Measuring Innovation

Michael Cooper, Anne Marie Knott, and Wenhao Yang

Below is a summary of the main results and an explanation of the relevance to the Q-Group sponsors and to Q's mission—the advancement of the global practice of investment management.

In this paper we develop a new and better way to measure firm innovation. Since accurately measuring innovation is critical to estimating firm value, we view this work as important. We create a measure of innovation that is tied to R&D, and not to patents or citations. This is important, since R&D measures can be created for a much larger sample of firms than patent or citation data. Our measure of innovation, a firm's research quotient (RQ), is defined as the firm-specific output elasticity of R&D. We show that RQ is more strongly associated with current and future firm value than previous patent-based measures. In fact, in many of our tests, RQ subsumes the information in patent and citation based measures of innovation. We find the clear result that present and future firm value (measured in term of Tobin's Q and returns) is positively and statistically significantly related to RQ. Within the sample of firms with R&D, we find the very important result that RQ is the strongest predictor of the cross-section of future returns (stronger than BM, size, asset growth and momentum). The relevance to the Q-group would seem to be in term of the predictive ability of RQ; within the sample of firms that innovate, RQ is a strong and robust predictor of future returns.

Measuring Innovation

Michael Cooper, Anne Marie Knott, and Wenhao Yang*

June 23, 2015

Keywords: Innovation, Research Quotient, Firm Value, Patents, Citations

JEL Classification: 031, 032, G12, G14

---

* Cooper and Yang are with the Eccles School of Business, University of Utah. Knott is with the Olin Business School, Washington University in St. Louis. We would like to thank Vivian Fang, Radha Gopalan, Bronwyn Hall, Dietmar Harhoff, Jack He, Dongmei Li, Angie Low, Pierre Mohnen, Theo Sougiannis, Kandarp Srinivasan and Chad Syverson for helpful comments. All errors and omissions remain our own.

Measuring Innovation

Abstract

Innovation is vital to firm value. Accordingly a number of recent papers in the finance literature have begun examining the impact of governance and other firm measures on innovation. To date these papers have used patent-based measures to capture innovation. However such measures ignore the greater than 50% of firms who do R&D but don't patent their innovations. To address that concern, we examine the feasibility of using a more universal measure of innovation — the firm's research quotient (RQ), defined as the firm-specific output elasticity of R&D. We show that RQ is more strongly associated with firm value than previous patent-based measures. Given RQ's universality advantages and stronger relation with firm value relative to patents, we propose using RQ in future studies of innovation.

I. Introduction

Firm innovation and related investment are vital to firm growth. While the economics and management literatures have examined innovation since the watershed Solow (1957) paper, the study of firm innovation is a relatively new topic in the finance literature. The finance field's interest stems from the realization that innovation is linked to firm value and that governance and other firm characteristics play substantial roles in a firm's ability to innovate. Accordingly, research examines the impact on innovation of governance measures such as takeover provisions (Becker-Blease 2011), analyst coverage (He and Tian 2013), institutional ownership (Becker-Blease 2011, He and Tian 2013, Aghion, Van Reenan and Zingales 2013), illiquidity (He and Tian 2013, Gormley, Matsa, and Milbourn 2013, Fang, Tian and Tice 2014), and leverage (Becker-Blease 2011, He and Tian 2013, Gormley, Matsa, and Milbourn 2013).[1] An important concern with many of these papers is that they (like much of the historical work in economics and management) tend to rely on patent-based measures. However, as we discuss below, these measures have weaknesses that potentially lead to flawed or incomplete inferences regarding the factors affecting innovation.

In this paper, we propose a new measure of innovation, research quotient (RQ), as a better measure of firm innovation. The measure is not based on patent or citation data, so it avoids the potential flaws and limitations of such measures. As we will show, our RQ measure is better at capturing the value creation to a firm of innovation.

One of the primary concerns with patent-based measures is they are not universal. In our sample, fewer than 50% of the firms in COMPUSTAT who conduct R&D actually file patents for their innovations.[2] Thus over half the firms conducting R&D will appear to have no innovation. This limits the sample size and decreases the power in any test based on these measures for publically traded firms.

Beyond the practical problem of limited sample size, is the identification issue that the decision to patent is endogenous to firm innovation policy. Firms choose which (if any) innovations to patent. Because patents are costly both financially (the cost to file

[1] Other work examines the CEO and innovation, not only CEO compensation (Lerner and Wulf 2007, He and Tian 2013, Gormley, Matsa, and Milbourn 2013), but also CEO personality (overconfidence) (Galasso and Simcoe 2011, Hirschleifer, Low and Teoh 2012).

[2] In our sample, 37% of firms with R&D data have patents.

and defend) as well as competitively (they require disclosure of the fundamental knowledge underpinning the innovation), firms file patents only under certain circumstances. They do so primarily to prevent copying when their innovations are easy to invent around (Cohen, Nelson and Walsh 2000), though also for strategic reasons, such as to block other firms' patents (82% of surveyed firms), to prevent lawsuits (59%), to use in negotiations (47%), or to enhance their reputation (48%).

In a paper which explicitly tackles concerns with patent-based measures, Abrams, Akcigit and Popadak (2013) develop a model that jointly considers productive and strategic rationale for patenting. They expect and find empirically that the relationship between citations and patent value is non-monotonic—suggesting that firms are less likely to patent their most valuable innovations. Thus firms who patent and the innovations they patent may be structurally different from those who don't. If so, then policy recommendations emerging from these studies may be misguided.[3]

Perhaps the most important concern with patents however is that their economic value is not uniform. Scherer and Harhoff (2000) report that 10% of U.S. patents account for 81%-85% of the economic value of all U.S. patents. Recognizing this, many patent-based studies weight patents by the number of citations they receive, or use the total citations (rather than total patents) received by the firm. Indeed, Hall, Jaffe and Trajtenberg (2001) show patent citations are a better predictor of firm value than patent counts. However Abrams et al (2013), discussed above, show that even this correction may be problematic. Thus while citations may help mitigate the non-uniformity problem of patent value, they don't solve it.[4]

Given these concerns, the economics and management literatures have each introduced alternative firm-level measures of innovation. The economics measure is total factor productivity (TFP); the management measure is research quotient (RQ). (See

[3] Further, patent counts may not accurately measure innovation even within the set of firms who do patent. This is because a single patent can be assigned to multiple assignees, each with potentially different contributions to the patent. These contributions cannot be quantified or reflected in the data.

[4] Another practical problem is that citations take years to materialize. Mechanically, an older patent can receive more citations than a newer patent, even if the older citation has only marginal value and the new citation is path breaking. This is known as "truncation bias" (Hall, Jaffe and Trajtenberg 2001). One solution is to use an adjustment factor that predicts future citations. Another solution is use a sample from an earlier period and to create a later-in-time hold-out period for citations to materialize. Both methods partially address the truncation bias, but to a limited degree and at a high cost (i.e., the first method requires restrictive assumptions, and the second method further limits the sample size).

Syverson 2011 for discussion of firm-level estimation of TFP; see Knott 2008 for discussion of firm-level estimation of RQ). Both TFP and RQ capture the source of technological change in the firm's production function, but they differ in the source of technological change. TFP maps onto Solow's (1957) theory of exogenous technological change—thus implicitly assumes R&D is being done outside the firm, e.g., in universities and/or government labs. Accordingly it is modeled as a shifter in the firm's production function, and empirically captured as a fixed effect or the residual in the firm's production function (Hulten 2000).

In contrast RQ, defined as the firm-specific output elasticity of R&D, maps onto models of endogenous technological change (Romer 1990, Thompson 1996, Lentz and Mortensen 2008) - thus treating R&D as investments made by profit-maximizing agents. Accordingly it is imbedded in the firm's production function as a continually evolving stock of knowledge, and is captured empirically as the output elasticity of R&D.

Industry-level versions of these measures have been used extensively in studies at the industry and economic level (See Hulten 2000 for a review of TFP studies; see Hall, Mairesse and Mohnen 2010 for a review of returns to R&D). The firm-level innovation in these measures stems from the recognition there is tremendous heterogeneity among firms even within the same industry

Both TFP and RQ offer two significant advantages over patent-based measures. First, they are *universal*—they can be estimated for any firm doing R&D, whereas patent studies restrict the sample size to the 50% of firms who patent their R&D. Second, the fact they are unitless (TFP is a shifter, RQ is an elasticity) means their interpretation is *uniform* across firms, whereas patent-value varies substantially.

Both universality and uniformity offer the potential for more robust tests. However of the two measures, RQ is the more appropriate for tests of governance and innovation because its foundational assumption (firms choose R&D to maximize firm value) matches the assumption in the governance studies that CEOs (and firm governance) affect the level of innovation.

In this paper, we investigate the feasibility of adopting RQ as an alternative measure of innovation within the finance literature. To examine feasibility, we subject both RQ and patent-based measures to a fairly standard battery of tests of their ability to capture things we expect from an innovation measure: correlations with firm characteristics related to innovation effort, the ability to predict firm value (as measured by both contemporaneous and subsequent Market-to-Book Ratios) (MTB), as well as Fama-MacBeth regressions of future stock returns.

We find first that RQ is significantly correlated with contemporaneous as well as subsequent firm value as measured by MTB or Tobin's Q. Moreover, the positive correlation between RQ and firm value remains robust to samples of firms with and without patents or citations. Patent based measures, on the other hand, have a rather fragile relation with firm value. In particular, the previously documented positive relationship between patent-based measures and firm value (Hirshleifer, Hsu and Li, 2013) disappears if we exclude firms that don't have any patents.

Second, we find that RQ is positively and significantly related to firm future returns after controlling for well know return predictors such as market capitalization, book-to-market ratio of equity, prior six months returns, asset growth and net stock issuance. This is not true for patents. Coefficients for citation weighted patents are insignificant, while those for unweighted patents are negative.

Finally, we conduct a simple validity check of RQ by examining its correlation with firm characteristics related to innovation. We find that RQ is higher in younger firms, more profitable firms, firms with higher R&D intensity, and firms that are less capital intensive.

Overall, given the structural advantages of RQ (universality, uniformity), its strong theoretical foundation, as well as its strong positive correlation with firm value, we recommend that researchers consider using RQ in future innovation studies.

Moreover, given the practical limitations of patents (endogenous choice to patent, highly variable value, and sample restrictions (less than 50% of firms patent their R&D)), as well as their lack of significance in tests of firm value and their negative coefficients in Fama-MacBeth regressions, we recommend replicating prior patent-based tests of innovation using the RQ measure.

The remainder of the paper is organized as following. In section II, we provide more motivation for the research quotient (RQ) measure of innovation and detail its construction. In section III, we construct the sample of firms used in our study and provide detail on the other measures of firm innovation and related control variables. In section IV, we present our main empirical results. We conclude in section V.

II. Research Quotient (RQ)

RQ is the firm-specific output elasticity of R&D (Knott 2008). It is exponent γ in firm i's production function (Equation 1). Accordingly, RQ represents the percentage increase in revenues from a 1% increase in R&D, when other inputs and their elasticites are held constant.

$$Y = A_{i,} K_{i,t}^{\alpha} L_{i,t}^{\beta} R_{i,t-1}^{\gamma} S_{i,t-1}^{\delta} D_{i,t}^{\phi} e_{i,t} \quad (1)$$

where $Y_{i,t}$ is output, $A_{i,}$ is a firm fixed effect, $K_{i,t}$ is capital , $L_{i,t}$ is labor, $R_{i,t-1}$ is lagged R&D, $S_{i,t-1}$ is lagged spillovers (defined below in equation 3), $D_{i,t}$ is advertising.[5]

The way to interpret RQ is a firm's ability to generate revenue from its R&D investment. Thus a firm can have high RQ either by generating a large number of innovations and being reasonably effective exploiting them, or by generating a smaller number of innovations and being extremely effective exploiting them.

Because RQ is estimated entirely from financial data, it can be derived for any firm engaged in R&D. Thus, it is universal. In addition, RQ is unitless. Thus its interpretation is uniform across firms within an industry as well as across industries. Perhaps most importantly, RQ is a reasonable empirical proxy for theoretical innovation constructs. For example, studies such as Knott and Vieregger (2014), indicate that RQ matches empirical predictions of firm R&D investment, firm value and growth from endogenous growth theory (Thompson 1996, Lentz and Mortensen 2008).

We derive RQ for each firm-year by estimating Equation 1 using a random coefficients model that allows for heterogeneity in the output elasticity for R&D (as well as all other inputs). A random coefficients model (Longford 1993) represents a general functional form model which treats coefficients as being non-fixed (across members of a cross-section or over time) and potentially correlated with the error term. Random coefficient models are those in which each coefficient has two components: 1) the direct

[5] We lag R&D one year relative to output. See Appendix A for a further discussion of this and other modeling decisions for RQ estimation.

effect of the explanatory variable and 2) the random component that proxies for the effects of omitted variables.

Use of a random coefficients (RC) specification follows from the need to capture firm specific estimates for R&D elasticity. Note that RC is a general functional form of which fixed effects is a restricted form where only the intercept is treated as a random coefficient. The random coefficients model we estimate is given in equation 2, where the $\beta_{\_}$ and $\beta_{\_i}$: represent the direct effect and the firm-specific error, respectively for each of the exponents in equation 1, e.g, $(\beta_3 + \beta_{3i})$ corresponds to γ in equation 1.

$$\ln Y_{it} = (\beta_0 + \beta_{0i}) + (\beta_1 + \beta_{1i}) \ln K_{it} + (\beta_2 + \beta_{2i}) \ln L_{it} + (\beta_3 + \beta_{3i}) \ln R_{i,t-1} + (\beta_4 + \beta_{4i}) \ln S_{i,t-1} + (\beta_5 + \beta_{5i}) \ln D_{it} + \varepsilon_{it} \quad (2)$$

We construct RQ for each firm-year by estimating Equation 2 using rolling 10-year windows of the COMPUSTAT North American Annual database from 1965-2006. Firm level data items include (in $MM unless otherwise stated): *revenues* (Y_{it}), *capital* as net property, plant and equipment (K_{it}), *labor* as full-time equivalent employees (L_{it}),in units of 1000, *advertising* (D_{it}), and *R&D* (R_{it}). From these primary data, we calculate an additional input into equation 2 of firm-specific *spillovers* (S_{it}) which is computed as the sum of the differences in knowledge between focal firm *i* and rival firm *j* for all firms in the four digit SIC industry with more knowledge (R&D) than the focal firm:

$$S_{it} = \sum_{j \neq i} R_{jt} - R_{it} \ \forall \ R_{jt} \geq R_{it} \quad (3)$$

This construction mimics the spillover construct in endogenous growth models (Jovanovic and Rob 1989, Jovanovic and MacDonald 1994, Eeckhout and Jovanovic 2002). It is a density measure that takes into account the number of firms with superior knowledge as well as the amount of each firm's surfeit of knowledge relative to the focal firm. In essence it represents the likelihood and extent of discovering superior knowledge in a random encounter with a rival firm.

We require all firms to have a minimum of six years of non-missing R&D data within each ten-year window. For the other inputs to equation 2, conditional on having non-missing R&D, the incidents of missing values are quite low (less than 0.1%), other than for advertising. If a firm-year observation for advertising is missing, we assign a value of zero. We delete observations with any other missing values in equation 2. Thus each RQ estimate compares revenues to inputs using up to 10 firm-year observations matching revenues to inputs. We perform random coefficients estimation of Equation 2 using Stata command, xtmixed. Xtmixed fits linear mixed models using maximum likelihood estimation. The random effects, b_i, are not directly estimated, but we form best linear unbiased predictions (BLUPs) of them using xtmixed post estimation. We then define the RQ for each firm-year as the sum $(\beta_3 + \beta_{3i})$ in the last year of each window. For example, the 2001 RQ for each firm is the estimate formed using data from the 1992 to 2001 window (Further estimation details and the full set of coefficient estimates for that window are provided in Appendix A).

To provide a sense of RQ values, Figure 1 Panel A plots the RQ's average, minimum and maximum value across the Fama French 48 industries in fiscal year 2001. Panel B shows the histogram of RQs in fiscal year 2001. The mean value is 0.104 (standard deviation= 0.120). This implies the mean firm generates a 0.12% increase in revenues for a marginal 1% increase in R&D. Note that some RQ estimates are below zero. Because this implies revenues are decreasing in R&D, we deem these observations invalid and drop them from the sample.[6] Panel A indicates that the mean RQ across industries is relatively constant, but there is greater variance of RQ within industries than across industries.[7] Thus very little of RQ is driven by industry effects.

III. Sample Construction and Summary of Statistics

A. Sample Construction

We merge firms' estimated RQs with data from Compustat and the CRSP monthly stock file. Our main sample period is a 26 year panel over 1981 to 2006, which is determined as a function of the data availability of the patent/citation data used in most

[6] Approximately 8 percent of firm-year observations have negative RQ values. Our results are qualitatively similar when we include negative RQ firms in our sample. The results of these tests are available upon request.

[7] In separate analysis available from the authors we examined the impact of industry mean R&D intensity on industry RQ. We found mean R&D intensity has no significant impact on mean RQ, but significantly increases the variance in industry RQ.

of our tests.[8] All financial firms (SIC between 6000 and 6999) are excluded. We then compute all the accounting related variables from COMPUSTAT and return (or stock price) related variables from the CRSP monthly stock file. Detailed variable definitions can be found in Appendix B.

To calculate patent-based measures used in the prior literature, we use the NBER patent database maintained by Hall, Jaffe and Trajtenberg (2001). The database provides detailed information about patent application and grant date, patent assignee, citation information for each patent and a link between patent assignee and COMPUSTAT GVKEY. This database starts from 1976 and ends at 2006. Following Hirshleifer et al. (2013), we use patent grant date as the effective date to compute any patent based variable. For all citation-based variables, we exclude all self-citations and adjust citations by the adjustment factor to account for the well-known truncation bias in citation data.[9] We estimate the innovation efficiency measures (IE) used in Hirshleifer, Hsu and Li (2013). There are two IE measures: IE based on patents and IE based on citations. IE(Patents) is the number of patents scaled by R&D capital:

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

IE(Citations) is an effort to solve the patent uniformity problem discussed in the introduction. IE(Citations) creates a measure, C, of each patent's citations that is scaled by the number of citations for all other patents in its class. These citations are then summed over the prior five years for each patent, then summed over all the firm's patents, before scaling by R&D capital:

$$IE_t = \frac{\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}}{R\&D_{i,t-3}+R\&D_{t,t-4}+R\&D_{i,t-5}+R\&D_{i,t-6}+R\&D_{i,t-7}} \quad (5)$$

where C_{ik}^{t-j} is the number of citations received in year t by patent k, granted in year $t-j$, scaled by the average number of citations received in year t by all patents of the same

[8] Our RQ measure is available before and after the patent/citation data. Our main results are qualitatively similar in periods before and after the 1981 to 2006 period. We do not report these extended results, but they are available upon request.

[9] See Hall, Jaffe and Trajtenberg (2001) for a detailed patent citation discussion. The adjustment factor is variable "hjtwt" in dataset "pat76_06_assg."

subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm i.

We find that less than half the firm-year observations can be matched to any patents. This reinforces one of the concerns with patent-based measures—that they can only be applied to about half the firms doing R&D. The normal treatment in the literature is to assign zero patents to these firms. This is reasonable because these firms indeed were not granted any patents. But in terms of an innovation proxy, assigning zeros implies that firms without patents aren't innovating, which we know to be false, because 63% of firms in our sample are doing R&D (innovating), but aren't patenting. Further, assigning zeroes is problematic because firms who patent may be fundamentally different from those who don't. Accordingly, inferences based on patent-based measures may be driven by these differences, rather than by innovation itself. To account for this problem, we conduct our tests with two samples, one in which firm-year observations with no patents are assigned zeros and the other sample in which firm-year observations without patents are dropped. Before we impose any filters, in our sample from 1981 to 2006, there are 106,023 firm-years in Compustat with non-missing and non-negative R&D. For these 106,023 firm-years, there are 35,145 firm-years of patent data. Thus, approximately 33% of the R&D firm-years have patents, or stated differently, approximately 67% of firm-year observations cannot be matched to patent data. When we examine the firm level, not firm-years, we find that 37% of firms with R&D data have patents.[10] After merging Compustat accounting data with CRSP financial data, and computing RQ from equation 2, we have 26,808 firm-years of observations. For the patent and citation based IE measures from equation 5, there are 14,512 firm-years with non-missing IE measures.

[10] Our measure will under-represent firms who innovate through means other than R&D. However, we note that the percent of firms with missing R&D data that engage in innovation (as measured by patents) is relatively small. For all firms on Compustat over the 1981 to 2006, we are able to match 44,159 firm-years of patent observations with Compustat data. Of these observations, 9014 firm-years (20.4%) do not engage in R&D. This translates to 2,225 (10.11%) of firms without R&D data that engage in patent activities. In related work, Koh and Reeb (2014) find that 10.5% of missing R&D firms exhibit patent activities.

B. Summary Statistics

Table 1 provides summary statistics for the main variables used in our study. In Panel A, we report sample characteristics where we assign missing patents and citations to be 0, consistent with some of the previous literature. In Panel B, we exclude the firm-years without any patents. For firms in Panel A, the average book value of total assets is $2564 million and average market capitalization is $2987 million. The average RQ is 0.14 (standard deviation of 0.076) with a 95th percentile of 0.28. In Panel B, the sample size drops to less than half of the sample size in Panel A. Firms in Panel B are significantly larger (average total assets is $4223 million, market equity is $4948 million) and spend more on R&D ($150 million on average) than firms in Panel A. This validates our concern that firms who patent are structurally different from those who don't.

In Table 2, we sort the sample into RQ quintiles and report the pooled average of all major variables for each quintile. The quintile sorting is based on contemporaneous RQ. In Panel A, we assign missing patents and citations as 0, while in Panel B, we exclude these observations. The pattern we observe across quintiles are similar across both panels.

In Panel A, the mean RQ ranges from 0.05 for the lowest quintile to 0.25 for the highest quintile. In contrast the Innovation Efficiency (IE) measures are essentially constant across the quintiles, for both patent based and citation based IE. Thus, the first observation from the quintile sorts is that RQ and IE appear uncorrelated. The second thing to note across the quintiles is that RQ is non-monotonic in measures of firm scale; total assets, revenue, market equity and book equity first increase from quintile 1 to quintile 4, but drop in quintile 5. R&D expense shows a similar pattern. Finally, valuation measures such as Market-to-Book ratio of equity and Tobin's Q increase with RQ. Thus, RQ appears correlated with firm value.

Overall, the summary statistics shows some preliminary univariate evidence that RQ is positively correlated with firm valuations, non-monotonically correlated with firm scale, and uncorrelated with IE.

IV. Empirical Results

A. Innovation and Contemporaneous Firm Value

Our first test formally examines whether RQ is related to firm value. Specifically, we run the following panel regression model:

$$\ln(MTB_{i,t}) = \alpha + \beta_1 * RQ_{i,t} + \beta_2 * IE_{i,t} + \gamma * Controls + \varepsilon_{i,t} \quad (6)$$

This regression model captures the contemporaneous relation between a firm's innovation and firm value as measured by the market to book ratio of equity (MTB). We run each innovation measure (RQ and IE) separately as well as together. We include other standard control variables and variables from accounting-based asset valuation models of corporate R&D.[11] We cluster standard errors at the firm level for all specifications and include industry fixed effects.

Table 3, Panel A reports the results of the regressions when we include firms without patents and assign their IE a value of zero. Models 1 through 3 examine the impact of innovation on MTB, using each innovation measure in isolation. The coefficient on RQ in model 1 is positive and statistically significant (t-statistic = 5.99). In models 2 and 3, the coefficients on both IE(Patents) and IE(Citations) are also positive and significant. However the economic effects of the coefficients are higher for RQ than IE. A 10% increase in RQ at the mean translates to a 0.75% increase in MTB, whereas a 10% increase in IE(Patents) translates to a 0.16% increase in MTB, and a 10% increase in IE(Citations) translates to a 0.27% increase in MTB. In model (4) and (5), we test the RQ and IE measures in combination. The results show that the coefficients for both RQ and IE are essentially unchanged when combining the measures - further evidence that IE and RQ are measuring different things.

In Panel B of Table 3, we repeat the analysis for a sample that excludes firms without patents in the observation year. This removes the implicit assumption in Panel A that firms without patents aren't innovating. Model 1 indicates that RQ remains positive and significant. However, model 2 indicates that IE(Patents) is now negative and statistically significant, while model 3 indicates that IE(Citations) is positive but insignificant. In models 4 and 5, where we combine RQ with the IE measures, the

[11] See Ohlson (1995) and Sougiannis (1994)

coefficient on RQ remains positive and significant, while the coefficients on both IE measures retain their signs and significance levels from the model 2 and 3 regressions.[12]

To summarize, RQ is positive and significant across all specifications and samples, while both IE measures are sensitive to excluding firms without patents. This implies that prior studies using IE measures are largely capturing differences between firms who patent versus those who don't, rather than differences in innovation across the subset of firms who patent.

B. RQ and Subsequent Firm Value

Having demonstrated the positive relationship between RQ and contemporaneous MTB, we move to investigate whether the market fully prices the effects of RQ into current valuations, or if RQ has predictive power for future firm valuation. In Table 4, we re-estimate all regression specifications from Table 3 replacing contemporaneous MTB with subsequent (i.e., one year ahead) MTB as the dependent variable. The results are similar to those in Table 3. RQ remains positive and significant in all specifications while IE measures become insignificant or even significantly negative if we exclude firms without patents.

Thus, it appears the market doesn't fully impound the effects of RQ into future valuations. This seems reasonable given the fact that the measure didn't exist previously, and that it is uncorrelated with the primary measure that has been used to capture innovation (patents). Thus, investors may only slowly impound the information in RQ into future firm valuation (Hong and Stein, 2007).

[12] In separate analyses we also test an innovation "ability" measure from Cohen, Diether and Malloy (2013) designed to capture a firm's past track record of converting R&D spending into sales growth. We estimate the CDM measure (defined as the average of coefficients from rolling regressions of sales growth regressed on lagged R&D scaled by sales), and find a near-zero correlation between RQ and the CDM measure. Further we replicated the analysis in Table 3 and Table 4 firm value regressions using the CDM measure. We find that the coefficient on the CDM measure is positive and statistically significant if we include all firms (Panel A), but it becomes insignificant if we delete firms without patents (Panel B). The inclusion of the CDM measure has no impact on either RQ or IE in those regressions. Detailed results are available upon request.

C. RQ and Firm Characteristics

Given that RQ appears to be a reliable measure of innovation, we next examine the characteristics of high and low RQ firms. While the true firm level predictors of RQ likely require understanding the internal operations of firms, we examine readily available firm characteristics that have been associated with innovation in prior literature. Table 5 presents regressions of RQ on those characteristics. All the right hand side variables are lagged by one year. We use firm fixed effects to capture any constant unobservable effect on RQ and report firm clustered standard errors. Note that firm fixed effects capture all semi-permanent factors differentiating firms' innovative capability, such as hiring and resource allocation practices.

Table 5 indicates that RQ increases with R&D intensity (RD/assets) and ROA. These results match expectations from endogenous growth theory that higher R&D productivity implies higher optimal R&D investment and that higher R&D productivity yields higher profits (Lentz and Mortensen 2008). The table also indicates RQ decreases with age. This matches recent findings that economic growth is coming principally from young firms (Haltiwanger, Jarmin and Miranda 2013). Interestingly RQ is negatively correlated with physical capital stocks (i.e., PPE/Assets). While there is no prior theory or empirics anticipating the result, it suggests exploiting new products and processes is more difficult in the presence of vintage capital.

Beyond these standard firm characteristics, we include two governance related measures that have been examined in the finance literature. The first is analyst coverage. He and Tian (2013) find a negative effect of analyst coverage on firms' innovation, when using patent based measures. Here however, we find analyst coverage is positively and significantly correlated with innovation when using RQ. The second measure is liquidity. Fang, Tian and Tice (2014) document a negative effect of liquidity on innovation when using patent-based measures. They postulate this is because more liquid firms face higher chances of hostile takeover and are more likely to have institutional investors who may hinder the investment in innovation. While we use a different measure of liquidity (i.e., Amihud's 2002 illiquidity measure), we find that the coefficient on illiquidity is not statistically significant in the RQ regressions.[13]

While no casual interpretation can be drawn from the results in Table 5, we hope to provide some idea of what type firms have high (or low) RQ. We also provide some indication that prior innovation studies may yield different results if patent-based measures are replaced with RQ.

D. RQ and Stock Returns

In this section, we examine whether RQ is able to predict monthly future returns.[14] We rely on Fama and MacBeth (1973) (FM) regressions to conduct the tests. We control for well-known return predictors such as firm market capitalization, book-to-market equity, past return performance, asset growth, and stock net issuance. Following Fama and French (1992), for each firm we use predictors constructed from fiscal year data ending in calendar year t to predict stock returns from July in year t+1 to June in year t+2. The annual fiscal year data is from 1981 to 2006 and the final monthly return dataset is from July 1982 to June 2009. Past performance is measured as the compounded buy and hold return of the prior six months returns from month t-7 to t-2 for month t. We then estimate monthly cross sectional regressions for each month and report the time series average of coefficients and R-squares. Standard errors are computed based on the time series variation of coefficients. In all specifications, we exclude stocks whose prices are below $5 at the end of the previous month to help ensure that the results are not driven by small cap stocks.

The results in Table 6 indicate RQ reliably predicts monthly stock returns across all specifications after controlling for other return predictors. The coefficient on RQ is positive and significant in all models. Moreover the significance levels for RQ are higher than for other standard predictors such as B/M and past returns. IE measures on the other

[13] Our innovation measure is different than the measure used in Fang, Tian and Tice (2014). Fang et al note that "We acknowledge that, although we follow prior literature and capture innovation using patents and citations, our results may not extend to other types of nonpatentable innovation, such as process innovation."

[14] We are agnostic to the sources of future return predictability due to RQ or other measures of innovation. For example, Hirshleifer, Hsu and Li, 2013, discuss both mispricing and risk-based stories for innovation related return predictability. Their mispricing story is based on investor underreaction to innovation due to limited attention and their risk-based story is related to *q*-theory.

hand don't reliably predict stock returns. In fact, the coefficient on IE(Patents) is significantly negative in a sample that excludes firms without patents.

V. Conclusion

In this study we propose replacing patent-based measures of innovation with a new measure from the management literature, firm RQ. Part of the motivation to consider RQ is the very practical consideration that less than 50% of firms who do R&D patent their innovations, and that those patents have highly variable value. However before introducing a new measure, it is important to validate it against existing measures. Thus, to determine whether RQ is a suitable measure of innovation, we run it through a fairly standard battery of tests comparing it to existing patent-based measures of innovation.

We find that RQ is significant in explaining contemporaneous as well as subsequent market-to-book ratios. All results are robust to alternative samples and to the inclusion of existing patent based measures of innovation. Moreover, we find that RQ predicts future stock returns. In fact, in our sample, its effect is more significant than many of the other established predictors of the cross-section of stocks returns (such as B/M and past returns).

While further research is necessary to understand how high RQ firms differ from their low RQ counterparts, simple correlations suggest RQ maps onto factors that ring true for innovative firms. Thus RQ has face validity as well as predictive validity.

Perhaps our most surprising result is that patent-based measures become insignificant or significantly negative in both MTB and Fama-MacBeth regressions when we exclude firms without patents. Excluding these firms is important because otherwise we are defining firms without patents to have no innovative output. We know this to be false because greater than 50% of our firms are innovating (investing in R&D), but not patenting. Patents are costly to file and maintain, and require that fundamental knowledge be publicly divulged. Accordingly, firms patent only under certain circumstances.

Our results suggesting patent-based measures may be problematic for gauging innovation's impact on firm value is reminiscent of results from Abrams et al (2013) showing patent citations are problematic in gauging patent value itself. They find patent citations are increasing then decreasing with patent value, and attribute the non-monotonicity to firms who patent strategically rather than productively. The fact that patent counts are negative and significant in both MTB and FM regressions suggests strategic patenting likely drives our results as well.

In summary, what began as an effort to validate a more universal and uniform measure of innovative output, became an exercise in demonstrating that patent-based measures of innovation, because they miss a large number of firms that conduct R&D, may be problematic as general measures of innovation. Accordingly, we recommend replicating prior studies of innovation using RQ in lieu of patent-based measures.

References

Abrams, D. S., U. Akcigit, and J. Popadak. 2013. Patent Value and Citations: Creative Destruction or Strategic Disruption? *NBER Working Paper No. 19647*

Adams, J. D., A. B. Jaffe. 1996. Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data. *RAND Journal of Economics* 27, 700–721.

Aghion, P. J. Van Reenen and L. Zingales. 2013. Innovation and Institutional Ownership. *American Economic Review* 103(1), 277-304

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

Becker-Blease, JR. 2011. Governance and Innovation. *Journal of Corporate Finance* 17, 947–958.

Bloom, N., M. Schankerman and J. Van Reenan 2007. Identifying Technology Spillovers and Product Market Rivalry, *Econometrica* 81(4), 1347-1393

Cohen, W.M., R.R. Nelson And J.P. Walsh. 2000. Protecting Their Intellectual Assets: Appropriability Conditions And Why U.S. Manufacturing Firms Patent (Or Not). *NBER Working Paper 7552*

Cohen, L., K. Diether, and C. Malloy. Misvaluing Innovation. *Review of Financial Studies* 26 (3).

Cuneo, P. and J. Mairesse, 1984. “Productivity and R&D at the Firm Level in French Manufacturing”, in *R&D, Patents, and Productivity*. Z. Griliches, (ed.) University of Chicago Press, Chicago

Eeckhout, J. & B. Jovanovic. 2002. Knowledge Spillovers and Inequality. *American Economic Review*, 92: 1290-1307.

Fama, E., and K. French, 1992, The Cross-section of Expected Stock Returns. *Journal of Finance* 47, 427-465.

Fama, Eugene F., MacBeth, James D. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81 (3), 607–636.

Fang, V., X. Tian, and S. Tice. 2014. Does Stock Liquidity Enhance or Impede Firm Innovation? *Journal of Finance*, Forthcoming.

Galasso, A and T. Simcoe. 2011. CEO Overconfidence and Innovation. *Management Science* 57(8), 1469-1484

Gormley, T., D. Matsa, and T. Milbourn. 2013. CEO Compensation And Corporate Risk: Evidence From A Natural Experiment. *Journal of Accounting & Economics* 56, 79-101

Griliches, Z. and J. Mairesse. 1984. Productivity and R&D at the firm level, in *R&D, Patents, and Productivity*. Z. Griliches, (ed.) University of Chicago Press, Chicago

Hall, B. and J. Mairesse and P. Mohnen. 2010. Measuring the Returns to R&D. *Handbook of the Economics of Innovation*, B. H. Hall and N. Rosenberg (eds.), North-Holland.

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

Haltiwanger, J., R. Jarmin and J. Miranda. 2013. Who Creates Jobs? Small Versus Large Versus Young. *The Review of Economics and Statistics* 95(2), 347-361

He, J. and X. Tian. 2013. The Dark Side of Analyst Coverage: The Case of Innovation. *Journal of Financial Economics* 109(3), 856–878.

Hirshleifer, D., P. Hsu and D. Li. 2013. Innovative Efficiency and Stock Returns. *Journal of Financial Economics* 107(3), 632-654.

Hirschleifer, D., A. Low and S. Teoh. 2012. Are Overconfident CEOs Better Innovators. *Journal of Finance* 67(4), 1457-1498.

Hulten, C. 2000. Total Factor Productivity: A Short Biography. *NBER working paper* 7471

Hong, H. and J. Stein. 2007. Disagreement and the Stock Market. *Journal of Economic Perspectives* 21, 109–128

Jaffe, A. 1986. Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value. *American Economic Review*, 76(5), 984-1001.

Jovanovic, B. & G. MacDonald. 1994. Competitive Diffusion. *Journal of Political Economy*,102, 24-52.

Jovanovic, B. & R. Rob. 1989. The Growth and Diffusion of Knowledge. *Review of Economic Studies*, 56, 569-582.

Knott, A.M. 2008. R&D/Returns Causality: Absorptive Capacity or Organizational IQ. *Management Science* 54, 2054-2067.

Knott, A.M., H. E. Posen, and B. Wu. 2009. Spillover Asymmetry and Why It Matters, *Management Science*, 55, 373-388.

Knott, A.M. and C. Vieregger, 2014. An Empirical Test of Endogenous Firm Growth. SSRN working paper: http://ssrn.com/abstract=2382885

Koh, Ping-Sheng and Reeb, David M., 2015. Missing R&D. *Journal of Accounting & Economics* 60, 73-94.

Lentz, R. and D. Mortensen. 2008. An Empirical Model of Growth Through Product Innovation. *Econometrica* 76(6), 1317-1373.

Levin, R. and P. Reiss. 1988. Cost-Reducing and Demand-Creating R&D with Spillovers. *RAND Journal of Economics*, 19(4), 538-56.

Lerner, J. and J. Wulf. 2007. Innovation and Incentives: Evidence from Corporate R&D. *The Review of Economics and Statistics* 89(4), 634-644.

Ohlson, J., 1995. Earnings, Book Values, and Dividends in Equity Valuation. *Contemporary Accounting Research* 11, 661–687.

Romer, P. 1990. Endogenous Technological Change. *Journal of Political Economy*, 98(5), 71-102.

Scherer, FM and D. Harhoff. 2000. Technology policy for a world of skew-distributed outcomes. *Research Policy* 29, 559-566.

Solow, R. 1957. Technical Change and the Aggregate Production Function. *Review of Economics and Statistics* 39(3), 312-320.

Sougiannis, T. 1994. The Accounting Based Valuation of Corporate R&D. *The Accounting Review* 69, 44-68.

Spence, A.M. 1984. Cost Reduction, Competition, and Industry Performance, *Econometrica*, 52(1), 101-21.

Syverson, C. 2011. What Determines Productivity. *Journal of Economic Literature* 49(2), 326-365.

Thompson, P. 1996. Technological Opportunity and the Growth of Knowledge: A Schumpeterian Approach to Measurement. *Journal of Evolutionary Economics* 6(1), 77-97

APPENDIX A

Details of RQ estimation

There are a number of issues that researchers confront in estimating production functions: what to use as the output measure, how to deal with lack of data on materials, whether to use stocks or flows, with what lag, how to construct spillover pools, how to adjust prices, and how to deal with endogeneity. We address each of these issues in turn to explain the mechanics of our estimation of the research quotient (RQ).

Output: We use gross output in dollars (revenues) rather than units because we have multi-product firms across multiple industries (so units aren't comparable). Moreover, dollars allow us to capture product innovation (higher prices) as well as process innovation (lower input use for same output). We use gross output (revenues) rather than value-added because it allows for substitution between materials and other inputs.

Materials: Compustat data suffers from lack of materials reporting (something common across many studies of R&D productivity). This yields upward bias in R&D elasticity (Griliches and Mairesse 1984). We examine the extent of bias by running estimates with cost of goods sold, COGS (which includes materials), in lieu of capital and labor. Either version includes some double counting (since R&D comprises labor, capital and materials). This in turn yields downward bias in R&D elasticity, though this is of greater concern in cross-sectional estimation than with panel data (Cuneo and Mairesse 1984).

Functional form of R&D input: There is a debate in the empirical innovation literature on whether to use R&D stocks or flows. Clearly the technology input is some form of knowledge stock, just as the capital input is the capital stock rather than current period investment. The problem with stocks is that adjustment costs and depreciation rates vary across firms even within the same industry. Prior studies indicate that depreciation rates vary between 15% and 19%, though industry estimates for the output elasticity of R&D appear to be insensitive to rates from 8% to 25% (Hall, Mairesse and Mohnen 2010). The convention in the R&D literature is to construct stocks assuming a common depreciation rate (typically 15%) across all firms and industries. This approach introduces noise in addition to consuming years of observation.

An alternative approach, which we adopt, is to use flows, under the assumption that firms are essentially in steady state where their investments reflect the R&D demand from obsolescence plus expected growth. As long as these rates of obsolescence and growth change slowly, they will be captured in the firm effect, so elasticity will not depend on obsolescence. In robustness checks we test this against models using R&D stocks and find no significant difference. This finding of steady-state explains two empirical regularities: econometric equivalence between stock and flow models and econometric equivalence of models with different lags (Griliches and Mairesse 1984, Adams and Jaffe 1996).

A related question is that of appropriate lag, because R&D generates an intermediate input (product/process design) that then combines with capital, labor and materials to generate final goods output. Here we follow convention and model R&D with a lag of one year. In robustness checks we test alternative lags and find no significant difference. Again, this matches prior results indicating the lag structure has little impact on elasticity estimates (Griliches and Mairesse 1984). This is likely due in part to the fact that firms are in steady state so R&D doesn't vary much.

Spillover pools. Empirical tests of R&D productivity typically employ a pooled form of spillovers (Jaffe 1986, Adams and Jaffe 1996, Bloom, Schankerman and VanReenen 2007), which sums the R&D for all firms in industry *j* other than focal firm *i* (though the pool is adjusted for technical and geographic proximity). This pooled form is faithful to early IO models of the impact of spillovers on incentives for R&D (Spence 1984, Levin & Reiss 1984). Because these models assume homogeneous firms, functional form is innocuous in that all firms have equivalent knowledge. Firm heterogeneity introduces the possibility that spillovers are asymmetric. Following work elsewhere testing alternative functional forms of spillover pools (Knott, Posen, Wu 2009), we employ an asymmetric form that mimics endogenous growth models (Jovanovic and Rob 1989, Jovanovic and MacDonald 1994, Eeckhout and Jovanovic 2002). The spillover pool is constructed for each firm as the sum of the differences in knowledge between focal firm *i* and rival firm *j* for all firms in industry *k* with more knowledge than the focal firm. What this construction represents is the likelihood and extent of superior knowledge in a random encounter with a rival firm. It is a density measure that takes into account the number of firms with superior knowledge as well as the amount of each firm's surfeit. In all models the spillover measure is matched to the R&D measure with regard to stocks versus flows and with regard to lags.

Price adjustment. In industry specific studies, it is common to use price deflators for both inputs and output. Because we examine multi-product firms whose outputs cross many industries, use of a uniform price deflator would introduce noise. Accordingly we rely on the fact that prices for the multi-industry inputs move roughly in tandem with those for the multi-industry outputs.

Simultaneity. The simultaneity concern in production functions is potential correlation between input levels and unobserved firm-specific shocks. In particular that firms will respond to a positive productivity shock by increasing inputs. Failure to account for this will positively bias coefficients. Random coefficients (RC) estimation inherently accounts for stable firm-specific shocks through the firm-specific error term on each input. Tests conducted elsewhere (Knott and Vieregger 2014) of RC versus IV approaches to solving the simultaneity problem indicate that RC yields more precise and reliable coefficient estimates for inputs than these other approaches.

As an example, estimation results for the random coefficients (RC) specification of the R&D production function (equation 1) are presented in Table A1 (Model 1) for the year 2001. Other year estimations are similar. The random effects for all inputs are significant (in addition to the main effects shown in the table). Similarly, a likelihood ratio test of the RC specification versus a linear specification is significant at the 0.001 level. For comparison we present results for a fixed effects (FE) specification (Model 2), as well as an OLS specification (Model 3). Note that the coefficients for R&D are similar across the RC and FE models at 0.095 and 0.098, while the coefficient for R&D in the OLS model is significantly different at 0.008. The OLS estimate is well outside historical estimates for the returns to R&D (Hall, Mairesse and Mohnen 2010).

Table A1. Summary Statistics

Model 1 in this table reports estimation results for the random coefficients (RC) specification of the R&D production function (equation 1) for the year 2001. For comparison we present results for a fixed effects (FE) specification (Model 2), as well as an OLS specification (Model 3).

Dependent variable: ln(revenue)	1 Random Coefficients	2 Fixed Effects	3 OLS
ln(capital)	0.168***	0.148***	0.178***
	(0.010)	(0.009)	(0.006)
ln(labor)	0.793***	0.746***	0.892***
	(0.013)	(0.012)	(0.008)
ln(advertising)	0.000	-.003***	0.015***
	(0.002)	(0.001)	(0.001)
ln(R&D), *t*-1	0.095***	0.098***	0.008***
	(0.007)	(0.006)	(0.004)
ln(spillovers), *t*-1	0.015***	0.027***	0.007***
	(0.003)	(0.003)	(0.002)
Constant	4.077***	4.076***	4.442***
	(0.037)	(0.034)	(0.024)
Observations	20386	20386	20386
No. of firms	2427	2427	
R-squared		0.907	0.912
R2-within	.	0.484	
R2-between		0.929	
Wald chi2	10583	.	.

***Signficant at 0.1 percent level

**Signficant at 1 percent level

*Significant at 5% level

Appendix B: Variable Definitions

RQ	Research Quotient. Firm-specific output elasticity of R&D (Knott 2008).
IE(Patents)	Innovation Efficiency based on Patents. Defined as firm's patents granted in a given year scaled by R&D capital: $IE_t = patents_{i,t}/(R\&D_{i,t-2} + 0.8R\&D_{i,t-3} + 0.6R\&D_{i,t-4} + 0.4R\&D_{i,t-5} + 0.2R\&D_{i,t-6})$
IE(Citations)	Innovation Efficiency based on Citations. Defined as adjusted number of citations received by patents granted in past 5 years, scaled by number of corresponding R&D expenses: $IE_t = \frac{\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}}{R\&D_{i,t-3} + R\&D_{t,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}}$
Firm Characteristics:	
Total Assets	Total Book Assets
BE	Book value of Equity, calculated as stockholders' equity plus balance sheet deferred taxes and investment tax credit, minus book value of preferred stock
Market-to-Book Equity (MTB)	Market value of Equity divided by Book value of Equity for the fiscal year ending in calendar year t. Market value of Equity is calculated as number of shares outstanding times the share price at December of year t.
Revenue	Total Revenue
R&D Expense	Research and Development expenditures

Age	Number of years since the first appearance on Compustat
R&D/Assets	Research and Development expenditures scaled by Total Book Assets
Advertising	Firm's Advertising Expense
PPE/Assets	Property, Plant and Equipment scaled by Total Book Assets
Leverage	Book value of Debt divided by Total Book Assets
Capex/Assets	Capital Expenditure scaled by Total Book Assets
Tobin's Q	Market Value of Equity plus Book Value of Assets minus Book Value of Equity minus Balance Sheet Deferred Taxes, divided by Total Book Assets.
Market to Book	MTB
R&D/BE	Research and Development Expenditures scaled by Total Book Equity
Abnormal Earnings	$\left(E_{i,t}^{B}\left(1-\tau_{i,t}\right)-r_t BE_{i,t-1}\right)/BE_{i,t}$, Where $E_{i,t}^{B}$ is earnings after extraordinary items before expensing R&D expense and less preferred dividends from firm i in year t, $\tau_{i,t}$ is tax rate and r_t is one year Treasury Bill rate

Operating Performance:

ROA	Operating Income before Depreciation scaled by Total Book Assets
Cash Flow	Operating Income before Depreciation minus Interest Expense, minus Taxes, minus Preferred Dividends, minus Common Dividends and divided by Total Book Assets
Others:	
Analyst Coverage	Rolling average of the last 12 months of the number of monthly earnings forecasts from IBES summary file
Illiquidity	Amihud Illiquidity measure. The yearly average of the daily value of 1,000,000 times the absolute value of daily returns scaled by the daily dollar volume for each trading day.
R&D tax shield	R&D Expenditure multiplied by the Tax Rate, scaled by Book Equity
Asset growth	1 year percentage change in Total Firm Assets
BHRET6	Compounded buy and hold return of the prior six months returns from month t-7 to t-2 for month t.
Net Issuance	Change of the logarithm of split adjusted shares outstanding, where the split is determined using the COMPUSTAT adjustment factor.

Table 1. Summary Statistics

This table reports pooled mean, standard deviation, 5 percentile, median and 95 percentile points of the main variables used in the paper. The innovation measures are Research Quotient (RQ), Innovation Efficiency based on patents (IE(Patents)), and Innovation Efficiency based on citations (IE(Citations)). The sample period is from 1981 to 2006. Detailed variable definitions can be found in Appendix B. Panel A reports summary statistics for the sample where IE is set to zero for firms with no patents or citations. Panel B reports summary statistics where firms without patents are excluded from the sample.

	Panel A. Sample includes IE=0						Panel B. Sample excludes IE=0					
variable	**N**	**Mean**	**Stdev**	**5 Pct**	**Median**	**95 Pct**	**N**	**Mean**	**Stdev**	**5 Pct**	**Median**	**95 Pct**
Innovation Variables:												
RQ	26806	0.140	0.076	0.033	0.130	0.280	14512	0.145	0.072	0.042	0.136	0.276
IE (Patent)	26806	0.129	0.226	0.000	0.024	0.576	14512	0.238	0.261	0.018	0.149	0.801
IE (Citation)	26806	0.144	0.250	0.000	0.039	0.643	14512	0.216	0.271	0.000	0.128	0.773
Firm Characteristics:												
Total Assets	26806	2564.751	16936.341	7.380	160.939	8616.000	14512	4223.021	22590.534	18.451	426.474	15463.000
Revenue	26806	2164.977	10055.806	6.778	166.414	8470.000	14512	3489.820	13054.965	15.048	434.911	13985.000
R&D expense	26806	89.026	425.742	0.251	7.500	318.331	14512	149.959	565.097	0.789	18.829	615.800
Advertising Expense	9834	88.053	370.362	0.066	2.600	415.000	5378	140.630	466.529	0.175	7.246	692.814
Market Equity	26806	2987.414	16251.690	5.653	176.130	10383.279	14512	4947.938	21459.369	16.226	472.801	19751.377
Book Equity	26806	1005.266	4889.626	3.450	87.452	3655.200	14512	1625.957	6345.971	9.463	222.902	6492.000
log (Age)	26806	2.904	0.592	1.946	2.890	3.850	14512	3.039	0.609	2.079	3.091	3.932
R&D/Assets	26806	0.076	0.089	0.004	0.047	0.242	14512	0.074	0.086	0.006	0.047	0.227
PPE/Assets	26806	0.242	0.155	0.039	0.217	0.547	14512	0.259	0.153	0.049	0.242	0.555
Leverage	26711	0.180	0.158	0.000	0.158	0.481	14471	0.186	0.151	0.000	0.173	0.460
Capex/Assets	26483	0.054	0.042	0.008	0.043	0.138	14316	0.057	0.041	0.011	0.048	0.138
Tobin Q	26798	1.941	1.472	0.791	1.450	4.878	14509	1.985	1.477	0.840	1.491	4.860
Acquired Assets	26806	0.019	0.057	0.000	0.000	0.122	14512	0.019	0.056	0.000	0.000	0.123
Market to Book	26806	0.708	0.847	-0.584	0.644	2.221	14512	0.753	0.815	-0.459	0.686	2.202
R&D/BE	26806	0.171	0.299	0.009	0.089	0.565	14512	0.163	0.280	0.012	0.092	0.495
Abnormal Earnings	26633	-0.030	0.708	-0.716	0.066	0.415	14447	0.018	0.591	-0.435	0.076	0.401
Operating Performance:												
ROA	26785	0.089	0.189	-0.214	0.122	0.274	14501	0.104	0.171	-0.173	0.132	0.273
Cash Flow	26785	0.038	0.150	-0.238	0.070	0.180	14501	0.049	0.138	-0.192	0.074	0.177

Table 2. Summary Statistics by Research Quotient (RQ) Quintiles

In Panel A, for each fiscal year t, we sort firms into quintiles based on RQ in year t. We do this every year from 1981 to 2006 and report the pooled average of each variable. In Panel B, we exclude firms with missing patents. We report the average of each variable for each quintile. The innovation measures are Research Quotient (RQ), Innovation Efficiency based on patents (IE(Patents)), and Innovation Efficiency based on citations (IE(Citations)). Detailed variable definitions can be found in Appendix B.

	RQ Ranks									
	Panel A. With IE=0					Panel B. Without IE=0				
	1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Innovation Variables:										
N	5353	5367	5364	5367	5355	2893	2908	2906	2908	2897
RQ	0.051	0.103	0.133	0.164	0.248	0.060	0.112	0.139	0.167	0.246
IE (Patent)	0.136	0.130	0.127	0.123	0.129	0.320	0.236	0.210	0.201	0.225
IE (Citation)	0.148	0.144	0.142	0.139	0.149	0.251	0.213	0.196	0.193	0.224
Firm Characteristics:										
Total Assets	822.307	1736.951	3822.687	3924.286	2513.569	1689.929	3146.323	6593.226	5951.458	3720.832
Revenue	702.657	1659.907	2850.322	3229.874	2379.172	1403.029	3019.465	4594.458	4855.678	3566.759
R&D expense	26.063	62.668	114.270	137.268	104.748	55.051	118.910	191.042	217.556	166.839
Advertising Expense	15.625	40.907	123.408	158.697	86.764	34.198	75.328	200.405	243.159	120.407
Market Equity	928.546	2040.210	4032.637	4568.628	3363.102	1885.626	3788.693	6582.780	7153.691	5315.618
Book Equity	388.213	794.548	1308.158	1460.775	1073.348	810.632	1380.220	2064.088	2233.856	1637.125
log (Age)	2.795	2.915	3.013	2.977	2.821	2.918	3.105	3.166	3.093	2.913
R&D/Assets	0.078	0.069	0.067	0.071	0.096	0.081	0.065	0.062	0.067	0.096
PPE/Assets	0.243	0.258	0.259	0.244	0.204	0.271	0.278	0.275	0.258	0.215
Leverage	0.185	0.184	0.186	0.183	0.160	0.190	0.192	0.198	0.188	0.163
Capex/Assets	0.053	0.056	0.056	0.055	0.050	0.059	0.059	0.058	0.058	0.053
Tobin Q	1.841	1.778	1.842	1.928	2.319	1.870	1.790	1.878	1.994	2.394
Acquired Assets	0.017	0.019	0.020	0.021	0.018	0.018	0.019	0.022	0.020	0.018
Market to Book	0.627	0.616	0.678	0.733	0.887	0.661	0.670	0.721	0.775	0.938
R&D/BE	0.180	0.156	0.148	0.157	0.214	0.188	0.144	0.136	0.140	0.209
Abnormal Earnings	-0.105	-0.036	0.019	0.015	-0.042	-0.059	0.016	0.060	0.049	0.023
Operating Performance:										
ROA	0.045	0.094	0.120	0.115	0.069	0.061	0.115	0.131	0.130	0.085
Cash Flow	0.004	0.045	0.062	0.058	0.019	0.017	0.059	0.068	0.068	0.030

Table 3. Regressions of Market-to-Book on Contemporaneous Research Quotient (RQ)

This table reports panel regressions of MTB on Research Quotient (RQ), Innovation Efficiency based on patents (IE(Patents)), Innovation Efficiency based on citations (IE(Citations)), and other control variables. All dependent variables are contemporaneous to MTB. The sample period is from 1981 to 2006. T-statistics are reported in parenthesis. Standard errors are clustered at the firm level. In all models, we include industry (Fama-French 48) fixed effects. Detailed variable definitions can be found in Appendix B. Panel A reports regression results where IE is set to zero for firms with no patents or citations. Panel B reports regression results where firms without patents or citations are excluded from the sample.

	Panel A. With IE=0					Panel B. Without IE=0				
	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
RQ	0.5371***			0.5315***	0.5322***	0.3680***			0.3581***	0.4164***
	(5.995)			(5.943)	(5.950)	(3.217)			(3.127)	(3.521)
IE (patent)		0.1258***		0.1230***			-0.0827**		-0.0785**	
		(4.022)		(3.934)			(-2.186)		(-2.069)	
IE(citation)			0.1904***		0.1891***			0.0591		0.0618
			(5.690)		(5.687)			(1.560)		(1.638)
Abnormal Earnings	0.0462***	0.0482***	0.0489***	0.0467***	0.0473***	0.0969***	0.0987***	0.0913***	0.0969***	0.0893***
	(4.398)	(4.578)	(4.667)	(4.447)	(4.535)	(5.746)	(5.863)	(5.334)	(5.744)	(5.211)
RD Tax Shield	0.0896***	0.0911***	0.0919***	0.0897***	0.0904***	0.0884***	0.0869***	0.0788***	0.0870***	0.0799***
	(4.740)	(4.813)	(4.857)	(4.752)	(4.797)	(3.168)	(3.106)	(3.009)	(3.123)	(3.054)
Ln(1+Capex/Assets)	-4.8004***	-4.8198***	-4.8057***	-4.8139***	-4.8000***	-5.0450***	-5.0239***	-5.0610***	-5.0286***	-5.0644***
	(-48.872)	(-48.641)	(-48.863)	(-48.761)	(-48.984)	(-38.873)	(-38.785)	(-39.602)	(-38.851)	(-39.674)
RD/BE	0.9538***	0.9683***	0.9648***	0.9581***	0.9547***	1.0129***	0.9965***	1.0512***	0.9929***	1.0464***
	(30.596)	(30.918)	(30.960)	(30.736)	(30.771)	(24.099)	(23.776)	(23.829)	(23.615)	(23.771)
1/BE	0.1524**	0.1476**	0.1557**	0.1553**	0.1634**	0.2650**	0.3527***	0.1286	0.3477***	0.1342
	(2.299)	(2.218)	(2.354)	(2.350)	(2.488)	(2.229)	(2.932)	(1.026)	(2.888)	(1.074)
Constant	0.7951***	0.8532***	0.8412***	0.7801***	0.7678***	0.8893***	0.9611***	0.9164***	0.9093***	0.8570***
	(42.989)	(61.524)	(59.820)	(41.554)	(40.515)	(37.860)	(50.945)	(48.823)	(35.774)	(33.810)
Observations	26,313	26,313	26,313	26,313	26,313	14,252	14,252	15,108	14,252	15,108
R-squared	0.424	0.423	0.425	0.425	0.427	0.482	0.481	0.467	0.482	0.468
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table 4. Regressions of Market-to-Book on lagged Research Quotient (RQ)

This table reports panel regressions of MTB on RQ, IE and other control variables. The sample period is from 1982 to 2006. All dependent variables are lagged one year relative to MTB. T-statistics are reported in parenthesis. Standard errors are clustered at the firm level. In all models, we include industry (Fama-French 48) fixed effects. The detailed variable definitions can be found in Appendix B. Panel A reports regression results where IE is set to zero for firms with no patents or citations. Panel B reports regression results where firms without patents or citations are excluded from the sample.

	Panel A. With IE=0					Panel B. Without IE=0				
RQ	0.4967***			0.4925***	0.4945***	0.3557***			0.3415***	0.2975**
	(4.973)			(4.939)	(4.957)	(2.767)			(2.650)	(2.219)
IE (patent)		0.0835**		0.0807**			-0.1174***		-0.1136***	
		(2.532)		(2.457)			(-2.851)		(-2.756)	
IE(citation)			0.1584***		0.1577***			0.0384		0.0406
			(4.457)		(4.471)			(0.934)		(0.992)
Abnormal Earnings	0.0194	0.0220*	0.0233*	0.0198	0.0210*	0.0702***	0.0726***	0.0437**	0.0706***	0.0418**
	(1.525)	(1.723)	(1.833)	(1.551)	(1.660)	(3.311)	(3.418)	(2.190)	(3.325)	(2.097)
RD Tax Shield	0.0108	0.0125	0.0134	0.0109	0.0119	0.0227	0.0204	0.0103	0.0202	0.0109
	(0.463)	(0.536)	(0.577)	(0.471)	(0.512)	(0.705)	(0.633)	(0.327)	(0.628)	(0.348)
Ln(1+Capex/Assets)	-3.7392***	-3.7480***	-3.7391***	-3.7483***	-3.7397***	-4.0798***	-4.0477***	-4.0790***	-4.0553***	-4.0842***
	(-35.257)	(-35.121)	(-35.183)	(-35.304)	(-35.368)	(-29.577)	(-29.053)	(-30.386)	(-29.142)	(-30.456)
RD/BE	0.8532***	0.8665***	0.8636***	0.8566***	0.8537***	0.9278***	0.8985***	0.9662***	0.8955***	0.9630***
	(24.315)	(24.381)	(24.407)	(24.364)	(24.384)	(19.915)	(19.042)	(19.572)	(19.063)	(19.610)
1/BE	0.0100	0.0039	0.0129	0.0119	0.0209	-0.1063	0.0388	-0.3108*	0.0331	-0.3050*
	(0.136)	(0.053)	(0.175)	(0.161)	(0.285)	(-0.718)	(0.255)	(-1.956)	(0.219)	(-1.931)
Constant	0.7709***	0.8278***	0.8153***	0.7607***	0.7476***	0.8643***	0.9428***	0.9000***	0.8937***	0.8577***
	(38.691)	(54.362)	(53.001)	(37.343)	(36.422)	(34.355)	(46.068)	(44.383)	(32.638)	(31.248)
Observations	23,227	23,227	23,227	23,227	23,227	12,789	12,789	13,478	12,789	13,478
R-squared	0.311	0.310	0.311	0.312	0.313	0.369	0.369	0.353	0.370	0.353
Ind FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES

Table 5. Regression of Research Quotient (RQ) on Firm Characteristics

This table reports panel regressions of RQ on various firm characteristics. The sample period is from 1981 to 2006. All dependent variables are lagged one year relative to RQ. The first column is a regression where we set IE to zero for firms with no patents or citations. In the second column, we exclude firms without patents or citations. We cluster standard errors at the firm level and include firm fixed effects.

	With IE=0	Without IE=0
Ln(Assets)	-0.006**	-0.007**
	(-2.241)	(-2.120)
RD/Assets	0.091***	0.098**
	(2.902)	(2.314)
Ln(Age)	-0.034***	-0.043***
	(-5.654)	(-5.134)
ROA	0.065***	0.059***
	(6.443)	(4.278)
PPE/Assets	-0.086***	-0.092***
	(-4.843)	(-4.223)
Leverage	-0.012	-0.013
	(-1.415)	(-1.171)
Capex/Assets	0.001	0.011
	(0.044)	(0.374)
Analyst Coverage	0.002***	0.002***
	(5.822)	(5.731)
Illiquidity	-0.001	-0.001
	(-1.399)	(-1.161)
Constant	0.274***	0.316***
	(18.349)	(16.031)
Observations	17,617	11,221
R-squared	0.600	0.596
Firm Fixed Effect	YES	YES

Table 6. Fama-MacBeth Regressions of Stock Returns on Research Quotient (RQ)

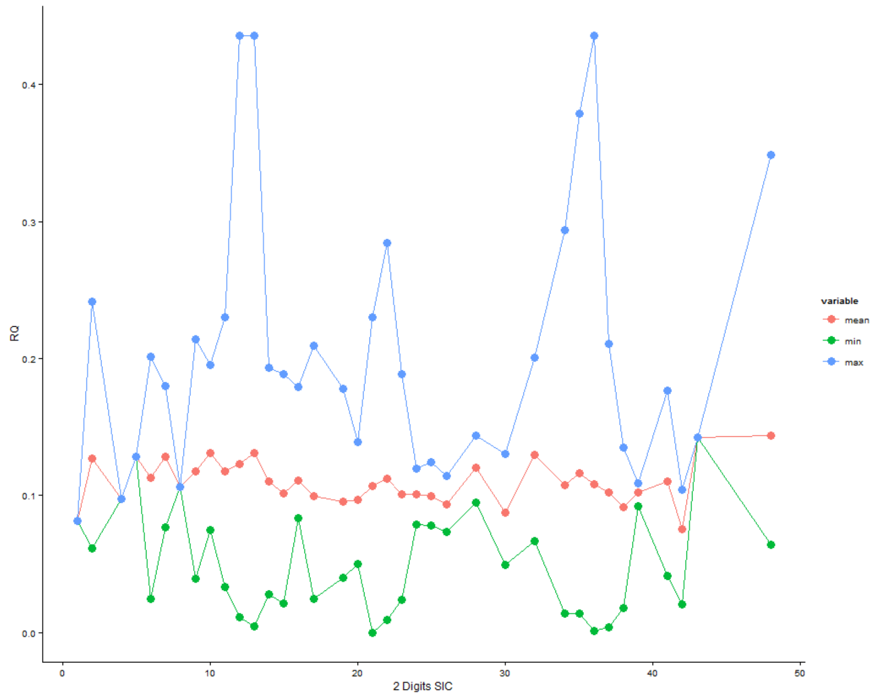
This table reports Fama-MacBeth regressions of firm excess monthly stock returns on lagged values of RQ and other variables. The sample period is from July 1982 to June 2009. The reported coefficients and t-statistics are time-series averages of cross-sectional regression coefficients and time-series t-statistics. We regress monthly excess returns from July of year t to June of year t+1 on RQ and other variables in fiscal year ending in year t-1. IE measures are defined in Appendix B. Ln(Mktcap) is the natural logarithm of market value of equity in June of year t. Ln(B/M Equity) is the natural logarithm of Book-to-Market ratio of equity. Book equity is defined as shareholder's equity plus balance sheet deferred taxes and investment tax credits minus book value of preferred stock in fiscal year ending in t-1. Market equity is defined as the number of shares outstanding times the price at the end of year t-1. BHRET6 is the rolling previous 6 months cumulative return computed by skipping the most recent lagged month. Asset growth is the one year percentage change in total firm assets. Net Issuance is the change of natural logarithm of split adjusted shares outstanding, using the COMPUSTAT adjustment factor to control for splits. Panel A reports regression results where IE is set to zero for firms with no patents or citations. Panel B reports regression results where firms without patents or citations are excluded from the sample.

	Panel A. With IE=0					Panel B. Without IE=0				
RQ	0.018***			0.018***	0.019***	0.019***			0.018**	0.0193***
	(3.045)			(2.994)	(3.046)	(2.735)			(2.564)	(2.595)
IE(Patents)		-0.001		-0.001			-0.006**		-0.005**	
		(-0.377)		(-0.218)			(-2.491)		(-2.124)	
IE(Citations)			0.001		-0.001			-0.001		-0.002
			(0.134)		(-0.556)			(-0.914)		(-1.466)
Ln(Mktcap)	0.001	0.001	0.001	0.001	0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.225)	(0.395)	(0.415)	(0.245)	(0.190)	(-1.106)	(-1.425)	(-0.404)	(-1.373)	(-0.517)
Ln(B/M equity)	0.002*	0.001	0.002	0.002*	0.002*	0.002	0.001	0.001	0.001	0.001
	(1.870)	(1.519)	(1.567)	(1.844)	(1.843)	(1.459)	(0.713)	(0.946)	(1.267)	(1.414)
BHRET6	0.006**	0.007**	0.006**	0.007**	0.006**	0.006*	0.006*	0.006*	0.006*	0.005
	(2.257)	(2.516)	(2.330)	(2.419)	(2.232)	(1.659)	(1.803)	(1.684)	(1.704)	(1.568)
Asset growth	-0.003*	-0.003**	-0.003**	-0.003*	-0.003**	-0.001	-0.001	-0.004*	-0.001	-0.003
	(-1.822)	(-2.118)	(-2.168)	(-1.840)	(-2.011)	(-0.214)	(-0.424)	(-1.785)	(-0.253)	(-1.486)
Net Issuance	-0.001	-0.001	0.00	-0.001	-0.001	-0.002	-0.003	0.001	-0.003	0.001
	(-0.244)	(-0.063)	(0.007)	(-0.261)	(-0.167)	(-0.354)	(-0.368)	(0.136)	(-0.437)	(0.000)
Constant	0.009**	0.011**	0.011**	0.009**	0.009**	0.013***	0.018***	0.013***	0.015***	0.011**
	(2.152)	(2.528)	(2.572)	(2.112)	(2.201)	(2.794)	(3.526)	(2.835)	(3.042)	(2.460)
Observations	239,412	239,412	239,412	239,412	239,412	146,632	146,632	152,706	146,632	152,706
R-squared	0.043	0.043	0.043	0.045	0.046	0.059	0.059	0.056	0.063	0.060
Number of months	312	312	312	312	312	312	312	312	312	312

Figure 1. RQ across Industries and the Histogram of Research Quotient (RQ)

Panel A plots the average, minimum and maximum RQ value of firms within each of the Fama French 48 industries in fiscal year 2001. Panel B shows the histogram of RQ in fiscal year 2001.

Panel A. RQ across Fama-French 48 Industries



Panel B. Histogram of RQ in fiscal year 2001