Decoupling Research from Development*

Ma Hou-Ning, David M. Reeb

November 24, 2022

Reporting research and development together implies their allocation provides limited insight to investors. We construct and corroborate unique measures of development and research to evaluate this aggregation. These measures combined correlate with reported R&D expenditures at 99% (industry) and 81% (firm). In the individual measure validation tests, our research measure correlates with scientific publications (77%), while the development measure correlates with patents (71%). Our R&D measures cover 76% of NYSE firms, providing broader coverage than patents (30%), new product announcements (20%), and reported expenditures (46%).

Using these measures, we investigate whether a firm's choice to emphasize either research or development provides relevant information to investors. We find that development and research differ significantly in predicting future earnings and cash flows. The allocation between research and development also helps predict whether a firm discloses its R&D expenditures. Development intensity, but not research intensity, forecasts increased product market concentration. Investors require higher risk premia for research relative to development activity. Overall, our results are inconsistent with the hypothesis that development and research load similarly in investor-prediction models. It is puzzling that firms do not explicitly disaggregate R&D disclosures to give investors insight into their allocation decision.

JEL Classification: M40, M41, O30

Keywords: R&D Disclosure, R&D Informativeness, R&D Allocation, Missing R&D, Disaggregating Disclosures

*Email: houning.ma@austin.utexas.edu and dmreeb@nus.edu.sg. We thank Oliver Li, Yupeng Lin, Xinjie Ma, Tim Martens, Peter Pope, Clemens Slaim, Laura Starks, Stephen Taylor, Sheridan Titman, Robert Vollon, Rachel Zhang, Wanli Zhao, and seminar participants at Bocconi University, the 2022 FIRN Corporate Conference, and the National University of Singapore for their comments and suggestions. We thank Philip Mattera at Good Jobs First for providing government subsidy data.

1 Introduction

Research and development expenditures are value relevant to investors (Lev and Sougiannis, 1996). Firms decide their allocation between development and research activities based on their internal capabilities and outside opportunities. US firms, however, typically report the sum of research and development expenditures, suggesting that the allocation between these two activities is immaterial or too costly to disclose separately. Theory suggests that managers consider providing research and development details by weighing the benefits of lower capital costs versus the damages from giving this information to competitors (Bhattacharya and Ritter, 1983). Others emphasize managers consider the impact of detailed disclosures on their personal reputations in evaluating the benefit to shareholders (Beyer and Dye, 2012). Potentially mitigating the demand for R&D expenditure details, investors can draw inferences from voluntary disclosures of the firm, such as patents and R&D-Narratives (Glaeser et al., 2020; Huang et al., 2021). Still, R&D spending is one of the most investor-relevant measures of corporate innovation (Ettlie, 1998).

We investigate whether research gives investors similar information as development to evaluate the potential benefits, if any, of disaggregating these expenditures. Historically, merging research and development into a single measure arose because of the difficulty of determining and tracking these concepts (Nix and Nix, 1992). Today, regulators, stock exchanges, and standard setters emphasize that their disclosure recommendations and requirements provide decision-useful information to shareholders and creditors (Barth, 2008). Standard setters typically group items together that give investors similar inferences about future cash flows, separating items that load differently in investor prediction models (Schipper, 2007).

Decoupling research from development would inform capital market participants about the manager's selections or relative emphasis on research versus development. We develop and test six hypotheses to assess the speculative gains of disaggregating research and development expendi-

tures for investors. Our empirical analysis compares the relative predictive powers of development and research to evaluate whether they load similarly in investor prediction models. These tests reveal whether a firm's allocation between research and development predicts future earnings, R&D disclosure decisions, product market concentration, or investor risk premia demands.

To examine their informational comparability, we create separate measures of research and development using comprehensive occupational classifications in firms' online job postings. Our analysis relies on job postings in listed firms. Burning Glass Technologies (BGT) provides detailed online job posting data which reflects individual firms' demand for human capital. Specifically, we use standard occupational classifications (SOCs) of job postings to categorize firms' human capital demand for research and development. We measure a firm's annual research intensity by the number of job postings and wages in 23 different SOCs that focus on scientists; similarly, we estimate a firm's development intensity by the number of job postings and wages in 22 SOCs that focus primarily on engineers. We gauge development and research flows separately and create size-scaled R&D intensity measures based on their cumulative job posting. Occupational wage data from the Census Bureau allows us to develop supplemental, dollar-based research and development measures.

A weakness of our job-posting-based approach is that it only captures the human capital component of firm-level research and development activities. Thomson (2017) reports that capital comprises 14% of R&D, with employees representing the remaining 86%. Others highlight that labor and capital in R&D endeavors are complements, with investments in human capital and equipment moving together (Funke and Strulik, 2000). Although employees comprise a substantial portion of R&D, a vulnerability is that job postings could provide a poor proxy of R&D.

We validate both our job-posting approach to measuring R&D and our decomposition of these SOCs into research and development occupations. Our firm-level corporate research and develop-

ment measures exhibit an 81.1% correlation with reported R&D spending when combined into a single metric among firms that disclose it.[1] Plotting the cumulative distributions of reported R&D and our measures of R&D reveals they have strikingly similar patterns, suggesting that job postings closely track R&D spending.

We also corroborate our STEM-occupational decomposition into research and development in several ways. First, we compare our distinct development and research measures to those based on patents and scientific publications (Arora et al., 2021). While only available for 39% of firms, patents and research publications are highly associated with our development and research measures. Specifically, development postings exhibit a high correlation with patenting (71%), while research postings overlap closely with scientific publications (77%). Empirical cumulative density plots (CDPs) reveal comparable patterns between the job-posting and patent-publication measures of development and research.

Second, we use the stand-alone references to "research" or "development" (outside of the phrase research and development) in company 10-K filings to compute a research-to-development ratio for each firm. We then compare this 10-K-based ratio to the same ratio using the posting-based measures. While our job-posting measure provides more comprehensive coverage than the 10-K approach, comparing their cumulative density plots shows substantial overlap. Decile-based-assignment groupings reveal an 89% correlation between the 10-K and job-posting ratios of research-to-development.

We evaluate specific firms and industries to further substantiate our SOC division into separate development and research measures. Most firms hire scientists and engineers, but the relative number of job postings in research and development indicates strong firm patterns. For instance, while Micron Technologies and Deere & Co hire research scientists, they repeatedly exhibit some

[1] The correlations between our measures and reported R&D increase to 84.3% if we restrict the sample to postings that require graduate degrees. We use this approach for robustness testing but note that 40% of job postings do not include education requirements, creating a missing data problem when restricting to graduate degrees.

of the highest demand for engineers. Similarly, Pfizer and Monsanto engage many engineers but are consistently in the top 10 firms seeking to hire research scientists.

Next, we analyze industry patterns of research and development allocations to ratify our decomposition. Pharmaceutical companies emphasize research intensity, while automotive firms stress development intensity. Graphing the relative demands for development and research talent at the industry level (NAICS-2-digits level) shows substantial variation in research versus development allocations across different sectors of the economy.

Our empirical tests examine the putative benefits of decoupling research from development. Our first set of tests focuses on whether a firm's allocation between research and development relates to their decision to report R&D spending. R&D expenditures are available for roughly 46% of the Compustat universe, but many firms have substantial innovation without disclosing it (Koh and Reeb, 2015). Leveraging our measures, we analyze R&D activities across reporting and non-reporting R&D firms. Using a linear probability model with industry-by-year fixed effects, we find that development intensity positively correlates with R&D reporting propensities. In contrast, research-heavy firms are less likely to report their R&D expenses than their industry peers. We interpret these findings to suggest that managers view research and development as having different informational properties, influencing the decision to report R&D.

Tax credit usage provides a second method to test for potential differences in research versus development and gives another avenue to corroborate our STEM-occupation assignments. Our tests reveal that development- and research-intensive firms each garner different R&D tax credits. Across various specifications, we consistently find that research-intensive firms attract basic research grants, whereas development-intensive firms exhibit greater propensities in receiving development grants. Thus, along two distinct dimensions (incentives to report and tax credits), research and development intensities give differing information to investors.

Our next set of tests seeks to assess the benefits of decoupling research and development by evaluating their relative predictive power on future firm performance. We document that research and development intensities provide significantly different predictions about future cash flows, accounting earnings, market capitalization, and Tobin's Q. For instance, research is positively associated with future profits but not development intensity. These tests provide evidence inconsistent with the idea that research and development offer similar inferences to outside investors.[2] From these results, it is difficult to derive a one-size-fits-all formula for how firms should allocate research and development spending. However, these tests provide evidence of investors' benefit from disclosing this allocation decision.

Still, a potential concern is that these tests only include information about the firm that operational ratios can easily codify or capture and used as controls in the analyses (or with firm-fixed effects). In additional testing, we extend this analysis using industries as the unit of analysis. We aggregate our firm-level measures to the industry level. We find that research and development provide significantly differing predictions about future industry profits and cash flows.

Our next test relies on our industry-level measures to assess the relevant informational content of research versus development intensity in the broader market. We measure industry concentration using a Herfindahl-Hirschman Index (Hirschman, 1964). Development intensity strongly predicts industry concentration. In contrast, research intensity foretells greater competition. One interpretation of these results centers on how allocating research and development within an industry provides investors with life-cycle information. They also highlight the informational differences between research and development.

Capital market return predictability offers another avenue to examine whether research and de-

[2]The short window of our analyses suggests the performance results likely stem from selection. Investors care about information that reveals managers' perceptions about future firm performance. While agency, production, and proprietary cost arguments offer different interpretations of these results, they all highlight the distinction between research and development to investors.

velopment provide similar information. The International Financial Reporting Standards (IFRS) discussion on research and development emphasizes that they have different risk profiles and valuation uncertainties. We focus on cross-sectional return models to investigate potential risk premia differences between research and development. Across multiple specifications, we find that research intensity predicts excess returns while development intensity does not. These tests provide evidence inconsistent with the hypothesis that development and research have similar risk profiles.

Our measures capture corporate demand for research and development employees, but we do not observe if the firm hires someone. Prominent firms may fill their jobs more efficiently than obscure ones or even post them anonymously. Dividing firms into S&P 500 and non-S&P 500 allows inferences about whether prominent firms' posting practices distort our development and research measures. We also find that anonymous job openings do not have higher percentages of R&D occupations than those with company names. Another potential threat to these results is heterogeneity in measurement errors across the two measures (i.e., research vs. development). Additional tests, with differing scaling variables, various subsets of development, and focusing only on firms with reported R&D provide similar results. We document investor-relevant informational differences in research and development in a series of additional robustness tests.

Our analysis makes three substantive contributions. First, we build on the literature on the disclosure of R&D (Chen et al., 2017; Merkley, 2014). Prior research establishes that firms often engage in substantive R&D but do not report it (Koh and Reeb, 2015). This work often relies on sparse measures of corporate innovation to detect firms who engage in R&D without disclosing it (Pseudo-R&D firms), allowing inferences about their existence but not their prevalence (Koh et al., 2018). Our research and development intensity measures suggest we could classify half of the firms without reported R&D as Pseudo-R&D firms (compared to just 5% in-patent-based tests). Reporting R&D depends on a firm's allocation between research and development. These

Pseudo-R&D firms receive 35.8% of the R&D tax credits awarded to public firms over the last decade.

Second, under US GAAP, firms disclose the combined amount of R&D in their financial statements. Development differs dramatically from research in predicting firm profits, cash flows, market value, and excess stock returns. Accounting scholars and regulators recognize the potential for differences in the informativeness of research and development and the benefit of increasing disclosure detail (Francis et al., 2008; Chen et al., 2015). Our analysis speaks to the magnitude of this benefit, suggesting it is substantive across a wide range of tests and specifications. Consequently, it is puzzling that some managers do not provide detailed disclosures in the notes or supporting text about their allocation between research and development.

The frictions preventing disclosure of the precise division between research and development likely center on the costs of more detailed disclosures. Standard accounting information systems often track research and development separately, so production costs seem an unlikely explanation. Chen et al. (2017) report that IFRS rules on expensing research and capitalizing development lead firms to increase voluntary disclosures about innovation, implying the proprietary costs of separating them could also be low. While we rarely see US firms explicitly disclose their development and research allocations, we observe substantial variation in their discussions about the firm. For instance, in their recent 10-K filings, Pfizer defines themselves as a "research-based" firm, while Micron Technologies describes their R&D as focusing on the "development of industry leading memory." Voluntary disaggregation of R&D occurs in many forms and depths. Arguably, standardizing the allocation disclosure of research and development could level the playing field between large and small investors.

Third, our results relate to the literature on measuring corporate innovation. We develop replicable measures for both research and development based on job postings that are densely pop-

ulated for both public and private firms. As for the listed firms, our measures cover 3,958 US-headquartered firms, comprising over 70% of listed firms in NYSE, AMEX, and NASDAQ. In contrast, patents and scientific publications combined cover 39% of the market, new product announcements 30%, and reported R&D of about 46%. These distinct research and development measures also allow researchers to study how firms engage in and monitor each component. Setting aside our decomposition of R&D into separate categories, our approach to measuring R&D provides new information. Our combined job-posting-R&D intensity measure allows researchers to study corporate innovation in a broad range of firms, such as private firms or listed firms that do not report R&D. At a minimum, our measures provide a second measure of R&D, complementing reported expenditures by providing an additional signal about corporate innovation to investors or empirical researchers (see Holmstrom (1979) on the Informativeness Principle).

2 Aggregating R&D

Conceptually, research is often described as basic or applied (Mansfield, 1980), while development often focuses on either exploitation or exploration (Swift, 2016). One of the earliest references to R&D reporting in the US occurs in the Federal Reserve Bulletin of 1917, allowing firms to designate it as a deferred charge (Nix and Nix, 1992). Alexander (1954) observes that a general definition of what they then called research and experimental expenditures did not exist. In 1954, tax legislation allowed firms to deduct R&D costs, regardless of whether they capitalized or expensed this investment (Blake, 1959). However, there were no standard definitions of research or development, allowing firms to use their own classifications (Sougiannis, 1994). In 1974, the *US Statement of Financial Standards No. 2: Accounting for Research and Development Costs* (SFAS2) mandated the disclosure of research and development and explicitly defined their measurement (Elliott et al., 1984). They described research as an "investigation aimed at the dis-

covery of new knowledge," while development focuses on the "translation of research findings" into products (SFAS #2).

Disclosure rules often focus on parsimonious presentations of the data. Penman (2016) observes that financial statements should be cohesive and convey critical information utilitarianly, with disaggregation to enhance communication about future cash flows. Guay et al. (2016) emphasize that complex financial statements negatively affect the information environment of investors. Others emphasize communicating helpful information to shareholders and creditors by separating items that behave differently in forecasting cash flows, have differing effects on valuation multiples, or exhibit different risk factor weights (Schipper, 2007). In short, an essential standard for reporting subtotals or the components of an expenditure, such as spending on both research and development, is whether it helps investors evaluate and judge corporate activity.

2.1 R&D Disclosure

Extensive literature documents that firms face a trade-off with disclosures about research and development. Detailed disclosures inform investors about the firm's prospects so that they can decide about investing at the cost of revealing this information to competitors. Beyond proprietary costs to the firm or agency costs from the manager, another issue faced by regulators in evaluating required disclosures: the cost of producing the information (Gonedes et al., 1976). Reporting development and research separately or providing a breakdown of the two charges in financial statement notes arguably have limited incremental production costs. This information is often collected and classified internally; a critical issue is whether separately disclosing research from development meets the goal of enhancing the usefulness of information to investors.

One approach to evaluating whether research expenditures provide similar information to investors as development spending centers on whether they give managers differing incentives to

report R&D. Koh and Reeb (2015) document that firms often have substantial R&D expenditures they choose not to report, labeling these firms as "Pseudo R&D firms." Firms may be motivated to report R&D expenditures when specializing in development tasks. In comparison, research-focused companies may face higher proprietary costs disclosing their R&D expenditure because of greater spillover risks (e.g., Arrow 1962, Dasgupta and David 1994). Research expenditures could also have differing agency costs than development if they raise expectations about future managerial performance. Consequently, managers balance a trade-off between investor valuation, proprietary costs, and agency concerns in deciding whether to disclose R&D.

We conjecture that research-focused firms potentially find R&D disclosure less attractive because of higher proprietary costs. In contrast, development-focused firms can better protect their knowledge through patenting and trade secrets. Yet, combining research and development into a single disclosure implies the allocation between research and development should be unrelated to the decision to disclose it. The decision to report or not report R&D leads to our first hypothesis.

H1: The relative mix of research and development is unrelated to the decision to report R&D.

2.2 R&D Tax Credits

Tax credits provide another avenue to investigate the aggregation of research and development into a single measure. Studies in accounting and economics examine the value relevance and impact of R&D tax credits. Brown and Krull (2008) report that R&D tax credits can offset managerial incentives to reduce R&D spending to meet earnings benchmarks. Others emphasize that R&D tax credits substantially impact future R&D spending and profits (Rao, 2016; Thompson, 2017). Investors observe reported R&D expenditures, seeking to understand future R&D and innovation. One aspect for investors is making inferences about the firm's tax subsidies and credits to appreciate how it motivates and interacts with other incentives for R&D spending (Brown and Krull, 2008;

Cheng et al., 2021). Combining corporate expenditures in research and development into a single disclosure implies they provide similar inferences about the firm's tax credits and subsidies. Alternatively, tax subsidies in research-intensive firms could differ from those in development-intensive firms, making a combined disclosure less informative about the firm's taxes and incentives. Our second hypothesis focuses on research versus development-intensive firms' innovation tax subsidies.

H2: Research-intensive and development-intensive firms have similar R&D tax subsidies.

2.3 R&D and Future Performance

The joint disclosure of research development implies these items are highly correlated and that separating would not aid outside investors. One concern of investors is the ability to predict future firm cash flows and accounting profits. A broad spectrum of studies across accounting, economics, and strategy report that R&D is related to future company performance. Ciftci and Cready (2011) report that R&D intensity is positively associated with future earnings. Boler et al. (2015) emphasize that cheaper R&D costs stimulate technological innovation and higher future performance. Similarly, management scholars report that R&D leads to higher future profitability (Hill and Snell, 1988).

Our research question centers on whether research expenditures provide similar predictive power as development expenditures. Aggregating R&D together builds on the notion that they offer comparable predictive power for future firm performance. We exploit our separate research and development measures to test if they give similar information to outside investors. This disaggregation benchmark on providing similar inferences about future profits and cash flows leads to our third hypothesis.

H3: Research and development equally predict future firm performance.

A potential confounding issue is that firms provide other information about corporate innovation activity that is difficult to codify (Erkens, 2011), making separate disclosures redundant. While firm fixed-effects can partially address this information environment concern, another approach centers on industry-based tests (Fairfield et al., 2009). Firm-level operational ratios are easy to assess, but other information can be challenging to collate and measure for empirical research. These difficulties in codifying measures can also affect investors undertaking industry-level analysis. Consequently, another approach to testing information relevance between development and research focuses on industry-level performance predictions. This leads to our fourth testable hypothesis:

H4: Industry research and development equally predict future industry performance.

Our next hypothesis focuses on changes in industry composition. Glaeser and Landsman (2021) observe that product market competition influences the disclosure of corporate innovation. Conceptually, development activities emphasize better utilization and exploitation of existing knowledge. If development activities promote firms' market power, one could expect development-intensive industries to become more concentrated over time. In contrast, arguments for aggregating research and development expenditures into a single disclosure suggest they should have similar predictive power for future industry concentration. Thus, our fifth hypothesis centers on the relative predictive power of research and development in foretelling changes in industry composition.

H5: Industry research intensity and development intensity equally predict future industry composition.

2.4 R&D and Stock Returns

Comparing the relative predictive power of research and development for future profits and cash flows provides evidence of their similarity or dissimilarity. Yet, research and development could

also have different risk profiles, with research focusing on relatively broad targets and development centered on narrower product-related targets. Chambers et al. (2002) emphasize that R&D leads to higher future excess stock returns because investors demand a risk premium for this activity. We argue that investors potentially view research as riskier than development, leading to higher future excess returns. IFRS rules on research and development explicitly highlight this potential risk difference in developing their R&D disclosure rules. Items that load similarly in excess stock return models motivate grouping these items for disclosure. Consequently, our final testable hypothesis on the appropriateness of aggregating R&D into a single measure focuses on comparing their future excess stock returns.

H6: Research and development equally predict future excess stock returns.

3 Measuring Research and Development

This section elaborates on how we construct distinct occupation-based corporate research and development measures. We exploit granular occupational classifications in online job postings to gauge a firm's particular human capital demand for development and research roles. Combining our separate research and development measures into a single measure reveals that it encompasses 76% of domestic NYSE firms, allowing us to gauge corporate innovation activities across a broad spectrum of companies and industries.

Our primary sample is the intersection of Compustat and the Burning Glass Technologies (BGT) data. Burning Glass Technologies is a labor market information company that examined over 40,000 online job boards and company websites to assemble job postings. The BGT data arguably contains the entire universe of posted online jobs (Hershbein and Kahn, 2018) and covers detailed information about the US labor market. Next, we use a fuzzy matching algorithm coupled with hand-matching to link Burning Glass data with Compustat. Overall, we match over 30% of

job postings with company names to US-listed firms. The proportion of job postings in Compustat firms is similar to other studies, such as Deming and Kahn (2018).

Using standard occupational classifications (SOCs), we construct distinct research and development measures from firms' job postings. We rely on the science, technology, engineering, and math (STEM) occupational groups described by the Bureau of Labor Statistics (BLS) to construct our research and development occupation directory.[3] The BLS classifies 100 STEM occupations that span a broad range of research and development roles, but also includes technicians, post-secondary teachers, and sales positions.

We take several steps to evaluate and select the research and development roles. First, we exclude the sales and teacher occupations, such as Sales Representatives, Sales Engineers, and Chemistry Teachers. Second, we remove all technicians and drafters. We note our results are robust to including these occupations. Third, we exclude architectural-related occupations such as Architectural Managers and Landscape Architects. Fourth, we exclude analysts, administrators, and user support jobs in the computer category. Last, we exclude Actuaries, Conservation Scientists, Foresters, Epidemiologists, Geoscientists, and Hydrologists. Again, we run extensive robustness checks, finding that our inference is insensitive to these selection steps.

We classify 23 occupations as research scientists and 22 occupations as development engineers. For example, research roles include Animal Scientists, Biochemists and Biophysicists, Chemists, Food Scientists, and Materials Scientists. Development occupations include Biomedical Engineers, Chemical Engineers, Materials Engineers, Petroleum Engineers, Software Developers, and Web Developers. In addition, we exploit the education requirement data in the job postings and classify jobs that require STEM Ph.D. degrees as part of the research category.

Table A.1 lists these research and development occupations' descriptions with their six-digit

[3] Source: https://www.bls.gov/opub/btn/volume-3/an-overview-of-employment.htm.

standard occupational classification (SOC) codes.[4] Across all 37.2 million full-time job postings, research occupations comprise 1.96% of total job postings, while development occupations account for 6.06% of total job postings. We examine several alternative classifications of job postings for development and research, finding robust results. In separate tests, we propose several alternative measures for firm-level R&D capacities (for combined measures of R&D). For instance, firms' demand for STEM master's and Ph.D.s closely correlates with their reported R&D expenditures.

We create both job-count and dollar-based measures of development and research. Focusing on job counts, we describe corporate research and development activities' flow, stock, and intensity measures. First, we aggregate the number of research and development job postings at firm-year levels to form the flow measures. The flow measures reflect firms' demand for research scientists (engineers) at each point in time and are the building block for the stock and intensity measures. Second, we apply a 15% depreciation rate to aggregate the flow variables into the stock measures.[5] The equations below specify the detailed computation of each stock variable for research and development.

$$
\begin{aligned}
&Research\ Stock_{i,t} = \\
&\qquad No.\ of\ Scientists_t + 0.85 \times No.\ of\ Scientists_{t-1} + \\
&\qquad 0.85^2 \times No.\ of\ Scientists_{t-2} + 0.85^3 \times No.\ of\ Scientists_{t-3} \ldots
\end{aligned} \tag{1}
$$

[4]Research occupations concentrate on SOC family 19 (Life, Physical, and Social Science Occupations). Development occupations span the SOC family 17 (Architecture and Engineering Occupations). Both research and development contain six-digit SOC codes from SOC family 15 (Computer and Mathematically Occupations).

[5]The R&D stock literature typically uses a 15% depreciation rate (Glaeser et al., 2020; Tseng, 2022). We evaluate the robustness of our findings using different depreciation rates and discounting windows. Other approaches include the varying the deprecation rate across each industry using employee turnover rates.

$$Development\ Stock_{i,t} =$$

$$No.\ of\ Engineers_t + 0.85 \times No.\ of\ Engineers_{t-1} + \tag{2}$$

$$0.85^2 \times No.\ of\ Engineers_{t-2} + 0.85^3 \times No.\ of\ Engineers_{t-3} ...$$

Finally, we construct research and development intensity measures by scaling the R stocks and D stocks with total assets/employment to account for the influence of firm size.

$$Research\ Intensity_{i,t-1} = \frac{Research\ Stock_{i,t-1}}{Total\ Assets_{i,t-1}\ or\ Total\ Employment_{i,t-1}} \tag{3}$$

$$Development\ Intensity_{i,t-1} = \frac{Development\ Stock_{i,t-1}}{Total\ Assets_{i,t-1}\ or\ Total\ Employment_{i,t-1}} \tag{4}$$

We also compute dollar-based flow, stock, and intensity measures for both development and research. Using occupational wage data from the Bureau of Labor Statistics at the occupation-state-year level, we calculate the dollar amount investments for both research and development human capital.[6] We compute the dollar-based stock measures of research and development (Equations 1 and 2) by replacing the number of scientists and engineers with the dollar value of a firm's science and engineering job postings each year. Our empirical results are based on both job-count and dollar-based research and development measures. We denote the dollar-based intensity measures as *Research-Dollar Intensity* and *Development-Dollar Intensity*. The job-count measures center on the number of job openings in each category, while the dollar-based allow for remuneration

[6]In particular, we multiply each job opening by its surveyed occupational wage reported in the Occupational Employment and Wage Statistics (OEWS) program. We use the 75th percentile wage to account for the potential wage premium of listed companies. We note that our results are robust to the choices of median and mean wage statistics. Because the OEWS program fails to provide wage estimates for certain occupations, we obtain dollar measures for about 34.5 million job postings. As a result, our dollar-based measures cover 25,685 firm-year observations.

differences across individual R&D positions.

4 Data and Sample

Our primary sample uses 37.2 million full-time online job postings by US-headquartered firms. We exclude part-time and contractor postings in our primary analysis.[7] Our matching steps between BGT and Compustat are three-fold. First, we clean and standardize BGT and Compustat company name strings. Second, we use fuzzy matching between BGT and Compustat company names and manually examine the matched results to avoid false-positive matches. Third, we focus on the unmatched Compustat companies and perform another round of manual searches in the BGT database.

Our matching process identified 4,964 public firms, containing both US-headquartered and cross-listing firms. Because our BGT data concentrates on job vacancies in the United States, we include only US-headquartered companies.[8] After this step, our sample consists of 4,330 firms and 37.2 million full-time job openings from 2010 to 2020. Our sample covers nearly all major firms in the US, comprising over 93% of the total market capitalization in the Compustat universe (US-headquartered firms). In comparison, among these 4,330 firms, only 2,239 firms consistently report their R&D expenditures (including both positive and zero R&D figures). Our human capital-based measures potentially double the R&D measures coverage for public firms.

Next, we remove observations with missing market capitalization, total assets, revenue, or employment to facilitate our regression analyses. We also require firms with total assets of over 10 million to enter the sample. After these steps, we arrived at a firm-year panel that covers 25,935 observations, covering 3,976 unique firms. In some of our tests, we include control variables such

[7] We include the 3 million part-timers in a robustness test, finding similar results.

[8] Nevertheless, our methodology extends to a broader range of international firms.

as cash flow growth and leverage. After removing missing items in these control variables, we arrive at our final sample, which contains 25,772 firm-year observations and spans 3,958 unique companies.[9]

Last, we construct a monthly stock return panel using the CRSP database. Following Fama and French (1992), we focus on US-based common stocks (share codes 10 and 11) and use only NYSE, AMEX, and NASDAQ firms. We adjust for delisting returns following Shumway (1997). Our CRSP sample contains 214,462 unique firm-month observations.

4.1 Descriptive Statistics

Table 1 presents the summary statistics. We winsorize all variables at the 1st and 99th percentile to reduce the influence of extreme outliers.[10] The table features firm characteristics such as accounting and stock market variables. A typical firm posts 1,136 online job vacancies each year. Firms' average research stock is 62 (scientist postings), and average development stock is 219 (engineer postings). Focusing on the dollar-based measures, we find the research-to-development ratio declines to 1:3.

4.2 Measure Validation

This section takes various approaches to validate our human-capital-based measures against several corporate innovation markers. First, we explore how industries' R&D human-capital-demand maps to their R&D expenditures. Our mapping exercise spans various levels of industrial specifications, such as NAICS-2-digits (19 industries), Fama-French 48 industries, and SIC-3-

[9] We perform predictive regressions with granular fixed effects in most of our analyses. This requires us to lag our independent variables, research intensity and development intensity as well as control variables by one year, resulting in most firm-level regression tables reporting 21,718 firm-year observations.

[10] In an untabulated test, we perform sensitivity analysis on the winsorization process. Our results hold for additional winsorizing of our dependent variables at the 2.5%, 5%, and 10% levels.

digits (199 industries).[11] Starting with the job counts, Figure 1A reports that our human-capital-based measures exhibit a 99.43% correlation with reported R&D spending at the NAICS-2-digits level. Focusing on Fama-French 48 and SIC-3-digits industries, our human capital-based measures correlate with reported R&D expenditure at 93.31% and 92.75%, respectively. Across different settings, our measure exhibits high correlations with aggregate R&D spending at the industry level.

Next, we expand the industry analyses at the firm level, focusing on the positive R&D firms (about 46% of our sample). Figure 1B presents the cumulative distributions plots of firm-level R&D jobs and R&D expenditures. Notably, we find that companies' R&D job postings closely map with reported R&D expenses throughout the distribution. While the baseline measure correlated with reported R&D by 81.1%, some of our alternative measures correlate with reported R&D by 84.3%. Jointly, we document that our human-capital-based R&D measure closely parallels firms' reported R&D activities.

We undertake a series of tests to validate our SOC-based division into research and development occupations. Table 2 compares our distinct development and research measures against different benchmarks. Statement of Financial Accounting Standards *(SAFS No. 2)* mandates that research should create new knowledge, while development primarily translates and explores the existing knowledge. Arora et al. (2021) focus on the differential outputs of corporate development and research activities, arguing that research activities primarily produce scientific articles, while development leads to more patents. Although firms' disclosure choices influence patents and scientific journal publications, they provide a practical approach to corroborate our distinct research and development measures against different benchmarks.

We obtain patent and scientific publications data from Arora et al. (2021).[12] Their data focuses on US manufacturing firms that report at least one year of R&D expenditure and have a patent. This

[11] There are 247 sub-industries; however, 48 are solely populated with missing R&D firms.

[12] Source:https://www.openicpsr.org/openicpsr/project/120550/version/V1/view?path=/openicpsr/120550/fcr:versions/V1/data&type=folder

data overlaps with our BGT sample between 2010 to 2015. Overall, 20.4% of our sample exhibit these innovation markers. Table 2 reports that our job-count research measure exhibits a 77% correlation with scientific publications. Interestingly, its correlation with patents is significantly lower at 50.6%.In contrast, our job-count development measure is highly correlated with patents (71.1%) but exhibits a low correlation (51.6%) with journal publications. These differences are statistically significant. We find similar patterns with the dollar-based research and development measures.

We extend the exercise in Table 2 by demonstrating the complete cross-firm distributions of research postings (scientific publications) and development postings (patents). Figure 2A depicts the empirical CDPs of firm-level research job openings and scientific publications among firms with reported R&D and patents. Similarly, Figure 2B plots the CDPs for development job openings and patents. Moreover, Figure A.2 presents the density plots using these innovation markers available for patenting firms.

We employ a second approach to validate our classifications into research and development occupations. Specifically, we use the text in each firm's 10-K filings to gauge their relative emphasis on research or development. Using the Edgar 10-K filings, we separately count the stand-alone references to either "research" or "development" after excluding the phrase "research and development." We compute the Research-to-Development ratio (10-K text) as the total stand-alone "research" divided by the sum of stand-alone "research" and stand-alone "development." Similarly, we construct the Research-to-Development ratio using the job postings in research occupations and development occupations. Notably, both measures range from 0 to 1 and capture firms' differential emphases between research and development. Moreover, they are based on two independent sources, allowing us to compare firm-specific innovation characteristics.

Figure 3A maps the distributions of text-based and posting-based Research-to-Development

ratios, showing a close relationship across ranks despite differences in levels. These two Research-to-Development ratios often assign firms to the same decile based on their research/development intensities. Figure 3B demonstrates the overlaps in their density functions. However, because of data sparsity issues (many firms do not mention research or development in their annual reports), we note that the 10-K measure has a narrow range and a heavy left tail. In contrast, our job posting measure exhibits comprehensive coverage. Grouping firms into deciles based on their text-based and posting-based Research-to-Development ratios, we find an 89% correlation across these two measures.

Overall, the validation exercises in this section deliver two key messages. First, when combined, our job posting research and development measures closely track reported R&D expenditures, providing over 81.1% cross-firm correlation and 99.4% cross-industry correlation. Our job-posting-based method gives R&D information for both firms that report and do not report R&D expenditures. Second, our distinct research and development measures exhibit close correlations to different innovation metrics. In particular, we find that research job postings closely parallel scientific publications and development job postings highly correlate with patents.[13]

4.3 Missing R&D Firms

This section documents an important disclosure consequence under the current combined reporting scheme. Recent literature finds that firms exhibit substantial differences in their R&D reporting practices, with about half of US firms not reporting R&D spending. Meanwhile, researchers also document considerable innovation heterogeneities within firms that do not report their R&D spending (i.e., differences in patenting). We investigate three related questions about

[13] We use a logarithm transformation to ease concerns on firm size differentials. We note that the firm-level correlations drop to 79.1% while the industry-level correlations drop to 93.2% (Fama-French 48) and 92.5% (SIC3). In all of our regression analyses, we use intensity measures to orthogonalize the potential influence of size effects.

firms' R&D practices and disclosure choices. First, we investigate the heterogeneities in R&D human capital demand among the missing R&D firms. In particular, what fractions of their job openings are scientists and engineers among firms that do not report their R&D expenses? Evaluating their human capital compositions, do missing R&D firms exhibit similar R&D intensities as firms that reported zero R&D expenses? Or do missing R&D firms also demand a material fraction of scientists and engineers? Second, we investigate whether the current combined R&D reporting scheme creates distinct reporting incentives between research- and development-focused firms. Finally, we test whether the combination of research and development intensity (both job-count and dollar-based) predict future firm performance without reported R&D.

We begin by evaluating innovation-related human capital in the missing R&D firms. First, we use firms that report positive and zero R&D expenditures as benchmarks because these firms represent two distinct cases of corporate R&D engagements. R&D intensities in these two types of firms provide reference points gauging the relative R&D capacities for missing-R&D firms. Second, we apply a 5% threshold to parse missing R&D firms into the pseudo-R&D firms. Pseudo R&D indicates that these firms demand a persistent proportion of R&D personnel but do not report their R&D expenses separately.

In contrast, unreported-zero firms neither engage nor report R&D activities, potentially exhibiting similar R&D activities as the reported-zero firms. To validate our classification of pseudo-R&D firms, we combine our research and development measures to assess materiality. We find that R&D in the pseudo firms predicts future performance (see Table A3 in the appendix).

Panel A of Table 3 summarizes the R&D posting fractions in all four types of firms. We present the mean of the firm-level R&D intensities, measured as the ratio of a firm's R&D job postings to its total job postings. We require firms to have total job postings (the denominator) over 100 to ensure a reliable ratio computation. Among the 2,574 firms with over 100 job postings, 1,021 are

positive R&D firms, 522 are pseudo-R&D firms (based on a 5% cutoff), 542 are unreported-zero firms, and 309 are reported-zero R&D firms.[14] We find pseudo-R&D firms consistently demand comparable R&D personnel to a significant fraction of the reporting R&D firms. Across the board, we find that a typical pseudo-missing firm requires 12.72% of scientists and engineers from their total job postings. Notably, their innovation-related human-capital demands differ significantly from zero R&D firms and are even higher than the bottom tercile of positive-reporting-R&D firms.

In contrast, these unreported-zero firms exhibit similar R&D intensities to firms that explicitly reported zero R&D. Panel B of Table 3 highlights that pseudo-missing firms are prevalent using the dollar-based measure of R&D. Pseudo-R&D firms represent 49% to 56% of missing R&D firms. The number of firms with significant R&D job postings which do not report R&D spending is substantial.[15]

Hypothesis 1 predicts that research- and development-intensive firms have similar R&D disclosure practices. Table 4 finds that companies' research and development intensities are associated with different R&D disclosure practices. This pattern is especially striking when we jointly test the coefficient loadings in research and development intensities. In sum, a one-standard-deviation higher research intensity reduces firms' R&D reporting likelihood by 1.6%. However, a one-standard-deviation higher development intensity increases firms' R&D reporting likelihood by 3.4%. These results are within estimates derived from a tight-fixed-effect model, which controls for unobservable heterogeneities at industry-by-year levels. Moreover, we find that the results in Table 4 are robust to different choices of scaling factors in computing the R&D intensity measures. This evidence rejects the hypothesis that research- and development-focused firms have

[14] Among the 2,681 firms with over 100 BGT job postings, we exclude 107 firms that report R&D spending inconsistently (fail to report in some years). Therefore, this sample has 2,574 firms that either consistently reported R&D spending or never reported R&D spending between 2010 to 2020.

[15] One consideration is that our classification of software engineers potentially inflates development intensity in software and information firms. To ease this concern, we exclude the software industry (NAICS code 51) and find pseudo-R&D firms represent 47% to 53% of missing-R&D firms.

similar likelihoods of reporting R&D. One potential interpretation is that the proprietary costs of development spending numbers are lower than research expenditures.

4.4 R&D Tax Credits and Subsidies

Hypothesis 2 predicts that both research and development-intensive firms exhibit similar R&D tax credits or subsidies. Nevertheless, it is common to find that the federal and state governments award research and development grants to different firms. Conceptually and empirically, given firms' specializations in various fields, their propensities to garner each type of government subsidy might vary. Therefore, the awarding history of government R&D subsidies provides another opportunity to evaluate whether research and development carry different information. We exploit the Good Jobs First dataset that contains federal and state government subsidies. After merging with our sample, we obtained 43,219 comprehensive subsidy histories for 1,455 unique firms, of which 7,513 subsidies are classified as R&D tax credits/grants.

We perform keyword searching in subsidies' program names and descriptions to identify research from development grants. In specific, we classify grants that contain "basic research," "applied research," "scientific/science," and "invention and innovation" as *research grants*, with the rest categorized as *development grants*. Overall, we identify 2,198 research-focused grants and 5,315 development-focused grants.

Table 5 investigates whether companies' differential research and development intensities affect their R&D subsidy results. Our firm-level regressions in Table 5 consistently find that higher research capacities positively correlate with the propensity and dollar amount of basic research grants. In contrast, development intensity predicts firms' likelihood of obtaining development grants. Firms with high development capacities are more likely to garner development grants than their industry peers. Yet, these tests have some of the lowest R-squares in all of our tests, so the

inferences are less pronounced. Still, these tests provide additional corroboration of the research and development measures and provide evidence that is inconsistent with similar informativeness from research and development.

5 Predictive Specifications

This section introduces the predictive regressions that estimate the individual effects of research and development on firm performance. Equation (5) below delineates our baseline firm-level regression for testing hypothesis 3.

$$Firm\ Performance_{i,t} = Research\ Intensity_{i,t-1} + Development\ Intensity_{i,t-1} + \delta_i + \lambda_t + X_{i,t-1} + \varepsilon_{i,t} \tag{5}$$

The dependent variable $Firm\ Performance_{i,t}$ represents a host of performance-related variables such as ROA, cash flows and Tobin's Q. i denotes individual firms and t denotes each year. $Research\ Intensity_{i,t-1}$and $Development\ Intensity_{i,t-1}$ as the key explanatory variables. Notably, $Research\ Intensity_{i,t-1}$and $Development\ Intensity_{i,t-1}$ establish low correlation with each other,[16] alleviating the multicollinearity concerns. Vector $X_{i,t-1}$ includes our control variables: $Leverage_{i,t-1}, ROA_{i,t-1}, ROA\ Growth_{i,t-1}, Cash\ Flow_{i,t-1}, Cash\ Flow\ Growth_{i,t-1}, CAPX/ME_{i,t-1}$. We include year fixed effects λ_t to absorb spurious relationships along the time dimension and use firm fixed effects δ_i to obtain the within estimates.[17] We cluster the standard errors at Fama-French 48 industry level.

Our next two tests explore the differential effects between research and development at the

[16]Our primary research and development intensity measures are scaled by firms' total assets. Their correlation is 26.7%.

[17]An alternative specification uses industry fixed effects, allowing cross-sectional variations when estimating the coefficients of interests

industry level (Hypotheses 4 and 5). We aggregate our firm-level data into an industry-year panel (based on Fama-French 48 industries). Equation (6) specifies our first industry-level regression:

$$Industry\ Performance_{j,t} = Research\ Intensity_{j,t-1} + Development\ Intensity_{j,t-1} + \delta_j + \lambda_t + X_{j,t-1} + \varepsilon_{j,t} \tag{6}$$

Equation (7) specifies our second industry-level regression:

$$Concentration_{j,t} = Research\ Intensity_{j,t-1} + Development\ Intensity_{j,t-1} + \delta_j + \lambda_t + X_{j,t-1} + \varepsilon_{j,t} \tag{7}$$

We investigate product market concentration to test hypothesis 5. We compute the Herfindahl index based on companies' total sales to proxy for product market concentration. Industry-level $Research\ Intensity_{j,t-1}$and $Development\ Intensity_{j,t-1}$ are ratios of industry total research or development job postings scaled by total industry assets. To test hypothesis 4, we compute market capitalization, cash flows, and ROA for the Fama-French 48 industries. Again, industry-level $Research\ Intensity_{j,t-1}$and $Development\ Intensity j_{i,t-1}$ are ratios of industry total research or development job postings scaled by total industry assets. Control variables are the number of firms in the industry, average industry revenue, and industry profitability.[18] We include industry fixed effects δ_j and year fixed effects λ_t. Finally, we rely on Fama-Macbeth regressions with monthly stock returns to investigate risk premia differences for research and development.

[18]Excluding the software or information industries provides qualitatively similar results. Adding control variables, including R&D spending, patents, firm size, or index status also leads to similar inferences.

6 Predictive Evidence: Research versus Development

The combined reporting of R&D expenses in the financial statement implies that development and research activities contain similar enough information to outweigh the costs of reporting them separately or with more detail. To evaluate this combination, we empirically examine the independent predicting power of research and development measures over an array of investor value-relevant variables. First, we focus on firm performance measures, such as future market valuation, ROA, Tobin's Q, and cash flows. Next, we conduct industry-level analyses, gauging the differential effects of research and development on industry dynamics. Finally, we focus on the cross-section of stock returns and investigate the different slopes of research intensity and development intensity using Fama-Macbeth regressions. To facilitate the quantitative interpretation of the point estimates, we normalize research and development intensities with means of 0 and standard deviations of 1.

6.1 Future Performance

Panel A of Table 6 tests if research intensity and development intensity exhibit differential predictive power for investor value-relevant variables such as future market size, Tobin's Q, ROA, and cash flows. This panel uses the job-count R&D measures. In column 1, we find consistent patterns that a one-standard-deviation increase in research intensity is associated with 8.0% higher future market capitalization. Column 2 shows that a one-standard-deviation increase in research foretells a 0.135 higher Tobin's Q (6.5% of the mean). The same increase in development predicts 0.027 higher Tobin's Q, yet the coefficient is statistically insignificant. In column 3, ROA is 0.9-percentage-points higher with a one-standard-deviation increase in research, while development is again uncorrelated with future profitability. Similarly, column 4 reports that research but not development predicts future cash flows.

Panel B of Table 6 repeats the firm performance tests using the dollar-based research and devel-

opment intensity measures. Again, we find inconsistent evidence across multiple firm performance metrics that research and development activities cast similar effects on firms' future performance. In sum, this evidence indicates that research and development provide distinct inferences about future firm performance, rejecting hypothesis 3.

To test hypothesis 4, we aggregate our firm-level measures to the industry level. Table 7 reports the results for both job-count and dollar-based measures. At the industry level, research positively predicts increases in market capitalization, Tobin's Q, ROA, and cash flows. In contrast, industry development is either negatively correlated with these performance measures (Tobin's Q and Cash Flow) or has zero predictive power. Again, we draw similar inferences; research and development provide divergent predictions about future performance, rejecting hypothesis 4.

We empirically test the industry composition hypothesis (H5) by constructing industry concentration proxies. We use Fama-French 48 industries as our unit of analysis. To measure industry concentration, we use the Herfindahl-Hirschman Index (HHI). Our concentration measures require a meaningful size of firms in each industry. Therefore, we focus on industries with over 15 firms in this analysis. As a result, we exclude industries such as Textiles, Fabricated Products, Shipbuilding/ Railroad Equipment, Defense, Shipping Containers, and Tobacco Products. None of the excluded industries are innovation active. Firms in these industries account for 1.34% of the Compustat universe.

Table 8 presents the industry concentration effects. We document that industry development intensity, but not research intensity, correlates with an increase in product market concentration. Focusing on the job-count tests, we find that a one-standard-deviation increase in industry development intensities corresponds to 147 higher Herfindahl index, which is equivalent to 16.7% of the sample mean (unconditional mean of the Herfindahl index is 899). Higher development intensities foster market power at the industry level. In contrast, a one-standard-deviation increase in industry

development intensities corresponds to 124 lower Herfindahl index (13.8% of the mean). Overall, Table 8 reports the distinct effects of development and research at the industry level, rejecting hypothesis 5.

6.2 Excess Stock Returns: Risk Premia Differences

This section examines whether development and research exhibit differing risk premia using the cross-section of stock returns. We use Fama-Macbeth regressions with monthly excess stock returns to estimate the various slopes on research and development (and again for the dollar-based intensity measures). We control for a battery of return-relevant firm characteristics, such as market-to-book ratio, CAPX/ME, asset growth, net stock issues, ROA, cash flow, ROA growth, and cash flow growth. In an untabulated subsample test using R&D reporting firms, we additionally control for R&D spending growth, log of firm age, and patents.

Panel A of Table 9 presents our capital market evidence with the job-count-based measures. Across the board, we find consistent evidence that research, but not development, explains the cross-section of stock returns. Quantitatively, a one-standard-deviation higher research yields 0.163% to 0.173% monthly excess returns, equivalent to 1.956% to 2.076% annualized returns. Controlling for the Fama-French 48 industry dummies, research is associated with 1.4% future, annualized returns. These excess loadings are significant at the 1% level, with t-statistics often larger than 3. Panel B of Table 9 repeats the risk premia analysis using dollar-based research and development intensity measures. Again, our tests suggest investors perceive research risks differently than development risks.

6.3 Robustness Tests

A potential concern with job postings is that prominent firms may have different hiring outcomes than their less prominent peers. Of particular concern is that either prominent or obscure firms drive the results in some unanticipated manner. Tables A4 and A5 examine the potential heterogeneous effects of well-known and less prominent firms in the firm performance and risk premia tests. We use S&P 500 inclusion to delineate prominent firms. In both tables, we find that research differs substantially from development. Yet, we note that development in S&P 500 firms has more predictive power than among non-S&P 500 firms in Table A4. In contrast, the risk premia differences in Table A5 show similar patterns in both S&P 500 and non-S&P 500 firms. These results highlight that selection plays an important role in the predictive power of research and development intensities.

Using both job-count and dollar-based methods, we document that the distinct effects between research and development are robust to selecting scaling factors. Tables A6 and A7 consider alternative scaling approaches in measuring R&D intensities. Table A6 uses the natural logarithms of research and development intensities. Table A7 replaces total assets with total employment in Compustat as the scaling variable. We find similar results in both sets of tests. Specifically, we find that research and development provide differing predictions about future firm performance.

Tables A8 and A9 repeat the alternative scaling approaches to the risk premia tests. Table A.8 repeats the analysis with total employment (Compustat variable EMP) as the denominator in computing research and development intensities. Table A.9 exploits the natural-algorithm-transformed research and development measures. Notably, Table A.8 finds an excess annualized return of 1.90% (with industry dummies) and 2.56% (without industry dummies), while Table A.9 demonstrates that the loading on research ranges from 1.75% (with industry dummies) and 2.39% (without industry dummies). Both alternative specifications yield higher excess returns than our

baseline constructs. In sum, this evidence rejects the hypothesis that research and development provide similar inferences about future stock returns, rejecting hypothesis 6.

US GAAP treats software R&D differently. Table A10 repeats the performance regressions after excluding the software industry. We find marginally stronger results after excluding NAICS code 51 relative to the main results reported in Table 6. Table A11 adds reported R&D expenditures as a control variable. Again, we find very similar results to the main results reported in Table 6. Results in Table A11 reveal that understanding the allocation between R&D provides additional explanatory power after including total R&D expenditures. Collectively, these tests suggest that our research and development measures provide investors with additional information about corporate innovation.

7 Conclusion

We examine whether research and development give shareholders and creditors similar information about value-relevant topics to evaluate reporting them without disaggregation. We use online job-posting data to develop distinct research and development intensity measures based on firms' human capital demand. Importantly, our R&D measures do not rely on managerial disclosure choices and allow for separate research and development analyses. Our human-capital-based measures capture over 76% of US-listed NYSE firms and even shed light on innovation activities in private firms. Moreover, our analyses make an important methodological point: researchers can use firm demand for unique human capital to gauge corporate capabilities.

Next, we document that firms' differential development or research focus affects their R&D disclosure decisions. Research-intensive firms are more inclined to keep R&D expenditures private. Our analysis indicates that roughly half of the missing R&D firms actively engage in R&D. Despite the non-reporting choices, these companies exhibit substantial human capital demand for

innovation-related occupations, exceeding the intensities of the bottom tercile of positive-R&D firms.

Moreover, we investigate whether corporate research and development have similar predictive powers over firm performance. We find they differ substantially in forecasting future market valuation, Tobin's Q, ROA, and cash flows. They also exhibit different risk profiles; research intensity predicts stock returns but not development intensity. Last, we document that industry development intensity fosters product market concentration.

This study makes several contributions. First, we document that the decision to report R&D depends on a firm's allocation between research and development capacities. Pseudo-R&D firms are prevalent, surpass the bottom tercile of reporting R&D firms, and receive 35.8% of the R&D tax credits. Second, US GAAP mandates reporting research and development as a combined measure. Our analysis shows that investors could often benefit from firms disaggregating these expenditures. It is perplexing that firms do not voluntarily disclose their relative mix of research and development spending. Finally, we develop replicable measures for research and development based on job postings that allow researchers to investigate corporate innovation in a broad range of firms, including private firms or public ones that do not report R&D spending.

References

[1] Donald C. Alexander. Research and experimental expenditures under the 1954 code. *Tax Law Review*, 10(4):549–568, 1954.

[2] Ashish Arora, Sharon Belenzon, and Lia Sheer. Knowledge spillovers and corporate investment in scientific research. *American Economic Review*, 111(3):871–98, 2021.

[3] Mary E. Barth, Wayne R. Landsman, and Mark H. Lang. International accounting standards and accounting quality. *Journal of Accounting Research*, 46(3):467–498, 2008.

[4] A. Beyer and R.A. Dye. Reputation management and the disclosure of earnings forecasts. *Review of Accounting Studies*, (17):877–912, 2012.

[5] Sudipto Bhattacharya and Jay R. Ritter. Innovation and Communication: Signalling with Partial Disclosure. *The Review of Economic Studies*, 50(2):331–346, 04 1983.

[6] John Blake. *Accounting Standards*. Pitman Publishing, 1959.

[7] Esther Ann Boler, Andreas Moxnes, and Karen Helene Ulltveit-Moe. R&D, international sourcing, and the joint impact on firm performance. *American Economic Review*, 105(12):3704–39, 2015.

[8] Jennifer L. Brown and Linda K. Krull. Stock options, r&d, and the r&d tax credit. *The Accounting Review*, 83(3):705–734, 2008.

[9] D. Chambers, R. Jennings, and R.B Thompson. Excess returns to r&d-intensive firms. *Review of Accounting Studies*, (7):133–158, 2002.

[10] Ester Chen, Ilanit Gavious, and Baruch Lev. The positive externalities of ifrs rd capitalization: enhanced voluntary disclosure. *Review of Accounting Studies*, (22):677–714, 2017.

[11] Shuping Chen, Bin Miao, and Terry Shevlin. A new measure of disclosure quality: The level of disaggregation of accounting data in annual reports. *Journal of Accounting Research*, 53(5):1017–1054, 2015.

[12] C. S. Agnes Cheng, Peng Guo, Chia-Hsiang Weng, and Qiang Wu. Innovation and corporate tax planning: The distinct effects of patents and r&d*. *Contemporary Accounting Research*, 38(1):621–653, 2021.

[13] Mustafa Ciftci and William M. Cready. Scale effects of r&d as reflected in earnings and returns. *Journal of Accounting and Economics*, 52(1):62–80, 2011.

[14] David Deming and Lisa B. Kahn. Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369, 2018.

[15] John Elliott, Gordon Richardson, Thomas Dyckman, and Roland Dukes. The impact of sfas no. 2 on firm expenditures on research and development: Replications and extensions. *Journal of Accounting Research*, 22(1):85–102, 1984.

[16] David H. Erkens. Do firms use time-vested stock-based pay to keep research and development investments secret? *Journal of Accounting Research*, 49(4):861–894, 2011.

[17] John E. Ettlie. R&d and global manufacturing performance. *Management Science*, 44(1):1–11, 1998.

[18] Patricia M. Fairfield, Sundaresh Ramnath, and Teri Lombardi Yohn. Do industry-level analyses improve forecasts of financial performance? *Journal of Accounting Research*, 47(1):147–178, 2009.

[19] Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2):427–465, 1992.

[20] Jennifer Francis, Dhananjay Nanda, and Per Olsson. Voluntary disclosure, earnings quality, and cost of capital. *Journal of Accounting Research*, 46(1):53–99, 2008.

[21] Michael Funke and Holger Strulik. On endogenous growth with physical capital, human capital and product variety. *European Economic Review*, 44(3):491–515, 2000.

[22] S. Glaeser, J. Michels, and R.E. Verrecchia. Discretionary disclosure and manager horizon: evidence from patenting. *Review of Accounting Studies*, (25):597–635, 2020.

[23] Stephen A. Glaeser and Wayne R. Landsman. Deterrent Disclosure. *The Accounting Review*, 96(5):291–315, 02 2021.

[24] Nicholas J. Gonedes, Nicholas Dopuch, and Stephen H. Penman. Disclosure rules, information-production, and capital market equilibrium: The case of forecast disclosure rules. *Journal of Accounting Research*, 14(1):89–137, 1976.

[25] Wayne Guay, Delphine Samuels, and Daniel Taylor. Guiding through the fog: Financial statement complexity and voluntary disclosure. *Journal of Accounting and Economics*, 62(2):234–269, 2016.

[26] Brad Hershbein and Lisa B. Kahn. Do recessions accelerate routine-biased technological change? Evidence from vacancy postings. *American Economic Review*, 108(7):1737–72, 2018.

[27] Charles W. L. Hill and Scott A. Snell. External control, corporate strategy, and firm performance in research- intensive industries. *Strategic Management Journal*, 9(6):577–590, 1988.

[28] Albert O. Hirschman. The paternity of an index. *American Economic Review*, 54(5):761–762, 1964.

[29] Bengt Holmstrom. Moral hazard and observability. *The Bell Journal of Economics*, 10(1):74–91, 1979.

[30] Rui Huang, Leye Li, Louise Yi Lu, and Hai Wu. The impact of the leahy-smith america invents act on firms' r&d disclosure. *European Accounting Review*, 30(5):1067–1104, 2021.

[31] Ping-Sheng Koh and David M. Reeb. Missing R&D. *Journal of Accounting and Economics*, 60(1):73–94, 2015.

[32] Ping-Sheng Koh, David M. Reeb, and Wanli Zhao. Ceo confidence and unreported r&d. *Management Science*, 64(12):5725–5747, 2018.

[33] Baruch Lev and Theodore Sougiannis. The capitalization, amortization, and value-relevance of R&D. *Journal of Accounting and Economics*, 21(1):107–138, 1996.

[34] Edwin Mansfield. Basic research and productivity increase in manufacturing. *The American Economic Review*, 70(5):863–873, 1980.

[35] Kenneth J. Merkley. Narrative disclosure and earnings performance: Evidence from rd disclosures. *The Accounting Review*, 89(2):725–757, 2014.

[36] Paul E. Nix and David E. Nix. A historical review of the accounting treatment of research and development costs. *The Accounting Historians Journal*, 19(1):51–78, 1992.

[37] Dasgupta Partha and Paul A. David. Toward a new economics of science. *Research Policy*, 23(5):487–521, 1994. Special Issue in Honor of Nathan Rosenberg.

[38] Stephen Penman. Valuation: Accounting for risk and the expected return. *Abacus*, 52(1):106–130, 2016.

[39] Nirupama Rao. Do tax credits stimulate r&d spending? the effect of the r&d tax credit in its first decade. *Journal of Public Economics*, 140:1–12, 2016.

[40] Katherine Schipper. Required disclosures in financial reports. *The Accounting Review*, 82(2):301–326, 2007.

[41] Tyler Shumway. The delisting bias in crsp data. *The Journal of Finance*, 52(1):327–340, 1997.

[42] Theodore Sougiannis. The accounting based valuation of corporate R&D. *The Accounting Review*, 69(1):44–68, 1994.

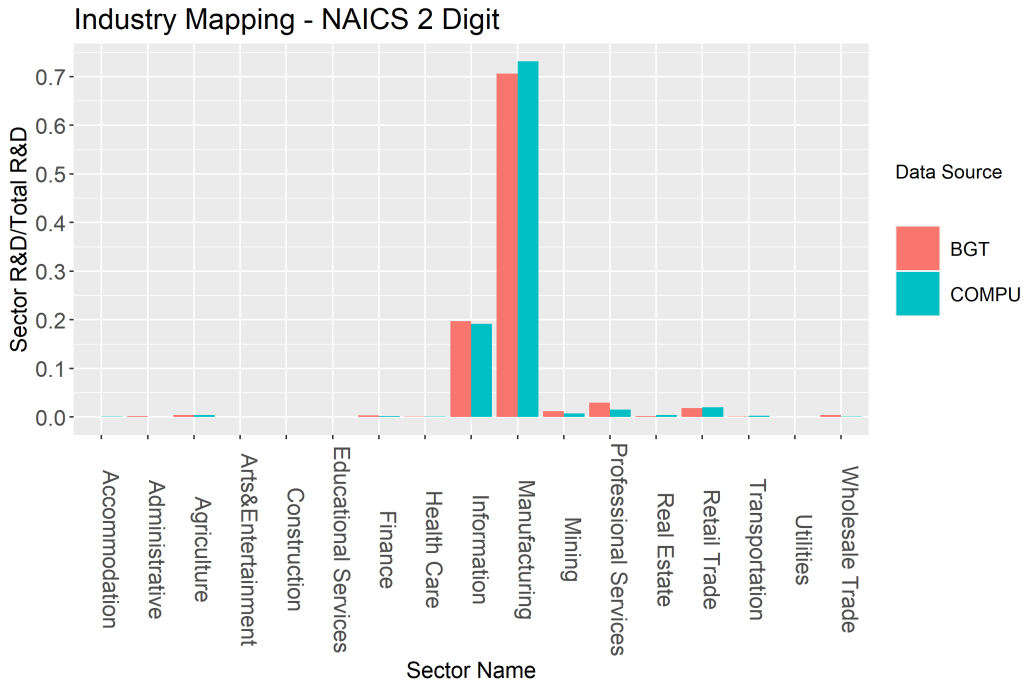
[43] Tim Swift. The perilous leap between exploration and exploitation. *Strategic Management Journal*, 37(8):1688–1698, 2016.

[44] Russell Thomson. The effectiveness of r&d tax credits. *The Review of Economics and Statistics*, 99(3):544–549, 07 2017.

Figure 1: Measure Validation - Reported R&D Expenses

These two subfigures compare the distribution of our occupational R&D measure with reported R&D (46% of the sample). Subfigure A reports the mapping at the NAICS-2-digit industry level. Subfigure B gives the CDPs of our R&D measure with reported R&D expenses (Compustat XRD) at the firm level. We take natural logarithm transformations in both measures.

(a) Industry Mapping - NAICS 2 Digit



(b) Empirical CDF - Firm Level

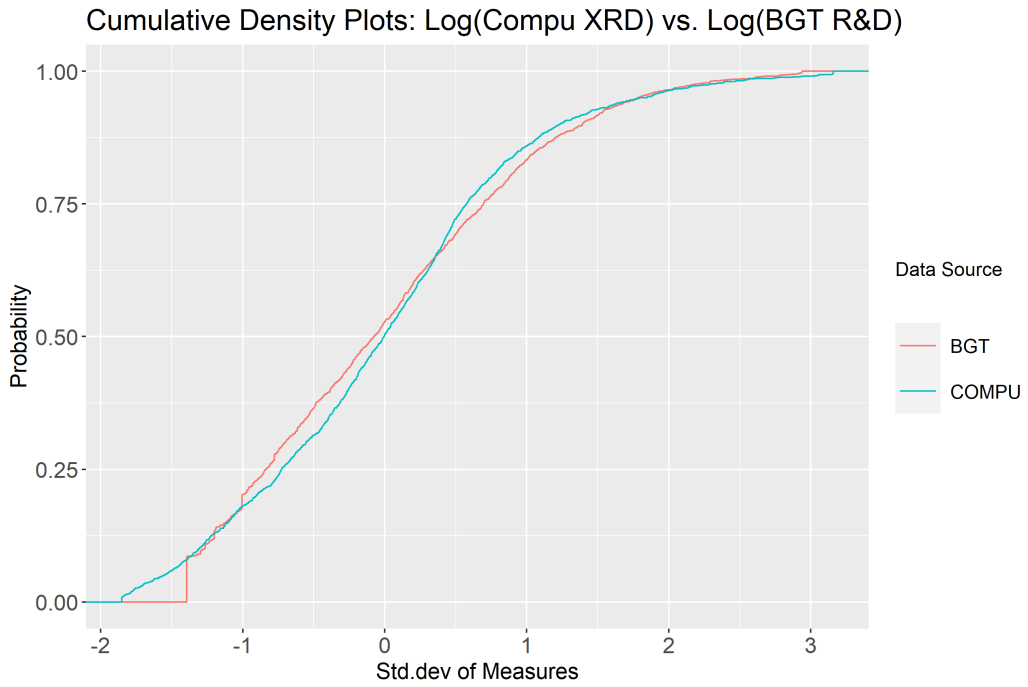
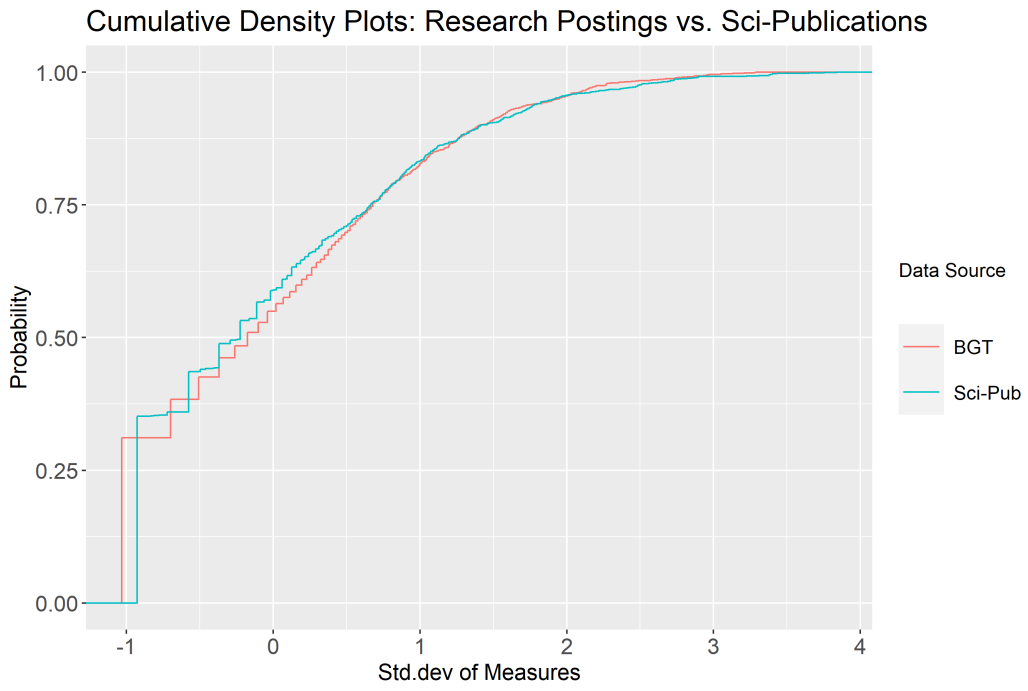


Figure 2: Sci-Pub & Patents among Patenting Firms

This figure validates our occupational classifications against different innovation markers focusing on patenting firms. We obtain the patents and scientific publications data from Arora et al. (2021). Subfigure A reports the firm-level mapping between research job postings and scientific publications. Subfigure B delineates the relationship between firms' development job postings and patents. We use natural logarithm transformations in both measures.

(a) Research Job Postings vs. Scientific Publications



(b) Development Job Postings vs. Patents

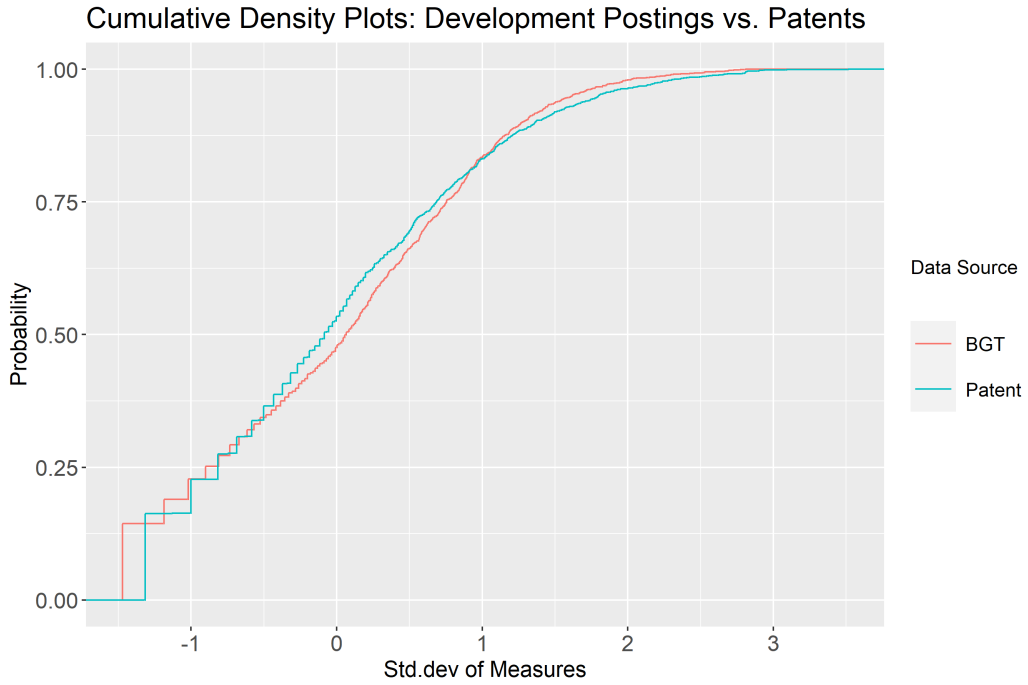
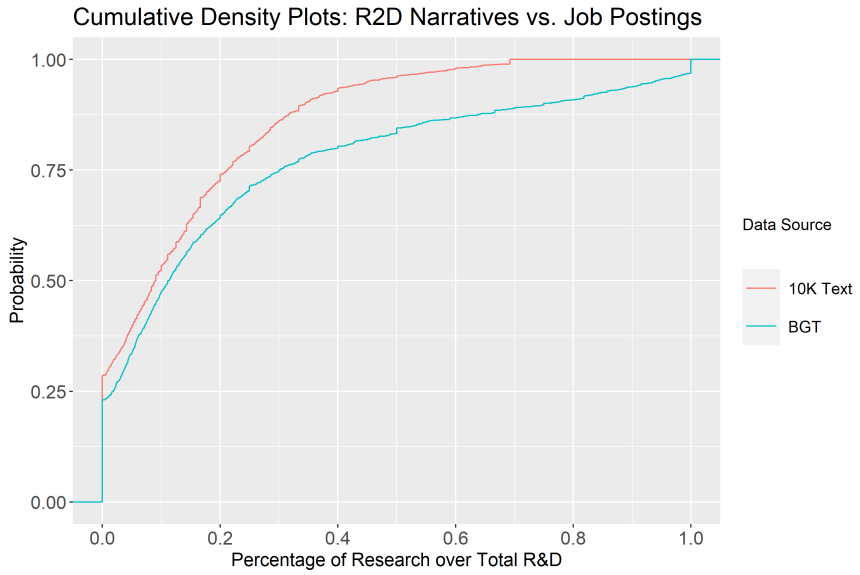


Figure 3: Measure Corroboration - R&D Narratives in 10-K Filings

These two subfigures examine our classifications into research and development occupations. Specifically, we exploit firms' text about research and development through simple keyword searches in their annual reports. We count the stand-alone references to "research" and "development" beyond the phrase "research and development." Next, we compute the R2D ratio as the total stand-alone "research" divided by the sum of stand-alone "research" and stand-alone "development." Therefore, the R2D (10-K Text) ranges from 0 to 1 and gauges firms' differential emphases on research or development. Similarly, we construct the R2D (Posting) using the job postings in research occupations and development occupations. Because both ratios span the same units, we can compare their distributions without standardization. Subfigure A reports the CDPs of firm-level R2D ratios constructed independently from 10-K filings and job postings. Subfigure B presents the density plots of both ratios.

(a) CDPs - R2D (10K Text) vs. R2D (BGT)



(b) Density Plot - R2D (10K Text) vs. R2D (BGT)

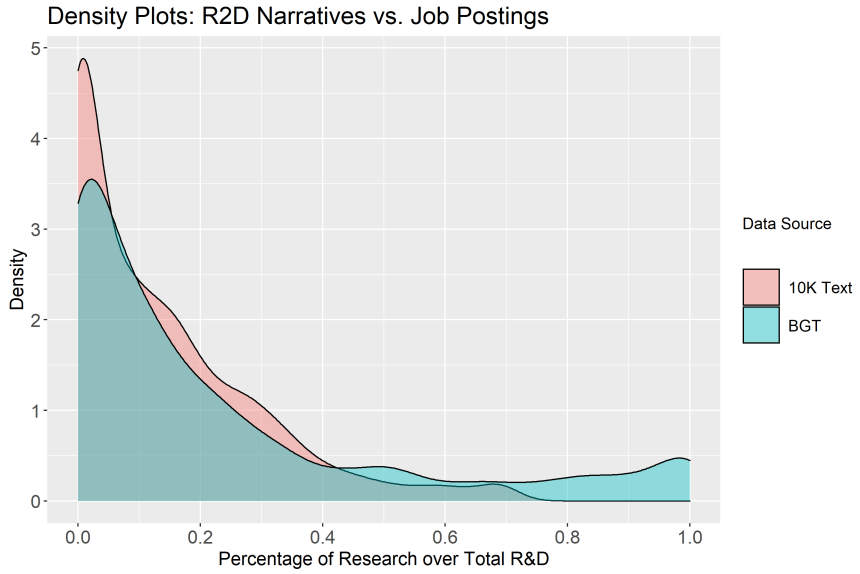


Figure 4: Sector R vs. D

This figure presents the industry-wide different capacities in research vs. development posting. The R ratio and D ratio are scaled by sectors' total job postings. We use NAICS-2-digit codes to parse industries.

Table 1: Summary Statistics

This table presents summary statistics for our firm-level sample. The sample is the cross-section of Compustat and BGT firms headquartered in the United States, ranging from 2010 to 2020. The final sample contains 25,772 unique firm-year observations and covers over 93% of the market capitalization of all US-listed firms. All variables are winsorized at 1% and 99% levels.

Statistic	N	Mean	St. Dev.	Lower Quartile	Median	Upper Quartile
Annual Job Postings	25,772	1,136	3,373	16	111	566
Annual Ph.D.	25,772	23	92	0	0	6
Annual STEM Ph.D.	25,772	14	60	0	0	2
Annual STEM Ph.D.&Master	25,772	51	193	0	2	18
Annual STEM Jobs	25,772	237	765	3	24	121
Research Stock (Count)	25,772	62	231	0	3	21
Development Stock (Count)	25,772	219	690	2	19	119
Research Stock (Dollar - millions)	25,685	7	25	0	0	2
Development Stock (Dollar - millions)	25,685	24	76	0	2	13
Total Assets (millions)	25,772	10,097	29,959	331	1,462	5,587
Total Emp (thousands)	25,772	12	30	0	2	8
Market Cap (millions)	25,772	7,740	21,471	299	1,228	4,616
Sales (millions)	25,772	4,705	13,184	188	781	2,918
Reported R&D Expenditure (millions)	14,647	184	683	1	20	76
ROA	25,772	0.023	0.172	0.013	0.052	0.101
Cash Flow	25,772	0.046	0.152	0.016	0.068	0.118
Revenue Growth	25,772	0.037	0.295	-0.024	0.053	0.141
Missing R&D	25,772	0.432	0.495	0	0	1
Leverage	25,772	0.253	0.228	0.055	0.213	0.389
Tobin's Q	25,772	2.07	1.582	1.104	1.501	2.338

Table 2: Measure Validation

This table presents our validation exercises of the count (Panels A and B) and dollar (Panels C and D) R&D measures. First, we validate our research and development measures against the reported R&D expenses. Subsequently, we a subsample of patenting firms and separately validate the research measure against scientific publications and the development measure against patents.

Panel A: Count Measures and Reported R&D Expenditures

N = 11,509 (Positive R&D Firms)	XRD Flow	XRD Stock	R&D Count (Flow)	R&D Count (Stock)
XRD Flow	1.000			
XRD Stock	0.957	1.000		
R&D Count (Flow)	**0.767**	0.765	1.000	
R&D Count (Stock)	0.761	**0.811**	0.930	1.000

Panel B: Count Measures Separate Validation

N = 5,630 (Patenting Firms)	#Research (Count)	#Development (Count)	#Patents	#Publications
#Research (Count)	1.000			
#Development (Count)	0.650	1.000		
#Patents	0.506	**0.711**	1.000	
#Publications	**0.770**	0.516	0.612	1.000

Panel C: Dollar Measures and Reported R&D Expenditures

N = 11,475 (Positive R&D Firms)	XRD Flow	XRD Stock	R&D Dollar (Flow)	R&D Dollar (Stock)
XRD Flow	1.000			
XRD Stock	0.957	1.000		
R&D Dollar (Flow)	**0.756**	0.751	1.000	
R&D Dollar (Stock)	0.762	**0.813**	0.911	1.000

Panel D: Dollar Measures Separate Validation

N = 5,623 (Patenting Firms)	#Research (Dollar)	#Development (Dollar)	#Patents	#Publications
#Research (Dollar)	1.000			
#Development (Dollar)	0.609	1.000		
#Patents	0.474	**0.711**	1.000	
#Publications	**0.759**	0.515	0.612	1.000

Table 3: Missing R&D Firms

This table focuses on firms' heterogeneous approaches in R&D reporting. We examine four types of firms based on their R&D reporting statuses and actual R&D capacities. Among reported firms, *Positive R&D Firms* and *Zero R&D Firms* disclose positive and zero R&D expenditures, respectively. As for missing R&D firms, *Pseudo R&D Firms* do not report R&D spending but post over 5% for R&D personnel, whereas *Unreported Zero Firms* neither report R&D spending nor hire a non-trivial amount of R&D personnel. We report R&D posting ratios in each type of firm.

Panel A: R&D Intensity (Count)

	Positive R&D Firms			Missing R&D Firms		Zero R&D Firms
Mean	Top tTercile	Middle Tercile	Bottom Tercile	Pseudo Missing	Unreported Zero	Reported Zero
No. of Firms	N = 400	N = 400	N = 401	N = 522	N = 542	N = 309
R&D Intensity	**37.06%**	**20.74%**	**9.26%**	**12.72%**	**1.81%**	**2.22%**
Research Intensity	11.70%	5.01%	1.90%	2.10%	0.38%	0.48%
Development Intensity	25.36%	15.73%	7.36%	10.62%	1.43%	1.74%
Diff. in R&D Intensity	Bottom Tercile (Positive R&D) - Pseudo Missing			Unreported Zero Firms - Reported Zero Firms		
ΔPercentage	-3.466***			-0.407*		
	(0.387)			(0.216)		

Panel B: R&D Intensity (Dollar)

	Positive R&D Firms			Missing R&D Firms		Zero R&D Firms
Mean	Top Tercile	Middle Tercile	Bottom Tercile	Pseudo Missing	Unreported Zero	Reported Zero
No. of Firms	N = 388	N = 389	N = 389	N = 585	N = 468	N = 324
R&D Intensity	**43.50%**	**24.89%**	**11.76%**	**14.67%**	**1.95%**	**3.24%**
Research Intensity	14.97%	5.24%	2.35%	2.47%	0.50%	0.81%
Development Intensity	28.53%	19.65%	9.40%	12.20%	1.44%	2.43%
Diff. in R&D Intensity	Bottom Tercile (Positive R&D) - Pseudo Missing			Unreported Zero Firms - Reported Zero Firms		
ΔPercentage	-2.913***			-1.296***		
	(0.439)			(0.295)		

Table 4: Predicting R&D Reporting

This table focuses on firms' heterogeneous approaches in R&D reporting. Reporting R&D is a dummy variable that takes the value of one if a firm has non-missing XRD in Compustat. We examine if research and development have similar predictive power over firms' R&D reporting choices by comparing within the industry-year grid. Standard errors in Panel B are clustered at the firm level. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A reports the count measure and Panel B reports the dollar measure. Standard errors clustered by firms are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent var:	Reporting R&D					
	Count Measure			Dollar Measure		
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Research}_{t\text{-}1}$	-0.007		-0.016**	-0.008		-0.018**
	(0.007)		(0.008)	(0.007)		(0.008)
$\text{Development}_{t\text{-}1}$		0.030***	0.034***		0.031***	0.036***
		(0.009)	(0.008)		(0.008)	(0.008)
Ind-by-Year FE	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	21,718	21,718	21,718	21,644	21,644	21,644
Adjusted R^2	0.499	0.501	0.502	0.499	0.502	0.502
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test			20.63***			24.76***

Table 5: R&D Tax Credits

This table examines if research and development-intensive firms receive similar or distinctive R&D tax credits and subsidies. We obtain information on firm-level subsidies from Good Jobs First. We parse innovation-related tax grants between those promoting basic research and grants that target development activities based on program names. We classify grants whose names contain "basic research," "scientific/science," and "invention and innovation" as Basic Research Grants. In contrast, we note the rest of the innovation-related grants as Development Grants. We identify 1,445 unique firms in the Good Jobs First subsidy tracker data. Panel A reports the count measure and Panel B reports the dollar measure. Standard errors clustered by Fama-French 48 industries are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

	Basic Research		Development	
Dependent var:	Probability	Log(1 + Grant Sum)	Probability	Log(1 + Grant Sum)
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.018*	0.253*	0.011	0.139
	(0.010)	(0.138)	(0.009)	(0.138)
$\text{Development}_{t\text{-}1}$	0.010	0.147	0.015**	0.256**
	(0.009)	(0.124)	(0.007)	(0.097)
Industry ×Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	12,003	12,003	12,003	12,003
Adjusted R^2	0.105	0.112	0.147	0.165

Panel B: R&D Dollar Measures

	Basic Research		Development	
Dependent var:	Probability	Log(1 + Grant Sum)	Probability	Log(1 + Grant Sum)
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.017*	0.244*	0.009	0.111
	(0.010)	(0.136)	(0.009)	(0.128)
$\text{Development}_{t\text{-}1}$	0.009	0.132	0.015**	0.257***
	(0.008)	(0.114)	(0.007)	(0.095)
Industry ×Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	11,979	11,979	11,979	11,979
Adjusted R^2	0.105	0.111	0.147	0.166

Table 6: Firm Performance

This table examines if research and development have similar or distinctive predictive power over firm performance metrics, such as market capitalization, Tobin's Q, ROA, and cash flows. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.080***	0.135***	0.009**	0.004*
	(0.015)	(0.023)	(0.004)	(0.003)
$\text{Development}_{t\text{-}1}$	0.003	0.027	-0.003	-0.002
	(0.016)	(0.022)	(0.004)	(0.003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,718	21,718	21,718	21,718
Adjusted R^2	0.943	0.764	0.792	0.765
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	19.49***	23.24***	2.88*	1.99

Panel B: R&D Dollar Measures				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.078***	0.130***	0.008**	0.005**
	(0.015)	(0.022)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	-0.001	0.019	-0.004	-0.003
	(0.016)	(0.023)	(0.003)	(0.003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,644	21,644	21,644	21,644
Adjusted R^2	0.943	0.764	0.785	0.764
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	22.91***	30.37***	4.17**	3.02*

Table 7: Industry Performance

This table examines if research and development have similar or distinctive predictive power over industry-level performance metrics such as market capitalization, Tobin's Q, ROA, and cash flows. We construct the same industry-level research and development intensity measures. Panel A uses the count measure of R&D. Panel B exploits the dollar measure of R&D. We describe the regression specifications and control variables in Equation (6). We include industry-fixed-effects to account for time-invariant industry-level characteristics. We do not control for lagged profitability in Column 3. Standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.115***	0.179***	0.012***	0.006***
	(0.032)	(0.029)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	0.014	-0.014	-0.0001	-0.002
	(0.027)	(0.025)	(0.002)	(0.002)
Industry FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	480	480	480	480
Adjusted R^2	0.977	0.838	0.698	0.766
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	3.90**	17.50***	7.50**	5.67**

Panel B: R&D Dollar Measures				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.121***	0.185***	0.011***	0.005**
	(0.032)	(0.029)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	0.018	-0.008	0.0001	-0.002
	(0.027)	(0.025)	(0.002)	(0.002)
Industry FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Observations	480	480	480	480
Adjusted R^2	0.977	0.840	0.696	0.76
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	4.12**	17.83***	6.29**	3.93**

Table 8: Product Market Concentration

This table examines the industry composition hypothesis (H5) at the Fama-French 48 industry levels. We use the Herfindahl-Hirschman Index (HHI) to measure industry concentration. We focus on industries with reasonable firm counts (at least 15) to ensure the stability of the concentration measure. As a result, we exclude industries such as Textiles, Fabricated Products, Shipbuilding/Railroad Equipment, Defense, Shipping Containers, and Tobacco Products. None of the excluded industries are innovation active. Firms in these industries account for 1.34% of the Compustat universe. We describe the regression specifications and control variables in Equation (7). Columns 1 and 3 use total assets to derive the industry-level research and development intensities, while columns 2 and 4 use total hiring as the scaling factor to compute the industry-level research and development intensities. Standard errors are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Dependent Var:	Herfindahl Index			
	R&D Posting Counts		R&D Dollar Measure	
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	-123.587***	-286.488***	-71.736**	-237.127***
	(32.987)	(55.668)	(30.208)	(56.230)
$\text{Development}_{t\text{-}1}$	146.718***	162.567***	136.239**	166.266***
	(50.814)	(33.327)	(52.761)	(34.060)
Industry FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Industry Characteristics	✓	✓	✓	✓
Observations	411	411	411	411
Adjusted R^2	0.927	0.931	0.925	0.930
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	15.07***	39.76***	8.93***	32.86***

Table 9: Stock Returns

This table examines if research and development intensities are associated with different risk premia. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A uses the count measure of R&D. Panel B exploits the dollar measure of R&D. We report Fama-Macbeth regressions with the specifications and control variables introduced in Equation (5). The dependent variable is the individual stock's monthly return excess of the risk-free rate. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.148***	0.110***			0.163***	0.117***
	(0.048)	(0.039)			(0.054)	(0.042)
Development_{t-1}			0.027	0.023	-0.012	-0.012
			(0.060)	(0.046)	(0.064)	(0.049)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	214,462	214,462	214,462	214,462	214,462	214,462
Adjusted R^2	0.214	0.263	0.215	0.263	0.217	0.264
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test					3.01*	2.79*

Panel B: R&D Dollar Measure

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.160***	0.119***			0.173***	0.122***
	(0.051)	(0.040)			(0.057)	(0.044)
Development_{t-1}			0.031	0.035	-0.007	-0.001
			(0.064)	(0.049)	(0.068)	(0.052)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	213,778	213,778	213,778	213,778	213,778	213,778
Adjusted R^2	0.214	0.263	0.215	0.263	0.217	0.264
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test					2.85*	2.33

For Online Publication

Figure A.1: Sector R2D

This figure presents the industry-wide distribution in R2D. The R2D ratio is calculated by sectors' total research posting over total development posting. We use NAICS-2-digit codes to parse industries.

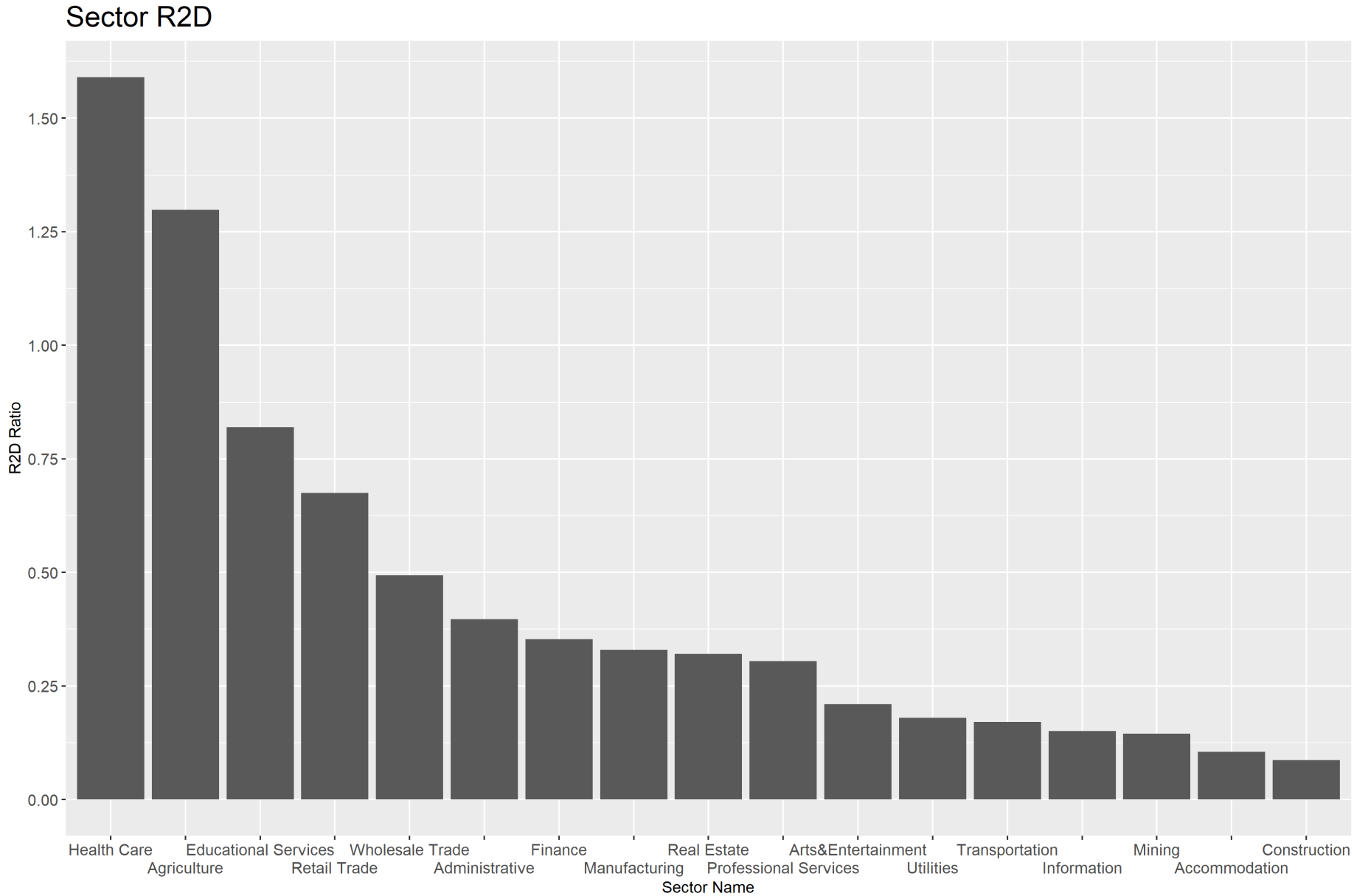
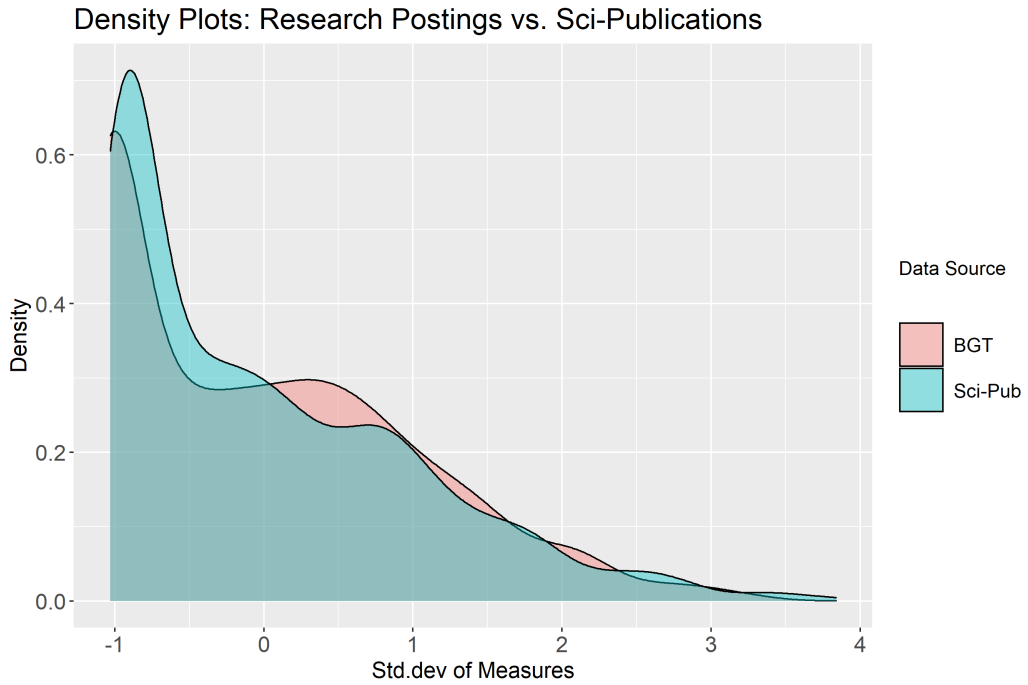


Figure A.2: Sci-Pub & Patents among Patenting Firms

This figure validates our occupational classifications against different innovation markers focusing on patenting firms. We obtain the patents and scientific publications data from Arora et al. (2021). Subfigure A reports the firm-level mapping between research job postings and scientific publications. Subfigure B delineates the relationship between firms' development job postings and patents.

(a) Research Job Postings vs. Scientific Publications



(b) Development Job Postings vs. Patents

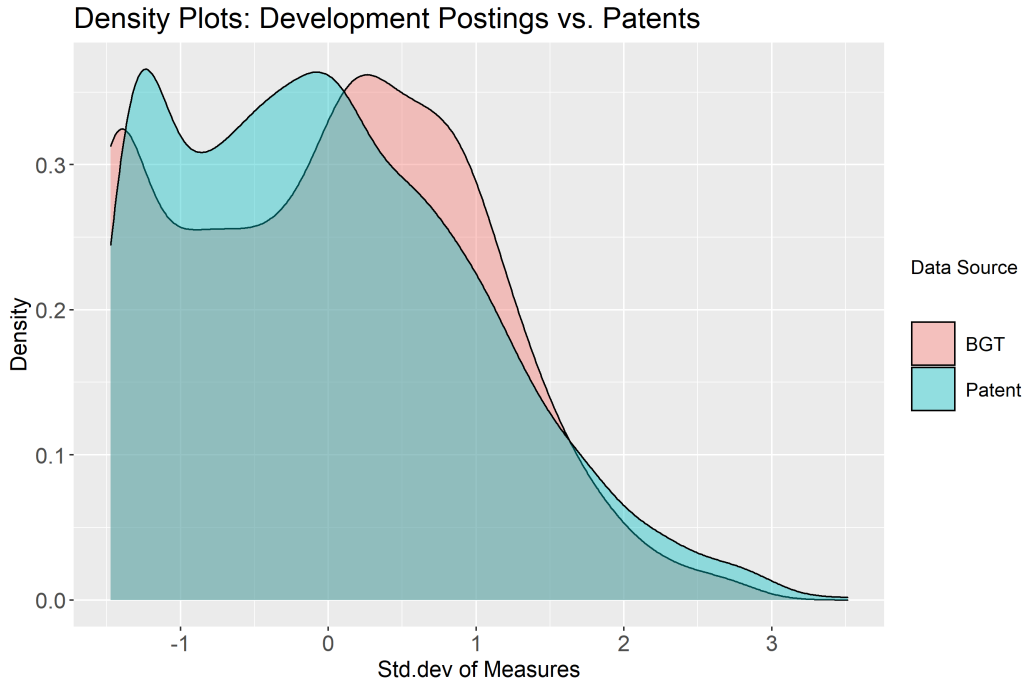
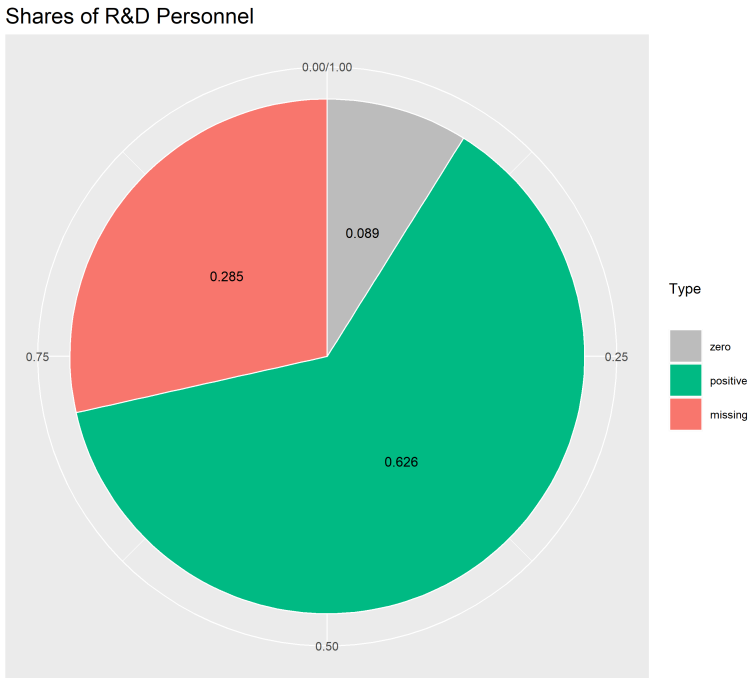


Figure A.3: Missing R&D Firms

These two subfigures depict the job postings for research scientists and development engineers by positive R&D firms, missing R&D firms, and zero R&D firms. Subfigure A reports distributions at the aggregate level. Among all R&D postings, 62.6% are posted by positive R&D firms, 28.5% are posted by missing R&D firms, and zero R&D firms post 8.9%. Subfigure B reports the proportion of R&D posting by missing R&D firms across all NAICS-2-digit industries.

(a) R&D Postings by Positive, Missing and Zero R&D Firms



(b) Industry Percentage of R&D Postings (by Missing R&D Firms)

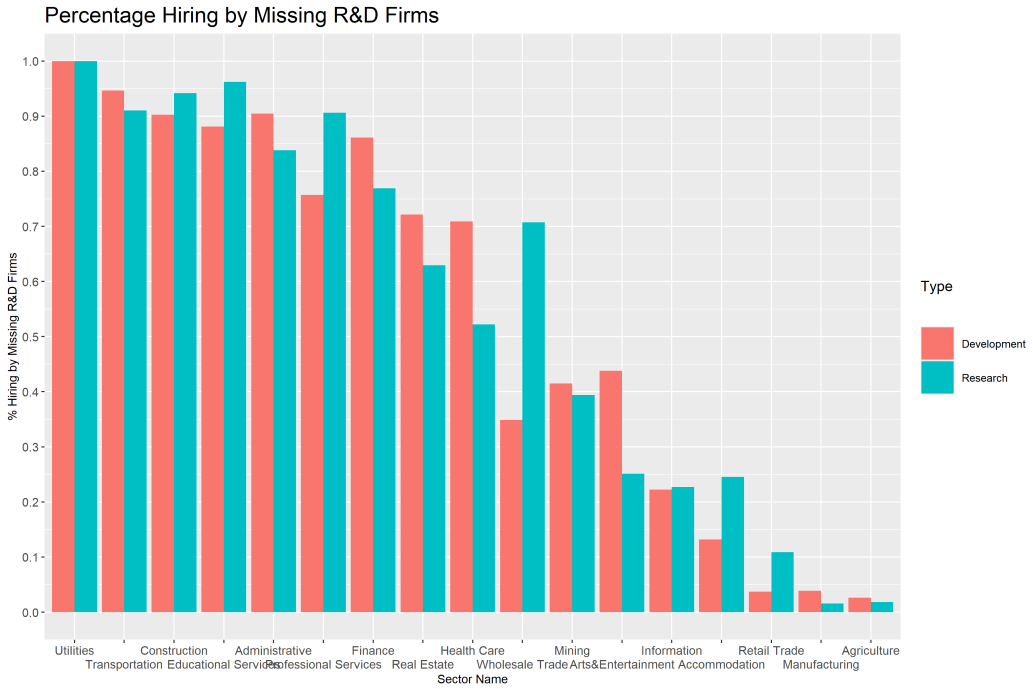


Table A.1: R&D Occupations

This table illustrates all 45 scientist and engineer occupations that comprise the research and development categories. Occupations are classified by six-digit Standard Occupational Codes (SOC) published by the Bureau of Labor Statistics (BLS). We develop separate research and development occupations based on the BLS STEM majors.

Research Category		Development Category	
SOC Code	**SOC Name**	**SOC Code**	**SOC Name**
11-9121	Natural Sciences Managers	11-3021	Computer and Information Systems Managers
15-1111	Computer and Information Research Scientists	15-1131	Computer Programmers
15-2021	Mathematicians	15-1132	Software Developers, Applications
15-2031	Operations Research Analysts	15-1133	Software Developers, Systems Software
15-2041	Statisticians	15-1134	Web Developers
15-2099	Mathematical Science Occupations, All Other	17-2011	Aerospace Engineers
19-1011	Animal Scientists	17-2021	Agricultural Engineers
19-1012	Food Scientists and Technologists	17-2031	Biomedical Engineers
19-1013	Soil and Plant Scientists	17-2041	Chemical Engineers
19-1021	Biochemists and Biophysicists	17-2061	Computer Hardware Engineers
19-1022	Microbiologists	17-2071	Electrical Engineers
19-1023	Zoologists and Wildlife Biologists	17-2072	Electronics Engineers, Except Computer
19-1029	Biological Scientists, All Other	17-2081	Environmental Engineers
19-1042	Medical Scientists, Except Epidemiologists	17-2111	Health and Safety Engineers
19-1099	Life Scientists, All Other	17-2112	Industrial Engineers
19-2011	Astronomers	17-2121	Marine Engineers and Naval Architects
19-2012	Physicists	17-2131	Materials Engineers
19-2021	Atmospheric and Space Scientists	17-2141	Mechanical Engineers
19-2031	Chemists	17-2151	Mining and Geological Engineers
19-2032	Materials Scientists	17-2161	Nuclear Engineers
19-2041	Environmental Scientists and Specialists	17-2171	Petroleum Engineers
19-2099	Physical Scientists, All Other	17-2199	Engineers, All Other
19-3011	Economists		

Table A.2: Alternative R&D Measures

This table examines several alternative R&D measures that exploit the interaction between education requirements and occupation classifications. The first method summarizes a firm's total demand for Ph.D.'s. The second method includes postings of graduate degrees (Ph.D./Master) in STEM majors. The third method uses job postings that require graduate degrees (Ph.D./Master) in the 45 specified R&D occupations. Panel A reports the count measure and Panel B reports the dollar measure. We validate these alternative measures against firms' reported R&D expenditures.

Panel A: Count Measures

N = 11,509 (Positive R&D Firms)	XRD	Total Ph.D.	STEM (Ph.D./Master)	R&D (Ph.D./Master)
XRD	1.000			
Total Ph.D.	**0.814**	1.000		
STEM (Ph.D./Master)	**0.834**	0.930	1.000	
R&D (Ph.D./Master)	**0.843**	0.954	0.990	1.000

Panel B: Dollar Measures

N = 11,475 (Positive R&D Firms)	XRD	Total Ph.D.	STEM (Ph.D./Master)	R&D (Ph.D./Master)
XRD	1.000			
Total Ph.D.	**0.814**	1.000		
STEM (Ph.D./Master)	**0.823**	0.915	1.000	
R&D (Ph.D./Master)	**0.843**	0.954	0.969	1.000

Table A.3: Pseudo R&D Firms

This table examines the materiality of R&D in Missing R&D Firms. We define Pseudo R&D Firms as firms with substantial R&D capacities yet fail to report R&D. We examine if Pseudo R&D Firms exhibit distinct performance and market evaluation than other Missing R&D Firms. Panel A examines missing R&D firms with over 5% R&D posting, and Panel B examines missing R&D firms with over 10% R&D posting. We use industry-by-year fixed effects to obtain inference in the cross-section of the same industry. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Pseudo 5				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
Pseudo 5	0.157**	0.186***	-0.003	0.001
	(0.077)	(0.044)	(0.004)	(0.003)
Industry×Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	9,374	9,374	9,374	9,374
Adjusted R^2	0.861	0.374	0.457	0.451

Panel B: Pseudo 10				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
Pseudo 10	0.191**	0.275***	-0.002	0.009*
	(0.081)	(0.085)	(0.004)	(0.005)
Industry×Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	9,374	9,374	9,374	9,374
Adjusted R^2	0.861	0.379	0.457	0.453

Table A.4: Prominent Firms

This table examines the potential heterogeneous effects in S&P 500 and non-S&P 500 firms. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). We interact the S&P 500 dummy with the research and development intensities. Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1} \times \text{SP500}$	0.071	0.456**	-0.012	-0.011
	(0.046)	(0.212)	(0.008)	(0.007)
$\text{Research}_{t\text{-}1}$	0.075***	0.109***	0.009**	0.005*
	(0.015)	(0.026)	(0.004)	(0.003)
$\text{Development}_{t\text{-}1} \times \text{SP500}$	0.216***	0.401	0.009**	0.013***
	(0.062)	(0.271)	(0.004)	(0.004)
$\text{Development}_{t\text{-}1}$	-0.004	0.014	-0.004	-0.003
	(0.017)	(0.022)	(0.004)	(0.003)
Firm FE, Year FE, Firm Controls	✓	✓	✓	✓
Observations	21,718	21,718	21,718	21,718
Adjusted R^2	0.943	0.766	0.785	0.765

Panel B: R&D Dollar Measures

Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1} \times \text{SP500}$	0.132***	0.594***	-0.014	-0.012*
	(0.037)	(0.192)	(0.009)	(0.006)
$\text{Research}_{t\text{-}1}$	0.073***	0.108***	0.009**	0.005**
	(0.014)	(0.023)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1} \times \text{SP500}$	0.186***	0.369	0.011***	0.014***
	(0.055)	(0.231)	(0.004)	(0.004)
$\text{Development}_{t\text{-}1}$	-0.008	0.004	-0.004	-0.003
	(0.016)	(0.023)	(0.003)	(0.003)
Firm FE, Year FE, Firm Controls	✓	✓	✓	✓
Observations	21,644	21,644	21,644	21,644
Adjusted R^2	0.943	0.766	0.785	0.764

Table A.5: Stock Returns - S&P Split

This table presents the robustness checks of risk premia in the S&P 500 and Non-S&P 500 subsamples. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A uses the count measure of R&D. Panel B exploits the dollar measure of R&D. We report Fama-Macbeth regressions with specifications and control variables introduced in Equation (5). The dependent variable is the individual stock's monthly return excess of the risk-free rate. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Monthly Excess Stock Return					
	S&P 500 Firms			Non-S&P 500 Firms		
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.254*		0.293**	0.101***		0.103**
	(0.133)		(0.139)	(0.039)		(0.043)
Development_{t-1}		-0.120	-0.251*		0.037	0.006
		(0.138)	(0.136)		(0.045)	(0.050)
Industry Dummies	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	38,795	38,795	38,795	175,667	175,667	175,667
Adjusted R^2	0.457	0.454	0.459	0.256	0.256	0.257

Panel B: R&D Dollar Measure

Dependent var:	Monthly Excess Stock Return					
	S&P 500 Firms			Non-S&P 500 Firms		
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.261*		0.269*	0.106***		0.105**
	(0.151)		(0.154)	(0.040)		(0.045)
Development_{t-1}		-0.142	-0.231*		0.049	0.018
		(0.143)	(0.139)		(0.048)	(0.053)
Industry Dummies	✓	✓	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	38,759	38,759	38,759	175,019	175,019	175,019
Adjusted R^2	0.458	0.455	0.459	0.256	0.256	0.256

Table A.6: Firm Performance - Log Measures

This table examines if research and development have similar or distinctive predictive power over firm performance metrics, such as market capitalization, Tobin's Q, ROA, and cash flows. We use the natural logarithm transformation of research intensity and development intensity before standardization as an alternative scaling approach. Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
Research_{t-1}	0.095***	0.160***	0.011***	0.005*
	(0.021)	(0.034)	(0.004)	(0.003)
Development_{t-1}	-0.024	0.014	-0.007*	-0.004
	(0.020)	(0.024)	(0.004)	(0.003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,718	21,718	21,718	21,718
Adjusted R^2	0.943	0.764	0.792	0.765
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test	32.65***	21.58***	5.87**	2.69

Panel B: R&D Dollar Measures				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
Research_{t-1}	0.097***	0.159***	0.010***	0.005**
	(0.021)	(0.031)	(0.003)	(0.002)
Development_{t-1}	-0.030	0.005	-0.008**	-0.004
	(0.021)	(0.029)	(0.004)	(0.003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,644	21,644	21,644	21,644
Adjusted R^2	0.943	0.764	0.791	0.764
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test	40.44***	36.98***	7.75***	3.72*

Table A.7: Firm Performance - Alternative Scaling Factors

This table examines if research and development have similar or distinctive predictive power over firm performance metrics, such as market capitalization, Tobin's Q, ROA, and cash flows. As a robustness check, we use total employment (COMPUSTAT EMP) as an alternative scaling factor. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total employment. Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.043***	0.094***	0.008***	0.002*
	(0.016)	(0.020)	(0.002)	(0.001)
$\text{Development}_{t\text{-}1}$	0.041**	0.057**	0.001	0.0001
	(0.019)	(0.024)	(0.003)	(0.002)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,718	21,718	21,718	21,718
Adjusted R^2	0.943	0.763	0.785	0.765

Panel B: R&D Dollar Measures				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.041***	0.089***	0.007***	0.002*
	(0.015)	(0.018)	(0.002)	(0.001)
$\text{Development}_{t\text{-}1}$	0.037*	0.051*	0.001	0.0004
	(0.019)	(0.028)	(0.003)	(0.001)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	21,644	21,644	21,644	21,644
Adjusted R^2	0.943	0.763	0.785	0.764

Table A.8: Stock Returns - Alternative Scaling Factors

This table examines if research and development intensities are associated with different risk premia. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). As a robustness check, we use total employment (COMPUSTAT EMP) as an alternative scaling factor. Panel A uses the count measure of R&D. Panel B exploits the dollar measure of R&D. We report Fama-Macbeth regressions with the specifications and control variables introduced in Equation (5). The dependent variable is the individual stock's monthly return excess of the risk-free rate. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Research}_{t\text{-}1}$	0.199***	0.156***			0.213***	0.158***
	(0.073)	(0.056)			(0.075)	(0.058)
$\text{Development}_{t\text{-}1}$			0.009	0.025	-0.016	-0.005
			(0.062)	(0.046)	(0.064)	(0.048)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	214,462	214,462	214,462	214,462	214,462	214,462
Adjusted R^2	0.216	0.264	0.215	0.263	0.218	0.264
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test					4.03**	3.74*

Panel B: R&D Dollar Measure

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
$\text{Research}_{t\text{-}1}$	0.202***	0.154***			0.213***	0.153***
	(0.075)	(0.057)			(0.077)	(0.059)
$\text{Development}_{t\text{-}1}$			0.007	0.029	-0.015	0.001
			(0.063)	(0.046)	(0.065)	(0.048)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	213,778	213,778	213,778	213,778	213,778	213,778
Adjusted R^2	0.216	0.264	0.215	0.263	0.218	0.264
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test					3.90**	3.17*

Table A.9: Stock Returns - Log Measures

This table examines if research and development intensities are associated with different risk premia. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). We take the natural logarithm transformation of research intensity and development intensity before standardization as a robustness check. Panel A uses the count measure of R&D. Panel B exploits the dollar measure of R&D. We report Fama-Macbeth regressions with the specifications and control variables introduced in Equation (5). The dependent variable is the individual stock's monthly return excess of the risk-free rate. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.175***	0.135***			0.191***	0.143***
	(0.058)	(0.041)			(0.065)	(0.046)
Development_{t-1}			0.031	0.037	-0.026	-0.023
			(0.067)	(0.049)	(0.073)	(0.053)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	214,462	214,462	214,462	214,462	214,462	214,462
Adjusted R^2	0.215	0.263	0.216	0.263	0.218	0.264
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test					3.44*	3.94**

Panel B: R&D Dollar Measure

Dependent var:	Monthly Excess Stock Return					
	(1)	(2)	(3)	(4)	(5)	(6)
Research_{t-1}	0.185***	0.141***			0.199***	0.146***
	(0.061)	(0.043)			(0.067)	(0.047)
Development_{t-1}			0.032	0.043	-0.022	-0.016
			(0.068)	(0.049)	(0.074)	(0.053)
Industry Dummies		✓		✓		✓
Firm Controls	✓	✓	✓	✓	✓	✓
Observations	213,778	213,778	213,778	213,778	213,778	213,778
Adjusted R^2	0.215	0.263	0.215	0.263	0.218	0.264
$\text{Research}_{t-1} \neq \text{Development}_{t-1}$ Wald Chi-Squared Test					3.40*	3.72*

Table A.10: Exclude Software Industry/AI

This table examines if research and development have similar or distinctive predictive power over firm performance metrics such as market capitalization, Tobin's Q, ROA, and cash flows. As a robustness check, we exclude the information/software industry (NAICS code 51). We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.088***	0.155***	0.010***	0.006***
	(0.016)	(0.022)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	-0.004	0.028	-0.008	-0.006
	(0.020)	(0.025)	(0.005)	(0.004)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	19,368	19,368	19,368	19,368
Adjusted R^2	0.942	0.759	0.790	0.772
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	27.52***	23.96***	6.98***	6.83***

Panel B: R&D Dollar Measures				
Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.084***	0.147***	0.010***	0.006***
	(0.016)	(0.020)	(0.002)	(0.001)
$\text{Development}_{t\text{-}1}$	-0.005	0.027	-0.008	-0.006
	(0.019)	(0.025)	(0.005)	(0.004)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	19,300	19,300	19,300	19,300
Adjusted R^2	0.942	0.759	0.790	0.771
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	25.33***	26.78***	7.06***	6.97***

Table A.11: Controlling for Reported R&D

This table examines if research and development have similar or distinctive predictive power over firm performance metrics such as market capitalization, Tobin's Q, ROA, and cash flows. As a robustness check, we control for the reported R&D expenses (COMPUSTAT variable XRD). Consequently, the sample size reduced by approximately 45%. We construct our research intensity and development intensity measures by scaling the stock of each variable with a firm's lagged total assets (AT). Panel A reports the count measure and Panel B reports the dollar measure. Standard errors, clustered by Fama-French 48 industries, are reported in parentheses. *, **, *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: R&D Posting Counts

Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.071***	0.124***	0.011***	0.007***
	(0.021)	(0.029)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	0.016	0.011	0.001	-0.001
	(0.019)	(0.023)	(0.005)	(0.004)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	12,335	12,335	12,335	12,335
Adjusted R^2	0.944	0.729	0.814	0.778
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	6.27**	9.00***	2.22	2.12

Panel B: R&D Dollar Measures

Dependent var:	Log(Mktcap)	Tobin's Q	ROA	Cash Flow
	(1)	(2)	(3)	(4)
$\text{Research}_{t\text{-}1}$	0.068***	0.121***	0.010***	0.006***
	(0.021)	(0.027)	(0.003)	(0.002)
$\text{Development}_{t\text{-}1}$	0.013	0.006	0.0002	-0.001
	(0.018)	(0.021)	(0.004)	(0.003)
Firm FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Firm Controls	✓	✓	✓	✓
Observations	12,303	12,303	12,303	12,303
Adjusted R^2	0.944	0.729	0.814	0.778
$\text{Research}_{t\text{-}1} \neq \text{Development}_{t\text{-}1}$ Wald Chi-Squared Test	8.49***	12.54***	2.56	2.46