-0.005***	-0.024***	-0.002*	-0.013***
	(0.001)	(0.004)	(0.001)	(0.006)	(0.001)	(0.004)
LN_AGE_{t+1}	0.002***	0.004**	0.003***	0.005	0.002***	0.004*
	(0.000)	(0.002)	(0.001)	(0.004)	(0.001)	(0.002)
$LN\_RD\_CAPITAL_t$	0.001***	0.005***	0.003***	0.011***	0.000	0.003***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.000)	(0.001)
$NASDAQ_t$	-0.001	-0.005**	-0.001	-0.005	-0.001*	-0.004*
	(0.001)	(0.002)	(0.001)	(0.005)	(0.001)	(0.002)
$MISSING\_RD_{t+1}$	-0.017***	-0.007***	-0.025***	-0.012***	-0.014***	-0.004***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Constant	0.027***	0.128***	0.031***	0.159***	0.019***	0.066***
	(0.003)	(0.010)	(0.004)	(0.016)	(0.003)	(0.012)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	105,037	94,894	55,184	48,365	49,060	44,968
N firms	4,053	3,657	2,460	2,160	1,738	1,586
Adj R2	0.802	0.857	0.756	0.801	0.796	0.876

Electronic copy available at: https://ssrn.com/abstract=3428896

Panel B: R&D Sample - Excluding Penny Stocks, Very Small Firms and Very Young Firms

	All Firms		High-Tech		Non High-Tech	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	One Quarter	**One Year**	**One Quarter**	**One Year**	**One Quarter**	**One Year**
	$t+1$	$t+1\ldots t+4$	$t+1$	$t+1\ldots t+4$	$t+1$	$t+1\ldots t+4$
Excluding Stocks with Price $<$ \$5 ***and Market Cap*** $<$ \$10***m***						
$AMBIGUITY_t$	-0.010**	-0.068***	-0.025***	-0.119***	-0.000	-0.034***
	(0.005)	(0.015)	(0.009)	(0.031)	(0.004)	(0.012)
$RISK_t$	0.150***	0.595***	0.216***	0.752***	0.066***	0.344***
	(0.021)	(0.084)	(0.033)	(0.131)	(0.025)	(0.103)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	83,963	77,004	41,468	37,003	41,944	38,931
N firms	3,278	3,017	1,926	1,721	1,456	1,361
Adj R2	0.809	0.873	0.765	0.822	0.791	0.887
Excluding Stocks with Price $<$ \$5, ***Market Cap*** $<$ \$10***m and Age*** $<$ 5 ***years***						
$AMBIGUITY_t$	-0.009**	-0.065***	-0.030***	-0.129***	0.002	-0.034***
	(0.005)	(0.016)	(0.009)	(0.034)	(0.004)	(0.013)
$RISK_t$	0.133***	0.485***	0.206***	0.642***	0.057*	0.299**
	(0.023)	(0.091)	(0.035)	(0.141)	(0.030)	(0.121)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	69,772	64,534	33,050	29,886	36,371	33,894
N firms	2,508	2,342	1,427	1,306	1,170	1,104
Adj R2	0.808	0.878	0.768	0.831	0.779	0.882

Electronic copy available at: https://ssrn.com/abstract=3428896

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 4: Subsample Analysis of R&D Investment in High-Tech Firms

The table presents OLS regression coefficients for R&D investment. The dependent variable is RD_ASSETS_{t+1}. The sample consists of all high-tech firms (3-digit SIC codes 283, 357, 366, 367, 382, 384, or 737) with at least four quarters of data for all variables of interest and at least one quarter of positive R&D expenditures in Compustat during the sample period, excluding penny stocks, very small firms and very young firms. *Penny stocks* are stocks with a price less than $5 at the end of quarter t. *Very small firms* are firms with a market capitalization less than $10 million at the end of quarter t. *Very young firms* are firms with less than 5 years in Compustat. Small (large) firms are firms with average sales below (above) the sample median. Young (old) firms are firms with average age below (above) the sample median. Low (high) leverage firms are firms with average leverage below (above) the sample median. The sample period is 1993-2016. All regressions include the following control variables: LN_SALES_t, Q_t, $LN\_K\_L_t$, $CASH\_FLOW_t$, $LEVERAGE_t$, LN_AGE_{t+1}, $LN\_RD\_CAPITAL_t$, $NASDAQ_t$ and $MISSING\_RD_{t+1}$. In Panel A, $MISSING\_RD_{t+1}$ is an indicator variable equal to 1 if the firm has missing R&D expenditures in Compustat in quarter $t+1$. In Panel B, MISSING_RD is the number of quarters with missing R&D in Compustat in the period $t+1 \ldots t+4$. All regressions include firm (*new gvkey*) fixed effects and quarter-year fixed effects. Standard errors are clustered by firm. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

	SIZE				**AGE**				**LEVERAGE**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	Small	**Large**	**All Firms**		**Young**	**Old**	**All Firms**		**Low**	**High**	**All Firms**	
Panel A: RD_AT one quarter ahead (quarter $t+1$)												
$AMBIGUITY_t$	-0.042***	-0.026**	-0.045***	-0.038**	-0.052***	-0.023**	-0.048***	-0.054***	-0.050***	-0.023**	-0.058***	-0.061***
	(0.014)	(0.011)	(0.014)	(0.015)	(0.015)	(0.011)	(0.013)	(0.014)	(0.017)	(0.010)	(0.017)	(0.018)
$RISK_t$	0.247***	0.225***	0.213***	0.182***	0.293***	0.171***	0.207***	0.281***	0.200***	0.198***	0.210***	0.261***
	(0.040)	(0.069)	(0.037)	(0.042)	(0.067)	(0.042)	(0.035)	(0.066)	(0.051)	(0.048)	(0.036)	(0.050)
$AMBIGUITY_t$ x $HIGH$			0.020	0.013			0.024	0.031*			0.043**	0.048**
			(0.017)	(0.019)			(0.015)	(0.017)			(0.019)	(0.020)
$RISK_t$ x $HIGH$				0.093				-0.103				-0.109
				(0.080)				(0.079)				(0.068)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	10901	22149	33050	33050	10974	22076	33050	33050	14777	18273	33050	33050
N firms	714	713	1427	1427	714	713	1427	1427	714	713	1427	1427
Adj R2	0.788	0.735	0.768	0.768	0.789	0.740	0.768	0.768	0.768	0.761	0.768	0.768
Panel B: RD_AT four quarters ahead (quarters $t+1 \ldots t+4$)												
$AMBIGUITY_t$	-0.197***	-0.109**	-0.175***	-0.172***	-0.194***	-0.109***	-0.147***	-0.183***	-0.209***	-0.100***	-0.257***	-0.270***
	(0.050)	(0.043)	(0.052)	(0.054)	(0.056)	(0.040)	(0.051)	(0.055)	(0.063)	(0.035)	(0.069)	(0.072)
$RISK_t$	0.825***	0.548*	0.662***	0.646***	1.105***	0.459***	0.643***	1.048***	0.574***	0.661***	0.662***	0.907***
	(0.150)	(0.283)	(0.148)	(0.168)	(0.249)	(0.169)	(0.141)	(0.260)	(0.199)	(0.186)	(0.141)	(0.200)
$AMBIGUITY_t$ x $HIGH$			0.062	0.059			0.023	0.066			0.194**	0.215***
			(0.064)	(0.071)			(0.059)	(0.067)			(0.076)	(0.079)
$RISK_t$ x $HIGH$				0.051				-0.559*				-0.523*
				(0.320)				(0.314)				(0.270)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9888	19998	29886	29886	10402	19484	29886	29886	13637	16249	29886	29886
N firms	653	653	1306	1306	654	652	1306	1306	653	653	1306	1306
Adj R2	0.837	0.819	0.831	0.831	0.842	0.814	0.831	0.831	0.820	0.835	0.831	0.831

Table 5: Determinants of R&D Investment: Robustness Tests

The table presents OLS regression coefficients for R&D investment. The dependent variable is RD_ASSETS_{t+1} in Panel A, $RD\_CAPEX\_ASSETS_{t+1}$ in Panel B and $RD\_ADJ\_ASSETS_{t+1}$ in Panel C. The sample period is 1993-2016. The sample consists of all firms with at least four quarters of data for all variables of interest and at least one quarter of positive R&D expenditures in Compustat during the sample period (*R&D Sample*), excluding penny stocks, very small firms and very young firms. *Penny stocks* are stocks with a price less than \$5 at the end of quarter t. *Very small firms* are firms with a market capitalization less than \$10 million at the end of quarter t. *Very young firms* are firms with less than 5 years in Compustat. All regressions include the following control variables: LN_SALES_t, Q_t, $LN\_K\_L_t$, $CASH\_FLOW_t$, $LEVERAGE_t$, LN_AGE_{t+1}, $LN\_RD\_CAPITAL_t$, $NASDAQ_t$ and $MISSING\_RD_{t+1}$. The denominator used to calculate Q_t is the book value of assets (Compustat item *atq* at the end of quarter t) in Panels A and B, and the book value of assets plus capitalized R&D (Compustat item *atq* at the end of quarter t plus $RD\_CAPITAL_t$) in Panel C. In columns (1), (3) and (5), $MISSING\_RD_{t+1}$ is an indicator variable equal to 1 if the firm has missing R&D expenditures in Compustat in quarter $t+1$. In columns (2), (4) and (6), $MISSING\_RD_{t+1}$ is the number of quarters with missing R&D in Compustat in the period $t+1 \ldots t+4$. All regressions include firm (*new gvkey*) fixed effects and quarter-year fixed effects. Standard errors are clustered by firm. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

	All Firms		**High-Tech**		**Non High-Tech**	
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
	One Quarter	**One Year**	**One Quarter**	**One Year**	**One Quarter**	**One Year**
	$t+1$	$t+1 \ldots t+4$	$t+1$	$t+1 \ldots t+4$	$t+1$	$t+1 \ldots t+4$
Panel A: Controlling for Institutional Ownership						
$AMBIGUITY_t$	-0.009**	-0.066***	-0.030***	-0.130***	0.002	-0.035***
	(0.005)	(0.016)	(0.009)	(0.034)	(0.004)	(0.013)
$RISK_t$	0.129***	0.478***	0.198***	0.629***	0.055*	0.291**
	(0.022)	(0.088)	(0.035)	(0.139)	(0.029)	(0.117)
$INSTOWN\_DED_t$	-0.000	0.005	-0.002	-0.003	0.001	0.006
	(0.001)	(0.005)	(0.002)	(0.009)	(0.001)	(0.004)
$INSTOWN\_TRA_t$	-0.001	-0.003	-0.002	-0.003	-0.001	-0.004
	(0.001)	(0.003)	(0.001)	(0.005)	(0.001)	(0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	69,772	64,534	33,050	29,886	36,371	33,894
N firms	2,508	2,342	1,427	1,306	1,170	1,104
Adj R2	0.809	0.878	0.768	0.831	0.779	0.882
Panel B: Total Investment (R&D plus CAPEX, $RD\_CAPEX\_ASSETS_{t+1}$**)**						
$AMBIGUITY_t$	-0.007	-0.063***	-0.028**	-0.097**	0.005	-0.045*
	(0.006)	(0.024)	(0.012)	(0.049)	(0.007)	(0.024)
$RISK_t$	0.121***	0.569***	0.188***	0.632***	0.032	0.400**
	(0.033)	(0.138)	(0.049)	(0.212)	(0.043)	(0.177)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	69,201	63,632	32,754	29,396	36,092	33,495
N firms	2,495	2,315	1,422	1,292	1,159	1,090
Adj R2	0.698	0.777	0.674	0.750	0.657	0.752
Panel C: Adjusting Total Assets For Capitalized R&D ($RD\_ADJ\_ASSETS_{t+1}$**)**						
$AMBIGUITY_t$	-0.003	-0.032***	-0.012**	-0.057***	0.002	-0.020**
	(0.003)	(0.010)	(0.006)	(0.022)	(0.003)	(0.010)
$RISK_t$	0.054***	0.198***	0.083***	0.253***	0.028	0.141*
	(0.015)	(0.059)	(0.022)	(0.089)	(0.021)	(0.082)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	69,631	64,102	32,978	29,597	36,310	33,756
N firms	2,500	2,333	1,424	1,299	1,166	1,102
Adj R2	0.814	0.890	0.776	0.837	0.766	0.892

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 6: Determinants of Patenting Activity in High-Tech Firms

The table presents estimation results for count models for patenting activity. In Panel A, the dependent variable is $PATENTS$, and the sample consists of all high-tech firms (3-digit SIC codes 283, 357, 366, 367, 382, 384, or 737) with at least four quarters of data for all variables of interest, four years in the presample period and at least one patent application filed during the sample period, excluding penny stocks, very small firms and very young firms. *Penny stocks* are stocks with a price less than \$5 at the end of quarter t. *Very small firms* are firms with a market capitalization less than \$10 million at the end of quarter t. *Very young firms* are firms with less than 5 years in Compustat. In Panel B, the dependent variable is $CITATIONS$, and the sample is further restricted to firms that have at least one *cited* patent applied for during the sample period (the *Citation Sample*). Marginal effects are calculated as differences in predicted counts at high (90th percentile of the estimation sample) and low (10th percentile of the estimation sample) $AMBIGUITY_t$ and $RISK_t$, while keeping all other variables at their sample means. The sample period is 1993-2008. All regressions include three-digit SIC code fixed effects, Blundell et al. (1999) presample firm fixed effects, and quarter-year fixed effects. Standard errors are clustered by firm. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

Panel A: Patents

	Poisson				**Negative Binomial**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	One quarter	**Year 1**	**Year 2**	**Year 3**	**One quarter**	**Year 1**	**Year 2**	**Year 3**
	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$
Coefficients								
$AMBIGUITY_t$	-4.742**	-5.659**	-5.780**	-5.383**	-1.980	-0.967	-0.671	-0.860
	(2.367)	(2.363)	(2.483)	(2.387)	(2.054)	(1.868)	(2.163)	(2.254)
$RISK_t$	-30.488**	-33.495**	-42.265***	-46.488***	-12.470**	-9.992*	-13.594**	-15.199***
	(13.013)	(13.356)	(14.485)	(15.247)	(5.795)	(5.330)	(5.690)	(5.804)
LN_SALES_t	0.241***	0.253***	0.274***	0.261***	0.189***	0.205***	0.222***	0.229***
	(0.076)	(0.077)	(0.078)	(0.079)	(0.047)	(0.046)	(0.051)	(0.059)
Q_t	-0.006	0.005	0.029	0.044*	0.034**	0.044***	0.056***	0.050***
	(0.028)	(0.025)	(0.024)	(0.023)	(0.016)	(0.015)	(0.017)	(0.018)
$LN\_K\_L_t$	0.701***	0.722***	0.791***	0.838***	0.340***	0.328***	0.369***	0.408***
	(0.154)	(0.144)	(0.134)	(0.122)	(0.063)	(0.058)	(0.064)	(0.071)
$CASH\_FLOW_t$	0.874	0.990	1.152	1.606**	-0.411	-0.047	0.221	1.646*
	(0.721)	(0.701)	(0.752)	(0.811)	(0.693)	(0.713)	(0.788)	(0.843)
$LEVERAGE_t$	-0.347	-0.342	-0.378	-0.492	-0.156	-0.155	-0.197	-0.203
	(0.289)	(0.289)	(0.312)	(0.343)	(0.209)	(0.212)	(0.244)	(0.284)
LN_AGE_{t+1}	-0.130	-0.144	-0.162	-0.197*	0.007	-0.045	-0.112	-0.199**
	(0.099)	(0.099)	(0.106)	(0.117)	(0.074)	(0.078)	(0.087)	(0.098)
$LN\_RD\_CAPITAL_t$	0.318***	0.312***	0.290***	0.291***	0.380***	0.359***	0.364***	0.364***
	(0.073)	(0.071)	(0.072)	(0.078)	(0.045)	(0.043)	(0.050)	(0.057)
$NASDAQ_t$	-0.207	-0.209	-0.216	-0.212	0.106	0.108	0.102	0.075
	(0.140)	(0.137)	(0.133)	(0.131)	(0.114)	(0.112)	(0.121)	(0.129)
Constant	-6.473***	-5.208***	-5.481***	-5.643***	-5.907***	-4.012***	-3.625***	-3.725***
	(0.707)	(0.683)	(0.703)	(0.727)	(0.530)	(0.515)	(0.572)	(0.635)
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	16,823	15,294	12,985	10,849	16,823	15,294	12,985	10,849
N firms	819	819	768	699	819	819	768	699
Pseudo R-squared					0.208	0.176	0.171	0.165
Marginal Effects								
(1) Low Ambiguity	1.452	6.198	6.213	6.153	1.568	6.689	7.048	7.362
(2) High Ambiguity	1.242	5.123	5.165	5.210	1.468	6.475	6.898	7.169
Marginal Effect (2)-(1)	-0.211**	-1.075**	-1.048**	-0.943**	-0.099	-0.214	-0.149	-0.193
	(0.102)	(0.433)	(0.435)	(0.404)	(0.102)	(0.412)	(0.480)	(0.504)
(3) Low Risk	1.573	6.760	7.180	7.422	1.619	6.932	7.496	7.913
(4) High Risk	0.982	3.980	3.552	3.329	1.335	5.919	5.978	6.089
Marginal Effect (4)-(3)	-0.591**	-2.779***	-3.627***	-4.092***	-0.284**	-1.013*	-1.518**	-1.825***
	(0.243)	(1.057)	(1.147)	(1.196)	(0.127)	(0.523)	(0.609)	(0.671)

Electronic copy available at: https://ssrn.com/abstract=3428896

Panel B: Citations

	Poisson				Negative Binomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	One quarter	**Year 1**	**Year 2**	**Year 3**	**One quarter**	**Year 1**	**Year 2**	**Year 3**
	$t+1$	$t+1\ldots t+4$	$t+5\ldots t+8$	$t+9\ldots t+12$	$t+1$	$t+1\ldots t+4$	$t+5\ldots t+8$	$t+9\ldots t+12$
Coefficients								
$AMBIGUITY_t$	-5.915***	-6.681***	-5.849**	-4.897*	-1.521	-1.641	-1.239	-1.510
	(2.197)	(2.252)	(2.464)	(2.579)	(2.323)	(2.010)	(2.364)	(2.439)
$RISK_t$	-36.567***	-37.451***	-47.887***	-49.462***	-13.225**	-9.910	-15.072**	-16.482***
	(12.333)	(12.720)	(14.386)	(15.394)	(6.445)	(6.047)	(6.557)	(6.379)
LN_SALES_t	0.214***	0.233***	0.232***	0.211***	0.177***	0.204***	0.218***	0.230***
	(0.067)	(0.068)	(0.071)	(0.076)	(0.045)	(0.045)	(0.052)	(0.059)
Q_t	-0.002	0.014	0.033	0.043*	0.032*	0.051***	0.070***	0.060***
	(0.026)	(0.025)	(0.024)	(0.024)	(0.018)	(0.017)	(0.019)	(0.022)
$LN\_K\_L_t$	0.606***	0.624***	0.666***	0.706***	0.270***	0.267***	0.297***	0.352***
	(0.105)	(0.098)	(0.094)	(0.088)	(0.060)	(0.061)	(0.068)	(0.076)
$CASH\_FLOW_t$	1.088	0.823	1.510**	1.875**	0.124	-0.525	0.504	1.578
	(0.682)	(0.641)	(0.677)	(0.785)	(0.738)	(0.787)	(0.865)	(1.005)
$LEVERAGE_t$	-0.227	-0.209	-0.236	-0.374	-0.110	-0.215	-0.307	-0.465*
	(0.275)	(0.272)	(0.298)	(0.324)	(0.232)	(0.230)	(0.260)	(0.266)
LN_AGE_{t+1}	-0.104	-0.125	-0.168	-0.220*	-0.011	-0.062	-0.152	-0.253**
	(0.094)	(0.097)	(0.110)	(0.125)	(0.087)	(0.091)	(0.102)	(0.110)
$LN\_RD\_CAPITAL_t$	0.279***	0.275***	0.267***	0.273***	0.287***	0.270***	0.271***	0.273***
	(0.058)	(0.058)	(0.066)	(0.076)	(0.040)	(0.041)	(0.049)	(0.056)
$NASDAQ_t$	-0.034	-0.034	-0.056	-0.072	0.237*	0.186	0.155	0.111
	(0.128)	(0.128)	(0.138)	(0.145)	(0.123)	(0.121)	(0.135)	(0.143)
Constant	-5.774***	-4.529***	-4.452***	-4.497***	-5.093***	-3.205***	-2.504***	-2.795***
	(0.681)	(0.652)	(0.689)	(0.739)	(0.577)	(0.614)	(0.740)	(0.781)
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	16,186	14,731	12,548	10,516	16,186	14,731	12,548	10,516
N firms	775	775	729	669	775	775	729	669
Pseudo R-squared					0.144	0.132	0.127	0.126
Marginal Effects								
(1) Low Ambiguity	2.052	8.753	8.694	8.438	2.071	8.901	9.329	9.389
(2) High Ambiguity	1.685	6.980	7.195	7.242	1.969	8.419	8.962	8.956
Marginal Effect (2)-(1)	-0.367***	-1.773***	-1.499**	-1.196*	-0.102	-0.481	-0.367	-0.432
	(0.134)	(0.589)	(0.621)	(0.618)	(0.156)	(0.588)	(0.695)	(0.693)
(3) Low Risk	2.250	9.590	10.320	10.388	2.159	9.134	9.922	10.077
(4) High Risk	1.279	5.308	4.649	4.425	1.760	7.811	7.720	7.583
Marginal Effect (4)-(3)	-0.971***	-4.282***	-5.672***	-5.962***	-0.399**	-1.323*	-2.203**	-2.494***
	(0.315)	(1.387)	(1.565)	(1.676)	(0.187)	(0.782)	(0.924)	(0.945)

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 7: Determinants of Patenting Activity in Patent-Intensive High-Tech Firms

The table presents estimation results for count models for patenting activity. The dependent variable is PATENTS in Panel A, and CITATIONS in Panel B. The sample is the same as in Table 6, but restricted to firms above the sample median in terms of the average number of patents applied for during the sample period. Marginal effects are calculated as differences in predicted counts at high (90th percentile of the estimation sample) and low (10th percentile of the estimation sample) $AMBIGUITY_t$ and $RISK_t$, while keeping all other variables at their sample means. The sample period is 1993-2008. All regressions include three-digit SIC code fixed effects, Blundell et al. (1999) presample firm fixed effects, and quarter-year fixed effects. Standard errors are clustered by firm. Sample construction is explained in detail in Section 3.1. For variable definitions see Appendix B. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

Panel A: Patents

	Poisson				**Negative Binomial**			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	One quarter	**Year 1**	**Year 2**	**Year 3**	**One quarter**	**Year 1**	**Year 2**	**Year 3**
	$t+1$	$t+1\ldots t+4$	$t+5\ldots t+8$	$t+9\ldots t+12$	$t+1$	$t+1\ldots t+4$	$t+5\ldots t+8$	$t+9\ldots t+12$
Coefficients								
$AMBIGUITY_t$	-4.792**	-5.602**	-5.527**	-4.963**	-3.980*	-3.918**	-4.213**	-3.430
	(2.333)	(2.334)	(2.497)	(2.418)	(2.053)	(1.867)	(2.020)	(2.136)
$RISK_t$	-24.008	-27.669*	-35.639**	-39.168**	-5.907	-3.061	-4.299	-6.485
	(14.826)	(15.229)	(16.474)	(17.291)	(7.186)	(6.465)	(6.715)	(6.845)
LN_SALES_t	0.274***	0.285***	0.308***	0.292***	0.218***	0.227***	0.250***	0.260***
	(0.082)	(0.083)	(0.084)	(0.088)	(0.043)	(0.042)	(0.047)	(0.056)
Q_t	-0.009	0.004	0.029	0.046**	0.026	0.037**	0.058***	0.057***
	(0.027)	(0.025)	(0.023)	(0.023)	(0.017)	(0.016)	(0.017)	(0.019)
$LN\_K\_L_t$	0.686***	0.706***	0.776***	0.824***	0.344***	0.344***	0.391***	0.420***
	(0.156)	(0.145)	(0.134)	(0.122)	(0.073)	(0.068)	(0.075)	(0.082)
$CASH\_FLOW_t$	0.579	0.684	0.792	1.238	-0.932	-0.708	-0.405	0.418
	(0.713)	(0.706)	(0.760)	(0.832)	(0.652)	(0.672)	(0.714)	(0.759)
$LEVERAGE_t$	-0.330	-0.323	-0.347	-0.461	-0.236	-0.271	-0.323	-0.493*
	(0.289)	(0.290)	(0.315)	(0.346)	(0.205)	(0.204)	(0.231)	(0.270)
LN_AGE_{t+1}	-0.137	-0.154	-0.171	-0.203*	-0.056	-0.086	-0.138*	-0.215**
	(0.098)	(0.099)	(0.107)	(0.120)	(0.073)	(0.074)	(0.080)	(0.092)
$LN\_RD\_CAPITAL_t$	0.279***	0.276***	0.251***	0.253***	0.331***	0.313***	0.302***	0.292***
	(0.077)	(0.075)	(0.077)	(0.085)	(0.044)	(0.043)	(0.051)	(0.059)
$NASDAQ_t$	-0.222	-0.228*	-0.238*	-0.233*	-0.061	-0.083	-0.113	-0.144
	(0.138)	(0.135)	(0.133)	(0.132)	(0.110)	(0.110)	(0.115)	(0.123)
Constant	-3.938***	-2.686***	-2.813***	-2.960***	-3.715***	-2.135***	-2.004***	-1.916***
	(0.728)	(0.697)	(0.776)	(0.765)	(0.619)	(0.596)	(0.686)	(0.725)
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,581	8,738	7,518	6,392	9,581	8,738	7,518	6,392
N firms	409	409	382	352	409	409	382	352
Pseudo R-squared					0.170	0.152	0.150	0.144
Marginal Effects								
(1) Low Ambiguity	4.793	20.797	21.439	21.865	5.382	23.369	25.127	26.273
(2) High Ambiguity	4.054	17.000	17.777	18.662	4.684	20.295	21.784	23.549
Marginal Effect (2)-(1)	-0.739**	-3.797**	-3.662**	-3.203**	-0.698*	-3.073**	-3.342**	-2.724
	(0.346)	(1.514)	(1.595)	(1.517)	(0.362)	(1.476)	(1.609)	(1.694)
(3) Low Risk	4.797	20.833	22.220	23.328	5.175	22.296	24.056	25.699
(4) High Risk	3.971	16.587	16.246	16.193	4.939	21.741	23.164	24.191
Marginal Effect (4)-(3)	-0.827	-4.246*	-5.974**	-7.134**	-0.235	-0.555	-0.892	-1.507
	(0.511)	(2.327)	(2.703)	(3.016)	(0.283)	(1.165)	(1.381)	(1.578)

Electronic copy available at: https://ssrn.com/abstract=3428896

Panel B: Citations

	Poisson				Negative Binomial			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	One quarter	**Year 1**	**Year 2**	**Year 3**	**One quarter**	**Year 1**	**Year 2**	**Year 3**
	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$
Coefficients								
$AMBIGUITY_t$	-6.100***	-6.688***	-5.389**	-4.074	-5.776**	-6.417***	-4.618**	-4.262*
	(2.274)	(2.350)	(2.601)	(2.686)	(2.295)	(2.020)	(2.226)	(2.312)
$RISK_t$	-27.648**	-27.280**	-37.436**	-35.963**	-3.941	0.454	-4.859	-4.985
	(13.611)	(13.830)	(15.814)	(16.439)	(7.494)	(6.628)	(6.850)	(6.854)
LN_SALES_t	0.233***	0.253***	0.253***	0.230***	0.195***	0.228***	0.241***	0.260***
	(0.070)	(0.071)	(0.074)	(0.082)	(0.040)	(0.041)	(0.045)	(0.054)
Q_t	-0.002	0.014	0.033	0.045*	0.039**	0.054***	0.073***	0.073***
	(0.026)	(0.024)	(0.024)	(0.024)	(0.018)	(0.018)	(0.019)	(0.021)
$LN\_K\_L_t$	0.589***	0.605***	0.646***	0.685***	0.291***	0.310***	0.355***	0.393***
	(0.106)	(0.099)	(0.095)	(0.089)	(0.065)	(0.066)	(0.073)	(0.078)
$CASH\_FLOW_t$	0.893	0.645	1.286*	1.590**	-0.043	-0.677	0.545	0.852
	(0.668)	(0.639)	(0.684)	(0.803)	(0.676)	(0.699)	(0.710)	(0.839)
$LEVERAGE_t$	-0.199	-0.169	-0.196	-0.332	-0.123	-0.170	-0.406*	-0.713***
	(0.280)	(0.278)	(0.305)	(0.332)	(0.236)	(0.228)	(0.241)	(0.241)
LN_AGE_{t+1}	-0.109	-0.132	-0.178	-0.227*	-0.020	-0.048	-0.123	-0.212**
	(0.095)	(0.098)	(0.113)	(0.130)	(0.085)	(0.086)	(0.093)	(0.098)
$LN\_RD\_CAPITAL_t$	0.252***	0.249***	0.239***	0.245***	0.228***	0.205***	0.186***	0.169***
	(0.060)	(0.060)	(0.069)	(0.082)	(0.037)	(0.038)	(0.045)	(0.053)
$NASDAQ_t$	-0.063	-0.067	-0.096	-0.117	0.054	0.016	-0.037	-0.097
	(0.126)	(0.126)	(0.137)	(0.146)	(0.114)	(0.113)	(0.121)	(0.122)
Constant	-3.611***	-2.416***	-2.147***	-2.052***	-3.206***	-1.773***	-1.260**	-1.400**
	(0.607)	(0.583)	(0.611)	(0.653)	(0.497)	(0.528)	(0.623)	(0.594)
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,042	8,240	7,090	6,019	9,042	8,240	7,090	6,019
N firms	387	387	359	327	387	387	359	327
Pseudo R-squared					0.108	0.111	0.107	0.106
Marginal Effects								
(1) Low Ambiguity	6.745	29.351	29.824	30.026	7.384	32.312	33.443	34.325
(2) High Ambiguity	5.416	22.970	24.752	26.285	5.999	25.540	28.505	29.864
Marginal Effect (2)-(1)	-1.329***	-6.381***	-5.072**	-3.741	-1.385**	-6.773***	-4.937**	-4.461*
	(0.486)	(2.195)	(2.406)	(2.434)	(0.552)	(2.144)	(2.356)	(2.389)
(3) Low Risk	6.666	28.778	30.920	31.802	6.870	29.402	31.846	33.004
(4) High Risk	5.422	23.325	22.918	23.580	6.670	29.505	30.632	31.664
Marginal Effect (4)-(3)	-1.245**	-5.454**	-8.002**	-8.223**	-0.199	0.103	-1.214	-1.341
	(0.616)	(2.775)	(3.343)	(3.693)	(0.377)	(1.504)	(1.705)	(1.842)

Electronic copy available at: https://ssrn.com/abstract=3428896

Table 8: Determinants of Patenting Activity in High-Tech Firms: Robustness

The table presents estimation results for count models for patenting activity. The sample and methods are the same as in Table 6, but augmented to include institutional ownership variables. All regressions include the following cotrol variables: LN_SALES_t, Q_t, $LN\_K\_L_t$, $CASH\_FLOW_t$, $LEVERAGE_t$, LN_AGE_{t+1}, $LN\_RD\_CAPITAL_t$, $NASDAQ_t$. *, ** and *** denote significance at the 10%, 5% and 1% levels respectively.

	Poisson				Negative Binomial			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	One quarter	Year 1	Year 2	Year 3	One quarter	Year 1	Year 2	Year 3
	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$	$t+1$	$t+1\dots t+4$	$t+5\dots t+8$	$t+9\dots t+12$
Panel A: Patents								
Coefficients								
$AMBIGUITY_t$	-4.083*	-4.914**	-4.963**	-4.402*	-1.971	-0.963	-0.638	-0.755
	(2.180)	(2.201)	(2.414)	(2.281)	(2.063)	(1.874)	(2.165)	(2.231)
$RISK_t$	-36.719***	-39.696***	-46.197***	-47.362***	-12.430**	-9.382*	-11.376**	-13.129**
	(12.632)	(12.889)	(13.656)	(14.643)	(5.591)	(5.130)	(5.588)	(5.928)
$INSTOWN\_DED_t$	1.280**	1.243**	1.176***	1.353***	0.031	-0.031	0.136	0.407
	(0.555)	(0.493)	(0.361)	(0.328)	(0.346)	(0.334)	(0.355)	(0.377)
$INSTOWN\_TRA_t$	-0.858**	-0.831**	-0.615	-0.419	0.005	0.114	0.413	0.357
	(0.383)	(0.393)	(0.404)	(0.434)	(0.263)	(0.262)	(0.277)	(0.323)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	16,823	15,294	12,985	10,849	16,823	15,294	12,985	10,849
N firms	819	819	768	699	819	819	768	699
Pseudo R-squared					0.208	0.176	0.171	0.165
Marginal Effects								
(1) Low Ambiguity	1.422	6.050	6.057	5.954	1.568	6.688	7.033	7.338
(2) High Ambiguity	1.243	5.128	5.168	5.197	1.469	6.475	6.891	7.169
Marginal Effect (2)-(1)	-0.180*	-0.922**	-0.888**	-0.757**	-0.099	-0.213	-0.142	-0.169
	(0.093)	(0.400)	(0.420)	(0.380)	(0.103)	(0.414)	(0.480)	(0.498)
(3) Low Risk	1.600	6.869	7.218	7.303	1.619	6.910	7.399	7.810
(4) High Risk	0.908	3.667	3.345	3.227	1.336	5.958	6.122	6.227
Marginal Effect (4)-(3)	-0.693***	-3.202***	-3.873***	-4.076***	-0.283**	-0.953*	-1.276**	-1.582**
	(0.227)	(0.977)	(1.036)	(1.108)	(0.122)	(0.505)	(0.603)	(0.689)
Panel B: Citations								
Coefficients								
$AMBIGUITY_t$	-5.514***	-6.206***	-5.177**	-3.936	-1.530	-1.624	-1.217	-1.446
	(2.108)	(2.170)	(2.442)	(2.514)	(2.331)	(2.015)	(2.363)	(2.415)
$RISK_t$	-40.684***	-41.274***	-48.471***	-47.067***	-12.689*	-8.579	-11.849*	-14.398**
	(12.252)	(12.728)	(14.237)	(15.226)	(6.483)	(5.987)	(6.585)	(6.664)
$INSTOWN\_DED_t$	0.755*	0.769*	0.926**	1.279***	-0.244	-0.489	-0.318	0.225
	(0.421)	(0.419)	(0.396)	(0.448)	(0.335)	(0.347)	(0.384)	(0.418)
$INSTOWN\_TRA_t$	-0.559	-0.522	-0.238	-0.097	0.129	0.342	0.696**	0.417
	(0.351)	(0.365)	(0.412)	(0.445)	(0.296)	(0.277)	(0.317)	(0.382)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Presample Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry (SIC3) FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter-year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	16,186	14,731	12,548	10,516	16,186	14,731	12,548	10,516
N firms					775	775	729	669
Pseudo R-squared					0.144	0.132	0.127	0.126
Marginal Effects								
(1) Low Ambiguity	2.031	8.640	8.525	8.184	2.071	8.888	9.297	9.367
(2) High Ambiguity	1.690	7.001	7.210	7.237	1.968	8.412	8.938	8.953
Marginal Effect (2)-(1)	-0.340***	-1.639***	-1.315**	-0.946	-0.103	-0.476	-0.359	-0.413
	(0.128)	(0.564)	(0.611)	(0.594)	(0.156)	(0.589)	(0.693)	(0.686)
(3) Low Risk	2.282	9.706	10.238	10.062	2.153	9.064	9.729	9.948
(4) High Risk	1.217	5.057	4.567	4.467	1.770	7.916	7.986	7.760
Marginal Effect (4)-(3)	-1.065***	-4.648***	-5.671***	-5.594***	-0.383**	-1.149	-1.742*	-2.188**
	(0.305)	(1.359)	(1.518)	(1.630)	(0.187)	(0.779)	(0.939)	(0.988)

Electronic copy available at: https://ssrn.com/abstract=3428896