JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 57, No. 5, Aug. 2022, pp. 1649–1694
© THE AUTHOR(S), 2021. PUBLISHED BY CAMBRIDGE UNIVERSITY PRESS ON BEHALF OF THE MICHAEL G. FOSTER SCHOOL OF BUSINESS, UNIVERSITY OF WASHINGTON
doi:10.1017/S0022109021000417

RQ Innovative Efficiency and Firm Value

Michael Cooper
University of Utah Eccles School of Business
mike.cooper@utah.edu (corresponding author)

Anne Marie Knott
Washington University in St. Louis Olin Business School
knott@wustl.edu

Wenhao Yang
School of Management and Economics and Shenzhen Finance Institute, The Chinese University of Hong Kong, Shenzhen
yangwenhao@cuhk.edu.cn

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

Abstract

We introduce and test a firm-level innovation-efficiency measure new to the finance literature. The measure, termed the *research quotient* (RQ), defined as the firm-specific output elasticity of research and development (R&D), was first developed in the management literature. RQ has a low correlation with existing innovation input, output, and efficiency measures. We test RQ in a number of innovation tests common to the finance literature and find that RQ is robust in all tests of firm value, even after controlling for previous innovation measures. The results suggest that RQ may serve as a relevant complementary measure of a company's innovation.

I. Introduction

Innovation and related investment are vital to firm growth and performance. In the finance literature, the study of innovation is steadily growing in importance.[1] In these articles, the innovation measures encompass three types: input measures (e.g., research and development (R&D) spending); output measures (e.g., patents or citations); and efficiency measures, which indicate a firm's ability to convert inputs to outputs (e.g., patents scaled by R&D). Researchers have studied how these innovation measures influence current and future firm valuations, future stocks returns, future operating performance, and other firm aspects.[2] Overall, these

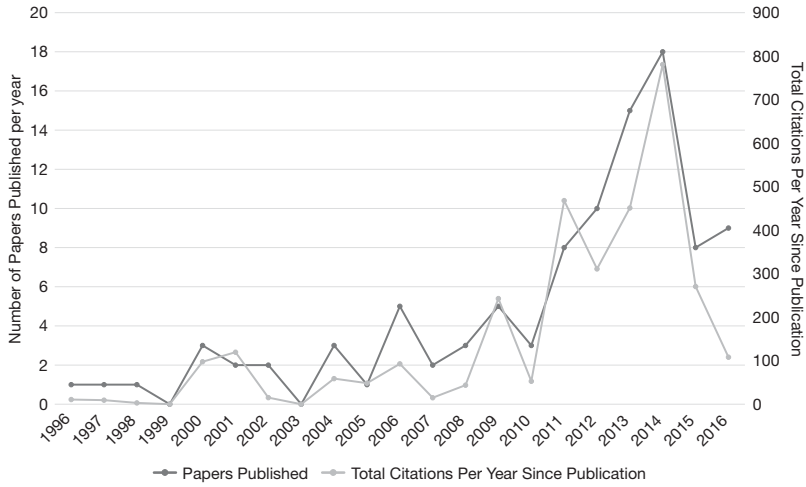
[1]In Figure 1, we plot the number of innovation-related articles published per year, the cumulative number over time, and citations to those articles in the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, and the *Journal of Financial and Quantitative Analysis.* There are clear increases on average over time in the number of innovation-related articles published per year and the number of citations. Summing up over time since 1996, there have been 100 innovation-related papers published in these top-four finance journals, which have received a combined total of 18,648 citations.

[2]For example, see Lev and Sougiannis (1996), Chan, Lakonishok, and Sougiannis (2001), Penman and Zhang (2002), Gu (2005), Lev, Sarath, and Sougiannis (2005), Hall, Jaffe, and Trajtenberg (2005), Eberhart, Maxwell, and Siddique (2008), Pandit, Wasley, and Zach (2011), Hirshleifer, Hsu, and Li (2013), and Cohen, Diether, and Malloy (2013).

FIGURE 1

Number of Innovation-Related Articles Published in Top Finance Journals and Citations to Them

Figure 1 presents the number of publications per year (dark-gray line), and citations per year (light-gray line) on Google Scholar related to innovation research and published in the *Journal of Finance*, the *Journal of Financial Economics*, the *Review of Financial Studies*, or the *Journal of Financial and Quantitative Analysis*. To be included in the publication counts, an article must have "Patents" in the title, or "Patent Citations" in the body of the article, or "R&D" (or "research and development") in the title, or "innovate" (or "innovative" or "innovation") in the title. The left vertical axis is the cumulative number of articles published each year. The right vertical axis is the total citations for articles published in that year received divided by the years since publication.



articles reveal important links among all three types of innovation measures and firm value, operating performance, and stock returns.

In this article, we introduce and test a firm-level innovation-efficiency (IE) measure new to the finance literature, one that relies exclusively on financial data. Accordingly, this novel measure can be formed for any firm reporting R&D. The measure, first developed in the management literature by Knott (2008), is the research quotient (RQ), defined as the firm-specific output elasticity of R&D. RQ is the exponent on R&D from estimating a firm's production function. Thus, we can interpret RQ as a firm's ability to generate revenue from its R&D investments. Measured this way, a firm can have high RQ either by generating a large number of innovations and exploiting them reasonably effectively or by generating a smaller number of innovations and exploiting them spectacularly.

Our goal is to empirically explore the relevance of the RQ measure for the finance literature. We find that RQ is not highly correlated with other commonly used measures of innovation. The correlations among RQ, R&D input, and patent-based output measures are close to 0. The correlations among RQ and IE measures from Hirshleifer et al. (2013) (constructed using patents or citations scaled by R&D) are also close to 0. Thus, RQ potentially contains useful information beyond established innovation measures.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

We first perform panel regressions of annual firm market value on RQ. We control for combinations of innovation input (R&D/BE), two forms of innovation output (patent count and citation-weighted patent count),[3] two forms of innovative efficiency (scaling patents by R&D and scaling patent citations by R&D),[4] and other controls typical in firm-value regressions. These tests employ industry and year fixed effects, and we cluster the standard errors at the firm level.[5] We find that RQ is significantly and positively correlated with contemporaneous as well as subsequent firm value as measured by market-to-book equity (MTB). Moreover, these findings are robust to sample periods, to specifications with and without previous innovation measures, and to Fama–MacBeth (1973) cross-sectional regressions.

The panel regressions demonstrate strong and robust associations between RQ and firm value. However, a concern is that RQ may be endogenous: Unobserved firm characteristics, correlated with both RQ and firm value, may drive the positive association between RQ and firm value. To alleviate these endogeneity concerns for RQ, we implement a 2-stage least squares (2SLS) estimation. We instrument RQ with a plausibly exogenous instrument, industry-level R&D_CENTRALIZATION, a measure related to work by Argyres and Silverman (2004), Arora, Belenzon, and Rios (2011), and Cummings (2019). R&D_CENTRALIZATION measures the location of R&D decision-making (headquarters versus divisions), with the idea that more centralized decision making may be related to greater innovative efficiency (Cummings (2019)). We argue that the organizational form of R&D only influences firm value through the innovation process, thus satisfying the exclusion restriction. Using industry R&D_CENTRALIZATION as an instrument, 2SLS results confirm the positive association between RQ and both contemporaneous and future firm value documented in the OLS panel-regression results.[6]

Next, we perform Fama and MacBeth (1973) regressions of future monthly returns on RQ, previous innovation measures, and established determinants of the cross section of returns, such as market capitalization, the book-to-market ratio of equity, returns from the prior 6 months, and asset growth. We find that the coefficient on RQ is positive and statistically significantly related to firm future returns in models with RQ as the sole innovation measure, as well as in models that include controls for previous innovation measures.

We examine the robustness of our 3 firm-value tests (contemporaneous MTB, future MTB, and future returns) to sample construction. Our main tests follow the common convention of inserting a value of 0 into missing patent-related variables. We reproduce our 3 value tests using a sample that excludes observations with

[3]See, for example, Hirshleifer, Low, and Teoh (2012), He and Tian (2013), and Fang, Tian, and Tice (2014).

[4]See, for example, Hirshleifer et al. (2013).

[5]We do not use firm-fixed effects in our panel regressions because most of the variation in RQ occurs in the cross section instead of in the own-firm time series. In our sample, the correlation between RQ and lagged RQ is 0.87. See related discussions in Zhou (2001).

[6]We also perform 2SLS tests using the industry median (Gormley and Matsa (2014), Grieser, Hadlock, and Charles (2019)) of RQ as an instrument. The results are robust using this alternative instrument.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

missing patents, and we find that the RQ results are highly robust across both samples and all 3 tests: The coefficients on RQ are uniformly positive and significant in all models.

Although our primary interest concerns the impact of innovation on market value, to show that the choice of innovation measure has broader implications, we also use RQ in a few representative studies on the determinants of corporate innovation (i.e., innovation as the dependent variable). Here we look at 2 determinants from prior studies: analyst coverage (He and Tian (HT) (2013)) and leverage (Atanassov (2013)). HT show that when using patents as the measure of innovation, analyst coverage is negative and significant in explaining innovation. Regarding firm leverage, Atanassov documents a negative relation between leverage and innovation, again using patents as the measure of innovation. Using patent- and citation-based dependent variables, we are able to replicate the findings in HT and Atanassov. In models with RQ as the dependent variable, we find an insignificant association between RQ and analyst coverage and a negative association between RQ and leverage. These determinants-of-innovation tests are designed to replicate prior studies, so we follow the methodologies of those studies and use level values of innovation measures as the dependent variables. However, these regressions may be misspecified because of high autocorrelation of the dependent variables, resulting in autocorrelated residuals and potentially biasing the coefficient standard errors. To address this issue, we reestimate the regression models using dynamic panel methods (Arellano and Bond (1991)). We find that analyst coverage and leverage become insignificant using the dynamic panel approach. Overall, the results show that the effects of innovation on firm behavior and valuation are sensitive to the type of innovation measure (input, output, or efficiency), the sample construction, and the estimation approach, suggesting that researchers should carefully consider these differences in their research designs.

To summarize our main empirical findings, across all of the market-to-book and return tests, across samples with and without previous innovation measures, and across samples with and without missing patents data, the results using RQ are highly consistent; RQ is positively and strongly statistically associated with current and future firm-value measures. In contrast, as we detail in the results sections, other common innovation measures are less robust across the 3 value tests, samples, and models.

Our work adds to a growing literature identifying and testing new measures of corporate innovation. For example, Kogan, Papanikolaou, Seru, and Stoffman (2017) use firm stock returns around the time of patent granting as a way to measure the value of patents. Mukherjee, Singh, and Žaldokas (2017) examine stock returns in response to innovation-related press releases. Bellstam, Bhagat, and Cookson (2017) use textual analysis of analyst reports to extract innovation measures, and Hirshleifer, Hsu, and Li (2018) use innovative originality, or the number of unique technological classes of patents cited by a firm's patents.

Overall, the main contribution of this article is to introduce to the finance literature a measure of firm innovation, RQ, from the management literature (Knott (2008)). RQ has features that, we believe, will be attractive to researchers working in the innovation field. First, because RQ is constructed without using patents, and

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

because only 50% of firms conducting R&D patent their innovations, the RQ measure increases coverage relative to previous patent-based innovation measures. Second, RQ's low correlation with previous input, output, and efficiency innovation measures suggests that it may provide researchers with an alternative metric of firm-level innovation not captured by established innovation measures. Finally, and perhaps most importantly, RQ is the only measure we test that is consistently positively related to contemporaneous and future firm value, as well as to future returns. These results suggest that RQ may serve as an additional measure for researchers needing innovation measures related to firm value. Finally, given that the data to estimate RQ are based on R&D and other commonly used data in Compustat, it is straightforward to estimate. Moreover, a database of RQ values is publicly available with Compustat subscriptions through Wharton Research Data Services (WRDS).[7]

The remainder of the article is organized as follows: In Section II, we construct the sample of firms used in our study and review the data, including the construction of RQ, previous innovation measures, and other control variables. In Section III, we evaluate RQ in common firm-value tests from the finance literature. Section IV concludes.

II. Sample Construction

A. Innovation Measures

Research Quotient (RQ)

Our focal measure of innovation is RQ. RQ, the firm-specific output elasticity of R&D (Knott (2008)), is the exponent γ_i in firm i's production function:

$$Y_{it} = A_i K_{it}^{\alpha_i} L_{it}^{\beta_i} R_{it-1}^{\gamma_i} S_{it-1}^{\delta_i} D_{it}^{\phi_i} e_{it}, \tag{1}$$

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 in equation (3)), and $D_{i,t}$ is advertising.[8] RQ thus matches the most common means in economics to measure returns to R&D at the industry and economy levels (Hall, Mairesse, and Mohnen (2010)). RQ maps most intuitively onto product innovation (increased revenue for the same level of inputs), but it also captures process (cost-reducing) innovation (the same revenues for reduced inputs).

RQ enjoys some coverage and uniformity advantages relative to patent-based measures. First, as we will detail in Section II.B, it provides increased coverage relative to patent-based measures. Second, because RQ is an elasticity, it is unitless and directly comparable across firms, whereas patents may vary widely

[7]We have posted firm-level RQ data on the University of Pennsylvania's WRDS database. The data are free to all WRDS users with valid Compustat subscriptions.

[8]We lag R&D by 1 year relative to output. See Appendix A for a further discussion of this and other modeling decisions for RQ estimation.

in value.[9] Having noted this, RQ has its own concerns. Most of these concerns are shared with measures of R&D (discussed in the next section): It is unavailable for firms that innovate without R&D and is only widely available for public firms (when using Compustat for the source of financial data).[10]

We compute RQ by estimating equation (1) using a random-coefficients (RC) model that allows for heterogeneity in the output elasticity for R&D (as well as all other inputs). The model we estimate is given in equation (2), where $\beta_$ and β_i represent the direct effect and the firm-specific error, respectively, for each of the exponents in equation (1); for instance, $(\beta_3 + \beta_{3i})$ corresponds to γ_i in equation (1).

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

We estimate the RC model in equation (2) in Stata, using the xtmixed command. Xtmixed fits linear mixed models using a maximum-likelihood estimator (MLE). Mixed models contain both fixed effects and random effects. The fixed effects, $\beta_$, are analogous to standard regression coefficients and are estimated directly. The random effects, β_i, allow for heterogeneity across observational units, i, in our case, firms. The β_i values are not directly estimated from the MLE procedure. Rather, we form best-linear unbiased predictions (BLUPs) of them (and their standard errors) from their estimated variances and covariance.[11]

We construct RQ for each firm-year by estimating equation (2) using rolling 10-year windows of the Compustat North American Annual database. The first estimation window is 1972–1981. We roll through the data each year, updating the window (dropping off the oldest year and adding the newest year), and create RQ values from 1981–2015. Firm-level data items used to estimate RQ include (in $millions 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 1,000; ADVERTISING (D_{it}); and *R&D* (R_{it}). From these primary data, we calculate firm-specific SPILLOVERS (S_{it}), a "free input," in equation (2). We follow Knott, Posen, and Wu (2009) in computing spillovers as the sum of the differences in knowledge between focal firm i and rival firm j for all firms in the 4-digit SIC industry with more knowledge (R&D) than the focal firm:

[9] Scherer and Harhoff (2000) report that 10% of U.S. patents account for 81% to 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 et al. (2001) show that patent citations are a better predictor of firm value than patent counts. 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 et al. (2001)).

[10] It can be estimated for private firms when private firm financial data are available, such as in the National Science Foundation (NSF) Survey of Industrial R&D (SIRD) and the Business Innovation and R&D Survey (BRDIS).

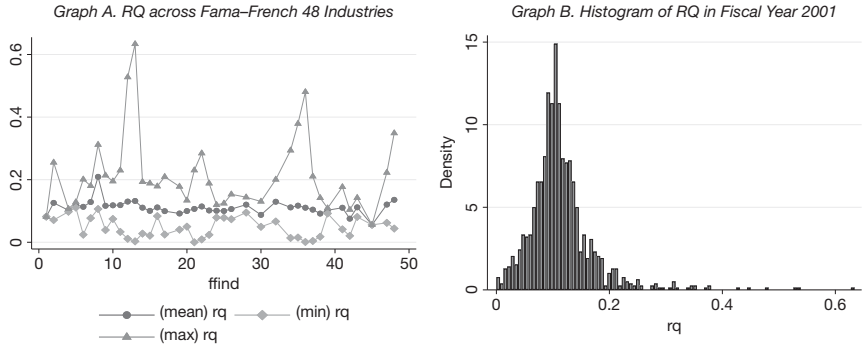
[11] We provide more details on the estimator for equation (2), including details of the standard errors and model assumptions, in Appendix A. Also, see https://www.stata.com/bookstore/pdf/xt_xtmixed.pdf for further details on the xtmixed procedure.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

FIGURE 2

RQ Across Industries and the Histogram of RQ

Graph A of Figure 2 plots the average, minimum, and maximum research quotient (RQ) value of firms within each of the Fama–French 48 industries in fiscal year 2001. Graph B shows the histogram of RQ in fiscal year 2001.



$$S_{it} = \sum_{j \neq i} R_{jt} - R_{it} \forall R_{jt} \geq R_{it}. \tag{3}$$

Knott et al. (2009) find that this functional form of spillovers better matches empirical outcomes than either pooled spillovers (all firms in an industry share knowledge equally) or a leader-distance measure (a firm's spillover pool equals its distance from the frontier firm).

We require all firms to have a minimum of 6 years of nonmissing R&D data within each 10-year estimation window.[12] For the other inputs to equation (2), conditional on having nonmissing R&D, the incidents of missing values are quite low (less than 0.1%), with the exception of advertising. If a firm-year observation for advertising is missing, we assign a value of 0.0001. 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.

We then define the RQ for each firm-year as the sum of $(\beta_3 + \beta_{3i})$ estimated over the prior 10-year window. For example, the 2001 RQ for each firm is the estimate formed using data from the 1992–2001 window. The full set of coefficient estimates for that window is provided in Appendix A.

To provide a sense of RQ values, Graph A of Figure 2 plots the average, minimum, and maximum values across the Fama–French 48 industries in fiscal year 2001. Graph B shows the histogram of RQs in this year. The mean value is 0.104 (standard deviation = 0.120), implying that the mean firm generates a 0.10% increase in revenues for a marginal 1% increase in R&D.[13] Because the

[12]The rolling-window method requires at least 6 years of nonmissing R&D data to compute RQ, which means we allow for up to 4 years of missing R&D data in a firm's estimation window. This helps alleviate potential survival-bias concerns because a firm could have missing R&D in the last year of the window (or even missing data in the last 4 years of the window). Thus, we are not imposing a constraint that a firm report R&D data in the year for which a dependent variable (e.g., firm value) is measured. Thus, there is no mechanical survival bias between our dependent variables and RQ.

[13]Note that approximately 8% of firm-year observations have negative RQ values. Because this implies revenues are decreasing in R&D, we deem these observations invalid and drop them from the

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

RQ measure is relatively new, studies of its correlates are limited. However, an NSF study matching RQ to companies' R&D practices indicates that RQ increases with company size (Knott and Vieregger (2020)) and with the extent of incremental innovation (Knott and Vieregger (2020)), whereas RQ decreases with the amount of R&D outsourcing (Knott (2019)) and the extent of decentralization (Cummings (2019)).

R&D Spending

Arguably the most basic measure of innovation is the input: R&D spending. Many innovation studies use R&D as either a dependent or independent variable.[14] Nonetheless, there are some concerns with its use. First, it misses firms that innovate without conducting R&D. Prior articles indicate that approximately 7% of firms reporting R&D values of 0 have introduced innovations (Boroush and Jankowski (2016), among others) and that 10.5% of firms reporting R&D values of 0 hold patents (Koh and Reeb (2015)). (This higher number reported by Koh and Reeb likely reflects firms that acquire patents in the technology markets rather than developing them internally.) Second, R&D data are generally not available for private firms (as they are for public firms). We include R&D in our analyses and follow the convention of normalizing it by company size (R&D/book equity).

Patent Counts

Patents are probably the most common innovation measure in the finance, economics, and management literatures. Patent measures take on several forms: patent counts, citation counts, and citation-weighted counts.[15] We create measures of both patent count and citation intensity (citations per patent). To calculate all patent-based measures, we use the National Bureau of Economic Research (NBER) patent database maintained by Hall et al. (2001). This database provides detailed information about patent application, grant date, and patent assignee; citation information for each patent; and a link between patent assignee and Compustat GVKEY. The database starts in 1976 and ends at 2006. Following Hirshleifer et al. (2013), we use the patent-grant date as the effective date on which to compute any patent-based variable. For all citation-based variables, we exclude all self-citations and adjust citations by an adjustment factor to account for the well-known truncation bias in citation data.[16] We construct ln(1 + PAT), which is the natural logarithm

sample. Our results are qualitatively similar when we include negative-RQ firms in our sample. The results of these tests are available from the authors.

[14]See, for example, Lerner and Wulf (2007), Becker-Blease (2011), Galasso and Simcoe (2011), Hirshleifer et al. (2012), Aghion, Van Reenan, and Zingales (2013), Atanassov (2013), Gormley, Matsa, and Milbourn (2013), He and Tian (2013), Fang et al. (2014), and Hombert and Matray (2018).

[15]The principal means to measure innovation in the social sciences (e.g., accounting, economics, finance, and management literatures) is with patents. A patent for an invention is a property right (typically for 20 years from date of application) granted to the inventor that "excludes others from making, using, offering for sale, selling or importing" the invention. The goal of patent policy is to stimulate innovation by granting this right in exchange for public disclosure of the underlying technology. Thus, the inventor appropriates the returns to the specific invention, and the public benefits through an expanded knowledge base.

[16]See Hall et al. (2001) for a detailed patent-citation discussion. The adjustment factor is variable "hjtwt" in data set "pat76_06_assg."

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

of 1 plus the number of patents applied for and eventually granted, and ln(1 + CIT/PAT), which is the natural logarithm of 1 plus mean citation intensity.

Innovative Efficiency (IE)

Our final measure is IE, which forms a ratio of patents or citation counts to multiperiod R&D. IE thus captures the conversion of inputs to outputs.

We construct 2 measures developed in Hirshleifer et al. (2013). The first measure, IE_PATENTS, is the number of patents scaled by R&D capital:

$$\text{(4)}\quad \text{IE_PATENTS}_t = \frac{\text{PATENTS}_{i,t}}{\text{R\&D}_{i,t-2} + \text{R\&D}_{i,t-3} + \text{R\&D}_{t,t-4} + \text{R\&D}_{i,t-5} + \text{R\&D}_{i,t-6}}.$$

The second measure is IE_CITATIONS. This variable defines a numerator C of each patent's citations scaled by the number of citations for all other patents in its class. These citations are then summed over the prior 5 years for each patent, and then summed over all the firm's patents, before scaling by R&D capital:

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

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

B. The Merged Samples

We merge firms' estimated RQs and the other innovation measures with data from Compustat and the CRSP monthly stock file. Our main RQ sample period is a 35-year panel covering 1981–2015. However, when using patent data, the sample is restricted to 1981–2006. All financial firms (SIC between 6000 and 6999) are excluded. We then compute other accounting-related variables from Compustat and stock price–related variables from the CRSP monthly stock file. Detailed variable definitions can be found in Appendix B.

As shown in Appendix C, after merging Compustat accounting data with CRSP stock return data, and imposing standard filters on our sample from 1981 to 2006, we identify 48,906 firm-years in Compustat with positive R&D. For these 48,906 firm-years, there are 22,701 firm-years of patent data (rows 3 and 4 of Appendix C, respectively).[17] After estimating RQ from equation (2), we have 26,967 firm-years (row 6) of RQ observations over 1981–2006 and 36,141 firm-years (row 7) of observations over 1981–2015.[18] For the patent- and citation-based

[17]The screens include the following: i) firms need to be in the database for at least 2 years, and ii) we exclude financial industry firms. From Appendix C, there are 22,701 firm-years (row 4) of data for patents and 106,396 firm-years in the merged CRSP–Compustat data set (row 2), resulting in 22,701/106,396 = 21.33% of firm-years in the merged Compustat–CRSP data with patents.

[18]The RQ measure will underrepresent firms that innovate through means other than R&D. However, we note that the percentage of firms with missing R&D data that engage in innovation (as measured

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

measures, the totals are 19,856 firm-years (row 8) with nonmissing IE_PATENTS, 19,790 firm-years (row 9) with nonmissing IE_CITATIONS, 22,701 firm-years (row 10) with nonmissing patents, and 20,301 firm-years with nonmissing citations (row 11). Thus, over the patent sample period, RQ has greater firm-year coverage than many of the previous innovation measures. For example, RQ covers approximately 19% more firm-years than patents.[19]

We follow common practice in the literature and assign a value of 0 to all patent-based measures in firm-years with missing patents. Note that this value implies that patent-less firms are not innovating, which we know is false because in our sample, 52% of firms are performing R&D without patenting.[20] The 52% number highlights a potentially important feature of RQ: that it provides an innovation measure for the large number of firms in Compustat that conduct R&D but do not patent.

We do not insert values of 0 for any other missing variables in our models. We merge the nonmissing values of RQ with these firms and conduct our main tests on this merged sample. Later in the article, we conduct robustness checks where we estimate our main tests using a sample of firm-years that excludes missing values of the patent-related innovation measures. RQ effects are robust across the 2 samples. However, we do find some variation across the 2 samples for the other innovation measures' relation with firm value. We document this variation and discuss the importance of sample selection in tests of innovation.

C. Summary Statistics

Table 1 provides summary statistics for the innovation measures. For the patent-based measures, we report summary statistics separately for the sample with missing values set to 0 and for the sample with missing values excluded. Note that for most tests throughout the article, we use the missing-patents-set-to-0 sample (with the exception being sample robustness tests in Table 7, where we exclude missing-patent data). As noted, we do not insert values of 0 into missing values of RQ or R&D. The mean RQ over the 1981–2015 period is 0.141, with a standard deviation of 0.086, and the mean RQ over the patent period 1981–2006 is 0.142 (with a similar standard deviation). The mean RQ implies that the average firm generates an increase of approximately 0.14% in revenues for a marginal 1% increase in R&D.[21] The mean value of RD/BE is 0.174 from 1981 to 2015 and

by patents) is relatively small. For the initially screened CRSP–Compustat merged sample over the 1981–2006 period, there are 57,490 firm-years with no R&D. Among these 57,490 firm-years, 3,665 (6.4%) firm-years have nonmissing patents. This translates to 1,011 (out of 8,095, or 12.5%) of firms without R&D data that engage in patent activity. In related work, Koh and Reeb (2015) find that 10.5% of firms with missing R&D exhibit patent activity.

[19]Computed from Appendix C using firm-years from (row 6)/(row 4) = [(26,967 RQ firm-years/22,701 patent firm-years) – 1] = 18.79% greater percentage of RQ firm-years relative to patents.

[20]In total, 52% of the firms in our sample report R&D without patenting. From Appendix C, the ratio of the count of firms on Compustat that conduct R&D and have patents divided by the total count of firms that conduct R&D (over the 1981–2006 sample) is (row 5/row 3) = 3,095 firms/6,499 firms = .476, or 47.6%. Thus, 1 – 0.476 = 0.524, or 52.4% of firms are performing R&D without patenting.

[21]Note that this mean for the entire sample period differs from that discussed in Section III.A for fiscal year 2001 because firms' R&D productivity has been declining since the mid-1980s.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 1

Summary Statistics

Table 1 reports pooled mean, standard deviation, 5th percentile, median, and 95th percentile points of the main variables used in the article. *N* is the number of firm-year observations. The sample period is from 1981 to 2015. The sample period for patent-based innovation measures (ln(1 + PAT), ln(1 + CIT/PAT), IE_PATENTS, and IE_CITATIONS) is from 1981 to 2006 because of data availability. RQ is the research quotient, defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as the firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2}+0.8R\&D_{i,t-3}+0.6R\&D_{i,t-4}+0.4R\&D_{i,t-5}+0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}$), scaled by the sum of corresponding R&D expenses ($R\&D_{i,t-3}+R\&D_{i,t-4}+R\&D_{i,t-5}+R\&D_{i,t-6}+R\&D_{i,t-7}$), 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 subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm *i*.

Variable	*N*	Mean	Std. Dev.	P5	Median	P95
RQ (1981–2015)	36,141	0.141	0.086	0.035	0.129	0.278
RQ (1981–2006)	26,967	0.1422	0.086	0.033	0.130	0.282
RD/BE (1981–2015)	132,309	0.1739	7.006	0.000	0.000	0.459
RD/BE (1981–2006)	106,396	0.1664	7.464	0.000	0.000	0.424
Missing Values Set to 0 (1981–2006)						
IE_PATENTS	106,396	0.061	0.241	0.000	0.000	0.355
IE_CITATIONS	106,396	0.062	0.242	0.000	0.000	0.358
ln(1 + PAT)	106,396	0.397	0.977	0.000	0.000	2.565
ln(1 + CIT)	106,396	0.494	1.080	0.000	0.000	3.070
Missing Values Excluded (1981–2006)						
IE_PATENTS	19,856	0.328	0.474	0.021	0.174	1.165
IE_CITATIONS	19,790	0.334	0.473	0.019	0.176	1.209
ln(1 + PAT)	22,701	1.862	1.321	0.693	1.386	4.635
ln(1 + CIT/PAT)	20,301	2.590	0.828	1.216	2.579	3.993

0.166 from 1981 to 2006. The mean values of ln(1 + PAT) and ln(1 + CIT/PAT) in the sample with values of 0 inserted are 0.397 and 0.494, respectively. The mean values of ln(1 + PAT) and ln(1 + CIT/PAT) in the missing-values-excluded sample are 1.862 and 2.590, respectively. The mean values of IE_PATENTS and IE_CITATIONS are 0.061 and 0.062, respectively, in the missing-values-set-to-0 sample, and they are 0.328 and 0.334, respectively, in the missing-values-excluded samples. The vast differences in means stem from the fact that approximately 80% of the observations for patents are missing and therefore set to 0.

In Table 2, we sort the sample into RQ quintiles each year and report the pooled average across the full sample of the year *t* variables for each quintile. The mean RQ ranges from 0.05 for the lowest quintile to 0.26 for the highest quintile. In contrast, ln(1 + PAT) and ln(1 + CIT/PAT) increase across the first through fourth quintiles and then decrease in the fifth quintile. The IE measures are essentially constant across the quintiles (for both patent-based and citation-based IE). Across increasing quintiles of RQ, we find that all measures of firm scale (revenue, R&D, advertising, market equity, and book equity) are nonmonotonic: increasing through quintile 4 and dropping in quintile 5. RQ is essentially uncorrelated with other firm factors, such as age and input intensity (R&D/ASSETS, PP&E/ASSETS, and CAPEX/ASSETS). Finally, valuation measures such as the market-to-book ratio of equity and Tobin's Q increase with RQ. Thus, at least univariately, RQ appears to be correlated with firm value.

In Panel B of Table 2, we report pooled pairwise correlations across the innovation measures. On average, the correlations among RQ and the previous

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 2

Summary Statistics by RQ Quintiles

For each fiscal year t in Table 2, we sort firms into quintiles based on the research quotient (RQ) in year t. We do this every year from 1981 to 2015 and report the equal-weighted average of each variable for each quintile. The sample period is from 1981 to 2015. The sample period for patent-based innovation measures ((ln(1 + PAT), ln(1 + CIT/PAT), IE_PATENTS, and IE_CITATIONS) is from 1981 to 2006 because of data availability. Missing firm-year values for patent-related measures are set to 0. RQ is defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as the firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2} + 0.8R\&D_{i,t-3} + 0.6R\&D_{i,t-4} + 0.4R\&D_{i,t-5} + 0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}$), scaled by the sum of corresponding R&D expenses ($R\&D_{i,t-3} + R\&D_{t,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}$), 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 subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm i. Other firm characteristics are defined commonly as in the literature, and the details can be found in Appendix B. Panel A reports summary statistics for each RQ quintile. Panel B reports the pooled pairwise correlations among the innovation variables.

Panel A. Summary of Statistics by RQ Quintiles

	RQ Quintiles					
	Low	2	3	4	High	High – Low
Innovation Variables						
RQ	0.053	0.103	0.132	0.162	0.257	0.204***
ln(1 + PAT)	0.575	0.965	1.229	1.287	1.123	0.548***
ln(1 + CIT/PAT)	0.747	0.992	1.169	1.231	1.122	0.375***
IE_PATENTS	0.142	0.134	0.127	0.126	0.133	−0.009*
IE_CITATIONS	0.158	0.147	0.145	0.140	0.153	−0.005
Firm Characteristics						
REVENUE	917.743	2,212.100	3,299.397	4,053.614	3,267.077	2,349.334***
R&D_EXPENSE	39.239	93.944	156.574	198.873	122.554	83.315***
ADVERTISING_EXPENSE	5.683	29.295	54.485	73.423	35.656	29.973***
MARKET_EQUITY	1,446.290	3,315.441	4,984.313	6,203.766	4,290.622	2,844.332***
BOOK_EQUITY	564.361	1,185.907	1,674.304	2,112.918	1,550.047	985.686***
AGE	20.355	23.383	25.517	24.169	20.636	0.281
R&D/ASSETS	0.086	0.073	0.069	0.074	0.110	0.024***
PPE/ASSETS	0.223	0.236	0.236	0.218	0.180	−0.043***
LEVERAGE	0.173	0.178	0.182	0.179	0.154	−0.019***
CAPEX/ASSETS	0.047	0.049	0.048	0.048	0.043	−0.004***
TOBIN_Q	1.898	1.845	1.926	1.976	2.484	0.586***
ACQUIRED_ASSETS	0.019	0.022	0.024	0.024	0.021	0.002
MTB	3.546	3.634	4.980	3.238	5.314	1.768***
RD/BE	0.252	0.198	0.568	0.191	0.334	0.082*
ABNORMAL_EARNINGS	−0.648	−0.052	0.190	−0.006	−0.273	0.375
ROA	0.037	0.092	0.114	0.110	0.051	0.014***
CASH_FLOW	−0.004	0.043	0.059	0.055	0.000	0.004

Panel B. Correlations Among Innovation Measures

	RQ	ln(1 + CIT/PAT)	ln(1 + PAT)	IE_PATENTS	IE_CITATIONS
RQ					
ln(1 + CIT/PAT)	0.077***				
ln(1 + PAT)	0.061***	0.741***			
IE_PATENTS	0.005	0.293***	0.283***		
IE_CITATIONS	0.000	0.264***	0.260***	0.418***	
RD/BE	0.004	0.003	0.002	0.002	0.019***

innovation measures are near 0. For example, the correlation between RQ and the RD/BE input measure is 0.004, that between RQ and patents is 0.061, and that between RQ and citations/patents is 0.077. Finally, the correlations between RQ and the 2 IE measures from Hirshleifer et al. (2013) are 0.005 and 0.000. The low correlations among RQ and the other measures show that the measures capture different aspects of innovation, suggesting that RQ potentially contains useful information beyond established innovation measures.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

III. Innovation Tests

Having defined the innovation measures and empirically characterized the relationship between them, we turn to our primary tests, those dealing with the relationships among the innovation measures and firm value. Our goal is to empirically explore the relevance of the RQ measure to firm value. Therefore, in this section, we perform tests to examine the relation of RQ with 3 forms of firm value: i) contemporaneous MTB, ii) subsequent MTB, and iii) future stock returns.

A. Innovation and Contemporaneous Firm Value

To test the link between innovation and firm value, we estimate the following general-form panel-regression model:

$$\begin{aligned}(6)\qquad \ln(\mathrm{MTB}_{i,t}) = {} & \alpha + \beta_1 \times \mathrm{RQ}_{i,t} \\ & + \beta_2 \times \mathrm{OTHER_INNOVATION_MEASURES}_{i,t} \\ & + \beta_3 \times \mathrm{CONTROLS} + \varepsilon_{i,t}.\end{aligned}$$

This regression model captures the contemporaneous relationship between a firm's RQ and its value as measured by the annual MTB. The expected sign on the RQ coefficient should be positive and significant; that is, increased RQ should be related to increased firm value. In all models, we include other standard control variables and variables from accounting-based asset-valuation models of corporate R&D.[22] We cluster standard errors at the firm level for all specifications and include year fixed effects. The RQ measure exhibits variation across firms but does not exhibit much firm-level variation over time (e.g., the average correlation between RQ and its lagged value is 0.87). Zhou (2001) shows that if an explanatory variable is slow moving, then regressions using firm fixed effects may fail to detect a relationship between the slow-moving explanatory variable and the dependent variable. Thus, we do not use firm fixed effects. Rather, in our panel regressions, we follow the approaches used in other articles with "slow" variables of interest (e.g., Hirshleifer et al. (2012), who test the effects of CEO overconfidence on innovation) and use panel regressions with time and industry fixed effects. As robustness tests, we implement cross-sectional regressions and instrument for RQ in 2SLS tests.[23]

Table 3 reports the contemporaneous firm-value regressions. Models 1 and 2 use RQ without controls for other innovation measures. Model 1 uses the entire data set of RQ firm-year values from 1981 to 2015. Model 2 uses 1981–2006 data, as do the remaining models. Models 3–6 estimate the firm-value regressions on the non-RQ measures in isolation to establish a base effect for these measures in the merged data set. Models 7–10 test RQ in bivariate combinations with each of the other innovation measures. Finally, model 11 combines RQ, IE_PATENTS, IE_CITATIONS, and ln(1 + PAT), and model 12 combines RQ, IE_PATENTS,

[22]See Ohlson (1995) and Sougiannis (1994).

[23]Gompers, Ishii, and Metrick (2003), Hirshleifer et al. (2013), and others use cross-sectional regressions to deal with slow-moving variables that exhibit variation across firms but little variation at a firm level across time.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 3

Regressions of MTB on Contemporaneous RQ and Other Innovation Measures

Table 3 reports panel regressions of contemporaneous market-to-book ratio (MTB) on the research quotient (RQ), innovation efficiency based on patents (IE_PATENTS), innovation efficiency based on citations (IE_CITATIONS), patent measures (ln(1 + PAT) and ln(1 + CIT/PAT)), and control variables. The sample period is from 1982 to 2015 for model 1 and 1982 to 2006 for all other models because of data availability for the patent-based measures. Missing firm-year values for patent-related measures are set to 0. RQ is defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2} + 0.8R\&D_{i,t-3} + 0.6R\&D_{i,t-4} + 0.4R\&D_{i,t-5} + 0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}$), scaled by the number of corresponding R&D expenses ($R\&D_{i,t-3} + R\&D_{t,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}$), 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 subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm i. RD/BE is research and development (R&D) expenditures divided by the book value of equity. ABNORMAL_EARNINGS is computed as $\left(E_{i,t}^{B}(1-\tau_{i,t}) - r_t \mathrm{BE}_{i,t-1}\right)/\mathrm{BE}_{i,t}$, 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 the tax rate, and r_t is the 1-year Treasury bill rate. R&D_TAX_SHIELD is R&D_EXPENDITURE multiplied by the tax rate, scaled by book equity. CAPEX/ASSETS is capital expenditure scaled by total book assets. All independent variables are contemporaneous to MTB. *t*-statistics are reported in parentheses. Standard errors are clustered at the firm level. In all models, we include industry (Fama–French 48) and year fixed effects (FE). Detailed variable definitions can be found in Appendix B. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: $\ln(\mathrm{MTB}_t)$											
	1	2	3	4	5	6	7	8	9	10	11	12
RQ	0.844***	0.917***					0.919***	0.919***	0.954***	0.941***	0.965***	0.950***
	(8.076)	(7.834)					(7.869)	(7.877)	(8.176)	(8.060)	(8.361)	(8.203)
IE_PATENTS			0.091***				0.093**				0.157***	0.074*
			(3.694)				(2.266)				(4.115)	(1.938)
IE_CITATIONS				0.097***				0.112***			0.149***	0.123***
				(4.190)				(3.162)			(4.260)	(3.511)
ln(1 + PAT)					−0.085***				−0.098***		−0.115***	
					(−11.626)				(−10.660)		(−12.385)	
ln(1 + CIT/PAT)						−0.029***				−0.034***		−0.045***
						(−4.962)				(−4.109)		(−5.508)
ABNORMAL_EARNINGS	−0.005***	−0.005***	−0.002***	−0.002***	−0.002***	−0.002***	−0.005***	−0.005***	−0.005***	−0.005***	−0.005***	−0.005***
	(−4.118)	(−3.990)	(−3.537)	(−3.537)	(−3.555)	(−3.542)	(−4.003)	(−4.005)	(−3.938)	(−3.952)	(−3.972)	(−3.968)
ADVERTISING_ME	−1.867***	−2.094***	−1.422***	−1.422***	−1.375***	−1.411***	−2.086***	−2.076***	−1.994***	−2.088***	−1.940***	−2.060***
	(−7.799)	(−8.612)	(−18.569)	(−18.579)	(−18.171)	(−18.479)	(−8.615)	(−8.595)	(−8.694)	(−8.657)	(−8.686)	(−8.658)
R&D_TAX_SHIELD	0.005***	0.005***	0.001	0.001	0.001	0.001	0.005***	0.005***	0.005***	0.005***	0.006***	0.005***
	(3.548)	(3.610)	(1.339)	(1.341)	(1.315)	(1.333)	(3.618)	(3.621)	(3.539)	(3.577)	(3.555)	(3.585)
ln(1 + CAPEX/ASSETS)	1.310***	1.213***	1.147***	1.147***	1.160***	1.159***	1.185***	1.155***	1.393***	1.290***	1.298***	1.226***
	(5.387)	(4.564)	(11.013)	(11.024)	(11.123)	(11.086)	(4.450)	(4.341)	(5.427)	(4.884)	(5.071)	(4.635)

(continued on next page)

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 3 (continued)

Regressions of MTB on Contemporaneous RQ and Other Innovation Measures

	Dependent Variable: $\ln(MTB_t)$											
	1	2	3	4	5	6	7	8	9	10	11	12
RD/BE	0.006***	0.005***	0.007**	0.007**	0.008**	0.007**	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(2.664)	(3.255)	(2.071)	(2.054)	(2.066)	(2.065)	(3.255)	(3.209)	(3.237)	(3.240)	(3.174)	(3.185)
ln(ME)	0.141***	0.140***	0.150***	0.150***	0.170***	0.156***	0.140***	0.139***	0.182***	0.148***	0.187***	0.150***
	(26.841)	(24.183)	(43.030)	(42.956)	(45.455)	(43.507)	(24.129)	(24.076)	(26.389)	(23.920)	(27.252)	(24.388)
Constant	−0.160***	−0.131***	−0.100***	−0.100***	−0.154***	−0.107***	−0.139***	−0.142***	−0.272***	−0.149***	−0.323***	−0.173***
	(−4.624)	(−3.561)	(−5.204)	(−5.219)	(−7.878)	(−5.536)	(−3.770)	(−3.853)	(−7.115)	(−4.013)	(−8.386)	(−4.641)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No of obs.	30,315	21,170	77,517	77,517	77,517	77,517	21,170	21,170	21,170	21,170	21,170	21,170
Adj. R^2	0.298	0.304	0.262	0.262	0.267	0.262	0.305	0.306	0.318	0.306	0.322	0.308

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

IE_CITATIONS, and ln(1 + CIT/PAT); we do not run all 5 of these innovation measures together because of the high correlation between ln(1 + PAT) and ln(1 + CIT/PAT). The number of firm-years in each model is determined by the constraints imposed by the independent variables in each model.[24] All models include the R&D input (RD/BE) as a control. Its coefficient is positive and significant in all Table 3 models.[25]

In model 1 of Table 3, with RQ and controls, the coefficient on RQ is positive and highly statistically significant (*t*-stat. = 8.08) over the full 1981– 2015 sample. In Model 2, over the shorter patent sample period (1981–2006), the coefficient on RQ is qualitatively similar to that of model 1 and again is statistically significant (*t*-stat. = 7.83). Thus, the positive relation between contemporaneous firm value and RQ is robust across these 2 sample periods.

In models 3 and 4 of Table 3, the coefficients on both IE_PATENTS and IE_CITATIONS are also positive and significant (*t*-stat. = 3.69 and 4.19, respectively), consistent with findings in Hirshleifer et al. (2013). In models 5 and 6, however, the coefficients on both ln(1 + PAT) and ln(1 + CIT/PAT) are negative and significant (*t*-stat. = –11.63 and –4.96, respectively). This is consistent with the findings of Hirshleifer et al. (2013), who find a negative and significant coefficient on a market-equity-scaled patent measure.

In the models in Table 3 combining RQ with each of the other 4 innovation measures (models 7–10), the significance of the coefficients on RQ is qualitatively unchanged from models 1 and 2 (*t*-stats range from 7.87 to 8.18). Thus, RQ survives controls for other measures of innovation. The coefficients on the other innovation measures are largely unchanged from models 3 through 6. The coefficients on IE_PATENTS and IE_CITATIONS remain positive and significant, and the coefficients on ln(1 + PAT) and ln(1 + CIT/PAT) remain significantly negative.

In model 11 of Table 3, we include RQ, IE_PATENTS, IE_CITATIONS, and ln(1 + PAT). Again, the coefficient on RQ retains its positive significance (*t*-stat. = 8.36), as do the coefficients for both IE_PATENTS and IE_CITATIONS (*t*-stat. = 4.12 and 4.26, respectively). The coefficient on ln(1 + PAT) remains negative and significant. In model 12, we substitute ln(1 + CIT/PAT) for ln(1 + PAT) in model 11. The results in model 12 are qualitatively the same as those in model 11.

B. Innovation and Subsequent Firm Value

Having tested RQ in contemporaneous MTB tests, we now test the ability of RQ to predict future firm value. In Table 4, we reestimate all regression specifications from Table 3, replacing contemporaneous MTB with subsequent (i.e., 1-year-ahead) MTB as the dependent variable. Thus, all independent variables are lagged 1 year relative to MTB.

[24]As noted, we do not insert 0 values for missing RQ observations in any of the models.

[25]We estimate the economic effects for the innovation measures in Table 3. A 1-standard-deviation change in RQ translates on average to a 7.53% change in MTB; for IE_PATENTS, it is a 2.22% change; for IE_CITATIONS, it is 2.38%; for ln(PAT) it is –7.97%; for ln(CIT/PAT), it is –3.08%; and for RD/BE, it is 4.29%.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 4

Regressions of MTB on Lagged RQ and Other Innovation Measures

Table 4 reports panel regressions of market-to-book ratio (MTB) on 1-year lagged values of the research quotient (RQ), innovation efficiency based on patents (IE_PATENTS), innovation efficiency based on citations (IE_CITATIONS), patent measures (ln(1 + PAT) and ln(1 + CIT/PAT)), and control variables. The sample period is from 1982 to 2015 for model 1 and 1982 to 2006 for all other models because of data availability for the patent-based measures. Missing firm-year values for patent-related measures are set to 0. RQ is defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2}+0.8R\&D_{i,t-3}+0.6R\&D_{i,t-4}+0.4R\&D_{i,t-5}+0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}$), scaled by the sum of corresponding R&D expenses ($R\&D_{i,t-3}+R\&D_{t,t-4}+R\&D_{i,t-5}+R\&D_{i,t-6}+R\&D_{i,t-7}$), 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 subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm i. RD/BE is research and development (R&D) expenditures divided by the book value of equity. ABNORMAL_EARNINGS is computed as $\left(E_{i,t}^{B}(1-\tau_{i,t})-r_t \mathrm{BE}_{i,t-1}\right)/\mathrm{BE}_{i,t}$, 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 the tax rate, and r_t is the 1-year Treasury bill rate. R&D_TAX_SHIELD is R&D expenditure multiplied by the tax rate, scaled by book equity. CAPEX/ASSETS is capital expenditure scaled by total book assets. All independent variables are lagged by 1 year relative to MTB. *t*-statistics are reported in parentheses. Standard errors are clustered at the firm level. In all models, we include industry (Fama–French 48) and year fixed effects (FE). The detailed variable definitions can be found in Appendix B. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: $\ln(\mathrm{MTB}_{t+1})$											
	1	2	3	4	5	6	7	8	9	10	11	12
RQ	0.862***	0.935***					0.937***	0.938***	0.955***	0.938***	0.965***	0.947***
	(7.704)	(7.362)					(7.394)	(7.405)	(7.527)	(7.376)	(7.658)	(7.475)
IE_PATENTS			0.144***				0.094**				0.115***	0.052
			(5.692)				(2.163)				(2.715)	(1.249)
IE_CITATIONS				0.117***				0.101***			0.114***	0.091**
				(4.588)				(2.740)			(3.082)	(2.478)
ln(1 + PAT)					−0.023***				−0.052***		−0.065***	
					(−3.086)				(−5.371)		(−6.629)	
ln(1 + CIT/PAT)						0.016***				−0.004		−0.012
						(2.780)				(−0.434)		(−1.366)
ABNORMAL_EARNINGS	−0.001*	−0.001*	−0.001**	−0.001**	−0.001**	−0.001**	−0.001*	−0.001*	−0.001*	−0.001*	−0.001*	−0.001*
	(−1.809)	(−1.760)	(−2.028)	(−2.028)	(−2.032)	(−2.023)	(−1.781)	(−1.790)	(−1.748)	(−1.754)	(−1.805)	(−1.783)
ADVERTISING_ME	−1.181***	−1.467***	−1.149***	−1.148***	−1.132***	−1.155***	−1.458***	−1.448***	−1.407***	−1.466***	−1.361***	−1.443***
	(−5.125)	(−6.162)	(−11.370)	(−11.368)	(−11.205)	(−11.409)	(−6.154)	(−6.113)	(−6.102)	(−6.165)	(−6.014)	(−6.123)
R&D_TAX_SHIELD	0.001	0.001*	0.001	0.001	0.001	0.001	0.001*	0.001*	0.001*	0.001*	0.001*	0.001*
	(1.584)	(1.774)	(1.065)	(1.074)	(1.055)	(1.067)	(1.789)	(1.805)	(1.745)	(1.769)	(1.791)	(1.796)
ln(1 + CAPEX/ASSETS)	0.869***	0.783***	0.798***	0.800***	0.811***	0.802***	0.753***	0.731***	0.885***	0.792***	0.814***	0.746***
	(3.512)	(2.936)	(7.114)	(7.141)	(7.213)	(7.146)	(2.817)	(2.738)	(3.383)	(2.976)	(3.114)	(2.798)

(continued on next page)

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 4 (continued)

Regressions of MTB on Lagged RQ and Other Innovation Measures

	Dependent Variable: $\ln(MTB_{t+1})$											
	1	2	3	4	5	6	7	8	9	10	11	12
RD/BE	0.002**	0.001**	0.002*	0.002	0.002*	0.002*	0.001**	0.001**	0.001**	0.001**	0.001**	0.001**
	(2.132)	(2.526)	(1.678)	(1.619)	(1.674)	(1.683)	(2.528)	(2.401)	(2.514)	(2.521)	(2.376)	(2.400)
ln(ME)	0.110***	0.108***	0.108***	0.108***	0.114***	0.105***	0.108***	0.108***	0.131***	0.109***	0.136***	0.110***
	(20.461)	(18.149)	(29.753)	(29.703)	(28.727)	(28.223)	(18.100)	(18.055)	(17.882)	(17.116)	(18.538)	(17.448)
Constant	0.025	0.054	0.092***	0.093***	0.079***	0.102***	0.046	0.044	−0.025	0.052	−0.066	0.034
	(0.676)	(1.374)	(4.446)	(4.486)	(3.691)	(4.919)	(1.166)	(1.118)	(−0.597)	(1.311)	(−1.562)	(0.862)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No of obs.	27,285	19,674	70,343	70,343	70,343	70,343	19,674	19,674	19,674	19,674	19,674	19,674
Adj. R^2	0.220	0.217	0.186	0.186	0.185	0.185	0.218	0.218	0.221	0.217	0.224	0.218

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

In Table 4, the coefficients on RQ remain statistically significant across all models. In models with just RQ (and controls), bivariate combinations of RQ and other innovation measures, and models with groups of other innovation measures, the RQ coefficients obtain highly significant *t*-statistics. In general, the coefficient magnitudes and significance across the other innovation measures are similar to Table 3, with the exception of model 6. Whereas the coefficient on ln(1 + CIT/PAT) is negative and significant in all contemporaneous MTB regressions in Table 3, it becomes positive and significant in isolation in subsequent MTB tests.

Overall, across the contemporaneous and predictive firm-value regressions, RQ is positively and reliably related to contemporaneous and future firm value. This conclusion holds across a full range of controls for the previous innovation measures. For the other innovation measures, we find that innovation input, measured by RD/BE, is consistently positively and significantly related to both contemporaneous and future firm value. The IE measures from Hirshleifer et al. (2013) are positively related to both measures of firm value. Finally, traditional patent-based output-innovation measures are mostly negatively (and significantly) related to firm current and future value.

C. An Instrumental-Variable Approach for RQ in the Firm-Value Tests

This section addresses the potential endogeneity concern that omitted variables correlated with both RQ and firm value could bias our results. A firm's RQ may be affected by a number of unobservable factors, such as firm culture, management, and CEO experience, and by other omitted variables that potentially affect firm value directly. To address the problem of endogenously omitted variables, we instrument for RQ using a plausibly exogenous instrumental variable (IV), R&D_CENTRALIZATION, previously shown to affect firm innovation (see, e.g., Argyres and Silverman (2004), Arora et al. (2011), and Cummings (2019)). R&D_CENTRALIZATION refers to the location of R&D decision making (headquarters vs. divisions). The potential advantage of decentralizing R&D is greater responsiveness to the customer. However, decentralization typically comes at the expense of both early-stage R&D and R&D that benefits multiple divisions. Both effects (customer responsiveness and the shift toward late-stage and parochial innovation) stem from the fact that division managers' compensation is generally tied to division profits rather than to the firm's market value. Indeed, empirical articles have found that firms with centralized R&D produce innovations with a larger and broader impact on subsequent technological evolution (Argyres and Silverman (2004), Arora, Belenzon, and Rios (2014)). Therefore, R&D_CENTRALIZATION satisfies the relevance condition for an IV. Theoretically, a firm's organizational form of R&D is unlikely to affect firm value other than through IE, so the measure would seem to satisfy the exclusion restriction as well. Moreover, previous work (see Hall et al. (2010)) assumes that IE is shared across firms within an industry. Accordingly, to further alleviate exclusion-restriction concerns, we use the industry median of R&D_CENTRALIZATION to serve as an instrument for RQ.

We perform our MTB regressions using a 2SLS regression that models RQ as a function of industry-median R&D_CENTRALIZATION in the first stage.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 5

2SLS Regression to Instrument RQ

Table 5 reports the 2-stage least squares (2SLS) regression results. The sample period is from 1981 to 2015. Columns 1 and 2 repeat the OLS regressions that regress the market-to-book ratio (MTB) on the research quotient (RQ) from Tables 3 and 4. Column 3 is the first-stage regression. Columns 4 and 5 are the second-stage regression. The instrumental variable is R&D_RD, defined as the industry median of the firm's R&D_CENTRALIZATION, which is the sum of the squared proportions of patent inventors in the top 5 of the firm's most frequent metropolitan areas. The remaining variables are defined exactly the same as in Tables 3 and 4, and their definitions can be found in Appendix B. Standard errors are clustered at the firm level. In all models, we include industry (Fama–French 48) and year fixed effects. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	OLS Regressions		2SLS		
Dependent Variable	$\ln(MTB_t)$	$\ln(MTB_{t+1})$	RQ_t	$\ln(MTB_t)$	$\ln(MTB_{t+1})$
	1	2	3	4	5
RQ	0.764***	0.764***		5.165***	4.126**
	(7.398)	(6.987)		(3.117)	(2.518)
R&D_CENTRALIZATION			0.039***		
			(5.77)		
ABNORMAL_EARNINGS	−0.005***	−0.001	0.00	−0.005***	−0.001
	(−4.018)	(−1.620)	(−0.920)	(−3.570)	(−1.253)
ADVERTISING_ME	−2.093***	−1.357***	0.026	−2.216***	−1.472***
	(−8.658)	(−5.853)	(1.292)	(−9.288)	(−6.398)
R&D_TAX_SHIELD	0.005***	0.001	0	0.005***	0.001
	(3.464)	(1.416)	(1.262)	(3.057)	(1.041)
ln(1 + CAPEX/ASSETS)	1.341***	0.893***	−0.055**	1.470***	1.004***
	(5.842)	(3.854)	(−2.270)	(5.714)	(4.068)
RD/BE	0.005***	0.002**	0	0.005***	0.002**
	(2.594)	(2.12)	(0.627)	(2.623)	(2.157)
ln(ME)	0.145***	0.109***	0.001***	0.140***	0.105***
	(29.307)	(21.685)	(3.319)	(26.103)	(18.754)
Constant	−0.169***	0.043	0.116***		
	(−4.945)	(1.207)	(24.108)		
F-statistic			76.99		
No. of obs.	30,316	27,286	30,223	30,223	27,208
Adj. R^2	0.274	0.191	0.034	−0.001	−0.011

The second stage then examines the impact of first-stage-estimated RQ on MTB. To capture R&D_CENTRALIZATION, we employ a measure developed by Cummings (2019),[26] a Herfindahl-like index of the geographic dispersion of the firm's patent inventors. This measure is constructed by first generating the proportion of the firm's patent inventors who live in each of the firm's 5 most frequent metropolitan areas (core-based statistical area). The R&D_CENTRALIZATION variable is then the sum of the squared proportion of patent inventors in each geographic region. Thus, if 50% of a firm's inventors are in region 1 and 50% are in region 2, the variable's value is $(0.5^2 + 0.5^2 = 0.50)$. Conversely, if 90% of inventors are in region 1 and 10% in region 2, the value is $(0.9^2 + 0.1^2 = 0.82)$.

We report the 2SLS regressions in Table 5 . For comparisons, models 1 and 2 of Table 5 repeat the OLS results (MTB regressed on RQ and controls, and year-ahead MTB regressed on RQ and controls, respectively) from Tables 3 and 4. Model 3 of Table 5 reports the first-stage regression of RQ regressed on the industry median of R&D_CENTRALIZATION and control variables. The control variables are the same as in the OLS baseline regressions from Table 3. Industry and year fixed effects are included, and the standard errors are clustered at the firm level.

[26]We thank the author for providing us with the centralization data.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

The coefficient on R&D_CENTRALIZATION from the first-stage regression in model 3 of Table 5 is positive and significant (t-stat. = 5.77). The F-test from the first-stage regression is 77, suggesting that our instrument passes the weak-IV test. In model 4, we perform the second stage and regress the contemporaneous MTB on the instrumented RQ. The instrumented RQ coefficient is 5.17, with a t-statistic of 3.12. This is larger than the OLS coefficient of 0.764, suggesting that once potentially omitted variables are accounted for in the 2SLS, the impact of RQ on firm value is more positive relative to OLS estimates. Finally, in model 5, we repeat the second-stage analysis but focus on the next-period MTB. Specifically, we estimate the same first-stage regression and then regress year-ahead MTB on the instrumented RQ and controls. The results here are similar to those of the contemporaneous MTB model, with a coefficient on the instrumented RQ of 4.13 and a t-statistic of 2.52. In sum, these 2SLS identification tests suggest a causal effect of RQ on both contemporaneous and future firm value.

Jiang (2017) surveys the use of IVs in the finance literature and reports that approximately 80% of published articles obtain IV coefficient estimates much larger (7 times larger on average) than their OLS estimates. In light of the Jiang article, we note that in our 2SLS results of Table 5, the ratios of IV-to-OLS coefficient estimates using R&D industry-level centralization as the instrument are also large, ranging from 5.4 to 6.8. As Jiang observes, the larger IV coefficients relative to OLS may arise from a number of issues. One concern is a "local average treatment effect," in our case that firms belonging to the more centralized R&D industries are more likely to be sensitive to RQ. Hirshleifer et al. (2013) show that investors tend to underreact to IE. However, if investors tend to react more efficiently (i.e., exhibit less underreaction) to the firms in industries with more centralized R&D, then we would expect the OLS coefficients, which capture the population average treatment effect, to be smaller than the 2SLS coefficients, which capture the local average treatment effect.

Jiang (2017) also proposes that weak instruments could lead to larger ratios of IV-to-OLS coefficients. Our first-stage regression yields an F-statistic of 77, which safely passes the weak-instrument test (Stock and Yogo (2005)). However, we find that the R^2 of our first-stage regression is a relatively low 3.4%. Thus, our relatively large IV-to-OLS coefficient ratio could in part reflect a weak-instrument issue.

Lastly, direct and indirect specification searches are potentially an important problem in all areas of finance (see Harvey, Liu, and Zhu (2016) for specification search issues in asset pricing or Heath, Ringgenberg, Samadi, and Werner (2019) for related concerns in corporate finance). We emphasize that we did not engage in an overt specification search to find our centralization instrument. Rather, we chose the centralization instrument because of compelling discussions in Argyres and Silverman (2004), Arora et al. (2011), and Cummings (2019) relating centralization to firm innovation.

Nonetheless, as an additional test of endogeneity, we introduce another IV, industry-level RQ. Both Gormley and Matsa (2014) and Grieser et al. (2019) discuss the possibility of using industry shocks from an explanatory variable of interest as an instrument in a panel context. We argue that the exclusion restrictions are likely to hold for this new instrument: Industry-level shocks to IE are more likely to capture exogenous changes to the industry-level dynamics that have no direct

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

impact on firm value other than through industry-level RQ. For example, consider an industry-level shock such as new regulation that changes the number of patients required for Phase III trials in the pharmaceutical industry. This rule would increase the cost per drug, which (all else equal) would reduce financial outcomes, but only through its impact on R&D and RQ. Similarly, declining technological opportunity (the ease of achievement of innovations and technical improvements within a technological domain) would depress firm value, but only through its impact on R&D.

For this second instrument, we estimate 2SLS using the industry median of RQ (excluding the firm's own RQ) as an IV. The first-stage coefficient on the RQ industry-median variable is positive and significant, with a coefficient of 0.75 and a t-statistic of 15. The F-statistic of the first-stage regression is 1,066, indicating a very strong IV. We find 2SLS industry RQ coefficients of 1.29 (t-stat. = 2.8) when regressed on contemporaneous MTB and 1.07 (t-stat. = 2.2) when regressed on next year's MTB. Thus, the ratio of the IV-to-OLS coefficients is lower than the ratio for the models using industry R&D_CENTRALIZATION and, on an absolute-level basis, should not be considered "implausibly large" relative to the numbers in Jiang (2017). These IV results are reported in Table D.1 of the Supplementary Material.

D. Innovation and Stock Returns

In this section, we examine the extent to which each innovation measure is able to predict the cross section of future monthly returns. Hirshleifer et al. (2013) propose two mechanisms to link IE with future returns, one based on mispricing and one on risk. Their mispricing explanation is centered on investor underreaction to innovation due to limited attention, and their risk-based story relates to q-theory. In our return tests, we control for established predictors of the cross section of returns, such as firm market capitalization, book-to-market equity, past performance, asset growth, and profitability. All models include controls for RD/BE. The definitions of the control variables may be found in Appendix B. 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 are from 1981 to 2015, and the final monthly return data set runs from July 1982 to June 2016 for tests using RQ. For tests using the patent-based measures, the sample ends in 2007. We estimate cross-sectional Fama and MacBeth (1973) regressions for each month and report the time-series average of coefficients and adjusted R^2. Standard errors are computed using a Newey and West (1987) estimator allowing for 3 lags of serial correlation. To control for microstructure and tiny-stock issues, we exclude stocks priced below \$5 at the end of the previous month.

In the return regressions, we also test an RD_ABILITY measure, modeled after Cohen et al. (2013) (CDM), designed to capture a firm's past track record of converting R&D spending into sales growth. CDM show that a binary measure of RD_ABILITY (high/low) interacted with R&D_INTENSITY positively predicts returns in both the United States and abroad when separately controlling for RD_ABILITY and R&D_INTENSITY. We estimate a continuous version of

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

RD_ABILITY, defined as the average of coefficients from rolling regressions of sales growth regressed on lagged R&D scaled by sales. RD_ABILITY is conceptually similar to RQ, with a few important differences: First, RD_ABILITY is estimated using sales *growth* (whereas because RQ is the output elasticity of R&D, it is estimated using sales *levels*). Second, RD_ABILITY is computed in a linear OLS regression of sales growth on lagged R&D, whereas RQ is estimated from a production function with a full set of firm-production function inputs (see our equation (2) and Appendix A). After estimating RD_ABILITY and multiplying it by R&D_INTENSITY, we find a correlation of nearly 0 between RQ and RD_ABILITY. Thus, empirically, the RQ measure is quite distinct from the RD_ABILITY measure, just as it is from other innovation measures.

We first run each innovation measure separately. We then combine RQ with each of the 5 other innovation measures: IE_PATENTS, IE_CITATIONS, ln(1 + PAT), ln(1 + CIT/PAT), and the CDM (2013) measure. Finally, we estimate models with larger sets of innovation measures.

The results are reported in Table 6. In model 1, with RQ and controls, the coefficient on RQ is positive and statistically significant (*t*-stat. = 2.60) over the full 1982–2016 sample. In model 2, over the patent sample period (1982–2007), the coefficient on RQ is also positive and statistically significant (*t*-stat. = 2.89). The coefficients on the innovation input measure, RD/BE, are positive and significant across models. In models 3 and 4, the coefficients on both IE_PATENTS and IE_CITATIONS are positive and significant (*t*-stats = 3.49 and 3.48, respectively), consistent with findings in Hirshleifer et al. (2013). In models 5 and 6, the coefficients on both ln(1 + PAT) and ln(1 + CIT/PAT) are positive and significant (*t*-stat. = 2.16 and 3.29, respectively), also consistent with Hirshleifer et al. (2013). In model 7, the coefficient on RD_ABILITY interacted with R&D_ INTENSITY is positive and significant, as found in Cohen et al. (2013). Thus, in models with single measures of previous innovation variables, our results are consistent with the related literature.

Table 6 results change in the models combining RQ with each of the other 5 innovation measures (models 8–12). While the coefficients on RQ remain positive and significant in all models, for the other innovation measures, only the coefficient on ln(1 + CIT/PAT) remains positive and marginally significant (*t*-stat. = 1.60). Finally, in models 13 and 14 (which include larger groups of the innovation measures), the coefficient on RQ remains positive and significant, whereas it is never positive and significant for any other innovation measure.

To summarize, across the 3 types of firm-value tests (current MTB, future MTB, and future returns), RQ is positively and significantly related to each measure. This holds in models with RQ as the sole innovation measure and in models that control for previous input, output, and efficiency measures of innovation. All other innovation measures fail to be significant in at least one specification.

E. Robustness Tests

Following the convention in the literature, our analysis so far has used a sample that inserts values of 0 into missing firm-year values of patent-based

TABLE 6

Fama–MacBeth Regressions of Stock Returns on RQ and Other Innovation Measures

Table 6 reports Fama–MacBeth regressions of firm excess monthly stock returns on lagged values of the research quotient (RQ) and other variables. Firm excess monthly return is the monthly return from CRSP, including dividends, minus the 1-month Treasury bill rate (from Ibbotson Associates). The sample period is from July 1982 to June 2016 for model 1 and July 1982 to June 2009 for other models because of data availability for the patent-based innovation measures. Missing firm-year values for patent-related measures are set to 0. RQ is defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PATENTS) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CITATIONS/PATENTS) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as the firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2} + 0.8R\&D_{i,t-3} + 0.6R\&D_{i,t-4} + 0.4R\&D_{i,t-5} + 0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{i,j}} C_{i,k}^{t-j}$), scaled by the sum of corresponding R&D expenses ($R\&D_{i,t-3} + R\&D_{t,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}$), 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 subcategory granted in year $t-j$, and N_{t-j} is the total number of patents granted in year $t-j$ to firm i. RD/BE is R&D expenditures divided by the book value of equity. ln(MKTCAP) is the natural logarithm of the market value of equity in June of year t. ln(B/M) is the natural logarithm of the book-to-market ratio of equity. Book equity is defined as shareholder's equity plus balance-sheet-deferred taxes and investment tax credits minus the book value of preferred stock in the 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 compounded buy and hold return from the previous 6 months computed by skipping the most recent lagged month. ASSET_GROWTH is the 1-year percentage change in total firm assets. OPERATING_PROFIT is measured with accounting data for the fiscal year ending in year $t-1$ and is revenues minus cost of goods sold; minus selling, general, and administrative expenses; minus interest expense, all divided by book equity. R&D_INTENSITY is the natural logarithm of 1 plus research and development (R&D) divided by REVENUE. HIGH_ABILITY is a dummy variable indicating whether the firm is in the top quintile in terms of innovation ability defined as in Cohen, Diether, and Malloy (2013) as the average of 5 coefficients from 5 rolling firm-by-firm regressions of firm-level sales growth on the natural logarithm of lagged R&D. The reported coefficients are time-series averages of cross-sectional monthly regression coefficients. The t-statistics (in parentheses) are computed using a Newey and West (1987) estimator allowing for 3 lags of serial correlation. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Dependent Variable: r_{t+1}													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
RQ	0.012***	0.019***						0.019***	0.019***	0.018***	0.018***	0.018***	0.018***	0.018***
	(2.604)	(2.886)						(2.881)	(2.874)	(2.841)	(2.831)	(2.783)	(2.766)	(2.755)
IE _PATENTS			0.003***					0.001					0.000	−0.000
			(3.487)					(0.710)					(0.112)	(−0.145)
IE _CITATIONS				0.003***					0.001				0.001	0.000
				(3.482)					(0.932)				(0.534)	(0.248)
ln(1 + PAT)					0.001**					−0.000			−0.000	
					(2.161)					(−0.141)			(−0.449)	
ln(1 + CIT/PAT)						0.001***					0.001			0.000
						(3.288)					(1.602)			(1.330)
HIGH_ABILITY × R&D_INTENSITY							0.060*					0.041	0.029	0.031
							(1.919)					(0.598)	(0.364)	(0.393)
HIGH_ABILITY							−0.001					−0.003*	−0.003*	−0.003*
							(−0.964)					(−1.910)	(−1.760)	(−1.760)
R&D_INTENSITY							−0.000					−0.004	−0.004	−0.004
							(−0.140)					(−0.571)	(−0.596)	(−0.612)

(continued on next page)

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 6 (continued)

Fama–MacBeth Regressions of Stock Returns on RQ and Other Innovation Measures

	Dependent Variable: r_{t+1}													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
RD/BE	0.006***	0.006**	0.006***	0.006***	0.006***	0.006***	0.006***	0.006**	0.006**	0.006**	0.006**	0.005*	0.005*	0.005*
	(3.668)	(2.533)	(3.462)	(3.445)	(3.362)	(3.307)	(3.165)	(2.558)	(2.568)	(2.467)	(2.507)	(1.862)	(1.857)	(1.858)
ln(ME)	−0.000	−0.000	0.000	0.000	−0.000	−0.000	0.000	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000	−0.000
	(−0.374)	(−0.550)	(0.385)	(0.388)	(−0.153)	(−0.029)	(0.329)	(−0.534)	(−0.542)	(−0.790)	(−0.863)	(−0.509)	(−0.583)	(−0.771)
ln(BM)	0.003***	0.003***	0.005***	0.005***	0.005***	0.005***	0.005***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***	0.003***
	(4.147)	(3.383)	(5.648)	(5.648)	(5.498)	(5.705)	(5.896)	(3.398)	(3.419)	(3.228)	(3.292)	(3.697)	(3.620)	(3.626)
BHRET6	0.004	0.004	0.005*	0.005*	0.005*	0.005*	0.005*	0.004	0.004	0.003	0.004	0.003	0.004	0.004
	(1.569)	(1.285)	(1.718)	(1.721)	(1.668)	(1.684)	(1.727)	(1.301)	(1.292)	(1.272)	(1.290)	(1.283)	(1.302)	(1.314)
ASSET_GROWTH	−0.002***	−0.003***	−0.000*	−0.000*	−0.000*	−0.000*	−0.000*	−0.003***	−0.003***	−0.003***	−0.003***	−0.003***	−0.003***	−0.003***
	(−2.908)	(−3.162)	(−1.789)	(−1.790)	(−1.790)	(−1.784)	(−1.693)	(−3.150)	(−3.137)	(−3.164)	(−3.096)	(−2.994)	(−2.995)	(−2.938)
OPERATING_PROFIT	0.003***	0.003*	0.003**	0.003**	0.003**	0.003**	0.003**	0.003*	0.003*	0.003**	0.003*	0.004**	0.004**	0.004**
	(2.771)	(1.923)	(2.285)	(2.282)	(2.272)	(2.272)	(2.197)	(1.841)	(1.957)	(2.036)	(1.953)	(2.349)	(2.373)	(2.367)
Constant	0.009**	0.008*	0.006	0.006	0.007	0.006	0.006	0.008*	0.008*	0.009*	0.008*	0.009**	0.009*	0.009*
	(2.557)	(1.901)	(1.367)	(1.370)	(1.537)	(1.450)	(1.524)	(1.827)	(1.861)	(1.929)	(1.908)	(1.976)	(1.870)	(1.921)
No of obs.	377,808	293,775	1,090,750	1,090,750	1,090,750	1,090,750	1,082,672	293,775	293,775	293,775	293,775	293,775	293,775	293,775
Avg. R^2	0.038	0.042	0.027	0.028	0.028	0.029	0.031	0.044	0.044	0.044	0.044	0.051	0.056	0.057
No. of months	402	324	324	324	324	324	324	324	324	324	324	324	324	324

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

measures. An alternate approach is to drop observations with missing patents. The logic here is that inserting values of 0 implies that firms without patents are not innovating. We know this is false, however, because 52% of the firms that conduct R&D never file patents, preferring to protect their innovations by other means (Cohen, Nelson, and Walsh (2000)). The advantages and disadvantages of sample-formation methods and related issues are discussed in Koh, Reeb, Sojli, and Tham (2016), Lerner and Seru (2017), and others. Given that researchers may consider both approaches, this section examines the robustness of previously reported results (using a sample that inserts values of 0 into missing firm-year values of patent-based measures) to an alternative sample constructed using only nonmissing patent data.

The results are reported in Table 7. Panel A reproduces Table 3 (contemporaneous MTB regressions) on the missing-patents-excluded sample, Panel B reproduces results for Table 4 (future MTB regressions), and Panel C reproduces results for Table 6 (return regressions). All regressions in these panels use the same controls and methodologies as the previous tables, but we suppress the output on the control variables to save space. We note that the coefficients on the controls are similar to those in the previous tables and are generally consistent with the findings in the related literature.

Before we discuss this section's main results, we should note that, across the missing-patents-set-to-0 sample and the missing-patents-excluded sample, we find differences in firm-year counts and firm-mean characteristics. For example, in the earlier Table 3 firm-value regressions, in model 5 using ln(1 + PAT), 77,517 firm-year observations emerge when we insert 0 values into missing patents. In contrast, for the same regression model in Panel A of Table 7, there are 17,177 firm-years of data for the sample with missing patents excluded. Thus, we lose 77% of observations. The 2 samples are also different in terms of common firm characteristics. We compute (but do not report in the tables) the means of common firm characteristics across the samples. We find that the firms in the sample with missing patents excluded, relative to the firms in the sample with missing patents set to 0, are significantly larger (average market equity is $3,634 million vs. $1,979 million), enjoy higher revenue ($2,684 million vs. $1,650 million), spend more on R&D relative to assets (0.085 vs. 0.052), are slightly older (20.4 years vs. 17.4 years), and are more profitable (ROA of 6.7% vs. 5.2%). All of these differences are statistically significant. Finally, we note that the mean RQ is similar across the 2 groups (0.147 for missing patents excluded vs. 0.141 for missing patents set to 0). Thus, the sample of firms with patents is qualitatively very different from the sample of all firms conducting R&D.

Moreover, we find that the RQ effects are highly robust across samples.[27] Across all 3 sets of tests in Panels A, B, and C (contemporaneous value, future firm value, and future returns) of Table 7 and across model types (RQ as the innovation measure, bivariate combinations of RQ and other innovation measures, and RQ with groups of other innovation measures), the coefficients on RQ are always

[27]Note that in Table 7, in each panel, model 1 uses RQ as the innovation measure and is the same regression (and has the same firm-year count) as in the earlier tables (and is included for comparison purposes). These RQ models have the same firm-year counts across the 2 samples because we do not insert 0 values (in any of the tables) into the missing RQ observations.

TABLE 7

Robustness Checks Using Alternative Sample

Table 7 reports the results from reestimating the tests in Tables 3, 4, and 6 using an alternative sample where we drop observations with missing firm-year values of patent-based measures. The sample period is from 1981 to 2006. Panels A, B, and C replicate the regressions in Tables 3, 4, and 6, respectively, using the alternative sample. All the control variables are included but not tabulated. RQ is the research quotient, defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in a given year. IE_PATENTS is the innovation efficiency based on patents, defined as the firm's patents granted in a given year scaled by R&D capital ($R\&D_{i,t-2} + 0.8R\&D_{i,t-3} + 0.6R\&D_{i,t-4} + 0.4R\&D_{i,t-5} + 0.2R\&D_{i,t-6}$). IE_CITATIONS is the adjusted number of citations received by patents granted in the past 5 years ($\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}$), scaled by the sum of corresponding R&D expenses ($R\&D_{i,t-3} + R\&D_{t,t-4} + R\&D_{i,t-5} + R\&D_{i,t-6} + R\&D_{i,t-7}$), where $C_{i,k}^{t-j}$ 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 subcategory granted in year $t - j$, and N_{t-j} is the total number of patents granted in year $t - j$ to firm i. R&D_INTENSITY is the natural logarithm of 1 plus R&D divided by revenue. HIGH_ABILITY is a dummy variable indicating whether the firm is in the top quintile in terms of innovation ability. Innovation ability is defined as in Cohen et al. (2013), which is the average of 5 coefficients from 5 rolling firm-by-firm regressions of firm-level sales growth (defined as $\ln(SALES_t/SALES_{t-1})$) on the natural logarithm of lagged R&D ($\ln(R\&D_{t-j}/SALES_{t-j})$, where $j = 1, 2, 3, 4, 5$). In all panels, the regressions use the same controls and methodologies as the previous tables, but we suppress the control-variable output. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Sample Robustness Tests of Table 3 Panel Regressions of MTB on Contemporaneous Innovation

	Dependent Variable: $\ln(MTB_t)$										
	1	2	3	4	5	6	7	8	9	10	11
RQ	0.917***					1.070***	0.966***	1.053***	0.858***	1.180***	0.951***
	(7.834)					(7.050)	(6.233)	(5.900)	(4.762)	(5.792)	(4.650)
IE_PATENTS		0.269***				0.370***				0.368***	0.218***
		(7.956)				(7.024)				(4.815)	(2.857)
IE_CITATIONS			0.254***				0.281***			0.303***	0.240***
			(9.489)				(6.779)			(5.389)	(4.052)
ln(1 + PAT)				−0.131***				−0.155***		−0.197***	
				(−12.161)				(−11.355)		(−12.149)	
ln(1 + CIT/PAT)					0.106***				0.093***		0.092***
					(8.443)				(5.303)		(4.525)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects (FE)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	21,170	15,914	16,026	17,177	14,919	11,286	12,216	9,580	8,259	7,517	6,548
Adj. R^2	0.304	0.311	0.313	0.333	0.326	0.346	0.332	0.374	0.360	0.436	0.409

(continued on next page)

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 7 (continued)

Robustness Checks Using Alternative Sample

Panel B. Sample Robustness Tests of Table 4 Panel Regressions of MTB on Lagged Innovation

	Dependent Variable: ln(MTB$_{t+1}$)										
	1	2	3	4	5	6	7	8	9	10	11
RQ	0.935***					0.963***	0.865***	1.055***	0.831***	1.131***	0.907***
	(7.362)					(5.645)	(4.919)	(5.488)	(4.454)	(5.043)	(4.109)
IE_PATENTS		0.201***				0.235***				0.205**	0.124
		(6.189)				(4.194)				(2.484)	(1.514)
IE_CITATIONS			0.182***				0.182***			0.221***	0.186***
			(6.129)				(4.241)			(3.770)	(3.062)
ln(1 + PAT)				−0.056***				−0.085***		−0.115***	
				(−4.971)				(−5.715)		(−6.406)	
ln(1 + CIT/PAT)					0.121***				0.114***		0.113***
					(9.144)				(6.392)		(5.253)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of obs.	19,674	14,902	14,977	16,911	14,690	10,679	11,493	9,464	8,159	7,434	6,478
Adj. R^2	0.217	0.222	0.224	0.234	0.243	0.245	0.233	0.258	0.261	0.293	0.290

Panel C. Sample Robustness Tests of Table 6 Fama–MacBeth Regressions

	Dependent Variable: r_{t+1}												
RQ	0.019***						0.014**	0.014**	0.017**	0.021***	0.019***	0.015*	0.020**
	(2.886)						(2.228)	(2.150)	(2.225)	(2.674)	(2.818)	(1.836)	(2.213)
IE _PATENTS		−0.001					−0.003*					−0.004	−0.003
		(−0.517)					(−1.683)					(−1.157)	(−0.762)
IE _CITATIONS			−0.000					−0.001				0.002	0.002
			(−0.079)					(−0.369)				(0.654)	(0.825)
ln(1 + PAT)				0.001*					0.000			0.001	
				(1.910)					(0.719)			(1.167)	
ln(1 + CIT/PAT)					0.002**					0.002**			0.002*
					(2.412)					(2.127)			(1.676)
HIGH_ABILITY × R&D_INTENSITY						0.061**					0.043	−0.033	−0.023
						(1.986)					(0.645)	(−0.314)	(−0.205)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No of obs.	293,775	220,309	218,544	262,927	235,528	1,082,672	161,814	170,760	146,328	130,604	293,775	113,985	102,506
R^2	0.042	0.051	0.050	0.055	0.064	0.031	0.067	0.063	0.081	0.089	0.051	0.129	0.143
No. of months	324	312	312	312	300	324	312	312	312	300	324	312	300

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

positive and statistically significant, with *t*-statistics ranging from 4.10 to 7.83 for the MTB regressions in Panels A and B to 1.84 to 2.89 in the Panel C Fama–MacBeth return regressions.

Conversely, the signs and significances for the coefficients of the other innovation variables change to varying degrees relative to the results with 0 values inserted shown in the earlier tables. The coefficients on RD/BE are positive and significant in 6 of 13 models. For the other IE measures, the coefficients on IE_CITATIONS and IE_PATENTS remain positive and significant in most of the firm-value tests of Table 7. In general, the IE coefficient *t*-statistics are higher in the samples without 0 values inserted than in the samples with 0 values inserted for both the contemporaneous (Panel A) and future (Panel B) MTB tests. However, in Panel C, which shows the return regressions, both IE measures drop in significance relative to the sample with 0 values inserted shown in Table 6.

For ln(1 + PAT), the results for the value regressions are consistent across samples (i.e., the coefficients are negative and significant in the majority of models). For ln(1+ CIT/PAT), the results flip across samples, and in the sample without 0 values inserted of Table 7, the coefficients on ln(1 + CIT/PAT) become positive and significant in all models (they were negative and significant in most models of Tables 3 and 4). Finally, in Panel C of Table 7, the return regressions, the coefficients on ln(1 + PAT) are qualitatively unchanged from Table 6 (only the univariate patent model coefficient is positive and significant). For the innovation measure from Cohen et al. (2013), RD_ABILITY interacted with R&D_INTENSITY, we see no change across samples (i.e., the coefficient is positive in a model without other innovation measures but loses significance in the models with RQ).

In addition to these sample robustness tests, we reproduce both our contemporaneous-MTB and future-MTB panel regressions from Tables 3 and 4, respectively, and the panel regressions without 0 values inserted reported in Panels A and B of Table 7 using Fama–MacBeth cross-sectional regressions with Newey–West standard errors with 3 lags.[28] We report these results in Table D.2 (contemporaneous-MTB regressions), Table D.3 (future-MTB regressions), and Table D.4 (versions of the MTB regressions without 0 values inserted) of the Supplementary Material. In these tests, the coefficients on RQ remain positive and significant across all specifications, with some minor variations on the coefficients for the other innovation measures. Overall, these new results are highly similar to the panel regressions.

To summarize, the sample robustness tests show that RQ is strongly robust across the 2 samples for all 3 types of firm-value tests.[29] In contrast, the results for

[28]These results are robust to variations in the number of lags used to construct the Newey–West standard errors.

[29]We also estimated sample robustness tests on models where we require that a given non-RQ innovation measure was populated and that RQ was nonmissing. We did this both for the sample with 0 values inserted and the sample with missing patents excluded. The results, available from the authors, are qualitatively similar to the main results in Tables 3, 4, 6, and 7. The only notable difference is that the coefficient on IE_PATENTS becomes insignificant in models where IE_PATENTS is the sole innovation measure.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

the other innovation measures exhibit varying degrees of sensitivity across samples. Thus, sample construction, and its influence on the innovation and firm-value relations, appears to be an important factor that researchers should carefully consider in innovation studies.

F. Using RQ in Determinants-of-Innovation Tests

Given that innovation drives firm value, it is not surprising that many prior articles have examined the determinants of innovation. Most of these studies use input (R&D) or output (patents and citations) measures to define innovation. For example, HT (2013) find that analyst coverage is negative and significant in explaining patent-based innovation. Atanassov (2013), studying the effects of takeovers on innovation, documents a negative relation between leverage and patent-based innovation.

In this section, we reexamine the innovation drivers from these 2 articles using RQ as the innovation measure. From our results thus far, RQ seems to capture different aspects of firm innovation than other established innovation measures, and thus it may be of interest to determine if drivers of other innovation measures also help determine RQ.

To implement these tests, we conduct a simple retest of the baseline analysis in HT (2013), which examines the impact of analyst coverage on innovation while controlling for the effects of leverage on innovation (both HT and Atanassov (2013) control for leverage).[30] We chose this retest because it is representative of important tests on the determinants of innovation and because the necessary data were available.

Panel A of Table 8 presents results from panel regressions of innovation measures (as the dependent variables) on analyst coverage, leverage, and other controls in a sample with missing firm-year patent values set to 0. Consistent with HT (2013), the models include firm fixed effects and year effects. The first 2 models reproduce the HT results from their Table 2. In model 1, the dependent variable is ln(1 + PAT); in model 2, it is ln(1 + CIT/PAT). In model 3, we replace the patent measures with RQ. In all 3 models, the dependent variables are lined up 3 years ahead of the analyst coverage (and the other independent variables), as in Table 2 of HT (2013). In models 4–6, we use the same dependent variables as in models 1–3, but now the innovation measures are 1 year ahead.

The results in models 1 and 2 of Panel A in Table 8, using ln(1 + PAT) and ln(1 + CIT/PAT) as the dependent variables, match the findings in HT (2013); the coefficients on analyst coverage in models 1 and 2 are negative and statistically significant. Thus, more analyst coverage corresponds to less innovation. However, for the RQ model in model 3, we see no evidence of a negative effect of analyst coverage; in this case, the coefficient on RQ is statistically insignificant.

For the 1-year-ahead innovation measures, only the coefficient on ln(1 + CIT/PAT) in model 5 of Panel A in Table 8 is negative and significant. For 1-year-ahead ln(1 + PAT) and RQ, the coefficients are insignificant. The logic underlying the negative effect of analyst coverage on output innovation in HT (2013) is that analyst

[30]Other articles, such as Becker-Blease (2011), include leverage as a control variable in their innovation regressions. These articles generally find a negative and significant coefficient on leverage in patent- and citation-based innovation regressions, consistent with Atanassov (2013) and HT (2013).

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

coverage exerts pressure on managers to meet short-term goals at the expense of the long-term projects typical of R&D. For the RQ measure, it would appear those managerial forces do not operate on RQ, or may take longer than 1–3 years to affect a firm's RQ. This represents an open and interesting question about the potential longer-term effects of analyst coverage on innovative efficiency.

Next, we examine the effects of leverage on innovation. Atanassov's (2013) study on the effects of hostile takeovers on innovation output finds negative and

TABLE 8

Panel Regressions of Innovation Measures on Firm Determinants

Table 8 reports panel regressions of innovation measures on lagged firm characteristics. Panel A includes models estimated using fixed effects (FE), and Panel B includes models estimated using a dynamic panel approach based on an Arellano and Bond (1991) estimator. In models 1, 2, and 3, the dependent variables are 3 years ahead of the independent variables. In models 4, 5, and 6, the dependent variables are 1 year ahead of the independent variables. The sample period is from 1993 to 2005. Missing firm-year values for patent-related measures are set to 0. RQ is the research quotient, defined as the firm-specific output elasticity (Knott (2008)). ln(1 + PAT) is the natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year. ln(1 + CIT/PAT) is the natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of the patents filed (and eventually granted) in a given year. R&D/ASSETS is research and development (R&D) expenditures scaled by total book assets. AGE is the number of years since the first appearance on Compustat. ROA is operating income before depreciation scaled by total book assets. PPE/ASSETS is property, plant, and equipment scaled by total book assets. LEVERAGE is the book value of debt divided by total book assets. CAPEX/ASSETS is capital expenditure scaled by total book assets. TOBIN_Q is the market value of equity plus the book value of assets minus the book value of equity minus balance-sheet-deferred taxes, divided by total book assets. HZ_INDEX is the firm's Kaplan and Zingales index measured at the end of fiscal year t, calculated as $-1.002 \times$ cash flow $+ 0.283 \times$ Q $+ 3.139 \times$ leverage $- 39.368 \times$ dividends $- 1.315 \times$ cash holdings. H_INDEX is the Herfindahl index of the 4-digit SIC industry where the firm belongs, measured at the end of the fiscal year. H_INDEX_SQR is the Herfindahl index squared. Detailed variable definitions can be found in Appendix B. t-statistics are reported in parentheses. In the Panel A regressions, we include firm and year FEs, with standard errors clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	$\ln(1+\text{PAT})_{t+3}$	$\ln\left(1+\frac{\text{CIT}}{\text{PAT}}\right)_{t+3}$	RQ_{t+3}	$\ln(1+\text{PAT})_{t+1}$	$\ln\left(1+\frac{\text{CIT}}{\text{PAT}}\right)_{t+1}$	RQ_{t+1}
	1	2	3	4	5	6
Panel A. FE Models						
ln(COVERAGE)	−0.050**	−0.095***	0.000	−0.026	−0.114***	0.001
	(−2.190)	(−3.514)	(0.059)	(−1.311)	(−4.496)	(0.245)
LEVERAGE	−0.248***	−0.231***	−0.023*	−0.362***	−0.357***	−0.030***
	(−4.821)	(−3.745)	(−1.735)	(−7.313)	(−6.082)	(−2.633)
ln(ASSETS)	0.015	−0.041**	0.005*	0.133***	0.071***	0.010***
	(0.996)	(−2.227)	(1.649)	(9.189)	(3.785)	(3.137)
R&D/ASSETS	0.247***	0.138	0.057**	0.349***	0.262**	0.099***
	(3.304)	(1.366)	(2.093)	(4.682)	(2.458)	(3.095)
ln(AGE)	0.351***	0.209***	−0.001	0.236***	0.135***	0.005
	(8.211)	(4.287)	(−0.161)	(6.447)	(3.037)	(0.432)
ROA	0.194***	0.151***	0.041***	0.084**	0.190***	0.050***
	(5.332)	(3.052)	(3.217)	(2.450)	(3.402)	(3.599)
PPE/ASSETS	0.387***	0.305***	−0.026	0.263***	0.288***	−0.060**
	(4.834)	(3.048)	(−0.876)	(3.751)	(3.044)	(−2.338)
CAPEX/ASSETS	−0.033	−0.011	−0.004	0.017	−0.099	0.030
	(−0.484)	(−0.124)	(−0.145)	(0.261)	(−1.012)	(0.878)
TOBIN_Q	0.009***	0.003	0.000	0.022***	0.020***	0.001
	(3.364)	(0.707)	(0.721)	(8.993)	(6.364)	(1.041)
HZ_INDEX	−0.000	−0.000	0.000	−0.000	−0.000	0.000
	(−0.204)	(−0.473)	(0.776)	(−1.525)	(−0.961)	(0.860)
H_INDEX	0.201	0.145	0.001	0.280*	−0.052	0.025
	(1.312)	(0.794)	(0.055)	(1.878)	(−0.272)	(0.977)
H_INDEX_SQR	−0.085	−0.001	0.001	−0.170	0.186	−0.015
	(−0.535)	(−0.005)	(0.036)	(−1.065)	(0.972)	(−0.535)
Constant	−0.590***	0.162	0.090***	−0.954***	−0.145	0.039
	(−4.353)	(1.020)	(3.150)	(−7.593)	(−0.940)	(1.143)
No. of obs.	36,679	34,282	11,659	43,215	40,504	13,582
R^2	0.797	0.591	0.746	0.801	0.599	0.742

(continued on next page)

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE 8 (continued)
Panel Regressions of Innovation Measures on Firm Determinants

	$\ln(1+\text{PAT})_{t+3}$	$\ln\left(1+\frac{\text{CIT}}{\text{PAT}}\right)_{t+3}$	RQ_{t+3}	$\ln(1+\text{PAT})_{t+1}$	$\ln\left(1+\frac{\text{CIT}}{\text{PAT}}\right)_{t+1}$	RQ_{t+1}
	1	2	3	4	5	6
Panel B. Dynamic Panel Models						
ln(COVERAGE)	−0.018	0.011	0.002	−0.027	0.001	−0.001
	(−0.605)	(0.315)	(0.551)	(−0.973)	(0.016)	(−0.285)
LEVERAGE	0.087	−0.031	0.005	0.013	0.128	0.002
	(1.236)	(−0.365)	(0.619)	(0.178)	(1.371)	(0.194)
ln(ASSETS)	−0.020	−0.016	0.003	−0.038	−0.018	0.002
	(−0.855)	(−0.612)	(0.964)	(−1.593)	(−0.556)	(0.597)
R&D/ASSETS	0.098	0.207	−0.021	−0.328***	−0.679***	−0.009
	(0.792)	(1.123)	(−1.246)	(−2.584)	(−2.862)	(−0.429)
ln(AGE)	0.172**	0.222**	−0.017	0.164**	0.063	−0.007
	(2.290)	(2.230)	(−1.420)	(1.979)	(0.655)	(−0.489)
ROA	0.071	0.069	0.003	0.051	0.044	−0.050***
	(1.243)	(0.827)	(0.298)	(0.848)	(0.434)	(−4.865)
PPE/ASSETS	0.218**	0.199*	−0.017	0.002	0.115	0.012
	(2.090)	(1.677)	(−0.966)	(0.020)	(0.875)	(0.713)
CAPEX/ASSETS	−0.092	−0.023	0.007	−0.093	−0.062	−0.006
	(−0.960)	(−0.187)	(0.325)	(−0.946)	(−0.404)	(−0.295)
TOBIN_Q	−0.007	0.003	0.001	0.011***	0.014**	0.001*
	(−1.274)	(0.602)	(0.781)	(2.715)	(2.057)	(1.646)
HZ_INDEX	−0.000***	−0.000***	−0.000	−0.000	−0.000***	0.000**
	(−4.577)	(−10.462)	(−0.794)	(−1.434)	(−3.112)	(2.376)
H_INDEX	0.172	0.454*	0.011	0.034	−0.056	0.015
	(1.033)	(1.736)	(0.708)	(0.180)	(−0.215)	(1.010)
H_INDEX_SQR	−0.144	−0.353	−0.012	0.028	0.013	−0.013
	(−0.908)	(−1.334)	(−0.772)	(0.162)	(0.052)	(−0.975)
Lag 1 of dependent variable	1.166***	0.497***	0.857***	1.213***	0.521***	1.002***
	(34.543)	(29.110)	(9.051)	(41.549)	(36.404)	(20.845)
Lag 2 of dependent variable	0.223***	0.239***	−0.228***	0.293***	0.314***	−0.225***
	(8.880)	(16.880)	(−4.024)	(15.276)	(22.335)	(−4.767)
Lag 3 of dependent variable	−0.032	0.064***	0.024	−0.003	0.126***	−0.007
	(−1.633)	(5.490)	(0.540)	(−0.151)	(10.100)	(−0.181)
Constant	−0.480**	−0.744***	0.050*	−0.598**	−1.009***	0.015
	(−2.349)	(−2.799)	(1.949)	(−2.531)	(−3.505)	(0.406)
AR1 (*p*-value)	0.000	0.000	0.000	0.000	0.000	0.000
AR2 (*p*-value)	0.241	0.868	0.881	0.726	0.590	0.194
AR3 (*p*-value)	0.823	0.115	0.489	0.082	0.636	0.934
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
R^2	21,223	19,173	5,906	26,796	24,453	7,421

statistically significant coefficients in regressions of patent-based innovation measures on various measures of leverage (as do HT (2013)). According to Atanassov, at least two mechanisms can generate a negative link between leverage and innovation. First, creditors are typically less tolerant of risky projects than shareholders and are thus more likely to exert pressures to reduce costly innovation. Second, leverage may reduce managerial flexibility (Graham and Harvey (2001)) and thus lower the firm's ability to engage in experimentation, creativity, and innovation.

We are able to replicate Atanassov's (2013) results. In Panel A of Table 8, using the patent measures in models 1 and 2 and models 4 and 5, we find that the estimated coefficients on leverage are all negative and statistically significant. Similarly, for RQ, the coefficients on leverage are also negative and significant in models 3 and 6. Thus, leverage does appear to exert a negative effect on all of the innovation measures we tested.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

The tests in this section are designed to replicate prior studies, so we follow the methodologies of those studies and use level values of innovation measures as the dependent variables. However, these regressions may be misspecified because of high autocorrelation of the dependent variables. For example, the first-order autocorrelation for patents is 0.92, that for citations 0.71, and that for RQ 0.87. Thus, the OLS assumption of noncorrelation in the regression residuals is likely violated, and accordingly, the standard errors of the coefficient estimates may be biased.[31] The fixed effects and clustered standard errors we used in these models may help reduce the autocorrelated error terms. However, we compute the autocorrelation of the residuals from each model in Panel A of Table 8 and find that residual terms are autocorrelated and statistically significant. For example, the residual from model 1 in Panel A has an autocorrelation of 0.58, the residual from model 2 has autocorrelation of 0.2, and the residual from model 3 has autocorrelation of 0.65. These numbers indicate that there is still significant autocorrelation in the regression residuals. Moreover, in these regressions, the lagged dependent variables should potentially be included as regressors. For example, RQ is generated from a rolling-window estimation. This overlapping estimation procedure necessarily makes the lagged RQ a predictor of RQ.

To address the autocorrelated errors and the overlapping-data issue, we estimate the models in Panel A of Table 8 using a dynamic-panel approach based on an Arellano–Bond estimator (Arellano and Bond (1991)). Similar approaches are used in Hoechle, Schmid, Walter, and Yermack (2012) and Wintoki, Linck, and Netter (2012). The Arellano–Bond estimator is a generalized method of moments estimator used to estimate dynamic panel data models, containing both level and first-difference equations, which helps to address serial-correlation problems in static panel data. We report the results in Panel B of Table 8. In this table, the coefficients on analyst coverage and leverage are insignificant across innovation models, suggesting that once the endogeneity in the dynamic panel is accounted for, the previously documented correlations between analyst coverage, leverage, and innovation are no longer robust. In the dynamic panel regressions, we include 3 lags of the dependent variables. The results are similar when we include different numbers of lags. We also test the serial correlation of the differenced error term from these models. The Arellano–Bond estimator requires the differenced error term to be first-order serially correlated, with no serial correlation in higher-ordered differences. This assumption is satisfied in our regression models. The second-order autoregression (AR2) and third-order autoregression (AR3) coefficients of the differenced error terms are insignificant in all models, except in model 4, the AR3 is marginally significant.

Overall, the results in this section on the determinants of innovation show that conclusions about the drivers of innovation change as a function of the type of innovation measure and the type of econometric model used. Researchers should carefully consider these variations when studying the determinants of innovation.

[31]These autocorrelations are for the sample with 0 values inserted for the patent and citation measures. For the sample without 0 values inserted, the autocorrelation for patents is 0.91, and for citations, it is 0.67.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

G. An Overview Discussion of RQ

In this section, we discuss how finance researchers working in the area of innovation might think about and use RQ. We do this by discussing RQ relative to previous innovation measures. In particular, we examine when RQ is important, what it captures that other measures do not, and when we might expect RQ to tell us something different (or something similar) relative to other innovation measures.

1. How Innovation Measures Relate to Firm Value

The simplest way to conceive of the link between innovation and firm value is through Romer's (1990) theory linking innovation to growth. Although Romer's theory pertains to the level of the economy, its basic structure is the production function, which is inherently a firm-level construct. Romer's final-goods production function has two elements: labor, L, and knowledge, A. Embedded within this function is an intermediate-goods production function for generating the knowledge. That knowledge-production function in turn has 2 inputs: the current stock of knowledge, A, and research labor, L_A. The main conclusion from Romer's model is that along a balanced growth path, final-goods output growth, g_Y, which matches the rate of knowledge growth, g_A, is a function of the level of research labor, L_A, and the productivity of that research labor, δ:

$$g_Y = g_A = \delta L_A. \tag{7}$$

Because firm value capitalizes growth, then firm value should also increase in the level of research labor, L_A, as well as the productivity of that research labor, δ, and the growth of the knowledge stock, g_A, but it should be independent of the knowledge stock itself, A.

Innovation measures seek to capture one or more of these constructs: research, L_A; knowledge, A; or research productivity. It is important to note, as we have mentioned earlier in the article, that each of these is a distinct innovation construct (e.g., input, output, and efficiency measures, respectively). The input constructs are typically captured in empirical articles by R&D. The output constructs are typically captured by patents or citations. The productivity measures can be captured by patents scaled by R&D (Hirshleifer et al. (2013)) or RQ, as in this article. Thus, the choice of measures depends on what aspect of innovation scholars are trying to capture.

2. When Is RQ Important?

RQ can be important in a number of instances. First and foremost, RQ is vital if a researcher needs an innovation measure that reliably maps onto firm value. As our results show, RQ is the only measure we test that is consistently positively related to contemporaneous and future firm value and to future returns. This result holds across OLS panel regressions, 2SLS tests, cross-sectional regressions, and variations in sample construction using observations with 0 values inserted and observations without 0 values inserted. We note that the consistency of the RQ results is especially strong relative to other innovation measures in the predictive cross-sectional return regressions of Table 6, a setting in which the issues related to identification are typically not a concern and an area within finance of immense

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

research activity. Thus, if a researcher desires to study innovation and its link to firm value, wishes to study the determinants of firm innovation using a value-relevant measure, or needs a robust IE control variable, RQ seems a prudent measure.

Second, RQ is important when researchers want greater coverage of innovating firms. As we discussed in Section II.B, 52% of the firms in our sample invest in R&D without ever patenting. This maps into greater RQ coverage relative to other innovation measures. For example, RQ covers approximately 19% more firm-years than patents, 33% more than citations scaled by patents, and 36% more than the IE_PATENTS measure of Hirshleifer et al. (2013).[32] For the firms in our sample that do not patent but have RQ data, our findings relating RQ to firm value are robust.[33] In addition, in markets where patent data are not easily obtainable (e.g., some international markets) or in the U.S. market before or after the period of the NBER patent database, RQ may serve as a straightforward-to-estimate innovation measure using only firm-level financial data.

3. When Might the Various Innovation Measures Tell Us Something Different?

Our results in Table 2 suggest unconditionally that RQ is not highly correlated with the other input-, output-, or efficiency-innovation measures (all of the correlations are close to 0). However, it is possible that conditionally, for some types of firms, RQ is correlated with other innovation measures or is correlated with a given firm characteristic in a different way than other innovation measures. In Section III.F, where we replicate previous articles examining the impact of changes in analyst coverage and leverage on innovation, we find some similarities (and differences) in innovation-measure responses to changes in those firm characteristics. Although a comprehensive study on the correlates of firm-level innovation remains beyond the scope of our article, this section asks, "When and how does the choice of innovation measures matter?" To answer the question, we review the existing literature on how RQ and other innovation measures relate to important firm characteristics.

Our review reveals 4 firm-level factors affecting company innovation for which there is empirical evidence for RQ and other innovation measures. RQ and other innovation measures have similar correlations for 2 of the factors but opposing effects for the other 2. We discuss all 4 factors.

a. Firm Size

Dating to Schumpeter (1942), the question of optimal firm size for innovation is one of the most examined in the economics and management literature. Arguments favoring large firms tend to rely on economies of scale: Larger scale allows firms to amortize the fixed cost per innovation over more units of output. Arguments favoring small firms typically rely on governance advantages, such as

[32] Using data from Appendix C, for patents, the calculation (using row 4 and row 6 firm-year data) is [(26,967 RQ firm-years/22,701 patent firm-years) – 1] × 100 = 18.79% greater coverage for RQ relative to patents. For citations scaled by patents, it is [(26,967 RQ firm-years/20,301citations/patents firm-years) – 1] × 100 = 32.83%, and for IE_PATENTS, it is [(26,967 RQ firm-years/19,856 IE_PATENTS – 1] × 100 = 35.81%.

[33] We estimate the firm-value regressions of Table 3 on the sample of firms with nonmissing RQ observations and missing patent observations. The coefficient on RQ is qualitatively similar to the full-sample Table 3 estimates and is highly statistically significant. The results are available from the authors.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

compensation systems that better align employee behavior with company goals. The consensuses from the literature, summarized in Cohen (2010), include the following: i) R&D rises proportionally with firm size, and ii) patent intensity (i.e., patents/R&D) declines with firm size. Thus, large firms appear to exhibit seemingly anomalous behavior: They invest more in R&D despite decreasing productivity. Knott and Vieregger (2020) propose that the anomaly stems from how one measures R&D productivity. In their Table 2, they replicate the firm-size puzzle: R&D spending increases with firm scale, whereas patent intensity decreases with firm scale. Their Table 3, however, shows that RQ increases with scale. The difference in results for the 2 productivity measures stems from the fact that RQ captures the returns to an innovation (which increase in the scale of the output market), whereas patent intensity merely captures the number of innovations. Summary: The correlation (firm size: RQ) is positive, and the correlation (firm size: patent intensity) is negative.

b. Radical Innovation

Radical innovations are new-to-the-world or new-to-the-market ones. Knott and Vieregger (2020) find that RQ decreases as self-reported "new to the market" innovations increase.[34] This decrease reflects that radical innovations require substantial market development in addition to technology development. As an example, the first home microwave was not sold until 20 years after Raytheon patented it. Other articles operationalize radical innovation through patents with high citation counts, diversity in patent citations, or the portion of a firm's patents outside of the firm's existing knowledge and expertise (Chen, Podolski, and Veeraraghavan (2015), Castaldi, Frenken, and Los (2015), among others). These articles find that radical innovation is associated with greater innovation output. Using a third approach to measuring radical innovation, Krieger, Li, and Papanikolaou (2017) develop a measure of a drug's radicalness using the novelty of its chemical composition. They find that more radical drugs have higher patent citations as well as higher firm value conditional on approval from the U.S. Food and Drug Administration (FDA). However, they also find that radical drugs are less likely to be approved by the FDA. Summary: The correlation (radical innovation: RQ) is negative, and the correlation (radical innovation: patents or citations) is positive.

c. Centralization

The first study of the impact of R&D_CENTRALIZATION on innovation was that by Argyres and Silverman (2004). They argued that centralization reduces the coordination costs across units and accordingly leads to innovations with larger and broader impact. Using data on 71 firms collected through an Industrial Research Institute (IRI) survey, they found weak support for their hypotheses. Their Table 4 indicates that centralized firms have more patent citations and that those patents come from a broader diversity of technologies. Knott (2017), who holds a similar view regarding the merits of centralization, believes that the benefits have less to do

[34]Knott and Vieregger (2020) capture radical innovation using a survey question from the NSF Business R&D and Innovation Survey. They classify the most radical innovations as goods/services that are new to the market.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

with coordination costs than they do with decision authority. In particular, division managers, whose compensation is tied to divisional profits, will prefer R&D projects that increase those profits. These projects may come at the expense of projects that might benefit multiple divisions, as well as projects with longer-term impact. Using data from Argyres and Silverman, Knott finds that companies with centralized R&D have 38% higher RQ than companies with decentralized R&D. Using data and a patent-based centralization measure from Arora et al. (2014), Knott finds that firms with centralized R&D have 64% higher RQ. Summary: The correlations (centralization: RQ) and (centralization: patents or citations) are both positive.

d. Outside CEOs

There are two potential views of the impact of externally hired CEOs on innovation. On the one hand, such CEOs might be more likely to introduce new ideas and therefore drive higher innovation. On the other hand, external managers might lack crucial firm-specific knowledge needed to manage the innovation process, which might result in lower innovation. The empirical evidence favors the latter view, whether using patent measures or RQ. Balsmeier and Buchwald (2015) find that the firms of externally hired CEOs file significantly fewer patent applications compared with the firms of their internally promoted colleagues (Tables 3 and 5), and they file significantly fewer citation-weighted patents (Tables 4 and 5). Cummings and Knott (2018) obtain similar results with RQ. They find that RQ decays during the tenure of outside CEOs relative to that of inside CEOs. They further find that this effect is more pronounced for firms with high R&D_INTENSITY and for firms employing outside CEOs with more remote experience, lending circumstantial support for the underlying assumption regarding lack of expertise. Summary: The correlations (outside CEOs: RQ) and (outside CEOs: patents or citations) are both negative.

Overall, innovation measures differ substantially. First, they differ in the innovation construct that they intend to capture: research labor, knowledge stock, or research productivity. However, even measures attempting to capture the same construct may differ in their correlation with one another and in how they relate to firm characteristics. It is important for scholars to understand all of this and to carefully match their measures to the underlying innovation construct they hope to capture. Otherwise, a result specific to a single innovation measure may be incorrectly interpreted as a broader innovation result.

IV. Conclusion

How to measure firm-level innovation and the drivers and consequences of innovation are important topics in the finance literature. This article has introduced a new-to-the-finance-literature innovation measure, RQ, defined as the firm-specific output elasticity of R&D. The measure was first developed in the management literature by Knott (2008). RQ measures a firm's ability to generate revenue growth from its R&D investments. The measure does not use patent information and thus provides coverage for a broader set of publicly traded firms relative to

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

measures that use patents. RQ also has a low correlation with other established innovation measures.

We test RQ in a number of common firm-value tests from the innovation literature. We show that firms with higher RQ have higher market valuations, both contemporaneously and in the subsequent year, controlling for other innovation measures and controlling for standard determinants of firm value. We also find, after controlling for other innovation measures and other determinants of the cross-section of returns, that higher-RQ firms earn higher future returns. Our results are robust to a variety of controls, variations in sample period and sample construction, and variations in testing methodologies. In contrast, many of the existing innovation measures obtain less consistent results across our battery of firm-value tests.

Overall, our results suggest that future researchers should carefully match the innovation measure in their empirics to the innovation construct of interest. If a researcher's desired goal is to identify a measure of firm-level innovation that reliably maps to firm-value outcomes, then, as our results show, RQ is an appropriate candidate based on its strong and consistent performance across a full spectrum of firm-value tests.

Appendix A. Details of RQ Estimation[35]

Researchers confront several issues when estimating production functions: what to use as the output measure; how to deal with a lack of data on materials; whether to use stocks or flows, and with what lag; how to construct spillover pools; how to adjust prices; and how to deal with endogeneity. We address each of these issues to explain the mechanics of our estimation of the RQ.

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

Materials: Compustat data suffers from a 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 the cost of goods sold (COGS; which includes materials) in lieu of capital and labor. Either version includes some double-counting (because R&D comprises labor, capital, and materials). This double-counting in turn yields downward bias in R&D elasticity, although 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 about 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

[35]An expanded discussion of RQ estimation is contained in the user's manual for the WRDS RQ database.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

and depreciation rates vary across firms even within the same industry. Prior studies indicate that depreciation rates vary between 15% and 19%, although industry estimates for the output elasticity of R&D appear to be insensitive to rates from 8% to 25% (Hall et al. (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. We do so under the assumption that firms are essentially in steady state; that is, 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. As a robustness check, we test this claim against models using R&D stocks and find no significant difference. This steady-state finding explains the empirical regularity of econometric equivalence between stock and flow models (Griliches and Mairesse (1984), Adams and Jaffe (1996)).

Because R&D generates an intermediate input (product/process design) that then combines with capital, labor, and materials to generate a final-goods output, a related question concerns the appropriate lag. Here, we follow convention and model R&D with a lag of 1 year. As a robustness check, we test alternative lags and find no significant difference. Again, our findings match prior results, indicating that 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 a steady state, so R&D does not vary much.

Spillover pools: Empirical tests of R&D productivity typically employ a pooled form of spillovers (Jaffe (1986), Adams and Jaffe (1996), and Bloom, Schankerman, and Van Reenen (2007)), which sums the R&D for all firms in industry j other than focal firm i (although the pool is adjusted for technical and geographic proximity). This pooled form is faithful to early input/output (IO) models of the impact of spillovers on incentives for R&D (Spence (1984), Levin and Reiss (1988)). 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 other work testing alternative functional forms of spillover pools (Knott et al. (2009)), we employ an asymmetric form that mimics endogenous growth models (Jovanovic and Rob (1989), Jovanovic and MacDonald (1994), and 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 multiproduct firms whose outputs cross many industries, the 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 the potential correlation between input levels and unobserved firm-specific shocks. In particular,

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

firms may respond to a positive productivity shock by increasing inputs. Failure to account for this will positively bias 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 (2020)) 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.

Estimation: The estimation of RQ in equation (2) is accomplished using a linear mixed model of the following form:

$$\mathbf{y} = \mathrm{X}\beta + \mathbf{Z}\mathrm{u} + \varepsilon,$$

where $\mathbf{y}$ is the $n \times 1$ vector of responses; $\mathbf{X}$ is an $n \times p$ matrix for the fixed effect, β; $\mathbf{Z}$ is the $n \times q$ matrix for the random effects $\mathbf{u}$; and $\boldsymbol{\epsilon}$ is the n × 1 vector of errors, which in our case we assume to be multivariate normal with a mean of 0 and the variance matrix $I_n\sigma_\epsilon^2$. We also assume that $\mathbf{u}$ has the variance-covariance matrix $\mathbf{G}$ and that $\mathbf{u}$ is orthogonal to $\boldsymbol{\epsilon}$, so that

$$\mathrm{Var}\begin{bmatrix}\mathbf{u}\\ \epsilon\end{bmatrix} = \begin{bmatrix}\mathbf{G} & 0\\ 0 & I_n\sigma_\epsilon^2\end{bmatrix}.$$

Viewing the combined error term as $\mathbf{Z}\mathrm{u}+\epsilon$, then y is multivariate normal with a mean of $\mathbf{X}\beta$ and an $n \times n$ variance-covariance matrix:

$$\mathbf{V}=\mathbf{ZGZ}' + I_n\sigma_\epsilon^2.$$

Defining θ as the vector of unique elements of G results in the following log likelihood:

$$L(\beta,\theta,\sigma_\epsilon^2) = -\frac{1}{2}\left\{n\log(2\pi) + \log|V| + (y-X\beta)'V^{-1}(y-X\beta)\right\},$$

which is maximized as a function of β, θ, and σ_ϵ^2. The random effects are not directly estimated; rather, they are predicted/computed from the estimated fixed effects and the variance-covariance matrix:

$$\begin{aligned}\mathrm{E}[\mathbf{u}|\mathbf{y}] &= \mathrm{E}[\mathbf{u}] + \mathrm{cov}[\mathbf{u},\mathbf{y}^{\mathrm{T}}]\mathrm{var}^{-1}[\mathbf{y}](\mathbf{y}\text{–}\mathrm{E}[\mathbf{y}])\\ &= \mathbf{GZ}^{\mathrm{T}}\left(\mathbf{ZGZ}^{\mathrm{T}} + \sum\right)^{-1}(\mathbf{y}\text{–}\mathbf{X}\boldsymbol{\beta}).\end{aligned}$$

The underlying "assumptions" are that $\boldsymbol{\epsilon}$ is multivariate normal with a mean of 0 and that the variance-covariance matrix, $\mathbf{G}$, is orthogonal to $\boldsymbol{\epsilon}$. These assumptions are imposed by the estimation process.

As an example of RQ estimation, results for the RC specification of the R&D production function (see equation (2) in the main text) 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 specification (model 2) and an OLS specification (model 3). Note that the coefficients for R&D are similar across the RC and fixed-effects models at 0.095 and 0.098, respectively, whereas 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 et al. (2010)).

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

TABLE A1

Summary Statistics

Model 1 in Table A1 reports estimation results for the random-coefficients specification of the research and development (R&D) production function (see equation (2) in the main text) for the year 2001. For comparison, we present results for a fixed-effects specification (model 2) and an OLS specification (model 3). *, **, and *** indicate significance at the 5%, 1%, and .1% levels, respectively.

	Random Coefficients	Fixed Effects	OLS
Dependent variable: ln(revenue)	1	2	3
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	−0.003***	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^2		0.907	0.912
R^2-within		0.484	
R^2-between		0.929	
Wald Chi2	10583		

Appendix B. Variable Definitions

HIGH_ABILITY: A dummy variable indicating that the firm is in the top quintile in terms of innovation ability. The innovation ability, defined as in Cohen, Diether, and Malloy (2013), is the average of 5 coefficients from rolling firm-by-firm regressions of firm-level sales growth (defined as $\ln(\text{SALES}_t/\text{SALES}_{t-1})$) on the natural logarithm of lagged R&D ($\ln(\text{R\&D}_{t-j}/\text{SALES}_{t-j})$, where $j = 1, 2, 3, 4, 5$) in the past 5 years.

IE_CITATIONS: IE based on citations. Defined as the adjusted number of citations received by patents granted in the past 5 years, scaled by the number of corresponding R&D expenses:

$$\text{IE}_t = \frac{\sum_{j=1}^{5}\sum_{k=1}^{N_{t-j}} C_{i,k}^{t-j}}{\text{R\&D}_{i,t-3} + \text{R\&D}_{t,t-4} + \text{R\&D}_{i,t-5} + \text{R\&D}_{i,t-6} + \text{R\&D}_{i,t-7}}.$$

IE_PATENTS: IE based on patents. Defined as the firm's patents granted in a given year scaled by R&D capital:

$$\text{IE}_t = \frac{\text{PATENTS}_{i,t}}{\left(\text{R\&D}_{i,t-2} + 0.8\text{R\&D}_{i,t-3} + 0.6\text{R\&D}_{i,t-4} + 0.4\text{R\&D}_{i,t-5} + 0.2\text{R\&D}_{i,t-6}\right)}.$$

ln(1 + CIT/PAT): Natural logarithm of 1 plus the firm's total number of non-self citations received on the firm's patents filed (and eventually granted), scaled by the number of patents filed (and eventually granted) in a given year.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

ln(1 + PAT): Natural logarithm of 1 plus the firm's total number of patents filed and eventually granted in a given year.

R&D_INTENSITY: Natural logarithm of (1 + R&D/REVENUE).

RQ: Research quotient. Firm-specific output elasticity of R&D (Knott (2008)).

Firm Characteristics

ABNORMAL_EARNINGS: Calculated as $\left(E^B_{i,t}(1-\tau_{i,t}) - r_t \text{BE}_{i,t-1}\right)/\text{BE}_{i,t}$, where $E^B_{i,t}$ is earnings after extraordinary items before expensing R&D expense and less preferred dividends from firm i in year t, $\tau_{i,t}$ is the tax rate, and r_t is the 1-year Treasury bill rate.

ADVERTISING_ME: Firm's advertising expense scaled by market equity.

AGE: Number of years since the first appearance on Compustat.

BE: Book value of equity, calculated as stockholders' equity plus balance-sheet-deferred taxes and investment tax credit, minus book value of preferred stock.

CAPEX/ASSETS: Capital expenditure scaled by total book assets.

LEVERAGE: Book value of debt divided by total book assets.

MTB: Market value of equity divided by book value of equity for the fiscal year ending in calendar year *t*. The market value of equity is calculated as the number of shares outstanding × share price in December of year *t*.

PPE/ASSETS: Property, plant, and equipment scaled by total book assets.

R&D/ASSETS: R&D expenditures scaled by total book assets.

RD/BE: R&D expenditures scaled by total book equity.

R&D_EXPENSE: R&D expenditures.

REVENUE: Total revenue.

TOBIN_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.

TOTAL_ASSETS: Total book assets.

Operating Performance

CASH_FLOW: Operating income before depreciation minus interest expense, minus taxes, minus preferred dividends, minus common dividends and divided by total book assets.

OPERATING_PROFIT: Revenues minus COGS; minus selling, general, and administrative expenses; minus interest expense, all divided by book equity.

ROA: Operating income before depreciation scaled by total book assets.

Others

ANALYST_COVERAGE: Rolling average of the last 12 months of the number of monthly earnings forecasts from Institutional Brokers' Estimate System (IBES) summary file.

ASSET_GROWTH: One-year percentage change in total firm assets.

BHRET6: Compounded buy and hold return of the prior-6-month returns from month $t-7$ to $t-2$ for month *t*.

H_INDEX: Herfindahl index of 4-digit SIC industry where firm belongs, measured at the end of the fiscal year.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

HZ_INDEX: Firm's Kaplan and Zingales index measured at the end of fiscal year *t*, calculated as $-1.002 \times$ cash flow $+ 0.283 \times$ Q $+ 3.139 \times$ leverage $- 39.368 \times$ dividends $- 1.315 \times$ cash holdings.

ILLIQUIDITY: Amihud (2002) 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_CENTRALIZATION: The industry median of the firm's R&D_CENTRALIZATION, the sum of the squared proportions of patent inventors in the top 5 of the firm's most frequent metropolitan areas.

R&D_TAX_SHIELD: R&D expenditure multiplied by the tax rate, scaled by book equity. The tax rate is defined as the tax expenses divided by the pretax book income.

Appendix C. Observation Counts

Conditions	Firms- Years	Firms	% Firm -Years[+]	% Firms[+]
1. CRSP–Compustat merged,* 1981–2015	132,309	14,760		
2. CRSP–Compustat merged,* 1981–2006	106,396	13,322		
3. (2) with positive R&D, 1981–2006	48,906	6,499	46%	49%
4. (2) with nonmissing patents (granted patents only)	22,701	3,901	21%	29%
5. (2) Conditional on positive R&D, with nonmissing patents	19,036	3,095	18%	23%
6. (2) with positive RQ, 1981–2006	26,967	3,526	25%	26%
7. (1) with positive RQ, 1981–2015	36,141	3,992	27%	27%
8. (2) with nonmissing IE_PATENTS	19,856	3,165	19%	24%
9. (2) with nonmissing IE_CITATIONS	19,790	2,787	19%	21%
10. (2) with nonmissing ln(1 + PAT)	22,701	3,901	21%	29%
11. (2) with nonmissing ln(1 + CIT/PAT)	20,301	3,659	19%	27%

* Excluding first 2 years in Compustat, excluding financial industry.
[+] Computed relative to row 2.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109021000417.

References

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

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

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

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

Arellano, M., and S. Bond. "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *Review of Economic Studies*, 58 (1991), 277–297.

Argyres, N., and B. Silverman. "R&D, Organization Structure, and the Development of Corporate Technological Knowledge." *Strategic Management Journal*, 25 (2004), 929–958.

Arora, A.; S. Belenzon; and L. Rios. "The Organization of R&D in American Corporations: The Determinants and Consequences of Decentralization." NBER Working Paper No. 17013 (2011).

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

Arora, A.; S. Belenzon; and L. Rios. "Make, Buy, Organize: The Interplay between Research, External Knowledge, and Company Structure." *Strategic Management Journal*, 35 (2014), 317–337.
Atanassov, J. "Do Hostile Takeovers Stifle Innovation? Evidence from Antitakeover Legislation and Corporate Patenting." *Journal of Finance*, 68 (2013), 1097–1131.
Balsmeier, B., and A. Buchwald. "Who Promotes More Innovations? Inside Versus Outside Hired CEOs." *Industrial and Corporate Change*, 24 (2015), 1013–1045.
Becker-Blease, J. R. "Governance and Innovation." *Journal of Corporate Finance*, 17 (2011), 947–958.
Bellstam, G.; S. Bhagat; and J. A. Cookson. "A Text-Based Analysis of Corporate Innovation." Working Paper, University of Colorado (2017).
Bloom, N.; M. Schankerman; and J. Van Reenan. "Identifying Technology Spillovers and Product Market Rivalry." *Econometrica*, 81 (2007), 1347–1393.
Boroush, M., and J. Jankowski. "Update on U.S. Business Innovation: Findings from 2011 Survey." *National Center for Science and Engineering Statistics Info Brief NSF* 16–308 (2016).
Castaldi, C.; K. Frenken; and B. Los. "Related Variety, Unrelated Variety and Technological Breakthroughs: An Analysis of US State-Level Patenting." *Regional Studies*, 49 (2015), 767–781.
Chan, L. K. C.; J. Lakonishok; and T. Sougiannis. "The Stock Market Valuation of Research and Development Expenditures." *Journal of Finance*, 56 (2001), 2431–2456.
Chen, Y.; E. Podolski; and M. Veeraraghavan. "Does Managerial Ability Facilitate Corporate Innovative Success?" *Journal of Empirical Finance*, 34 (2015), 313–326.
Cohen, L.; K. Diether; and C. Malloy. "Misvaluing Innovation." *Review of Financial Studies*, 26 (2013), 635–666.
Cohen, W. M. "Fifty Years of Empirical Studies of Innovative Activity and Performance." In *Handbook of the Economics of Innovation*, E. H. Hall and N. Rosenberg, eds. Amsterdam, Netherlands: North-Holland Press (2010).
Cohen, W. M.; R. R. Nelson; and J. P. Walsh. "Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not)." NBER Working Paper 7552 (2000).
Cummings, T. "Microfoundations of Innovation in Organizations." PhD Dissertation, Washington University in St. Louis (2019).
Cummings, T., and A. M. Knott. "Outside CEOs and Innovation." *Strategic Management Journal*, 39 (2018), 2095–2119.
Cuneo, P., and J. Mairesse. "Productivity and R&D at the Firm Level." In *French Manufacturing, R&D, Patents, and Productivity*, Z. Griliches, ed. Chicago: University of Chicago Press (1984).
Eberhart, A. C.; W. F Maxwell; and A. R. Siddique. "A Reexamination of the Tradeoff between the Future Benefit and Riskiness of R&D Increases." *Journal of Accounting Research*, 46 (2008), 27–52.
Eeckhout, J., and B. Jovanovic. "Knowledge Spillovers and Inequality." *American Economic Review*, 92 (2002), 1290–1307.
Fama, E., and K. French. "The Cross-Section of Expected Stock Returns." *Journal of Finance*, 47 (1992), 427–465.
Fama, E., and J. D. MacBeth. "Risk, Return, and Equilibrium: Empirical Tests." *Journal of Political Economy*, 81 (1973), 607–636.
Fang, V.; X. Tian; and S. Tice. "Does Stock Liquidity Enhance or Impede Firm Innovation?" *Journal of Finance*, 69 (2014), 2085–2125.
Galasso, A., and T. Simcoe. "CEO Overconfidence and Innovation." *Management Science*, 57 (2011), 1469–1484.
Gompers, P.; J. Ishii; and A. Metrick. "Corporate Governance and Equity Prices." *Quarterly Journal of Economics*, 118 (2003), 107–156.
Gormley, T., and D. Matsa. "Common Errors: How to (and Not to) Control for Unobserved Heterogeneity." *Review of Financial Studies*, 27 (2014), 617–661.
Gormley, T.; D. Matsa; and T. Milbourn. "CEO Compensation and Corporate Risk: Evidence from a Natural Experiment." *Journal of Accounting and Economics*, 56 (2013), 79–101.
Graham, J. R., and C. R. Harvey. "The Theory and Practice of Corporate Finance: Evidence from the Field." *Journal of Financial Economics*, 60 (2001), 187–243.
Grieser, William D.; C. Hadlock; and J. Charles. "Panel-Data Estimation in Finance: Testable Assumptions and Parameter (In)Consistency." *Journal of Financial and Quantitative Analysis*, 54 (2019), 1–29.
Griliches, Z., and J. Mairesse. "Productivity and R&D at the Firm Level." In *French Manufacturing, R&D, Patents, and Productivity*, Z. Griliches, ed. Chicago: University of Chicago Press (1984).
Gu, F., and W. Wang. "Intangible Assets, Information Complexity, and Analysts' Earnings Forecasts." *Journal of Business Finance and Accounting*, 32 (2005), 1673–1702.
Hall, B.; A. Jaffe; and M. Trajtenberg. "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools." NBER Working Paper No. 8498 (2001).

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

Hall, B. H.; A. Jaffe; and M. Trajtenberg. "Market Value and Patent Citations." *Rand Journal of Economics*, 36 (2005), 16–38.
Hall, B.; J. Mairesse; and P. Mohnen. "Measuring the Returns to R&D." In *Handbook of the Economics of Innovation*, B. H. Hall and N. Rosenberg, eds. Amsterdam, Netherlands: North-Holland Press (2010).
Harvey, Campbell R.; Yan Liu; and Heqing Zhu. "… and the Cross-Section of Expected Returns." *Review of Financial Studies*, 29 (2016), 5–68.
He, J., and X. Tian. "The Dark Side of Analyst Coverage: The Case of Innovation." *Journal of Financial Economics*, 109, (2013), 856–878.
Heath, D.; M. Ringgenberg; M. Samadi; and I. Werner. "Reusing Natural Experiments." Working Paper, Fisher College of Business (2019).
Hirshleifer, D.; P. Hsu; and D. Li. "Innovative Efficiency and Stock Returns." *Journal of Financial Economics*, 107 (2013), 632–654.
Hirshleifer, D.; P. Hsu; and D. Li. "Innovative Originality, Profitability, and Stock Returns." *Review of Financial Studies*, 31 (2018), 2553–2605.
Hirshleifer, D.; A. Low; and S. Teoh. "Are Overconfident CEOs Better Innovators?" *Journal of Finance* 67, (2012), 1457–1498.
Hoechle, D.; M. Schmid; I. Walter; and D. Yermack. "How Much of the Diversification Discount Can Be Explained by Poor Corporate Governance?" *Journal of Financial Economics*, 103 (2012), 41–60.
Hombert, J., and A. Matray. "Can Innovation Help U.S. Manufacturing Firms Escape Import Competition from China?" *Journal of Finance*, 73 (2018), 2003–2039.
Jaffe, A. "Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits, and Market Value." *American Economic Review*, 76 (1986), 984–1001.
Jiang, W. "Have Instrumental Variables Brought Us Closer to the Truth?" *Review of Corporate Financial Studies*, 6 (2017), 127–140.
Jovanovic, B., and G. MacDonald. "Competitive Diffusion." *Journal of Political Economy*, 102 (1994), 24–52.
Jovanovic, B., and R. Rob. "The Growth and Diffusion of Knowledge." *Review of Economic Studies*, 56 (1989), 569–582.
Knott, A. M. "R&D/Returns Causality: Absorptive Capacity or Organizational IQ." *Management Science*, 54 (2008), 2054–2067.
Knott, A. M. "The Record on R&D Outsourcing." U.S. Census Bureau, Center for Economic Studies Working Paper No. CES-WP-16-19 (2019).
Knott, A. M. *How Innovation Really Works: Using the Trillion-Dollar R&D Fix to Drive Growth*. New York, NY: McGraw-Hill (2017).
Knott, A. M.; H. E. Posen; and B. Wu. "Spillover Asymmetry and Why It Matters." *Management Science*, 55 (2009), 373–388.
Knott, A. M., and C. Vieregger. "Reconciling the Firm Size and Innovation Puzzle." *Organization Science*, 31 (2020), 245–534.
Kogan, L.; D. Papanikolaou; A. Seru; and N. Stoffman. "Technological Innovation, Resource Allocation, and Growth." *Quarterly Journal of Economics*, 132 (2017), 665–712.
Koh, P., and D. Reeb. "Missing R&D." *Journal of Accounting and Economics*, 60 (2015), 73–94.
Koh, P.; D. Reeb; E. Sojli; and W. Tham. "Measuring Innovation Around the World." Working Paper, available at https://www.law.northwestern.edu/research-faculty/clbe/events/innovation/documents/sojli_measuring_innovation_around_the_world.pdf (2016).
Krieger, J.; D. Li; and D. Papanikolaou. "Developing Novel Drugs." SSRN Working Paper, available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3095246 (2017).
Lerner, J., and A. Seru. "The Use and Misuse of Patent Data: Issues for Corporate Finance and Beyond." Working Paper, Harvard Business School (2017).
Lerner, J., and J. Wulf. "Innovation and Incentives: Evidence from Corporate R&D." *Review of Economics and Statistics*, 89 (2007), 634–644.
Lev, B.; B. Sarath; and T. Sougiannis. "R&D Reporting Biases and Their Consequences." *Contemporary Accounting Research*, 22 (2005), 977–1026.
Lev, B., and T. Sougiannis. "The Capitalization, Amortization, and Value-Relevance of R&D." *Journal of Accounting and Economics*, 21 (1996), 107–138.
Levin, R. C., and P. C. Reiss. "Cost-Reducing and Demand-Creating R&D with Spillovers." *RAND Journal of Economics*, 19 (1988), 538–556.
Mukherjee, A.; M. Singh; and A. Žaldokas. "Do Corporate Taxes Hinder Innovation?" *Journal of Financial Economics*, 124 (2017), 195–221.
Newey, W., and K. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press

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

Pandit, S.; C. Wasley; and T. Zach. "The Effect of R&D Inputs and Outputs on the Relation between the Uncertainty of Future Operating Performance and R&D Expenditures." *Journal of Accounting, Auditing, and Finance* 26 (2011), 121–144.

Penman, S. H., and X.-J. Zhang. "Accounting Conservatism, the Quality of Earnings, and Stock Returns." *Accounting Review*, 77 (2002), 237–264.

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

Scherer, F. M., and D. Harhoff. "Technology Policy for a World of Skew-Distributed Outcomes." *Research Policy*, 29 (2000), 559–566.

Schumpeter, J. A. *Capitalism, Socialism, and Democracy*. Urbana, IL: University of Illinois Press (1942).

Sougiannis, T. "The Accounting Based Valuation of Corporate R&D." *Accounting Review*, 69 (1994), 44–68.

Spence, A. M. "Cost Reduction, Competition, and Industry Performance." *Econometrica*, 52 (1984), 101–121.

Stock J., and M. Yogo. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models*, D.Andrews and J. Stock, eds. New York, NY: Cambridge University Press (2005).

Wintoki, M. B.; J. S. Linck; and J. M. Netter. "Endogeneity and the Dynamics of Internal Corporate Governance." *Journal of Financial Economics*, 105 (2012), 581–606.

Zhou, X. "Understanding the Determinants of Managerial Ownership and the Link between Ownership and Performance: Comment." *Journal of Financial Economics*, 62 (2001), 559–5571.

https://doi.org/10.1017/S0022109021000417 Published online by Cambridge University Press