Journal of Accounting and Economics 60 (2015) 73–94

Contents lists available at ScienceDirect

Journal of Accounting and Economics

journal homepage: www.elsevier.com/locate/jae

Missing R&D ☆

Ping-Sheng Koh [a], David M. Reeb [b,*]

[a] Hong Kong University of Science and Technology, Hong Kong
[b] National University of Singapore, ABFER, Singapore

ARTICLE INFO

Article history:
Received 15 November 2012
Received in revised form
25 March 2015
Accepted 25 March 2015
Available online 3 April 2015

JEL classification:
M41
O32
D83

Keywords:
Research and Development
R&D reporting
Disclosure
Blank R&D
R&D simulation

ABSTRACT

We investigate whether missing R&D expenditures in financial statements indicates a lack of innovation activity. Patent records reveal that 10.5% of missing R&D firms file and receive patents, which is 14 times greater than zero R&D firms. Pseudo-Blank R&D firms (missing R&D firms with patent activity) demonstrate patent filings analogous to the bottom 90–95% of the positive R&D population. Multivariate difference-in-differences tests indicate that Pseudo-Blank R&D firms are more likely to report R&D after an exogenous auditor change. Finally, we provide simple Monte Carlo simulations to evaluate different methods to handle missing R&D in empirical research.

1. Introduction

Over the past 50 years, both the popular and academic literatures have come to consider corporate research and development (R&D) as a primary measure of innovation and technological progress in the firm (Lerner and Wulf, 2007). Bushee (1998), for instance, suggests that evaluating investments in corporate R&D represents a critical component of evaluating a firm's long-run value. Others emphasize that R&D expenditures, while more volatile than capital expenditures, provide tangible long-term benefits to the firm (Kothari et al., 2002). Yet, the notion of what outlays are considered R&D, especially outside of dedicated R&D units, can be difficult to assess, and often represents the manager's discretionary choice (Horwitz and Kolodny, 1980). A perusal of a subsample of the 3000+ NYSE-listed firms in our sample shows that a substantial number fail to provide any information regarding their corporate R&D efforts. Specifically, 1,737 NYSE-listed firms do not report any information on R&D, while 373 of them report zero R&D. The US *Statement of Financial Accounting Standards No. 2 (SFAS2)* requires a firm to disclose in its financial statements material R&D expenditures, suggesting these

☆ We would like to thank Sudipta Basu, Betsy Gordon, Robert Holthausen (editor), Jayanthi Krishnan, Alastair Lawrence, Baruch Lev, Juhani Linnainmaa, Amore Mario, Mihir Mehta, Ram Mudambi, Nagpurnanand Prabhala, Eric Press, Florin Sabac, Srini Sankaraguruswamy, John Wald, Sheryl Winston-Smith, Giovanni Valentini, Bernard Yeung, an anonymous reviewer and workshop participants at National University of Singapore, Singapore Scholars Symposium, Temple University, and the University of Alberta for helpful suggestions. All remaining errors are the sole responsibility of the authors. David Reeb acknowledges the support of the Ministry of Educations Tier 1 Grant (R-315-000-095-133).

* Corresponding author.
E-mail addresses: ackoh@ust.hk (P.-S. Koh), dmreeb@nus.edu.sg (D.M. Reeb).

http://dx.doi.org/10.1016/j.jacceco.2015.03.004

blank or empty R&D fields occur in firms with zero or immaterial corporate R&D. Yet, another interpretation of blank R&D values centers on firms choosing not to disclose this information in their accounting reports. In this context, a blank R&D field could represent a firm's conscious decision not to separate R&D expenses from other reported expenses, such as expense shifting (e.g. McVay, 2006).

We investigate whether missing R&D expenditures in financial statements indicates a lack of innovation activity. Empirical research in business often interprets the blank R&D fields using different assumptions, with research in accounting typically recoding the blank R&D fields as firms with zero R&D. A review of recent articles in prominent association journals in accounting, finance, and strategic management is illustrative of the divergent interpretations across the business disciplines regarding these blank R&D values.[1] For instance, 19% of the studies in *The Accounting Review* use R&D in their analysis and code the missing values as zero, implicitly assuming that blank R&D is equal to zero R&D. Similarly, 30% of the articles in the *Journal of Finance* use R&D, coding the blank values as zero and including a dummy variable to indicate blank R&D firms. Tacitly, this approach also suggests that blank R&D represents zero or negligible R&D. In contrast, 42% of the studies in the *Strategic Management Journal* use R&D and they take a very different approach, often replacing the missing R&D values with either the industry average R&D, or a historical value from prior years. The *SMJ* approach indicates that management scholars implicitly view the omission as a disclosure event and that these blank R&D firms engage in substantive R&D.

To investigate innovation in blank R&D firms we rely on a standard output of corporate R&D, namely patents. We use the universe of non-financial/regulated utility Compustat firms from 1980 to 2006 to identify firms that fail to report any information on their corporate R&D. We investigate innovation in blank R&D firms by focusing on patent applications and grants. Hall et al. (2005) observe that patents commonly occur with corporate R&D and suggest they provide an estimate of the intensity of the R&D process. In contrast, firms that do not engage in R&D are often found to have few, if any, patent applications (Kortum, 1993). Moreover, prior research on patent incentives suggests that firms often avoid filing patents, because filing informs competitors of their exact development process (Lerner, 2002). Arundel and Kabla (1998) estimate that less than 40% of firms file patents for their technological breakthroughs. Empirically, among the universe of Compustat firms, less than half of the positive R&D firms file or receive patents. Intuitively, the incentives to under-report R&D seem potentially related to the decision to file or not to file patents, suggesting we will find limited patents in firms without any reported R&D. Consequently, firms with missing R&D and without patents could carry out considerable R&D but remain undetectable in our tests because they forego patent protection. In this context, tests involving patents provide an especially conservative test environment for examining firms that do not report corporate R&D. Notably, we find that 10.5% of non-reporting R&D firms receive patents, with several of these firms receiving dozens of patents each year.

Our first set of tests compare patent activity between non-reporting R&D firms and firms that report zero R&D. One potential concern arises from differing industry rates of patentability from base levels of R&D. Consequently, we use both full sample and propensity score matched samples. We match on several observable firm characteristics that prior studies suggest influence corporate disclosures and R&D policies (e.g. Healy and Palepu, 2001; Andersen et al., 2009). While full sample tests allow for greater external validity, the matched sample tests potentially improve the local treatment effect (improved accuracy). Using the full and propensity score matched samples we find significant differences in patent activity between zero and non-reporting R&D firms. Based on the propensity score matched sample, our multivariate tests indicate that, on average, non-reporting R&D firms file about 14 times more patent applications than the matched zero R&D firms. In addition, non-reporting R&D firms obtain significantly more patent approvals than zero R&D firms. Thus, on average, non-reporting R&D firms apply for and are granted more patents than similar firms that report zero R&D.[2] Patent incentives research (e.g. Gallini, 2002) and the fact that the majority of positive R&D firms do not seek patents suggest our evidence provides a lower bound estimate of R&D activity among non-reporting firms.

Although, non-reporting R&D firms file substantial numbers of patents, these patents could represent inconsequential improvements. Research in economics and management highlights that patents, like published research articles, differ in their incremental contribution. If these non-reporting R&D firms file patents with negligible influence, then we expect to find they receive patents with limited impact. In addition to lifetime patent citations, two of the most discussed measures of patent influence are "Generality" and "Originality" (Moser and Nicholas, 2004). Patent "Generality" captures the spread of future patent citations across differing subgroups (Hall et al., 2001). The analogy in accounting research is whether an article is only cited in accounting journals or is also cited in economics, finance, and management journals. Originality in contrast refers to the heterogeneity in the references to other fields included in a patent (backward looking). Using accounting research as an analogy again, does the article strictly draw on prior accounting research, or does it also incorporate research from other disciplines and fields? A patent/study that only cites prior work in its own subfield results in a lower originality

[1] We choose the association journals as being representative of each discipline without having to focus on identifying a set of top journals. Every article in 2011 was assessed to determine if it used R&D and how the authors coded or dealt with the blank R&D issue. A perusal of other top journals in accounting, finance, and management showed similar patterns. Studies in economics often exclude these blank firms from the study, or seek to fill in the blanks (Hall and Oriani, 2006).

[2] Focusing on those firms that have blank R&D fields for at least 20 years, we examine their patent activity, finding that many of them have substantial numbers of patents, year in and year out. For instance, MCI and Coca-Cola have not disclosed corporate R&D in at least 20+ years but have each filed over 500 patent applications during this same time frame. Among the firms that do not report any information on R&D for 20+ years of our sample, we find that 46% of them have applied for or received patents during this period.

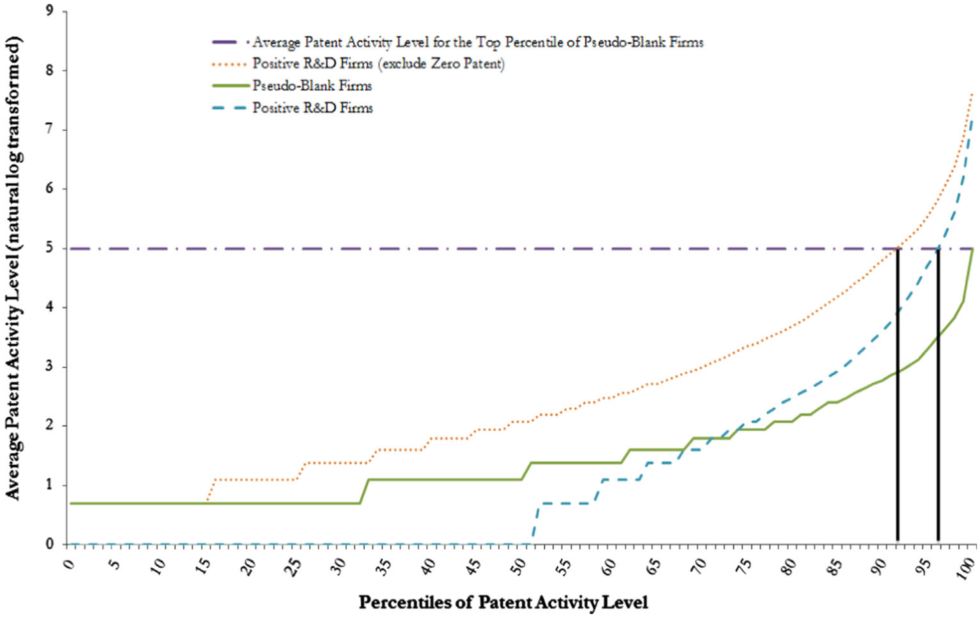

Fig. 1. Average patent activity level for Pseudo-Blank and positive R&D firms.

score. A fourth citation-based measure of patent influence centers on citation lag that captures competitor discovery time (Arundel, 2001; Hall et al., 2001). Intuitively, valuable innovations lead firms to pursue secrecy and seek lead-time advantages (Cohen et al., 2000). Beyond these citation measures, the patent itself can be used to assess the patentees' and patent examiners' assessment of contributions because they must explicitly declare the number of claimed developments. A patent with more claims engenders greater property rights, bears greater litigation risk, and attains greater market value (Lanjouw and Schankerman, 2004).

Using the full and propensity score matched samples we find significant differences in patent characteristics between zero and non-reporting R&D firms. Based on the propensity score matched sample, our multivariate tests indicate that, on average, non-reporting R&D firms have patents that are over 27% more influential than the matched zero R&D firms. The patents of non-reporting R&D firms also exhibit 25.9% longer competitor discovery periods than those found in the zero R&D firms. All told, the lifetime citations, citation breadth, declared contribution, and competitor discovery rate tests are inconsistent with the notion that the substantially greater number of patents filed by non-reporting R&D firms, relative to zero R&D firms, stem from them filing more trivial patents.

Conceptually, these patent results imply that many firms that fail to report R&D engage in innovation and R&D activities. In our second set of tests, we compare the patent characteristics of these missing R&D firms with patent data and those that report positive R&D. We label non-reporting R&D firms with patent activity as “Pseudo-Blank” R&D firms. Our analysis focuses on two separate sets of comparisons. First, we compare patents in Pseudo-Blank firms to the average patent profile across all positive R&D firms (both matched and full sample) in order to determine if Pseudo-Blank firms engage in inconsequential research and development. Intuitively, comparing the number and influence of patents in Pseudo-Blank firms to the group of positive R&D firms provides indirect evidence on the materiality of their R&D. Second, we compare Pseudo-Blank R&D firms to positive R&D firms with patent activity. This second comparison provides insights on the relative number and importance of individual patents in positive and Pseudo-Blank R&D firms.

We find that positive reporting R&D firms receive substantially more patents than Pseudo-Blank firms. Further analysis reveals that these differences in patents between Pseudo-Blank and positive reporting R&D firms stem from the high numbers of patents obtained by the top 5% of positive R&D firms. Fig. 1 graphically shows the relative number of patents in positive R&D and Pseudo-Blank firms. The patent activity profiles of Pseudo-Blank firms visibly intersect those of positive R&D firms, with Pseudo-Blank receiving more patents in the bottom three-quarters of the distribution and positive R&D firms prevailing in the top quartile (due in part to the numerous positive R&D firms that do not seek patents). The top 5% of positive R&D firms receive substantially more patents than the Pseudo-Blank R&D firms. To gauge the relative number of individual patents, we compare Pseudo-Blank R&D firms to positive R&D firms with patent activity, finding that Pseudo-Blank R&D firms correspond to the bottom 90th percentile of this subcategory. In sum, based on the number of patents, Pseudo-Blank R&D firms engage in less R&D than positive R&D firms, analogous to the bottom 90–95% of the positive R&D population.

Although, non-reporting R&D firms file more influential patents than zero R&D firms, the Pseudo-Blank designation could still capture negligible R&D. Consequently, we compare patent influence in Pseudo-Blank and positive reporting R&D firms. In the comparison with the average positive R&D firm, we find that in each of the patent influence measures noted above that Pseudo-Blank R&D firms receive more influential patents than the average matched sample of positive R&D firms. Of course about half of the positive R&D firms do not seek any patents. While this test provides evidence inconsistent with the hypothesis that Pseudo-Blank R&D firms have inconsequential R&D, it provides limited insight on the relative importance of individual patents in positive reporting and Pseudo-Blank R&D firms. Consequently, we repeat the analysis but exclude all positive reporting R&D firms without patent activity. We find that the average individual patent in positive reporting R&D firms exhibits statistically significant greater patent characteristics relative to individual patents in the Pseudo-Blank R&D firms. Specifically, the individual patents of positive reporting R&D firms exhibit about an 8% greater influence across the various patent characteristic measures relative to Pseudo-Blank R&D firms. Taken together, the patent profiles, citation breadth, declared contribution, and competitor discovery rate tests are inconsistent with the hypothesis that Pseudo-Blank R&D firms have trivial patents.

Engaging in R&D-type activities without reporting R&D could stem from either discretionary reporting choices or from the structure of corporate activities. Our third set of tests center on exploring the extent of the non-reporting decision in these firms. To investigate the potential for discretionary reporting choices, we rely on an exogenous shock – namely, the collapse of Arthur Anderson and the subsequent auditor change forced upon their clients. Under the standard view that blank R&D represents zero or negligible R&D, an exogenous change in external auditor should be unrelated to the decision by these previously blank R&D firms to report non-blank R&D the following year. Moreover, the forced auditor change would seem unlikely to cause firms to engage in corporate R&D. In contrast, if blank R&D represents a conscious disclosure decision aimed at obfuscating the R&D information environment, then we expect to observe that treated firms have a greater propensity to report R&D following this forced auditor change.

Using multivariate difference-in-differences type specifications, we find significant evidence that an exogenous auditor change led firms to report non-blank R&D. More specifically, the Pseudo-Blank R&D firms were about 6.1% more likely to report R&D after the forced auditor change. The results further indicate that when firms switch from reporting blank R&D to reporting non-blank R&D, the magnitude of the R&D seems substantive. Particularly, we find that firms that switched away from not reporting any information about R&D now disclose, on average, R&D of about 1.8% of total assets after the forced auditor change. This newly reported R&D, post auditor forced change, ranks among the 26th percentile of all firms that report positive R&D in the universe of Compustat firms.

To gauge the potential for structural attributes to allow blank R&D, we examine whether joint-venture activity is related to Pseudo-Blank R&D. Pseudo-Blank R&D firms should have greater numbers of joint ventures, because this organizational structure could allow firms to avoid separately reporting their corporate R&D. We find that both the level and addition of new joint ventures are positively associated with firms failing to disclose any information regarding corporate R&D. Overall, we find evidence that both discretionary reporting choices and the structure of corporate activities appears to facilitate firms' reporting of missing R&D.

Finally, we consider the question of how researchers in accounting and business who rely on reported R&D in their empirical tests should incorporate missing R&D. To gauge the relative benefits of the different empirical approaches to handling missing R&D values, we develop a simple Monte Carlo simulation. Simulating firm characteristics and missing R&D (for firms with either positive or zero R&D) we evaluate the impact of various remedies to blank R&D in simple multivariate regressions. One clear trend in the simulations is the importance of including a dummy variable to denote firms with missing R&D regardless of whether one focuses on replacing missing with industry average or zero. These two approaches provide marginally lower levels of improvement than the inclusion of a Pseudo-Blank dummy but are unrestricted by patent data availability. Surprisingly, incorporating patent data by itself (i.e. number of patents) offers only moderate gains over treating the missing R&D values as representing zero R&D and including a dummy variable for missing R&D. While our primary simulations suggest that replacing missing R&D with the industry average plus a dummy tends to generate good fitting models, we observe that replacing missing with zero plus a dummy works best when true R&D in blank firms is less than approximately 25% of average industry R&D. Unfortunately, for empirical studies the true level of R&D spending in firms with missing R&D remains unobservable, making it difficult to assess the distance between their R&D spending and the industry average of reported R&D.

One tractable implication for empirical researchers is to always incorporate a dummy variable for missing R&D and to alternatively replace missing with the industry average R&D and then with zeros (and if patent data is available using a Pseudo-Blank dummy as well). While the results of any series of simulations depend on the generation process of the underlying variables, our Monte Carlo analysis highlights the difficulties in simply treating missing R&D as zero R&D. In sum, our analysis suggests that empirical research that uses R&D must consider how R&D reporting may influence their research design and inferences.

2. Innovation and reported R&D

A common theme in the disclosure literature centers on reducing information asymmetries to reduce the cost of capital, while limiting the availability of proprietary cost information to competitors (Verrechia, 1983; Dye, 1985; Leuz and Verrecchia, 2000). Other studies on disclosure decisions draw attention to the fact that managers have substantial discretion in their disclosure choices. Fishman and Hagerty (1990) underscore the notion that a portion of managerial disclosure-discretion involves a level of detail or vagueness surrounding their disclosure, even with mandatory accounting statements. Gigler (1994) emphasizes that differing stakeholders or users of accounting information may have differing preferences over what is disclosed, leading to conflicts in the disclosure choices among firms. Focusing on managers' discretion in determining materiality regarding mandatory disclosures, recent literature implies that proprietary costs could influence this decision (Lo, 2010). Thus, prior research suggests that managers face conflicting pressures regarding disclosure decisions about corporate activities. Building on this influential literature, we consider the potential for various stakeholders in the firm to have differing preferences for disclosures about corporate R&D. Specifically, we posit that the R&D classification decision affords managers opportunities to classify corporate outlays as R&D (or not) depending on their particular objectives.

2.1. R&D disclosure rules

Although accounting guidelines exist for categorizing R&D expenses, they still involve a subjective component or decision by the managers of the firm. For instance, the modification of an existing product in hopes of securing a potential client might be considered R&D, while the modification of an existing product for a client would likely not be considered R&D. In addition, R&D can occur in formal R&D departments or in individual operational units, making the allocation decision even more subjective. R&D in smaller firms or in operational units often involves people who spend 50% of their time on research activity and 50% supporting current operations (Kleinknecht and Reijnen, 1991). By their nature, these fine distinctions between expense types appear subjective, even if they have the same impact on accounting profit. While the accounting guidelines for recognizing R&D are detailed, they appear to afford managers substantial discretion in recording outlays as operational or R&D expenses.

US GAAP requires that firms disclose material R&D in either their income statement or in a footnote.[3] We argue that the body of generally accepted accounting practices leaves considerable discretion regarding the classification of an outlay as R&D or as other types of expense. Since both classifications have the same impact on reported profits, the classification decision has often been viewed as cut and dried. However, for a variety of reasons, managers may prefer R&D expenses to be included with other expenses. For instance, if managers are interested in limiting information to competitors about their corporate R&D, then they may have incentives not to classify these outlays as R&D expenses.[4]

2.2. Missing or blank R&D values

Disclosure bias in corporate R&D engenders profit-neutral distortions that seem especially difficult for auditors and other information intermediaries to detect or monitor. A firm engaged in modest levels of R&D that fails to report information regarding their corporate R&D activities would seem to garner little attention. One approach to investigating innovation in firms with missing R&D focuses on the potential for firms to seek protection for the knowledge gained during the R&D process. Patents provide legal protection for a firm's knowledge capital (Gallini, 2002). Presumably, firms that file and receive patents have incurred costs in developing this knowledge capital.

Focusing more directly on the potential for under-reporting corporate R&D, we posit that the large number of firms that fail to report any information about R&D includes firms that actually engage in R&D, but choose not to disclose it separately in their financial statements. Examining patents in firms with blank R&D allows us to investigate under-reporting and firm incentives to create uncertainty regarding their corporate R&D. This leads to our first hypothesis, which is first stated in a manner to highlight the inherent argument for interpreting missing or blank R&D values as representing zero R&D.

H1. Firms that report no information on R&D expenditure have patent activity similar to the firms that report zero R&D.

The alternative condition that patent activity of non-reporting R&D firms is dissimilar to that of zero R&D firms would be inconsistent with the idea that blank R&D indicates a lack of innovation and R&D-type activities.

2.3. Research and development influence

Although patents can be a tangible outcome of corporate R&D activity, patent influence is not uniform. Layne-Farrar and Lerner (2011) suggest that firms are able to gauge the relative importance of their patents very early in the process. Reitzig (2003) emphasizes that patent applications themselves differ, suggesting that patents with more claimed contributions have greater subsequent influence. Markman et al. (2004) observes that patent claims define the patentee's view of the scope of

[3] Accounting Series Release 125 (1972) uses 1% of sales as the materiality threshold. In practice, firms report a wide range of R&D values. The emphasis in prior literature on the incentives to over report corporate R&D perhaps led many to conclude that missing or blank values represent zero or immaterial R&D. The difference between blanks and zeros in Compustat appears to stem from reporting differences by the firms themselves. If a firm reports a zero for R&D, Compustat records it as zero. Reportedly, if the firm provides no information about their R&D, then Compustat leaves the field blank. Our comparison of actual disclosures to Compustat disclosures seems to validate such treatments by Compustat.

[4] Competitors observe patent activity and possess other sources of information about research activities in firms with missing R&D. Yet, R&D is an input measure of innovation while patents are an output measure of innovation. In our tests we use this output variable to infer that firms with missing R&D appear to engage in some input (i.e. R&D). However, we cannot quantify how much R&D these firms undertake. In this context, patents can provide evidence that firms engage in R&D but do not provide sufficient information for competitors to develop a point estimate of their rivals R&D inputs. Moreover, as noted in Arundel and Kabla (1998) firms often choose to forego patent protection in order to keep their innovations secret, suggesting this proprietary cost information could be important information to corporate competitors.

the invention, while Lanjouw and Schankerman (2001) emphasize how these claims garner substantive scrutiny from the patent-office examiner. Consequently, the number of claimed contributions in a patent can be used to measure the firm's perception of the value or importance of their R&D efforts (Hall et al., 2001).

Extensive empirical research suggests that patent citation data can also be used to gauge the influence of corporate R&D (e.g., Hall et al., 2005). Schankerman (1998) observes that lifetime patent citations and the characteristics of these citations provide useful information to assess the value of corporate R&D. Bessen (2008) indicates that citation statistics, both forward- and backward-looking, provide the best metrics for evaluating the influence of the underlying R&D. Trajtenberg et al. (1997) developed three measures of patent importance that have gained widespread acceptance in studies on innovation, namely generality, originality, and citation lag. Generality captures how widespread an impact a patent has on future innovations, while originality refers to the breadth of the underlying technology used to develop the patent. Higher generality and originality indicate that a firm is engaged in more influential R&D. Citation lag, in contrast, captures competitor discovery periods and lead-time advantages in corporate R&D (Arundel, 2001; Hall et al., 2001). In sum, a substantial body of empirical research indicates that patent characteristics provide several potential metrics to evaluate the influence of corporate R&D.

We exploit these differences in patent influence to further test the intensity of innovation in blank R&D firms. The immateriality argument posits that missing R&D values represent immaterial spending on R&D and should be classified as zero R&D firms, suggesting that non-reporting R&D firms receive similar patents in comparison to firms with zero R&D. This leads to our second hypothesis:

H2. Firms that report no information on R&D expenditures have similar patent characteristics to those found in firms that report zero R&D.

2.4. Patents in Pseudo-Blank and positive R&D firms

Comparing patents in non-reporting R&D firms and those with zero R&D provides evidence on the applicability of classifying non-reporting firms as engaging in zero R&D. A natural extension of this analysis centers on comparing patents in Pseudo-Blank (i.e., non-reporting R&D firms with patent activities) and positive R&D firms. Expanding the arguments in Hypothesis 1 to encompass positive R&D reporting firms, we compare the number of patents in Pseudo-Blank R&D firms to those found in positive R&D firms. If Pseudo-Blank R&D firms occur across the entire spectrum of positive R&D firms, then we expect to observe similar numbers of patents in positive and Pseudo-Blank R&D firms. Alternatively, the non-reporting of corporate R&D may occur in firms with more modest R&D programs, suggesting a lower number of patents in Pseudo-Blank R&D firms relative to positive R&D firms. This leads to our third hypothesis:

H3. Firms that report no information on R&D expenditures but have patent activity exhibit similar patent activity level to the firms that report positive R&D.

As noted in Hypothesis 2, patent influence provides another approach to gauging the importance of corporate R&D. Failing to report corporate R&D expenditures centers may be driven by managerial attempts to obfuscate their true R&D intensity. Building on the framework in Hypothesis 2, we compare patent characteristics of Pseudo-Blank and positive R&D firms. Intuitively, these tests provide an avenue to gauge the influence or importance of patents in Pseudo-Blank firms. Immateriality arguments suggest that patents of Pseudo-Blank R&D firms are less influential than the patents of positive R&D firms. In contrast, arguments that center on creating uncertainty about the firm's true level of innovation and future growth prospects imply the potential for greater influence and competitor discovery periods for individual patents in Pseudo-Blank firms, relative to those in positive R&D firms. Thus, our fourth hypothesis focuses on comparing patent influence in Pseudo-Blank and positive R&D firms.

H4. Firms that report no information on R&D expenditures but have patent activity exhibit similar patent characteristics to those found in firms that report positive R&D.

2.5. Discretionary reporting choices

Our final two hypotheses focus on discretionary reporting choices and the structure of corporate activities. Understanding that a variety of incentives can lead managers to provide ambiguous financial statements, investors rely on external auditors to provide assurance and attestation of the firm's accounting reports. Healy and Palepu (2001) emphasize that the role of the auditor centers on confirming whether the firm's financial statements conform to generally accepted accounting practices. More recently, Cahan and Zhang (2006) report that forced auditor changes improve financial statement quality. Consequently, we seek to shed light on whether the non-reporting of R&D to some extent stems from discretionary reporting choices.

Under the hypothesis that blank R&D represents zero or negligible R&D, a forced auditor change should have no effect on the decision to report corporate R&D subsequent to the forced change. Neither should the forced auditor change influence firms to initiate R&D. However, to the extent that the forced auditor change leads to improved financial reporting quality, firms with missing R&D arising from discretionary reporting choices should be more likely to report R&D after a forced auditor change. This leads to our fifth testable hypothesis, again starting with the classic view that blank R&D represents zero R&D.

H5. A forced auditor change has no effect on the decision to report corporate R&D.

Although, the non-reporting of corporate R&D may arise from discretionary reporting choices, organizational structure potentially facilitates this omitted disclosure. For instance, a common approach to organizing R&D activities centers on joint ventures, where the firm engages in a project with another firm with joint overall control. Research on the governance of R&D joint ventures emphasizes how companies use different mechanisms to mitigate various contracting hazards, control spillovers, and assign profits (Kogut, 1998). Others observe that firms report within firm operations in greater detail than found in disclosures regarding joint ventures, potentially masking material important information (Kothavala, 2003). Pseudo-Blank firms potentially arise from R&D occurring in joint ventures, limiting the firms need to disclose this R&D in their accounting statements. Thus, our final hypothesis focuses on whether the use of joint ventures is associated with missing R&D for firms with patent activity.

H6. New joint ventures are unrelated to the Pseudo-Blank R&D firm designation.

3. Sample and research design

3.1. Sample, variable measurements and descriptive statistics

Our initial sample comprises Compustat firms with sales and total assets above $1 million between 1980 and 2006 inclusively. The patent data availability restricts our sample period to end on 2006. Regulated utilities (SIC between 4900 and 4999) and financial firms (SIC between 6000 and 6999) are excluded due to their different operating and reporting environment. We also exclude observations without sufficient information to construct our variables leading to an overall sample with 104,506 firm-year observations between 1980 and 2006 inclusively. Our overall sample ensures we have a representative sample for a large cross-section of firms over a long time horizon.

In this study, we use both full sample and matched samples to test our hypotheses. Firm-level R&D differs across firm size, type, and industry, highlighting the difficulty in identifying the correct counter-factual firm. A challenge in comparing firms with missing R&D to those with zero or positive R&D is that the firms are not randomly selected. Intuitively, the matched sample framework seeks to correct for this non-random treatment effect by matching to the appropriate counter-factual firm. The costs of this approach center on the lower sample sizes and the potential loss of external validity, while the benefits stem from greater precision in the comparison among firms with missing R&D, zero R&D, and positive R&D firms.

In addition to the full sample analysis, we use propensity scores to match treatment and control firms, without replacement, on the following dimensions:

$$\text{Treatment Group} = f(\text{PPE, Lagged Tobin's } Q,\ \text{Firm Age},\ \text{Leverage},\ \text{Proportion of Firms Reporting}\ \text{R \& D in the Industry},\ \text{Year Fixed Effects}) \quad (1)$$

Our matched samples for different tests differ according to their respective test settings as described in the relevant section later. Firms that rely heavily on property, plant and equipment (PPE) in their operations are, on average, less likely to engage in R&D activities. As such, we expect machinery-intensive firms to have low R&D outlays and hence less need to disclose R&D information in their financial statements. PPE is measured as property, plant and equipment divided by total assets. Firms with high growth options are more likely to be engaging in R&D and face greater information asymmetry between managers and investors (Gaver and Gaver, 1993). Thus, they are more likely to disclose their R&D efforts to reduce such information asymmetry. We measure firms' growth options, lagged Tobin's Q, as (market value of equity + total debt) divided by total assets. We also match on firm age, as Andersen et al. (2009) found that older firms exhibit greater opacity than younger firms and Chen et al. (2002) report a negative correlation between voluntary disclosure and firm age.

Prior research shows that firms have incentives to increase their disclosure activities regarding growth opportunities prior to accessing the equity market to reduce information asymmetry (Healy et al., 1999). In contrast, firms that heavily rely on debt have outside investors that presumably are less interested in growth options and face low information asymmetry between managers and lenders as the latter can have more privileged access to business information through their monitoring (e.g., through debt renewals and debt covenant compliance). As such, we expect more highly levered firms to be less inclined to disclose R&D expenditure information because lenders do not share the future benefits of successful R&D projects and because they have greater access to management. Further, intangibles created by R&D activities are unsuitable collateral for lenders, making loans difficult and expensive to secure. We measure leverage as total debt divided by total assets.

Similarly, if firms have incurred R&D expenditures during the year and are operating in an industry where a large proportion of their peers is reporting R&D expenditure (zero or positive values), then they are more likely to disclose their R&D expenditures (zero or positive values). Intuitively, firms in industries with greater R&D disclosure have a higher probability of disclosing their own R&D, because investors update their assessments about the firm's likelihood of engaging in R&D based on their competitors' disclosures (see Dye and Sridhar, 1995). As such, we also include the proportion of firms in the same industry (based on SIC two-digit industries) that have reported zero or positive R&D expenditure during the year. Finally, we include year fixed effects to control for time-specific differences in R&D disclosure choice and R&D expenditure level. We summarize the variable definitions in Table A1. All continuous variables are winsorized at the 1% and 99% levels.

Table 1 Panel A reports the descriptive statistics on the overall sample (before matching), while Panel B presents the same information partitioned by whether firms reported blank/missing, zero or positive R&D expenditures. Overall, around 58% of our observations reported non-missing R&D expenditure with an average R&D expenditure of 7.3% of total assets. Among these observations, around 82% (or 49,737) reported positive R&D outlays, averaging 8.8% of total assets.

Turning to Panel B, firms that reported missing R&D outlays have the highest mean and median PPE, leverage and the lowest mean and median proportion of industry peers reporting non-missing R&D, implying many of these firms are unlikely to engage in R&D activities. Firms with missing R&D on average have more patent applications and approvals than

Table 1
Descriptive statistics.

Panel A: Overall sample

Variable	*N*	Mean	Median	Std. dev.
R&D Expenditure	60,453	0.073	0.034	0.106
Reported R&D	104,506	0.578	1.000	0.494
PPE	104,506	0.304	0.248	0.226
Lagged Tobin's Q	104,506	1.657	1.126	1.576
Firm Age	104,506	14.270	10.000	14.008
Leverage	104,506	0.245	0.212	0.220
Proportion of firms reporting R&D in the Industry	104,506	0.554	0.610	0.282
Log(Total Assets)	104,506	4.993	4.824	2.094
Capital Expenditure	104,506	0.003	0.000	0.009
ROA	104,506	0.079	0.115	0.191
Institutional Ownership	104,506	0.285	0.208	0.273
No. of Patent Applications	104,506	7.851	0.000	71.389
No. of Patent Granted	104,506	8.590	0.000	72.702
Patent Citations	104,506	0.615	0.000	1.176
Generality	104,506	0.025	0.000	0.064
Originality	104,506	0.088	0.000	0.193
Citation Lag	104,506	0.484	0.000	0.766
Patent Claims	104,506	0.713	0.000	1.247
Analyst Following Indicator	104,506	0.613	1.000	0.487
No. of Analysts Following	104,506	5.102	2.000	7.865
Dispersion	48,362	0.132	0.033	0.323

Panel B: Subsamples by R&D reporting type

	Blank R&D ($n=44{,}053$)			Zero R&D ($n=10{,}716$)			Positive R&D ($n=49{,}737$)		
Variable	Mean	Median	Std. dev.	Mean	Median	Std. dev.	Mean	Median	Std. dev.
R&D Expenditure				0.000	0.000	0.000	0.088	0.050	0.111
PPE	0.370	0.324	0.256	0.362	0.313	0.248	0.233	0.200	0.163
Lagged Tobin's Q	1.356	1.009	1.161	1.385	1.011	1.189	1.981	1.310	1.877
Firm Age	13.943	10.000	12.984	11.114	8.000	10.387	15.240	10.000	15.380
Leverage	0.295	0.271	0.229	0.281	0.257	0.226	0.192	0.152	0.198
Proportion of Firms Reporting R&D in the Industry	0.364	0.333	0.253	0.609	0.631	0.200	0.711	0.815	0.211
Log(Total Assets)	5.081	4.983	1.990	4.846	4.846	1.905	4.947	4.660	2.216
Capital Expenditure	0.003	0.000	0.011	0.003	0.000	0.011	0.002	0.000	0.007
ROA	0.106	0.119	0.145	0.104	0.120	0.143	0.049	0.108	0.227
Institutional Ownership	0.265	0.177	0.269	0.306	0.232	0.290	0.298	0.230	0.272
No. of Patent Applications	0.361	0.000	4.132	0.011	0.000	0.170	16.175	0.000	102.768
No. of Patent Granted	0.421	0.000	4.262	0.023	0.000	0.237	17.671	0.000	104.558
Patent Citations	0.205	0.000	0.707	0.032	0.000	0.307	1.103	0.000	1.408
Generality	0.011	0.000	0.049	0.003	0.000	0.030	0.043	0.000	0.076
Originality	0.029	0.000	0.125	0.006	0.000	0.061	0.159	0.000	0.233
Citation Lag	0.268	0.000	0.663	0.061	0.000	0.334	0.767	0.631	0.811
Patent Claims	0.251	0.000	0.805	0.041	0.000	0.337	1.266	0.000	1.441
Analyst Following Indicator	0.569	1.000	0.495	0.615	1.000	0.487	0.652	1.000	0.476
No. of Analyst Following	4.381	1.000	6.953	5.078	2.000	7.277	5.746	2.000	8.654
Dispersion ($n=18{,}669$; 5,041; 24,652)	0.144	0.036	0.341	0.101	0.022	0.286	0.130	0.034	0.316
% of Observations with Patent Activity	10.54%			1.73%			48.73%		

Variables are as defined in Table A1.

zero R&D firms. Similarly, they also have more influential patents than zero R&D firms, as measured by Generality, Originality, Citation Lag, the number of patent claims and citations. Thus, non-reporting R&D firms, relative to firms with zero R&D, receive more patents and individually these patents have greater influence. However, in comparison to positive R&D firms, missing or blank R&D firms receive far fewer and less influential patents. Analysts appear to be less likely to follow firms with missing R&D and have greater forecast dispersions as compared with zero and positive R&D firms. These findings suggest a complex relation between firms' R&D disclosures and the information environment.

4. Analysis of results

4.1. Patent activity and corporate R&D

Our investigation of R&D disclosure bias relies on one of the outcomes of R&D activities – patent applications and patents granted. We focus on patent activities as they are a common outcome of R&D endeavors and can be used to measure the intensity of firms' R&D investments (Hall et al., 2005). We investigate the differences in patent activity between firms reporting blank and zero R&D expenditure.[5] If there is no under-reporting bias among firms reporting blank R&D, then we expect no significant differences in patent activity between these two types of firms. We use propensity scores to match non-reporting and zero R&D firms, without replacement, on the dimensions described in Eq. (1) over the period 1980–2006 where patent data is available. This leads to a matched sample of 14,430 observations (7,215 observations each for blank and zero R&D).

We find no significant differences between non-reporting firms and zero R&D firms across all our matching dimensions (Table 2 Column (4); p-values >0.10), suggesting covariate balance in our propensity-matching process. Table 2 also details the univariate results using the full sample and shows there are significant differences on all variables that are used to create the matched sample. Examining the results for the matched sample, firms with blank R&D have significantly more patent applications and approvals than firms reporting zero R&D (p-values$=0.000$). On average, we find non-reporting firms in the matched sample file about 16 times more patent applications ($0.209/0.012-1=16.42$) and are granted about seven times more patents ($0.217/0.026-1=7.35$) than zero R&D firms. The results for the non-matched sample indicate that non-reporting firms have around 32 times more patent applications ($0.361/0.011-1=31.82$) and 17 times more patent approvals ($0.421/0.023-1=17.30$) than zero R&D firms. Arguably, the magnitudes of these numbers stem from the small number of patents filed by zero R&D firms. Thus, these results provide preliminary evidence to suggest that non-reporting R&D firms, on average, appear dissimilar to firms with zero R&D.

We next perform multivariate analysis to examine whether patent activity differs between blank and zero R&D firms using Poisson regressions on the matched sample as follows:

$$\text{No. of Patents} = f(\text{Blank R \& D, PPE, Lagged Tobin's } Q, \text{ Firm Age, Leverage, Proportion of Firms Reporting R \& D in the Industry, Industry and Year Fixed Effects}), \quad (2)$$

where No. of Patents is either the number of patent applications or patents granted; Blank R&D is an indicator variable taking 1 if firms report blank R&D; 0 if firms report zero R&D. We control for the variables used in the propensity score matching process to reinforce the assimilation of their cross-sectional influence on patent applications. All other variables are as previously defined.

Table 3 presents the results. Consistent with our univariate results, among our matched sample, we find non-reporting firms have significantly more patent applications than zero R&D firms (Column (1); z-statistic$=7.312$). Translating the estimated coefficients to incident-rate ratios, non-reporting firms have around 14 times more patent applications than zero R&D firms ($e^{2.658}=14.27$). Similar results are found for patents granted, where blank R&D firms received significantly more patent approvals in a blank R&D year than zero R&D firms (Column (2); z-statistic$=7.749$) at a rate of around seven times ($e^{1.980}=7.24$). We find similar results using the non-matched sample (untabulated), where non-reporting firms have around 14 times more patent applications (eight times more patent approvals) than zero R&D firms ($e^{2.666}=14.38$ and $e^{2.102}=8.18$ respectively), all statistically significant at the 0.1% level. This evidence suggests that at least some firms that have missing R&D do in fact conduct substantive R&D activities, but choose not to report such information.

Taken together, our evidence rejects the hypothesis that patent activity, on average, is similar between blank R&D firms and zero R&D firms (H1). Our findings also indicate that the common treatment of coding non-reporting firms as zero R&D firms by accounting and finance scholars appears problematic.

[5] Monetizing patents involves developing the technology, licensing the patent, and/or selling the patent. The patent literature emphasizes that firms rarely sell or even license their strategic patents (Gambardella et al., 2007). However, an active market for acquiring patents does exist in the US. For instance, Ocean Tomo organizes patent auctions; Thomson-Reuters PatentWeb facilitates patent sales, and so-called patent trolls routinely purchase sets of patents (Abril and Plant, 2007). Our analysis centers on patent applications and grants to the original assignees rather than on the ultimate owner of a given patent. Limiting the analysis to firms without M&A activity provides similar results.

Table 2
Univariate results on tests of patent counts (applications and grants) and patent characteristics between firms reporting blank and zero R&D expenditure.

Variable	Matched Sample: Mean Blank (1)	Matched Sample: Mean Zero (2)	Matched Sample: Differences in Means (3)=(1)–(2)	Matched Sample: Differences in *p*-Value (4)	Overall Sample: Mean Blank (5)	Overall Sample: Mean Zero (6)	Overall Sample: Differences in Means (7)=(5)–(6)	Overall Sample: Differences in *p*-Value (8)
Patent Counts:								
No. of Patent Applications	0.209	0.012	0.197	0.000	0.361	0.011	0.350	0.000
No. of Patent Granted	0.217	0.026	0.191	0.000	0.421	0.023	0.398	0.000
Patent Characteristics:								
Patent Citations	0.137	0.038	0.099	0.000	0.205	0.032	0.173	0.000
Generality	0.008	0.003	0.005	0.000	0.011	0.003	0.008	0.000
Originality	0.022	0.006	0.016	0.000	0.029	0.006	0.023	0.000
Citation Lag	0.194	0.070	0.124	0.000	0.268	0.061	0.207	0.000
Patent Claims	0.168	0.045	0.123	0.000	0.251	0.041	0.021	0.000
Matching variables:								
PPE	0.334	0.329	0.005	0.177	0.370	0.362	0.008	0.002
Lagged Tobin's Q	1.350	1.336	0.014	0.451	1.356	1.385	−0.029	0.023
Firm Age	12.038	11.993	0.045	0.813	13.943	11.114	2.829	0.000
Leverage	0.297	0.299	−0.002	0.604	0.295	0.281	0.014	0.000
Proportion of Firms Reporting R&D in the Industry	0.537	0.532	0.005	0.174	0.364	0.609	−0.245	0.000
Observations	7,215	7,215			44,053	10,716		
% of Observations with Patent Activity	7.10%	1.91%			10.54%	1.73%		

The sample period is 1980–2006 inclusive due to patent data availability. Control firms are matched on the nearest propensity score based on the following logit model:
Treatment group=*f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year Fixed Effects).
Treatment group takes 1 if firms reported missing R&D expenditure; 0 if reported zero R&D expenditure. Variables are as defined in Table A1.

Table 3
Regression analysis on regressing the patent counts (applications and grants) and patent characteristics on R&D reporting choice among firms reporting blank and zero R&D expenditure (matched sample).

Variables	Applications (1)	Granted (2)	Patent citations (3)	Generality (4)	Originality (5)	Citation lag (6)	Patent claims (7)
Blank R&D	2.658***	1.980***	2.996***	0.182***	0.656***	1.562***	3.119***
	(7.312)	(7.749)	(8.103)	(6.229)	(8.403)	(7.453)	(8.358)
PPE	0.171	−0.529	−1.557**	−0.076	−0.246	−0.863*	−1.435*
	(0.352)	(−1.057)	(−2.067)	(−1.073)	(−1.438)	(−1.862)	(−1.906)
Lagged Tobin's Q	0.089	0.072	0.156	0.018**	0.030	0.084	0.161
	(1.525)	(0.772)	(1.273)	(2.010)	(0.980)	(1.375)	(1.245)
Firm Age	0.054***	0.048***	0.100***	0.009***	0.022***	0.072***	0.105***
	(5.809)	(5.463)	(8.232)	(10.021)	(8.365)	(10.394)	(8.618)
Leverage	−0.894*	−0.910*	−0.051	−0.013	0.067	0.440	0.065
	(−1.737)	(−1.725)	(−0.066)	(−0.247)	(0.377)	(1.259)	(0.083)
Proportion of Firms Reporting R&D in the Industry	2.923**	2.638*	2.032*	0.037	0.420*	0.226	2.128*
	(2.189)	(1.923)	(1.814)	(0.429)	(1.704)	(0.429)	(1.933)
Constant	−5.399***	−4.734***	−11.747***	−1.076***	−3.317***	−5.852***	−12.209***
	(−6.618)	(−8.383)	(−12.145)	(−9.089)	(−10.714)	(−11.073)	(−12.498)
Year & Industry Dummies	Y	Y	Y	Y	Y	Y	Y
Observations	14,430	14,430	14,430	14,430	14,430	14,430	14,430
Log Likelihood	−5,073	−5,633	−2,823	−1,425	−1,749	−5,172	−3,206

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ two-tailed. Test statistics in parentheses are based on firm-level clustering. The sample period is 1980–2006 inclusive due to patent data availability
Columns (1) and (2) use the following Poisson Regression Model:
No. of Patents=*f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year and Industry Fixed Effects).
No. of Patents is either No. of Patent Applications or Patents Granted; Blank R&D is an indicator variable 1 for firms reporting blank R&D expenditure; 0 if firms reporting zero R&D expenditure. All other variables are as defined in Table A1.
Columns (3) and (7) use the following Tobit Regression Model:
Dependent variable=*f*(Pseudo-Blank Firms, PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year and Industry Fixed Effects).
Dependent variables are: Generality, Originality, Citation Lag, and Patent Claims. All variables are as defined in Table A1.

4.2. R&D impact of non-reporting Firms

Next, we turn our attention to assess the R&D impact of firms reporting missing R&D relative to firms reporting zero R&D. Patents, like published research articles, can differ in their incremental contribution – some appear to receive only limited attention, while others appear to have substantial influence and impact across multiple fields. Consequently, scholars in economics and strategic management extensively analyze measures of patent influence. Using both the matched and unmatched samples described earlier, we run the following Tobit regression to test the influence argument.

$$\text{Dependent Variable} = f(\text{Blank R \& D, PPE, Lagged Tobin's } Q, \text{ Firm Age, Leverage, Proportion of Firms Reporting R \& D in the Industry, Industry and Year Fixed Effects}) \quad (3)$$

where the dependent variable includes our five measures of patent influence: Generality, Originality, Citation Lag, Patent Claims and Patent Citations. We use a Tobit regression as our patent influence measures are censored on the left with Originality and Generality also censored on the right.[6]

Table 2 reports the univariate results of our patent influence measures for both matched and unmatched samples. We find blank R&D firms have significantly more influential patents than zero R&D firms (p-values = 0.000). Our matched sample results indicate that blank firms on average received 2.61 times (0.137/0.038–1) more cites than zero R&D firms. On average, patent breadth of blank R&D firms is 1.7 times (Generality; 0.008/0.003–1) and 2.7 times (Originality; 0.022/0.006–1) greater than zero R&D firms. Citation lag of blank firms is 1.8 times (0.194/0.070–1) longer that of zero R&D firms; while they also make 2.7 times (0.168/0.045–1) more claims than zero R&D firms. The differences are more pronounced in the unmatched sample. While these magnitudes are likely to be affected by the low number of patents filed by zero R&D firms, they nonetheless provide evidence inconsistent with the notion that non-reporting R&D firms' patents are similar to the patents of zero R&D firms. We repeat the patent analysis on just the subset of firms in both groups with patent activity (i.e., exclude firms without patent activity in both groups). In untabulated results, we find that among this subset of firms, missing R&D firms, relative to zero R&D firms, have significantly greater patent applications, approvals, generality, citation lag, and claims.

Our regression analysis results are consistent with the univariate evidence discussed above (see Table 3 Columns (3)–(7)). Focusing on matched sample results, patents of blank R&D firms receive more citations than zero R&D firms (Column (3); t-statistics = 8.103), where on average the former receive 7.645 citations compare with an average of 5.529 citations for zero R&D firms,[7] a difference of around 38.3% (7.645/5.529–1). We also find that individual patents in non-reporting R&D firms exhibit a broader impact on other technological classes (as measured by Generality), than zero R&D firms (Column (4); t-statistics = 6.229). For an average firm, the associated marginal effect on the Generality scores of blank R&D firms are higher than zero R&D firms by 0.020 (0.147 vs. 0.127), or higher by around 15.7% (0.020/0.127). As for Originality, we similarly find patents of blank R&D firms show greater originality than zero R&D firms (Column (5); t-statistics = 8.403). For an average firm, the Originality scores for blank R&D firms are 0.032 higher than zero R&D firms (0.350 vs. 0.318) or higher by around 10.1% (0.032/0.318).

In addition, we find individual blank R&D firms' patents display longer citation lags in comparison to zero R&D firms (Column (6); t-statistics = 7.453), where for an average firm the former citation lag is 3.831 years ($e^{1.343}$) while the latter is 3.043 years ($e^{1.113}$), a difference of around 25.9% (3.831/3.043–1). Patents of blank R&D firms also exhibit greater contributions than zero R&D firms (Column (7); t-statistics = 8.358). Specifically, for an average firm, the patents of blank R&D firms reveal an average of 9.826 claims ($e^{2.285}$) vis-à-vis an average of 6.753 claims ($e^{1.910}$) for zero R&D firms, a 45.5% difference (9.826/6.753–1).[8] Taken together, this evidence appears inconsistent with the hypothesis that non-reporting R&D firms are similar to zero R&D firms.

4.3. Pseudo-Blank R&D firms' patents and R&D influence

Our next series of tests focus on comparing the subset of non-reporting R&D firms with patent activity (i.e., Pseudo-Blank R&D firms) to positive R&D firms.[9] Using our conservative approach of identifying missing R&D expenditure firms with R&D activities, we find a non-trivial number of missing R&D firms are Pseudo-Blank firms (~10.5% of blank firms). As a benchmark we note that about half of the positive R&D firms do not seek patents.[10] We compare the patent activity of these

[6] We also perform additional analysis using multi-year averages of the independent variables; dependent variables; or both for Eqs. (2) and (3). In untabulated results, we find evidence consistent with those reported in Tables 3 and 6 at the same significance levels using these alternative specifications.

[7] We compare the economic effects for the two groups based on their respective predicted values, holding other variables at the mean. Since patent citations, citation lag and claims are in logarithms, we further take the exponential on the predicted values to arrive at their original measurement unit. Using patent citations as an example, the predicted values for blank and zero R&D firms are respectively 2.034 and 1.170. These translate to 7.645 citations ($e^{2.034}$) and 5.529 citations ($e^{1.710}$) respectively.

[8] In additional untabulated analysis we conduct parallel tests for Table 3 using unmatched samples of blanks and zero R&D firms. We continue to find blanks firms patents have higher Patent Citations, Generality, Originality, Citation Lag, and Patent Claims than zero R&D firms, all significant at the 0.1% level.

[9] When examining the 2-digit SIC industry distributions of Pseudo-Blank R&D firms, we do not find any specific industry dominating the Pseudo-Blank firms.

[10] We also observe that among positive R&D firms above the median reported R&D, around half of them (51%) do not file or receive patents.

Table 4
Tests of patent counts and patent characteristics for Pseudo-Blank and positive R&D firms.

Panel A: Pseudo-Blank and positive R&D firms

Variable	Matched sample					Overall sample				
	Pseudo-Blank R&D (n=4,536)		Positive R&D (n=4,536)		Diff.	Pseudo-Blank (n=4,644)		Positive R&D (n=49,737)		Diff.
	Mean (1)	Std. dev. (2)	Mean (3)	Std. dev. (4)	(5)=(1)–(3)	Mean (6)	Std. dev. (7)	Mean (8)	Std. dev. (9)	(10)=(6)−(8)
Patent Counts:										
No. of Patent Applications	3.475	12.445	20.995	111.037	−17.520***	3.427	12.309	16.175	102.768	−12.748***
No. of Patent Granted	4.045	12.710	21.824	103.601	−17.779***	3.997	12.570	17.671	104.558	−13.674***
Patent Characteristics:										
Patent Citations	1.946	1.165	1.075	1.322	0.871***	1.943	1.165	1.103	1.408	0.840***
Generality	0.055	0.090	0.040	0.070	0.015***	0.056	0.091	0.043	0.076	0.013***
Originality	0.280	0.280	0.144	0.215	0.136***	0.280	0.280	0.159	0.233	0.121***
Citation Lag	1.240	0.790	0.859	0.869	0.381***	1.240	0.790	0.767	0.811	0.473***
Patent Claims	2.381	1.033	1.234	1.402	1.147***	2.384	1.034	1.266	1.441	1.118***
Matching Variables:										
PPE	0.338	0.201	0.340	0.195	−0.002	0.347	0.207	0.233	0.163	0.114***
Lagged Tobin's Q	1.266	0.981	1.223	0.920	0.043**	1.257	0.973	1.981	1.877	−0.724***
Firm Age	22.557	17.034	23.086	20.772	0.529	22.803	17.159	15.240	15.380	7.563***
Leverage	0.261	0.182	0.264	0.203	−0.003	0.264	0.183	0.192	0.198	0.072***
Proportion of Firms Reporting R&D in the Industry	0.464	0.245	0.467	0.239	−0.003	0.454	0.250	0.711	0.211	−0.257***
Percentage of Observations with Patents	100%		48.15%			100%		48.73%		

Panel B: Pseudo-Blank and positive R&D firms with patent Activity

Variable	Matched Sample					Overall Sample				
	Pseudo-Blank R&D (n=4,114)		Positive R&D with patent activity (n=4,114)		Diff.	Pseudo-Blank (n=4,644)		Positive R&D with patent activity (n=24,236)		Diff
	Mean (1)	Std. dev. (2)	Mean (3)	Std. dev. (4)	(5)=(1)−(3)	Mean (6)	Std. dev. (7)	Mean (8)	Std. dev. (9)	(10)=(6)−(8)
Patent Counts:										
No. of Patent Applications	3.543	12.487	30.719	150.719	−27.176***	3.427	12.309	33.194	145.290	−29.767***
No. of Patent Granted	4.171	13.271	30.645	130.451	−26.474***	3.997	12.570	36.264	147.518	−32.267***
Patent Characteristics:										
Patent Citations	1.946	1.173	2.224	1.092	−0.278***	1.943	1.165	2.263	1.200	−0.320***
Generality	0.056	0.090	0.066	0.078	−0.010***	0.056	0.091	0.072	0.078	−0.016***

Originality	0.278	0.279	0.298	0.242	−0.020***	0.280	0.280	0.326	0.238	−0.046***
Citation Lag	1.249	0.784	1.300	0.656	−0.051***	1.240	0.790	1.218	0.639	0.022**
Patent Claims	2.373	1.039	2.524	0.894	−0.151***	2.384	1.034	2.598	0.894	−0.214***
Matching Variables:										
PPE	0.329	0.195	0.330	0.175	−0.001	0.347	0.207	0.250	0.158	0.097***
Lagged Tobin's Q	1.275	1.012	1.242	0.942	0.033	1.257	0.973	1.942	1.811	−0.685***
Firm Age	23.154	17.392	23.525	19.921	−0.371	22.803	17.159	19.323	18.004	3.480***
Leverage	0.255	0.178	0.259	0.192	−0.004	0.264	0.183	0.192	0.184	0.072***
Proportion of Firms Reporting R&D in the Industry	0.495	0.234	0.496	0.234	−0.001	0.454	0.250	0.837	0.200	−0.383***
Percentage of Observations with Patents	100%		100%			100%		100%		

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ two-tailed. The sample period is 1980–2006 inclusive due to patent data availability. Matched sample firms are drawn from Pseudo-Blank firms and positive R&D firms and are matched on the nearest propensity score based on the following logit model:

Treatment group = *f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year Fixed Effects).

Treatment group takes 1 if firms reported positive R&D expenditure; 0 if Pseudo-Blank firms.

Variables are as defined in Table A1.

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ two-tailed. The sample period is 1980–2006 inclusive due to patent data availability. Matched sample firms are drawn from Pseudo-Blank firms and positive R&D firms with patent activity and are matched on the nearest propensity score based on the following logit model:

Treatment group = *f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year Fixed Effects).

Treatment group takes 1 if firms reported positive R&D expenditure that have patent activity; 0 if Pseudo-Blank firms.

Variables are as defined in Table A1.

Pseudo-Blank R&D firms with that of positive R&D firms. We use the full sample and also develop a propensity score matched sample between Pseudo-Blank firms and positive R&D firms based on Eq. (1) over the period 1980–2006. This leads to a matched sample of 9,072 observations. Table 4 reports the univariate analysis using our matched and unmatched samples of Pseudo-Blank and positive R&D firms.[11] We find that positive R&D firms exhibit significant greater patent activity levels than Pseudo-Blank R&D firms (for both matched and unmatched samples).

Given the significantly larger standard deviation in patent applications and approvals of positive R&D firms (see Table 4), we perform additional analysis on the patent activities of Pseudo-Blank and positive R&D firms. In order to evaluate how patent filings in Pseudo-Blank firms compare to the population of positive R&D firms, we visually examine their respective distributions. Specifically, in Fig. 1, we plot the average patent activity levels (natural log transformed) of 3 groups of firms: (1) Pseudo-Blank R&D firms; (2) positive R&D firms; and (3) positive R&D firms that filed or received patents. Fig. 1 suggests that the differences in patent activities between Pseudo-Blank and positive R&D firms stem from the high level of patent activities among the top 5% of positive R&D firms. In fact, Pseudo-Blank firms' patent activity level is visibly higher than positive R&D firms until about the 75th percentile (since approximately half of the positive R&D firms do not file for patents), where it falls below those of positive R&D firm and peaks at about the 95th percentile of positive R&D firms. Focusing on the subset of positive R&D firms with patents, pseudo-blank R&D firms correspond to the bottom 90% of these positive R&D firms.

Turning to the influence of the R&D in Pseudo-Blank R&D firms, the univariate results in Table 4 Panel A indicate that the patents in Pseudo-Blank firms exhibit more influence in terms of patent Generality, Originality, Citation Lag, Patent Claims, and Patent Citations than found in the average positive R&D firm (for both matched and unmatched samples; see Table 4). For instance, using both matched and unmatched samples, we find significant evidence (p-values = 0.000) that patents in Pseudo-Blank firms claim around 1.9 times as many claims as patents in positive R&D firms.[12] Intuitively, comparing the number and influence of patents in Pseudo-Blank firms to the group of positive R&D firms provides indirect evidence on the R&D materiality of Pseudo-Blank firms.

Table 4 Panel B repeats the above analysis but excludes positive reporting R&D firms without patent activity. This second set of tests provides insights on the relative importance of the individual patents in positive and Pseudo-Blank R&D firms. Across each measure of patent characteristics, patents of Pseudo-Blank firms exhibit roughly 4–15% lower values than found in positive reporting R&D firms. For instance, based on the matched sample, Pseudo-Blank firms trail positive reporting R&D firms in terms of patent citations by 12.5%, while citation lag is about 3.9% lower in Pseudo-Blank firms relative to positive reporting R&D firms. Thus, the relative importance of patents in Pseudo-Blank firms, on average, is about 9% less than that found in positive reporting R&D firms with patent activity. In sum, this evidence is inconsistent with the view that R&D activities of Pseudo-Blank R&D firms are negligible or inconsequential.

4.4. Discretionary reporting and corporate structure

Next, we take advantage of the rapid demise of Arthur Andersen (AA) as a quasi-natural experiment to further examine R&D disclosures and corporate innovation. To the extent that managers exercise discretion in their R&D disclosure with the approval of (or lack of dissent from) their auditors, then a change in auditor may alter the existing understanding between outgoing auditors and clients. The advantage of the AA demise is that it took place very rapidly and that the change in auditor was forced upon its clients, i.e., an exogenous event with respect to R&D disclosure choices. Prior research emphasizes that a forced auditor change alters the auditor–client relations and reduces firm discretion. Cahan and Zhang (2006), for instance, indicate that replacement auditors for former AA clients exercised greater scrutiny on financial reporting; suggesting that obfuscation of R&D spending will be less likely. In contrast, if missing R&D expenditure simply represents zero or negligible R&D, a forced auditor change would have no effect on firms' R&D reporting decisions. Moreover, a forced auditor change should not cause firms to engage in R&D activities. As such, if firms have been reporting blank R&D expenditure as a discretionary reporting choice to obfuscate their R&D expenditures, we would expect some of these firms to alter their R&D disclosure policy.

To construct our test samples, we first identify the set of non-AA Big N clients without a change in auditor between 2001 and 2002 to use for forming a matched sample. We then match these potential control firms to former AA clients based on the dimensions in Eq. (1). We also include a variable to indicate whether a firm has reported non-blank R&D and their reported R&D level as matching dimensions to ensure that treatment and control firms are similar in these dimensions. All matching characteristics are measured in 2001 (i.e., pre-AA collapse). The resulting matched sample consists of 1,152 observations, and we adopt the following logistic regression for hypothesis testing:

$$\begin{aligned}\text{Prob}(\text{Reporting R \& D}) = f(&\text{Former AA Clients, Pseudo} - \text{Blank R \& D, Former AA Clients}^{*}\text{Pseudo}-\\&\text{Blank R \& D, PPE, Lagged Tobin's } Q\text{, Firm Age, Leverage, Proportion of}\\&\text{Firms Reporting R\&D in the Industry, Industry Fixed Effects})\end{aligned} \quad (4)$$

[11] In unreported results, we observe that Pseudo-Blank firms attract more analysts following than firms with missing R&D and firms with positive R&D.

[12] We perform multivariate tests using a specification equivalent to Eqs. (2) and (3). Multivariate results using both matched and unmatched samples of Pseudo-Blanks and positive R&D firms (see Table A2 Panel A for matched sample results) are similar to those shown in the univariate results in Table 4 Panel A.

Table 5
Matched sample logistic regression tests of the likelihood of former pseudo-blank firms' reporting non-missing R&D after AA collapse.

Variables	Reported non-missing R&D	
	(1)	(2)
Former AA Clients	−0.187	−1.408*
	(−1.100)	(−1.752)
Pseudo-Blank R&D Firms	−19.005***	−15.455***
	(−17.603)	(−7.222)
Former AA Clients*Pseudo-Blank R&D Firms	15.514***	17.202***
	(11.377)	(6.502)
Blank R&D		−8.755***
		(−9.577)
Former AA Clients*Blank R&D		0.384
		(0.315)
PPE	0.236	1.625
	(0.524)	(1.180)
Lagged Tobin's Q	0.373***	0.134
	(4.345)	(0.412)
Firm Age	0.019***	−0.008
	(3.040)	(−0.469)
Leverage	−0.279	−2.679***
	(−0.643)	(−2.576)
Proportion of Firms Reporting R&D in the Industry	5.302***	5.016***
	(10.681)	(3.735)
Constant	−3.482***	2.907*
	(−7.933)	(1.789)
Industry Dummies	Y	Y
Observations	1,152	1,152
Log Likelihood	−461.8	−73.47

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ two-tailed. z-Statistics in parentheses are based on firm-level clustering. The post AA collapse period is 2002. Control firms are non-AA Big N clients without auditor change and are matched on the nearest propensity score based on the following logit model (all variables are measure in the year immediately prior to AA collapse, i.e., 2001):
Treatment group=f(Reported Non-blank R&D, Reported R&D, PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year Fixed Effects).
Treatment group takes 1 if firms are former AA clients; 0 otherwise. All other variables are as defined in Table A1.

Our multivariate difference-in-differences design (Abadie, 2005) compares the following two dimensions: (a) Pseudo-Blank R&D firms 0vs. non-Pseudo-Blank R&D firms; and (b) firms that experienced the forced auditor change vs. those that did not. Included in our comparison group are firms that reported R&D prior to the forced auditor change. Presumably, a forced auditor change should not lead these firms to discontinue their R&D activities. Intuitively, including firms that report R&D in the comparison group provides a conservative test setting. If former AA clients reported blank R&D to hide their corporate R&D, then we expect they would be more likely to report non-blank R&D expenditure after the forced auditor change, namely Former AA Clients*Pseudo-Blank R&D > 0.

Table 5 reports the results. In Column (1), we find former AA clients that are Pseudo-Blank R&D firms prior to the forced auditor change are more likely to report non-blank R&D expenditure after the forced auditor change (z-statistics for Former AA Clients*Pseudo-Blank R&D Firms = 11.377). Compared to non-AA Big N clients that are Pseudo-Blank R&D firms before the AA demise, former Pseudo-Blank AA clients are 6.1% more likely to report non-blank R&D after the forced auditor change. It is interesting to note that the unconditional probability in this sample of reporting non-blank R&D after the AA collapse is 55.4%, suggesting that former Pseudo-Blank AA clients prior to the forced auditor change explain about 11.0% of the unconditional probability of reporting non-blank R&D after the AA demise (0.061/0.554 = 0.110). Column (2) shows that Pseudo-Blank R&D firms, as opposed to blank firms, largely drive the change in reporting behavior. Thus, our evidence is consistent with the idea that Pseudo-Blank R&D firms under-report their R&D expenditure prior to the forced auditor change. Further results indicate that the magnitude of the reported R&D by former Pseudo-Blank AA clients seems substantive subsequent to the forced auditor change. In particular, we find that these firms that switched away from not reporting any information about R&D now disclose, on average, R&D of about 1.8% of total assets after the forced auditor change. This newly reported R&D, post forced auditor change, ranks among the 26th percentile of all firms that report positive (or non-zero) R&D in the universe of Compustat firms.

To further investigate the robustness of this research design, we perform a series of placebo tests using endogenous changes in auditor. We structure our placebo tests as follows: (a) we replace AA with another Big N as the placebo firm; (b) we use 2004–2006 as the placebo event years; (c) we then follow the matching procedures used in the paper to find match sample observations within each placebo event year; and (d) run the logistic regression specification as per Table 5 of the paper. We do this for each of the Big 4 auditors, leading to 4 sets of placebo test results. The untabulated results for each of the placebo tests indicate little evidence that the former Big 4 clients that previously reported blank (or Pseudo-Blank) changed their reporting behavior. We interpret this evidence to suggest our multivariate difference-in-differences research design captures the decision to begin reporting R&D after an exogenous change in auditor.

Classificatory shifting is not the only avenue available to firms that wish to obfuscate their corporate R&D disclosures. For example, firms may use different organizational structures such as joint ventures to hide their corporate R&D. In this test, we investigate whether firms with new joint ventures in a particular year are more likely to be Pseudo-Blank R&D firms and whether firms are more likely to be a Pseudo-Blank firm if they have new joint venture and as the number of new joint ventures increases. Again, we adopt a match sample design by matching firms that entered into new joint ventures with those that did not based on the Eq. (1) dimensions.[13] We report these results in Table 6. We find some evidence that firms with new joint ventures are more likely to be Pseudo-Blank firms (Column (1); z-statistic$=$3.435). In Column (2), instead of examining the existence of new joint ventures, we additionally examine whether the number of new joint ventures is associated with the likelihood of being Pseudo-Blank firms. We find that the number of new joint ventures is also positively associated with the likelihood of firms being Pseudo-Blank firms (z-statistics$=$3.463).

4.5. Simulation analysis

We next perform a series of simple simulations to assess and compare the remedies adopted across disciplines to deal with missing R&D values. Simulation of corporate disclosures of R&D involves creating firm-level innovation data, generating disclosure differences among firms, evaluating the impact of these disclosures choices on empirical tests, and then assessing the impact of the various remedies for these disclosure choices. Our evaluation analysis relies on a representative regression observed in cross-disciplinary research using R&D, namely regressing it against Tobin's Q. We expand on these issues below.

Our simulation analysis centers on attempting to use as much of the Compustat information as possible to allow assignment of firms into the blank R&D category using an empirical model. We only simulated R&D-related variables (namely, R&D rate, number of patents, and blank R&D) and Tobin's Q; with the remaining firm characteristics coming directly from Compustat. To begin the data-generation process, we start with randomly selecting a year from the Compustat database. After excluding missing observations, this led to a base sample of 5,464 firms in our two-digit SIC level analysis.

We simulate corporate R&D, Tobin's Q, and the number of patents using algorithms that include random components, industry effects, and/or other firm characteristics. For instance, we simulate the R&D rate (i.e., R&D expenditure divided by total assets) for each firm in an industry using a combination of random variables from an industry specific beta distribution and the uniform distribution.[14] Thus, the simulation assumes the firms' true R&D are based on an industry distribution. Similarly, we simulate firm-level patents based on firm-level R&D and a random component from the normal distribution. Tobin's Q is simulated using firm-level characteristics and the “runiform” random number generator. Our goal in the simulation exercise is to generate data that approximates their real life data as close as possible.

The next step in the simulation focuses on the disclosure of corporate R&D. Ideally, a complete model of the missing R&D assignment facilitates simulation analysis but requires an accepted analytical model of the R&D reporting process. Alternatively, we generate the missing R&D observations using an empirical model based on the associations between missing R&D and firm characteristics. This approach requires a three-step process: (1) estimate the correlations with firm characteristics and missing R&D in Compustat firms; (2) use these estimated correlations along with a random number to obtain the likelihood score of a firm reporting missing R&D; and (3) assign firms to have missing R&D according to the computed likelihood score and the percentage of firms in their industry with missing R&D in Compustat. We use the following prediction model of blank R&D with the coefficients estimated using Compustat data:[15]

$$\text{Prob(Blank)} = f(\text{ROA, PPE, Leverage, Log(Total Assets), HHI, No. of Patent, random number}) \tag{5}$$

In this context, the blank assignment approach above allows us to use empirical prediction models for assigning firms into the blank R&D category that takes into consideration firm characteristics in their R&D disclosure decisions. The resulting outcome is that the simulated percentage of firms reporting missing R&D varies across industries in a manner consistent with observations in Compustat. Moreover, this approach leads to some firms reporting their true R&D (either positive or zero) and other firms report R&D as missing (concealing either their zero or positive R&D). Importantly for our test procedure described below, for each firm with missing or blank R&D, we know their “true” R&D.

For our simulation exercise, we adopt two main approaches. Our primary approach, described above, uses Compustat data for firm characteristics and an empirical model-based assignment of blank R&D firms to the entire distribution of firms (referred to as the “Entire Distribution Assignment” approach). The second approach focuses on assigning blank R&D among

[13] By construction, half of the matched sample has new joint ventures during the year, with the number of new joint ventures during the year ranging from one to three.

[14] The starting value for each firms is created using the “rbeta (a,b)” random number generator with “a” varying by industry and “b” set at 5. The second step uses the “runiform” random number generator to lower the R&D for a varying degree for each firm.

[15] The model used in the matching process provides a natural alternative prediction specification. However, the matching process centers on a slightly different problem and uses Tobin's Q as one of the predictor variables, which is the dependent variable in the simulation tests. In addition, the scope of our tests in the prior section precludes including patents as a matching variable but it facilitates precision in the assignment of missing R&D in the simulation analysis. Consequently, the two procedures use different variations of predictors with similar underlying concepts to predict or assign firms into the missing R&D category based on their ultimate uses.

Table 6
Matched sample logistic regression tests of the likelihood of firms with new JV being firms with patent activities that reported missing R&D expenditure.

Variables	Pseudo-Blank Firms	
	(1)	(2)
Firms with New JV	1.208***	
	(3.435)	
No. of New JVs		0.968***
		(3.463)
PPE	−0.169	−0.121
	(−0.228)	(−0.163)
Lagged Tobin's *Q*	0.085	0.091
	(0.863)	(0.916)
Firm Age	0.044***	0.043***
	(4.933)	(4.782)
Leverage	−1.880***	−1.883***
	(−2.799)	(−2.833)
Proportion of Firms Reporting R&D in the Industry	0.928	1.002
	(1.145)	(1.240)
Constant	−4.003***	−3.919***
	(−4.614)	(−4.592)
Year & Industry Dummies	Y	Y
Observations	460	460
Log Likelihood	−159.827	−160.261

***$p < 0.01$, **$p < 0.05$, *$p < 0.1$ two-tailed. *z*-Statistics in parentheses are based on firm-level clustering. The sample period is 1980–2006 inclusive due to patent data availability. Matched sample firms are drawn from firms that reported missing R&D expenditure and are matched on the nearest propensity score based on the following logit model:
Treatment group=*f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year Fixed Effects).
Treatment group takes 1 if a firm has entered into new JV during the year; 0 otherwise.
Logistic Regression Model:
Prob(Pseudo-Blank Firms)=*f*(Firms with new JV (or number of new JVs), PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year and Industry Fixed Effects).
All other variables are as defined in Table A1.

the bottom 95% of the positive R&D population (based on our findings of the patent profile of Pseudo-Blank R&D firms in Fig. 1). For both of these simulation approaches, we base our analysis on 1,000 iterations of the simulations.[16]

Our test procedure relies on estimating a simple regression with Tobin's *Q* as the dependent variable and using R&D as an independent variable plus firm size, leverage and industry fixed effects as control variables. We then repeat these basic regressions to evaluate the relation between Tobin's *Q* and corporate R&D using the "true" R&D to establish a baseline comparison. In subsequent regressions we use the missing R&D following the common approaches in accounting, finance, and management. Our analysis is based on the results of repeating the above process 1,000 times. The difference in R^2 between the regressions with "true" R&D and each specified technique provides a direct measure of the benefit of each method for incorporating missing R&D.[17]

In Panel A of Table 7, we report the results of using two-digit industry samples for the "Entire Distribution Assignment" approach. We report the mean and standard deviation for each variable from the 1,000 iterations. Column (1) presents the baseline regression results, showing that the base model with "true R&D" has a mean R^2 of 0.5107. The remaining columns report the analysis after assigning some firms to conceal their positive or zero R&D (as described above) and using the various methods in the literature to incorporate missing R&D. By construction, each of the various approaches to deal with missing R&D will perform worse than the baseline true R&D regressions, because they are based on incomplete data.

[16] In untabulated results, we modify the likelihood score model used to assign blanks by replacing the "No. of Patent" in Eq. (5) with industry specific relations between R&D and the number of patents, and assign blanks to the entire distribution. In another approach, we randomly assign missing R&D using the Compustat data sample. Both of the alternative assignment approaches provide evidence qualitatively consistent with that reported in Table 7.

[17] In particular, for each iteration in the simulation, we compute the percentage of R^2 difference between the model of interest and the model that replaces blanks with zero. We then average these percentages across the 1,000 iterations and report them as the "Relative Improvements" in Table 7. For Table 7 analysis, we reach similar conclusions if we focus on adjusted-R^2 statistics rather than R^2 statistics. In addition, adding joint ventures to the test regressions leads to similar inferences as noted in Table 7.

Table 7
Summary of simulation results (model-based assignment simulation approach).

Variable	Baseline	Blanks as zeros	Blanks as zeros + Blank dummy	Blanks as industry average	Blanks as industry average + Blank dummy	Blanks as zeros + Pseudo-blank + Blank dummies	Blanks as industry average + Pseudo-blank dummy	Blanks as zeros + No. of patent + Blank dummy
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Blank assignment to entire distribution of true R&D								
Mean R&D	3.9012	3.1013	3.5522	3.7940	3.8102	3.5602	3.8237	3.3725
(Std. Dev.)	(0.0682)	(0.0767)	(0.0837)	(0.0832)	(0.0856)	(0.0836)	(0.0833)	(0.0916)
Mean R^2	0.5107	0.3901	0.4023	0.4127	0.4129	0.4037	0.4145	0.4055
(Std. Dev.)	(0.0096)	(0.0104)	(0.0102)	(0.0103)	(0.0103)	(0.0102)	(0.0102)	(0.0101)
Relative Improvements (*Mean% ΔR^2 cf. model 2*)$^{\Delta}$			*3.1402*	*5.8008*	***5.8460***	*3.5017*	***6.2600***	*3.9424*
Panel B: Alternative 1 – Restrict blank assignment to bottom 95 percentile of positive true R&D								
Mean R&D	3.9012	3.2179	3.6041	3.8884	3.8508	3.6128	3.9121	3.4826
(Std. dev.)	(0.0682)	(0.0711)	(0.0789)	(0.0800)	(0.0821)	(0.0789)	(0.0801)	(0.0841)
Mean R^2	0.5107	0.4273	0.4369	0.4445	0.4449	0.4380	0.4456	0.4391
(Std. dev.)	(0.0096)	(0.0101)	(0.0100)	(0.0100)	(0.0100)	(0.0100)	(0.0100)	(0.0099)
Relative Improvements (*Mean% ΔR^2 cf. model 2*)$^{\Delta}$			*2.2524*	*4.0419*	***4.1390***	*2.5108*	***4.2923***	*2.7770*

Reported coefficients are the average estimated coefficients from the 1,000 simulated iterations. Standard deviations are reported in parentheses.
$^{\Delta}$For each iteration we compute the percentage change in R^2 of Models (3)–(8) over Model (2). "Relative Improvements" reports the average of these percentages across the 1,000 iterations. An alternative is to compute the relative improvement based on mean R^2 reported in this table. This approach would lead to relative improvements of 3.1274%, 5.7933%, 5.8447%, 3.4863%, 6.2548%, and 3.9477% for Panel A Models (3)–(8) respectively; and 2.2467%, 4.0253%, 4.1189%, 2.5041%, 4.2827%, and 2.7615% for Panel B Models (3)–(8) respectively.
This table reports results using model-based assignment simulation approach (where non-R&D related variables are drawn from Compustat) where blank R&D firms are assigned based on likelihood scores derived from: Prob(Blank) = *f*(ROA, PPE, Leverage, Log(Total Assets), HHI, No. of Patent). Panel A reports results using blank assignment to the entire distribution of simulated true R&D. Panel B reports results using blank assignment restricted to the bottom 95 percentile of positive true R&D.
All Panels report coefficient estimate of R&D variable based on the following regression:
Tobin $Q = f$(R&D, Leverage, Log(Total Assets), Industry fixed effects).
Column (1) reports the baseline results; Column (2) report results where blank R&D are assumed to be zero; Column (3) assumes blank R&D as zero R&D and added a Blank R&D dummy variable; Column (4) replaces blank R&D by industry average R&D; Column (5) replaces blank R&D by industry average R&D and added a dummy variable for Blank R&D; Column (6) assumes blank R&D as zero R&D and added a dummy variable for Blank R&D and another dummy variable for Pseudo-Blank firms; Column (7) replaces blank R&D by industry average R&D and added a dummy variable for Pseudo-Blank firms; Column (8) assumes blank R&D as zero R&D and added a dummy variable for Blank R&D and log(1 + number of patents).

Column (2) in Panel A shows the results of treating the missing R&D as representing zero R&D; specifically we find that this regression gives a mean R^2 of 0.3901. Thus, the approach of replacing missing R&D with zero allows for roughly an average 24% decline in the R^2 statistic. Column (3) considers the approach of coding the missing R&D values as zero and including a dummy variable to denote blank R&D. The results show that this approach provides a modest average improvement of 3.14%, with a mean R^2 of 0.4023, relative to just coding the missing R&D as zero R&D. Column (4) shows that replacing the missing values with the industry average of reported R&D leads to a mean R^2 of 0.4127, an average improvement of 5.80% relative to treating the missing R&D as zero R&D. Among the three main approaches in the literature, the management approach provides the best results in this simulation.[18]

Next, we consider a few alternative approaches to deal with the missing R&D. First, we consider hybrid approaches to dealing with the missing R&D. Column (5) reports the results of combining the use of a dummy variable for missing R&D with replacing missing values with the industry average of reported R&D, while Column (6) shows the results of using patent data to identify Pseudo-Blank and a dummy for blank firms and replacing the missing values with zero. Column (7) shows the results of using patent data to identify Pseudo-Blank firms and the industry average replacement approach. These tests reveal the hybrid approach of replacing the missing values with the industry average plus a dummy variable to denote Pseudo-Blank R&D provides the greatest improvement relative to treating the missing values as zero in this simulation

[18] In contrast to business research, studies in economics typically drop all the firms with missing R&D from their analysis. A simulation analysis with random R&D disclosure assignment mechanically provides a higher R^2 than any of the approaches used in business research. In the scenario where missing R&D arises from deliberate discretionary choices or corporate structures designed to conceal true R&D, any simulation analysis using the economics approach is quite sensitive to the disclosure assignment process.

analysis (Column (7); an average 6.26% improvement).[19] The next-best alternative for incorporating missing R&D values consists of combining the blank dummy variable with the industry average approach (Column (5); an average 5.85% improvement). Finally, we consider setting the missing values to zero but also including patent data as an additional variable together with a dummy denoting missing R&D. In Column (8), we surprisingly find that this approach fares worse than the industry average approach, with an average improvement of 3.94%. Arguably, due to the limitation in using patent data, replacing missing with zero or the industry average and including a dummy variable provides a tractable approach. Unfortunately, none of the methods we investigate approach the point of fully correcting for this missing R&D problem.

For the alternative simulation approach, instead of assigning blank R&D to the entire distribution based on the likelihood scores, we restrict our blank assignments to the bottom 95% of the positive R&D population according to our findings on the patent profiles of Pseudo-Blank R&D firms in Fig. 1. Other than the restriction imposed on blank assignment, all the test model specifications remain the same. Panel B of Table 7 reports these results. Consistent with Panel A, we find the same qualitative results. Namely, these tests reveal the hybrid approach of replacing the missing values with the industry average plus a dummy variable to identify Pseudo-Blank firms (Column (7)) provides the greatest improvement in this simulation (an average 4.29% improvement relative to coding the blanks as zero R&D). The hybrid approach of combining the industry average and a dummy variable for blank R&D (Column (5)) provides the next best outcome with an average 4.14% improvement over coding blanks as zero R&D.[20]

Yet, the simulation results reported in Table 7 are likely to find that some version of industry average replacements instead of zero replacements (regardless of whether using blank or Pseudo-Blank dummies) is superior because the simulation develops the firms' true R&D based on an industry distribution. Across the simulated samples of Table 7 specifications, our model-based assignment process leads to average true R&D of blank firms of about 47% and 35% of the true R&D in positive reporting R&D firms for Panels A and B respectively. Comparing the results in Panels A and B illustrates that as the true R&D of blank firms gets lower, the performance of models with blank firms coded as zero improves. In untabulated analysis we find that replacing missing with zero works better than using the industry average when true R&D in blank firms is below 20–25% of industry average reported R&D. Unfortunately, the true spending of the missing R&D firms in practice is unobservable, suggesting that researchers should carefully consider their setting and determine whether their results are sensitive to eliminating the blank observations, treating them as zero or treating them with the industry average (both of the latter with blank dummies) and including a Pseudo-Blank dummy variable.

5. Conclusion

In this study, we investigated the discretionary nature of the corporate R&D disclosure decision. Our arguments focus on managers' discretion to decide which expenses are classified as R&D rather than as other types of expenses. Reported R&D represents an internal allocation of outlays for labor and equipment that has the same influence on accounting profit as other pre-tax expenses of the firm. US GAAP provides detailed guidelines for classifying expenses as R&D, but these guidelines still require managerial discretion, suggesting the potential for a disclosure bias in reported corporate R&D. In addition to expense-shifting techniques to provide biased R&D disclosures, firms can also use various organizational structures, such as joint ventures, to create R&D opacity.

To test if missing R&D indicates that firms do not engage in innovation and R&D-type activity, we collected a broad sample of firms and used patent activity to test the discretionary nature of the R&D reporting decision. First, we compared patent activity between firms that report zero R&D and those that failed to report any information on corporate R&D (typically assumed to have zero R&D). In a series of univariate and multivariate tests we consistently found that the non-reporting R&D firms file over 14 times as many patents as firms that report zero R&D expenditure. Moreover, we discovered that these non-reporting firms obtained patents with broader contributions and greater citation breadth than zero R&D firms. All told, the first set of tests uncover that over 10% of the Compustat universe of missing R&D cases display substantial evidence they engage in innovation and R&D activities.

Our second set of tests compared patents in Pseudo-Blank (i.e., firms with missing R&D but that have patent activities) and positive R&D firms. We found that positive R&D firms had substantially more patents than the

[19] We also use another specification that combines the industry average replacement approach with blank and Pseudo-Blank dummies. In untabulated results, we find this specification performs marginally better than Model (7), where the average relative improvements are 6.29% and 4.45% for Table 7 Panels A and B respectively.

[20] Another alternative approach models all firms with missing R&D and without patents as zero R&D firms, providing the conceptual maximum number of zero R&D firms in the population (in this case 48% zero R&D firms in the population). Alternatively, patent activity in positive R&D firms and research on patenting decisions provides additional estimates of the maximum percentage of zero R&D in the population (around 43% zero R&D firms). In our simulation analysis, the 1,000 simulated samples range from roughly 38–42% zero R&D which mimics the later alternative in conjunction with the assumption that non-reporting R&D firms are like to under-patent their innovation relative to positive R&D firms. Redoing the simulation analysis using both the 43% and 48% estimates for the mean of the percentage of zero R&D firms in the population provides qualitatively similar results to those reported in Table 7.

Pseudo-Blank R&D firms, with Pseudo-Blank firms corresponding to the bottom 90th to 95th percentiles of patents in positive R&D firms. Interestingly, we also find that the individual patents in Pseudo-Blank R&D firms appear exhibit roughly an 8% lower influence than those found in positive R&D firms with patent activity. These results are inconsistent with the idea that blank R&D firms have insignificant R&D. Focusing on forced auditor changes and joint ventures we then evaluated the roles of discretionary reporting choices and corporate structure in understanding Pseudo-Blank R&D firms. We found significant evidence that unexpected auditor changes led firms to alter their R&D reporting. These results are most pronounced in the set of Pseudo-Blank R&D firms. Additional tests revealed a positive relation between joint ventures and Pseudo-Blank R&D, suggesting that organizational structure also plays a role in R&D disclosures.

Finally, we perform a simple simulation analysis to compare the different approaches found across the business literature on R&D to deal with missing R&D. The results indicate that interpreting missing R&D to indicate zero R&D can lead to substantive bias in empirical tests. Regardless if one treats missing as zero or replaces with industry average, our Monte Carlo simulations highlight the importance of always including a blank dummy variable to denote missing R&D (and a Pseudo-Blank dummy if available). Further, our results suggest that researchers should ensure that their results are not sensitive to the treatment of R&D expenditures in any empirical work where R&D plays a substantial role. While our simple simulations strictly focus on cross-sectional tests, the underlying intuition implies that panel data could also benefit from a hybrid approach, alternating between treating missing with zero and using either the industry or historical firm average R&D and dummy variables to denote missing and Pseudo-Blank R&D firms. Perhaps there are also some cases (e.g. studies that examine just a single industry) where it could be surmised that the blank firms are more like their industry counterparts with respect to R&D.

In sum, this analysis investigated whether missing or blank R&D indicates a lack of innovation and R&D activities. Using patent data, we found that non-reporting R&D firms file more patents and more influential patents than firms that report zero R&D. Moreover, Pseudo-Blank R&D firms, relative to positive R&D firms, obtain individual patents with broader contributions, greater citation breadth, and lengthier competitor discovery periods despite having fewer patents. We interpret these results to suggest that missing R&D does not imply that these firms have no substantive innovation or R&D-type activities. Further tests indicated that Pseudo-Blank R&D is associated with both discretionary reporting choices and the structure of corporate activities. One interpretation of these results implies that missing R&D is not an accidental outcome but instead, a deliberate firm choice. Finally, we provided simple simulations that suggest researcher care is needed in dealing with missing R&D.

Appendix A

See Tables A1 and A2.

Table A1
Variable definitions.

Variable names	Variable definitions
R&D Expenditure	R&D expenditure divided by total assets
Report R&D	Indicator variable: 1 if a firm reported zero or positive R&D expenditure; 0 otherwise
PPE	Property, plant and equipment divided by total assets
Lagged Tobin's *Q*	Lagged of Tobin's Q, measured as (market value of equity + book value of total debt) divided by total assets
Firm Age	Age of the firm as appear on CRSP
Leverage	Total debt divided by total assets
Proportion of Firms Reporting R&D in the Industry	Proportion of firms reported zero or positive R&D within 2-digit SIC
Pseudo-Blank R&D Firms	Indicator variable for Pseudo-Blank R&D firms; 0 otherwise, where Pseudo-Blank firms are firms reporting blank R&D expenditure but have patent activity
Patent Citations	Log of the average lifetime patent citations
Generality	1 – concentration of citing patents' technology fields
Originality	1 – concentration of cited patents' technology fields
Citation Lag	Log of the average time lag between patent citations and the originating patents (in years)
Patent Claims	Log of the average number claims made by a firm's patents
Log(Total Assets)	Log of total assets
Capital Expenditure	Capital expenditure divided by total assets
ROA	Operating income divided by total assets
Institutional Ownership	13-F institutional ownership
Analyst Following Indicator	Indicator variable for firms with analyst following; 0 otherwise
No. of Analysts Following	The number of analyst following a firm
Dispersion	Analyst forecast dispersion scaled by absolute value of consensus forecast

Table A2
Tests of patent counts and characteristics.

Panel A: Matched Sample tests between pseudo-blank firms and positive R&D firms

Variables	Applications (1)	Granted (2)	Patent citations (3)	Generality (4)	Originality (5)	Citation lag (6)	Patent claims (7)
Pseudo-blank R&D	−1.430***	−1.267***	1.517***	0.034***	0.271***	0.577***	1.719***
	(−6.133)	(−5.997)	(21.387)	(5.409)	(16.584)	(12.887)	(24.873)
PPE	−0.383	−0.612	0.025	0.019	0.082*	0.154	0.321*
	(−0.541)	(−0.918)	(0.131)	(1.110)	(1.944)	(1.182)	(1.786)
Lagged Tobin's Q	0.045	0.021	0.124***	−0.001	0.014**	−0.069***	0.068***
	(0.641)	(0.337)	(4.149)	(−0.385)	(2.073)	(−3.470)	(2.594)
Firm Age	0.050***	0.051***	0.023***	0.003***	0.005***	0.022***	0.023***
	(8.006)	(9.059)	(12.667)	(18.161)	(13.615)	(19.851)	(13.041)
Leverage	0.311	0.436	0.130	0.006	0.022	−0.096	0.124
	(0.967)	(1.559)	(0.904)	(0.436)	(0.632)	(−1.006)	(0.903)
Proportion of Firms Reporting R&D in the Industry	2.297**	1.763**	0.612***	0.030	0.095**	0.216*	0.529***
	(2.567)	(2.424)	(2.982)	(1.579)	(1.962)	(1.709)	(2.895)
Constant	−1.236***	−1.001***	−0.860***	−0.221***	−0.622***	−0.614***	−0.661***
	(−3.085)	(−2.789)	(−5.102)	(−11.559)	(−13.482)	(−5.408)	(−4.165)
Year & Industry Dummies	Y	Y	Y	Y	Y	Y	Y
Observations	9,072	9,072	9,072	9,072	9,072	9,072	9,072
Log Likelihood	−166,606	−161,565	−13,821	−1,247	−5,544	−11,362	−14,211

Panel B: Matched Sample Tests between Pseudo-Blank Firms and Positive R&D Firms with Patent Activity

Variables	Applications (1)	Granted (2)	Patent citations (3)	Generality (4)	Originality (5)	Citation lag (6)	Patent claims (7)
Pseudo-Blank R&D	−1.831***	−1.714***	−0.278***	−0.038***	−0.055***	−0.097***	−0.166***
	(−8.864)	(−9.262)	(−7.250)	(−7.317)	(−5.297)	(−3.350)	(−5.642)
PPE	−0.321	−0.565	−0.510***	0.003	−0.003	0.114	−0.096
	(−0.637)	(−1.247)	(−4.382)	(0.195)	(−0.107)	(1.186)	(−1.079)
Lagged Tobin's Q	0.075	0.019	0.089***	−0.003	0.004	−0.077***	0.025
	(1.217)	(0.282)	(4.591)	(−1.292)	(0.654)	(−5.442)	(1.592)
Firm Age	0.048***	0.047***	0.004***	0.002***	0.002***	0.014***	0.004***
	(7.911)	(9.019)	(4.621)	(13.674)	(6.362)	(18.318)	(7.020)
Leverage	0.620*	0.560*	0.050	−0.004	0.023	0.027	0.099
	(1.719)	(1.716)	(0.564)	(−0.355)	(0.919)	(0.407)	(1.462)
Proportion of Firms Reporting R&D in the Industry	1.714**	1.537***	−0.042	0.015	−0.002	0.028	−0.100
	(2.497)	(2.781)	(−0.331)	(1.094)	(−0.057)	(0.334)	(−1.137)
Constant	−0.255	−0.069	1.959***	−0.090***	−0.120***	0.448***	2.197***
	(−0.541)	(−0.168)	(20.102)	(−6.070)	(−3.675)	(5.899)	(27.948)
Year & Industry Dummies	Y	Y	Y	Y	Y	Y	Y
Observations	8,228	8,228	8,228	8,228	8,228	8,228	8,228
Log Likelihood	−171,576	−158,839	−12,787	67.233	−4,387	−9,340	−11,972

$^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$ two-tailed. Test statistics in parentheses are based on firm-level clustering. The sample period is 1980–2006 inclusive due to patent data availability
Columns (1) and (2) use the following Poisson Regression Model:
No. of Patents=*f*(PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year and Industry Fixed Effects)
No. of Patents is either No. of Patent Applications or Patents Granted; Blank R&D is an indicator variable 1 for firms reporting blank R&D expenditure; 0 if firms reporting zero R&D expenditure. All other variables are as defined in Table A1.
Columns (3) and (7) use the following Tobit Regression Model:
Dependent variable=*f*(Pseudo-Blank Firms, PPE, Lagged Tobin's Q, Firm Age, Leverage, Proportion of Firms Reporting R&D in the Industry, Year and Industry Fixed Effects).
Dependent variables are: Generality, Originality, Citation Lag, and Patent Claims. All variables are as defined in Table A1.

References

Abadie, A., 2005. Semiparametric difference-in-differences estimators. Review of Economics and Statistics 72, 1–19.
Abril, P., Plant, R., 2007. The patent holder's dilemma: buy, sell or troll? Communications of the ACM 50, 37–44.
Andersen, R., Duru, A., Reeb, D., 2009. Founders, heirs, and corporate opacity in the United States. Journal of Financial Economics 92, 205–222.
Arundel, A., 2001. The relative effectiveness of patents and secrecy for appropriation. Research Policy 30, 611–624.
Arundel, A., Kabla, I., 1998. What percentage of innovations are patented? Empirical estimates for European firms. Research Policy 27, 127–141.
Bessen, J., 2008. The value of U.S. patents by owner and patent characteristics. Research Policy 37, 932–945.
Bushee, B., 1998. The influence of institutional investors on myopic R&D investment behavior. The Accounting Review 73, 305–333.
Cahan, S., Zhang, W., 2006. After Enron: auditor conservatism and ex-Andersen clients. The Accounting Review 81, 49–82.
Chen, S., DeFond, M., Park, C., 2002. Voluntary disclosure of balance sheet information in quarterly earnings announcements. Journal of Accounting and Economics 33, 229–251.
Cohen, W., R., Nelson, J., Walsh, 2000. Protecting their intellectual assets: appropriability conditions and why U.S. manufacturing firms patent (or not). NBER Working Paper No. 7552.

Dye, R., 1985. Strategic accounting choice and the effects of alternative financial reporting requirements. Journal of Accounting Research 23, 544–574.
Dye, R., Sridhar, S., 1995. Industry-wide disclosure dynamics. Journal of Accounting Research 33, 157–174.
Fishman, M., Hagerty, K., 1990. The optimal amount of discretion to allow in disclosure. Quarterly Journal of Economics 105, 427–444.
Gallini, N., 2002. The economics of patents: lessons from recent U.S. patent reform. Journal of Economic Perspectives 16, 131–154.
Gambardella, A., Giuri, P., Luzzi, A., 2007. The market for patents in Europe. Research Policy 36, 1163–1183.
Gaver, J., Gaver, K., 1993. Additional evidence on the association between the investment opportunity set and corporate financing, dividend, and compensation policies. Journal of Accounting and Economics 16, 125–160.
Gigler, F., 1994. Self-enforcing voluntary disclosures. Journal of Accounting Research 32, 224–241.
Hall, B., Oriani, R., 2006. Does the market value R&D investment by European firms? Evidence from a panel of manufacturing firms in France, Germany, and Italy. International Journal of Industrial Organization 24, 971–993.
Hall, B., A. Jaffe, and M. Trajtenberg, 2001. The NBER patent citation data file: lessons, insights, and methodological tools. NBER Working Paper 8498.
Hall, B., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. RAND Journal of Economics 36, 16–38.
Healy, P., Hutton, A., Palepu, K., 1999. Stock performance and intermediation changes surrounding sustained increases in disclosure. Contemporary Accounting Research 16, 485–520.
Healy, P, Palepu, K., 2001. Information asymmetry, corporate disclosure, and the capital markets: a review of the empirical disclosure literature. Journal of Accounting and Economics 31, 405–440.
Horwitz, B., Kolodny, R., 1980. The economic effects of involuntary uniformity in the financial reporting of R&D expenditures. Journal of Accounting Research 18, 38–74.
Kleinknecht, A., Reijnen, J., 1991. More evidence on the undercounting of small firm R&D. Research Policy 20, 579–587.
Kogut, B., 1998. Joint ventures: theoretical and empirical perspectives. Strategic Management Journal 9, 319–332.
Kothari, S.P., Laguerre, T., Leone, A., 2002. Capitalization versus expensing: evidence on the uncertainty of future earnings from capital expenditures versus R&D outlays. Review of Accounting Studies 7, 355–382.
Kothavala, K., 2003. Proportional consolidation versus the equity method: a risk measurement perspective on reporting interests in joint ventures. Journal of Accounting and Public Policy 22, 517–538.
Kortum, S., 1993. Equilibrium R&D and the patent-R&D ratio: U.S. evidence. American Economic Review 83, 450–457.
Lanjouw, J., Schankerman, M., 2001. Characteristics of patent litigation: a window on competition. Rand Journal of Economics 32, 129–151.
Lanjouw, J., Schankerman, M., 2004. Patent quality and research productivity: measuring innovation with multiple indicators. The Economic Journal 114, 441–465.
Layne-Farrar, A., Lerner, J., 2011. To join or not to join: examining patent pool participation and rent sharing rules. International Journal of Industrial Organization 29, 294–303.
Lerner, J., 2002. 150 years of patent protection. American Economic Review 92, 221–225.
Lerner, J., Wulf, J., 2007. Innovation and incentives: evidence from corporate R&D. Review of Economics and Statistics 89, 634–644.
Leuz, C., Verrecchia, R., 2000. The economic consequences of increased disclosure. Journal of Accounting Research 38, 91–124.
Lo, K., 2010. Materiality and voluntary disclosures. Journal of Accounting and Economics 49, 133–135.
Markman, G., Espina, M., Phan, P., 2004. Patents as surrogates for inimitable and non-substitutable resources. Journal of Management 30, 529–544.
McVay, S., 2006. Earnings management using classification shifting: An examination of core earnings and special items. The Accounting Review 81, 501–531.
Moser, P., Nicholas, T., 2004. Was electricity a general purpose technology? Evidence from historical patent citations. American Economic Review 94, 388–394.
Reitzig, M., 2003. What determines patent value – insights from the semiconductor industry. Research Policy 32, 13–26.
Schankerman, M., 1998. How valuable is patent protection: Estimates by technology field. Rand Journal of Economics 94, 77–107.
Trajtenberg, M., Henderson, R., Jaffe, A., 1997. University versus corporate patents: a window on the basicness of invention. Economics of Innovation and New Technology 5, 19–50.
Verrechia, R., 1983. Discretionary disclosure. Journal of Accounting and Economics 5, 179–194.