Do peer firms influence innovation?*

Michael Machokoto[a], Daniel Gyimah[b,*], Collins G. Ntim[c]

[a] *The Faculty of Business and Law, The University of Northampton, Waterside Campus, University Drive, Northampton, NN1 5HP, UK*
[b] *Business School, University of Aberdeen, Dunbar Street, Aberdeen, AB24 3QY, Scotland, UK*
[c] *Centre for Research in Accounting, Accountability and Governance (CRAAG), Department of Accounting, Southampton Business School, University of Southampton, University Road, Southampton, SO17 1BJ, UK*

Abstract

Using a large sample of 4,545 US firms over the period 1968–2018, we find robust and significant positive peer effects on corporate innovation. Consistent with the need to keep ahead or abreast of rivals, we document an increase in peer firms' influence with product market competition. Our further analyses show interesting leader–follower interactions with firms following or adopting innovation policies of counterparts perceived or likely to have superior information. This finding supports the information-based motives of mimicking. More importantly, we show that adopting peers' innovation policies is associated with improvements in long-term innovation outputs and product market performance. Our results suggest that peer effects are a critical determinant of corporate innovation in addition to other factors examined so far in the literature.

Keywords: Research and development (R&D), innovation, peer effects, product market competition, heterogeneity effects.

JEL classification: G30, G31, O30

Declarations of interest: none

*Corresponding author; Business School, University of Aberdeen, Dunbar Street, Aberdeen, AB24 3QY, Scotland, UK, **Tel:** +44 122 427 2210

Emails: michael.machokoto@northampton.ac.uk (M. Machokoto), daniel.gyimah@abdn.ac.uk (D. Gyimah), C.G.Ntim@soton.ac.uk (C. G. Ntim).

*We thank Nathan L. Joseph and Alan Lowe (editors), and two anonymous reviewers for useful comments. We also thank Brian Bolton (PFMC 2018 discussant), participants at the 2018 Paris Financial Management Conference, and seminar participants at the University of Northampton, Coventry University, University of Aberdeen. The authors thank the University of Aberdeen, University of Northampton, and University of Southampton for financial support.

1. Introduction

Research and development (hereafter R&D) investment has increased more than fourfold over the last three decades due to an economic-wide shift from the manufacturing sector towards technology and service-based sectors.[1] This change is occurring against a backdrop of rapid technological advancements and intensified product market competition, which has further incentivised firms to innovate.[2] While the extant literature identifies several determinants of R&D (see Aghion et al., 2013; Atanassov, 2015; He and Wintoki, 2016), it remains an empirical question whether or not the interplay of industry dynamics, more specifically, peer firms influence R&D. Yet, the literature relating to other corporate decisions, such as capital structure (Leary and Roberts, 2014; Kaustia and Rantala, 2015; Francis et al., 2016; Fairhurst and Nam, 2018), dividend policy (Adhikari and Agrawal, 2018; Grennan, 2019), cash holdings (Chen and Chang, 2012), and investment (Foucault and Frésard, 2014; Frésard and Valta, 2016; Frydman, 2015; Bustamante and Frésard, 2017), show that peer effects matter. Motivated by this growing literature, we examine whether and to what extent peer firms also influence innovation.[3]

However, due to two counteracting reasons, it is not clear why a firm would imitate the R&D of peer firms. On the one hand, a firm has to monitor and respond to peer firms, in particular, their investments in innovation, as this affects its competitive position. Investments in innovation have far-reaching effects on the firm that arise from the ever-increasing pressure to reduce production costs, enhance growth, and competitiveness (see Hart, 1983; Aghion et al., 2001, 2005). Hence, mimicking peer firms can be an effective way of achieving competitive parity or market dominance (Lieberman and Asaba, 2006). On the other hand, investments in R&D are irreversible, capital intensive, risky and require a significant commitment of resources over a long period, which could reduce or

[1]For example, Borisova and Brown (2013) report that US young and mature firms recorded a fourfold and a twofold increase in R&D between 1980 and 2001, respectively. Similarly, Brown and Petersen (2011) report R&D increases of 2% over the period 1970–1981 compared to 6.3% between 1982 and 1993, and 10.3% from 1994 to 2006.

[2]According to Bates et al. (2009), the incentive to innovate in order to keep abreast of market competition is increasing over time.

[3]We use the terms 'R&D' and 'corporate innovation' interchangeably throughout this study.

discourage imitation. Therefore, peer effects could be heterogeneous as small or young firms are less likely to mimic relative to large and mature firms that can more readily absorb the risks and costs associated with innovation. At the same time, there are instances where a firm can strategically reduce R&D even though peer firms are increasing their spending. For example, Mudambi et al. (2015) argue that a firm can reduce R&D to stimulate patent output. A notable example of this strategic reallocation of resources is when Cisco reduced research spending by USD1.5 billion from 2002 through 2004, but at the same time achieved significant patent output by strategically shifting from radical high-cost and risky new ideas (exploratory innovation) to low-cost and less-risky ideas (exploitative innovation) (see Mudambi et al., 2015). Due to these opposing or counteracting predictions, it is *a priori* unclear whether or not peer influence matters for corporate innovation.

In this study, we examine whether peer firms influence corporate innovation using a large sample of 4,545 US firms over the period 1968–2018. Observably, prior literature routinely controls for industry effects using industry fixed effects, which do not capture how firms interact within an industry. Industry fixed effects do not allow for an in-depth analysis of how and why peer effects might influence R&D (see Brown et al., 2009, 2012; He and Wintoki, 2016). Using instrumental variable regressions, which address the reflection problem surrounding the study of peer effects (Manski, 1993) and endogeneity issues in line with Leary and Roberts (2014), we find significant peer influence on corporate innovation.[4] Specifically, we find that, on average, a firm increases R&D by about 4% in response to a one standard deviation increase in peer firms' R&D. This finding, which is robust to using alternative measures of innovation, industrial definitions, subsampling and estimation methods, show that peer firms significantly influence the focal firm's R&D.

Next, we examine the channels through which peer firms influence corporate innovation by testing the predictions of the rivalry and information theories. In the first instance, the rivalry theory posits that firms adopt policies similar to their successful

[4]We estimate 2SLS and IV-Tobit models to examine whether or not peer firms influence R&D.

rivals in order to catch up with or keep abreast of competitors (Lieberman and Asaba, 2006; Leary and Roberts, 2014). To test this prediction, we create a dummy variable for firms facing high and low product market competition based on several measures of product market competition and interact these dummy variables with peer average R&D. Based on this analysis, we find significant and robust evidence that is in line with the rivalry theory as firms operating in industrial segments subject to intense product market competition are more responsive to their peers relative to those facing less competitive pressures. This finding is in line with Aghion et al. (2005), who find that direct competition only generates incremental profits for firms in neck-and-neck industries if they innovate. Our study provides new empirical evidence suggesting that product market competition is also an important channel through which peer firms influence innovation.

In the second instance, we explore whether or not firms respond to their counterparts for informational reasons. This analysis is premised on the information theory, which posits that firms tend to follow policies of peers perceived to possess superior information (see Leary and Roberts, 2014; Adhikari and Agrawal, 2018). The information theory also underpins the "learning motive" by which followers learn from leaders. We conjecture that the high information asymmetry associated with corporate innovation is likely to make mimicking for informational reasons unidirectional rather than bidirectional. Thus, we predict that mimicking is more pronounced for followers rather than leaders within a product market. However, unlike the other corporate decisions so far examined in the literature, we also expect leaders to respond to followers, as not doing so might entail a potential loss of competitive advantage in the long-run.[5]

To test our predictions, we categorise firms as followers (leaders) based on profitability, firm size, sales, and analyst followings. We find significant heterogeneity in peer effects on R&D, with followers being more responsive to their peers relative to industry leaders. Our results also show that followers mimic leaders and vice-versa, which we consider to be important bidirectional peer effects that have not yet been documented in the extant

[5]For studies on peer effects relating to capital structure and corporate investments, see, Leary and Roberts (2014), Frydman (2015), Frésard and Valta (2016) and Francis et al. (2016).

literature. While we establish this bidirectional response for both followers and leaders, our results, however, indicate that the response to peer firms' R&D is more pronounced for followers relative to leader firms. These results highlight the extent to which leaders perceived to possess superior information about the product market influence the innovation of their followers. By adopting the R&D policies of peer leader firms, followers can reduce the time, effort, and costs required to optimise their R&D investments. In doing so, the firm also reduces or minimises the risk of misalignment with industrial peers, which can be costly in the case of R&D that is characterised by high irreversibility, information asymmetry and long-investment horizons. Thus, our findings suggest that the "learning motive" dominate peer effects associated with corporate innovation.

Finally, even though we document significant peer effects on corporate innovation, we have so far not addressed an important and largely unanswered empirical question of whether following peer firms is beneficial or not. On the one hand, a firm might strategically play it safe or seek to learn new information by paying more attention to its peers. On the other hand, it might be detrimental for the firm to neglect its own information or competitive advantages, and instead seek to align with its peers by adopting sub-optimal policies. In line with the latter prediction, Kaustia and Rantala (2015) find no benefits of mimicking equity-share splits. Similarly, Fairhurst and Nam (2018) argue that mimicking the financing policies of peer firms could lead to sub-optimal decisions as managers tend to be less focused on their own fundamentals. Motivated by these opposing predictions, we examine the implications of mimicking R&D by running IV-2SLS and cross-sectional regressions that relate innovation outputs and product market performance to the peer firms' R&D. To ensure that we draw meaningful inferences that are not subject to omitted variables bias, we also control for several confounding factors from the literature and include both firm and time-fixed effects in our models.

Using several measures of innovation outputs and product market performance, we find that, on average, peer effects on innovation are associated with an increase in firm value and innovation output (the number and the value of patents). This is in line with "the playing it safe or learning new information-based motive" that firms can avoid the effort and search costs involved in optimising their decisions by closely following their peers. Our results indicate that the benefits of mimicking peer firms increase with the intensity of mimicking. These findings are important because they not only show that peer firms significantly influence innovation, but also that mimicking has significant implications on innovation outputs and product market performance.

Our study's contributions relate to two strands of the extant literature. First, prior studies have, for example, reported evidence of peer effects on capital structure (Leary and Roberts, 2014; Kaustia and Rantala, 2015; Francis et al., 2016), dividend policy (Adhikari and Agrawal, 2018; Grennan, 2019), cash holdings (Chen and Chang, 2012), investment decisions (Foucault and Frésard, 2014; Frésard and Valta, 2016; Frydman, 2015; Bustamante and Frésard, 2017), and trade credit (Gyimah et al., 2020). Our paper extends this literature to corporate innovation, which is unique in that it is more costly to mimic given that it is irreversible, has high information asymmetry and requires a significant commitment of capital over very long periods. Focusing on peer influence on corporate innovation is important as strategic interactions or industry dynamics can amplify both positive and negative firm-specific shocks within and across industries. This is particularly relevant in the case of R&D that has now surpassed physical investments with economies increasingly shifting from the manufacturing sectors toward technology and services sectors. To the extent that corporate innovation is increasingly becoming the main driver of economic growth in the digital economy, it is important to understand how industry dynamics influence further investments in innovation.

Second, we contribute to the literature by providing new evidence on the specific channels through which peer firms influence R&D. This line of inquiry has so far been overlooked in the extant literature, except for the contemporaneous study on R&D by Bui et al. (2019), which examines how peer effects impact innovation. However, our paper differs from Bui et al. (2019) along several dimensions. First, Bui et al. (2019) correlate the lagged average peer R&D with a firm's current innovation outputs. This only indicates an association between average industry R&D and each firm's future performance rather than the implications of mimicking. To robustly address this issue, we explore the long-term implications of mimicking instead of short-term performance and innovation outputs given the uncertainty surrounding outputs from innovation and its long investment horizon (gestation period). Using this approach is, thus, more in sync with the literature on innovation outputs (Chan et al., 2001; Hirshleifer et al., 2013; Bena and Li, 2014; Hsu et al., 2018). In doing so, we also evaluate several ways of identifying mimicking firms, which are not directly observable. Second, in examining heterogeneity in peer effects using firm characteristics and product market competition, Bui et al. (2019) estimate separate regressions by splitting the sample into high and low regime firms. However, in such regression specifications, the magnitude of the coefficients are difficult to interpret and are not directly comparable. To more robustly capture the asymmetry in mimicking and deliver meaningful economic sense, we estimate a single model with interactive dummies that enables direct comparisons to be drawn between firms in the low and high regimes. Using this framework, we document new empirical evidence on the asymmetry in peer effects and their impact on long-term innovation outputs and product market performance.

In addition, we also extend our analyses to closely examine the channels of mimicking through product market competition and follower–leader interactions. Our novel results show that peer effects on innovation increase with product market competition. At the same time, they also indicate significant bidirectional peer effects, where followers and leaders adopt the innovation policies and practices of one another. As our results show stronger responses for firms that lack informational advantages relative to those with superior information in the product markets, we conclude that the information theory dominates the rivalry theory as the main driver of peer effects. These results, in general, enrich our understanding of how and when firms strategically mimic their peers' R&D policies, and this has wider academic and policy implications. For academic researchers, the significant peer effects we document over and above firm-specific and macroeconomic factors highlight a new determinant of innovation that could enrich our understanding of industry dynamics and corporate decisions. On the policy front, peer effects matter as they amplify the propagation of economic shocks both within and across industries. This has a direct impact on the general welfare and economic growth, given the central role of innovation as a catalyst for boosting productivity and economic growth.[6]

The rest of our paper is organised as follows. Section 2 reviews the relevant literature and formulates the testable hypotheses. Section 3 describes the data, whereas Section 4 presents the methodology. Next, Section 5 presents and discusses the empirical results. Section 6 presents the robustness tests and Section 7 concludes the paper.

2. Related Literature and Hypotheses Development

Rivalry theory suggests that rational decision-makers mimic successful rivals to avoid expending effort or search costs involved in finding the optimal solution (Cyert and March, 1963; Lieberman and Asaba, 2006). In other words, there is herd behaviour associated with most economic decisions, especially during periods of heightened uncertainty or in environments characterised by high market imperfections. Following on these theoretical predictions, recent empirical studies document significant peer influence on corporate decisions. For example, Leary and Roberts (2014), Kaustia and Rantala (2015), Francis et al. (2016) and Fairhurst and Nam (2018) find significant peer effects on capital structure, whereas Adhikari and Agrawal (2018) and Grennan (2019) report similar peer effects on dividend policy. These studies show that peer firms significantly influence a firm's decisions beyond firm-specific and macroeconomic factors so far examined in the literature.

In particular, the extant literature identifies several determinants of R&D, such as cash holdings (Brown et al., 2009; Borisova and Brown, 2013; He and Wintoki, 2016), external financing (Brown et al., 2009; Atanassov, 2015), greater institutional ownership (Aghion et al., 2013), and product market competition (Aghion et al., 1999, 2005). Mimicking peer firms' R&D may not be as straightforward as other corporate policies. On the one hand, R&D investments are associated with significant funding constraints, which might prevent firms from mimicking their peer firms' R&D (Hoberg and Maksimovic, 2015). For

[6]See, Acemoglu et al. (2006), Brown et al. (2009), Gorodnichenko et al. (2010) and Atkeson and Burstein (2019).

example, Brown and Petersen (2011) argue that firms face high adjustment costs when investing in R&D. This arises as R&D investments tend to be irreversible, risky and require a significant commitment of resources over a long period. These unique features of corporate innovation are likely to discourage firms from imitating their peers. On the other hand, investments in innovation have far-reaching effects on the firm, which relate to reducing production costs, enhancing firm growth, and improving competitiveness (Hart, 1983; Aghion et al., 2001, 2005). Hence, mimicking peer firms can be a cost-effective way of achieving competitive parity or market dominance (Lieberman and Asaba, 2006). Against this background, we formulate and test these two competing hypotheses:

Hypothesis 1a (H1a): *Firms do not increase R&D in response to peer firms.*

Hypothesis 1b (H1b): *Firms increase R&D in response to peer firms.*

The relationship between product market competition and innovation is not clear, given the mixed theoretical and empirical evidence. Whereas industrial organisational theory predicts an inverse relationship between competition and innovation (Schumpeter, 1947; Grossman and Helpman, 1991), the empirical evidence suggests that product market competition drives innovation (Nickell, 1996; Blundell et al., 1999; Aghion et al., 1999). Thus, there are two opposing arguments for the effects of product market competition on innovation. First, because competition destroys monopoly rents, firms are likely to innovate if it increases their monopoly power so that they can appropriate the returns arising from such innovation (Schumpeter, 1947). In other words, intense competition tends to reduce the overall share of industrial profits, thereby discouraging firms from investing in R&D. Second, competition can act as a disciplinary tool that reduces managerial slack and encourages innovation and growth (Machlup, 1967; Scharfstein, 1988). In line with this reasoning, Aghion et al. (2005) argue that increased product market competition provides incentives for firms to innovate in a bid to "escaping competition", especially in neck-and-neck industries.

The significance of valuable innovation, therefore, is that it ensures the enjoyment of sustained superior corporate profits, especially when it serves new and unmet consumer demands (see Roberts, 1999). Whereas firms can still achieve high-profitability by avoiding product market competition, innovation provides the much-needed tool that facilitates the staunching of competitive pressures. This effect is more pronounced for firms operating in less established industries, in which the first to innovate ("first movers") increases market share and profitability (Geroski et al., 1993; Banbury and Mitchell, 1995). These findings reinforce the importance of corporate innovation to business survival, with the increased competitive edge stemming from innovation that brings in new customers, and thereby generates new and increased sales, profitability, and cash reserves. Thus, innovation creates growth opportunities and enhances the firm's competitive position.

Noticeably, although product market competition is a related concept to peer effects, the two concepts are significantly distinct (Bird et al., 2018). For instance, we know that some firms derive their market power from their ability to influence product prices, as well as the quality and nature of their products (Kubick et al., 2015). Further, according to Aghion et al. (2005) and Lieberman and Asaba (2006), firms operating in industries with intense competition have more incentives to innovate in order to keep abreast with the competition and achieve competitive parity. This argument is consistent with Grossman and Helpman (1991) and Aghion et al. (1997), who assert that a firm must adopt a step-by-step innovation in order to achieve a similar status as that of the leading firms before pursuing cost-leadership status in future through increased innovation. Thus, faced with intense product market competition, firms innovate to pull ahead of the competition and generate incremental profits at the expense of their non-innovating rivals (Aghion et al., 2001). As rivals innovate, the incentive to imitate peer firms' innovation policies to maintain competitive parity and preserve market share becomes larger. We, therefore, test the following hypothesis that competitive industrial pressures drive firms to imitate their peers, as our second hypothesis:

Hypothesis 2 (H2): *Peer effects on R&D increase with product market competition.*

Past studies suggest that the main reasons why firms adopt the innovation policies of their peers can be traced to two main motives Lieberman and Asaba (2006). First, the learning motive implies that firms learn from their more successful rivals by following their innovation policies (Scharfstein and Stein, 1990; Leary and Roberts, 2014). Second, the feedback theory, which is akin to the predation theory, emphasises that firms with surplus cash adopt similar financial policies as their cash-constrained competitors to drive them out of business (Brander and Lewis, 1986; Bolton and Scharfstein, 1990). This channel of peer effects becomes more apparent in competitive industries, where there is an increased need to consolidate competitive positions. Thus, leader–firms can adopt the innovation policies of their followers in order to reduce competitive pressures.

We define firms as *Leaders* and *Followers* using firm characteristics, such as profitability, firm size, market share (a firm's proportion of industry sales), and analyst followings. For example, Adhikari and Agrawal (2018) define followers (leaders) as firms that are smaller (larger) and Leary and Roberts (2014) classify firms as leaders/followers based on profitability and market share. In the context of corporate innovation, large and profitable firms are more likely to lead in R&D investments due to their ability to absorb the high costs of R&D investments, uncertain outcomes and significant moral hazard associated with innovation (Hall, 2009; Brown and Petersen, 2011; Hoberg and Maksimovic, 2015; Kogan et al., 2017). Moreover, analyst coverage reduces information asymmetry, provides an effective external monitoring mechanism (Brennan and Subrahmanyam, 1995; Hong et al., 2000), and influence investment decisions (Chang et al., 2006). Thus, whereas *Leaders* initiate innovation, *Followers* tend to have more incentives to mimic the innovation of their *Leaders*. This mimicking behaviour enables the *Followers* to reduce the uncertainty inherent in untested innovations and enhance their competitive position. Thus, we derive our third hypothesis as follows:

Hypothesis 3 (H3): *Followers are more likely to adopt R&D policies of leader-peer firms.*

Discernibly, the beneficial effects of corporate innovation on firm value have been explored in the existing literature. For example, innovation generates a positive stock market response and firm growth (Kogan et al., 2017), and positively predicts future stock returns (Hirshleifer et al., 2013, 2018; Mama, 2018). This positive effect of innovation on stock returns stems from the increase in post-innovation patent counts and citations. Ehie and Olibe (2010) argue that the positive effect of R&D on firm value is more pronounced for firms in the manufacturing sector than those in the service sector. Thus, whether or not innovation proves beneficial might depend on the type of firm and its sector. For example, technology firms that depend on R&D for growth might generate substantial benefits from innovation than would an otherwise non-technologically oriented firm. Consistent with this narrative, Kallunki et al. (2009) find that technology acquirers are more successful in converting their R&D into increases in current market value and future profitability relative to non-technology acquirers.

Observably, the empirical literature documents evidence suggesting that adopting a rival firm's R&D policies impact positively on firm value, operating performance and innovation output of the focal firm. For instance, Bui et al. (2019) find a linear relationship between a firm's R&D and firm value and risk. This evidence appears to suggest that there is no limit to the effect of innovation on corporate outcomes, contrary to Hartmann et al. (2006), who argue that R&D spending yields reward up to a cut-off point beyond which additional investment in innovation does not generate commensurate returns. On the one hand, a firm might play it safe or learn from additional information when it closely follows its peers. On the other hand, it might neglect its own information or competitive advantages, which might lead to over- or under-investment. In this case, blindly mimicking peers innovation policies might redirect resources to over-invest in R&D and distract managers from focusing on their own value-enhancing strategies (Fairhurst and Nam, 2018). Therefore, we test the implications of mimicking corporate innovation on firm value as our final hypothesis as follows:

Hypothesis 4 (H4): *Peer effects of R&D have a positive impact on innovation outputs and product market performance.*

3. Data and summary statistics

3.1. Data

Our sample is drawn from the Compustat North America Database over the period 1968–2018. Data availability restricts the sample period for our analysis. We apply several filters to the data as a standard practice in the literature. For example, we exclude firms in the utility, financial, and public sectors, which we consider to be regulated sectors. We also exclude firms with negative equity, missing data on key variables, and asset growth in excess of 100%. Furthermore, we winsorise all variables at the upper and bottom 1% in order to reduce the effects of outliers. We augment our dataset with product-market concentration, product similarity, and product fluidity data from Hoberg and Phillips's data library (see Hoberg and Phillips, 2010; Hoberg et al., 2014; Hoberg and Phillips, 2016).[7] These filtering results in 51,990 firm-year observations for 4,545 firms.

3.2. Variables

Our primary measure of innovation is R&D expenditure (He and Wintoki, 2016; Acharya and Xu, 2017) since it accounts for more than 50% of corporate innovation (Hall, 2009). According to Hall (2009) and Mairesse et al. (2005), R&D is the only measure of innovation that is frequently used over a long period, and has higher predictive power for firm performance relative to other proxies. The R&D dummy (RDD) is equal to one if a firm reports R&D during a particular year and zero otherwise. Following He and Wintoki (2016), we define the ratio of R&D expenditure to total assets (RD/TA), as R&D intensity.

We use other alternative measures, R&D to net assets (RD/NA) and change in R&D to total assets ($\Delta RD/TA$). Since firms sometimes report some R&D expenditure as part of selling, general and administrative expenses (SG&A), we also measure R&D spending as the sum of R&D expenditure and SG&A (Banker et al., 2011; Lévesque et al., 2012). These alternative proxies ($RDSGA/TA$ and $RDSGA/Sales$) are used to re-estimate the baseline results, as robustness checks. In line with prior studies, we set R&D to zero for firms with missing R&D in their annual reports (Hirschey et al., 2012; He and Wintoki, 2016). We then estimate peer averages by defining peer firms as those that fall within the same three-digit SIC code, where three-digit SIC code is the definition of an industry.[8] We also use several alternative definitions of the industry for our extensive robustness

[7] The data is available from http://hobergphillips.usc.edu/industryconcen.htm.

[8] Peer average variables are the variables with an overline. For example, peer RDD is $\overline{RDD}$; peer RD/TA is $\overline{RD/TA}$, etc.

tests. We describe in detail the construction of the other variables used in Appendix A.

3.3. Summary statistics

Table 1 summarises the basic statistics for all variables used and these are comparable to prior studies. Firms that invest in R&D constitute 77% (40,059) of the sample. The mean (median) *RD/TA* is 0.045 (0.000) and is within the ranges that have been reported by Brown and Petersen (2011) and He and Wintoki (2016). The mean (median) of the control variables; cash, market to book ratio (Q), debt, size, and return on assets (ROA) are 0.152 (0.030), 1.685 (1.005), 0.193 (0.044), 5.620 (4.142) and 0.079 (0.051), respectively. The basic statistics for the other measures of research and development (R&D), *RD/NA*, Δ*RD/TA*, *RDSGA/TA*, and *RDSGA/Sales*, and those of the control variables are in line with the literature.

PLEASE INSERT TABLE 1 HERE

PLEASE INSERT FIGURE 1 HERE

The time series plots in Figure 1 show that R&D (*RD/TA*) has been increasing from an average of 0.7% in the 1960s to a peak of 4.9% in the late 2000s, and thereafter, decreasing to 4.5% in the 2010s. This shows the increasing significance of R&D as a form of corporate investment (see Brown et al., 2009; Borisova and Brown, 2013; Falato and Sim, 2014; He and Wintoki, 2016).

4. Methodology

4.1. Estimation model

We investigate peer effects on R&D using the baseline model of Leary and Roberts (2014), which is specified as follows:[9]

$$y_{ijt} = \alpha + \beta \overline{y}_{-ijt} + \gamma' \overline{X}_{-ijt-1} + \lambda' X_{ijt-1} + \psi' \nu_j + \phi' \nu_t + \epsilon_{ijt} \quad (1)$$

[9]Several other studies which examine peer effects on corporate decisions use the same model (see, among others, Chen and Chang, 2012; Francis et al., 2016; Bustamante and Frésard, 2017; Park et al., 2017; Adhikari and Agrawal, 2018).

where Y_{ijt} is either R&D or the R&D dummy (RDD) for firm i in industry j at time t, α is a constant, β, γ' and λ' are the vectors of coefficients to be estimated, $\overline{y}_{-ijt}$ is peer firm average excluding firm i, $\overline{X}_{-ijt-1}$ and X_{ijt-1} are vectors of peer firm averages and firm-specific characteristics, respectively. ν_j and ν_t are industry and year-fixed effects, respectively. Finally, ϵ_{ijt} is the error term. In line with Leary and Roberts (2014), we assume that ϵ_{ijt} is correlated within firms and heteroskedastic. The vectors $\overline{X}_{-ijt-1}$ and X_{-ijt-1} include cash, market-to-book, leverage, size and profit.[10]

To examine the heterogeneity in peer effects on innovation, we estimated a modified version of our baseline model that includes an indicator variable as follows:[11]

$$\begin{aligned} y_{ijt} =& \alpha + \left[\beta_1 \overline{y}_{-ijt} + \boldsymbol{\gamma_1}'\overline{\boldsymbol{X}}_{-ijt-1} + \boldsymbol{\lambda_1}'\boldsymbol{X}_{ijt-1}\right]\mathbf{1}_{q_{ijt}\leq m} \\ &+ \left[\beta_2 \overline{y}_{-ijt} + \boldsymbol{\gamma_2}'\overline{\boldsymbol{X}}_{-ijt-1} + \boldsymbol{\lambda_2}'\boldsymbol{X}_{ijt-1}\right]\mathbf{1}_{q_{ijt}> m} \\ &+ \psi'\nu_k + \phi'\nu_t + \epsilon_{ijt} \end{aligned} \tag{2}$$

where $\mathbf{1}_{\{.\}}$ is an indicator variable that takes the value of one if a firm is categorised as being in the low (high) regime in a particular year and zero otherwise. To categorise the firms into the two regimes, we use the median (m) of several variables that proxy for product market competition (the Herfindahl-Hirschman Index (HHI) based on assets (HHI-Assets), sales (HHI-Sales), Top-Four Concentration Index based on assets (CR4-Sales), and sales (CR4-Assets) and leader/follower status (based on profitability, firm size, the logarithm of sales (LogSales), and analyst followings).[12] Using the Equation (2) with indicator variables enable us to directly test whether peer effects (β_1 and β_2) differ between firms in the low and high regimes. We estimate our empirical models using several approaches for robustness and to allow for comparisons with previous studies.

[10]These factors are informed by the literature (e.g. Aghion et al., 2004; Brown et al., 2009, 2012; Borisova and Brown, 2013; Brown and Petersen, 2015).

[11]The model is in the spirit of Bustamante and Frésard (2017), who use a similar approach to examine the asymmetry in peer effects on physical capital investments.

[12]We thank an anonymous reviewer for the helpful insights relating to factors that best proxy for leader and follower status.

4.2. Instruments for 2SLS estimation

A potential endogeneity problem associated with studies of peer effects is the "reflection problem" (Manski, 1993). This problem arises due to the difficulty of disentangling peer effects from common industry effects when industrial characteristics dictate corporate policies. That is, endogeneity arises in an attempt to infer whether the average group behaviour influences the behaviour of an individual who belongs to the group. Leary and Roberts (2014) also stress that this form of endogeneity arises from the selection of firms into peer groups or an omitted common factor and then attempting to identify whether the response to peer effects operates through their actions or characteristics. We address these potential endogeneity concerns by estimating instrumental variable models (IV-Tobit and 2SLS) using appropriate instruments. By addressing potential endogeneity issues, we are able to analyse the extent to which peer firms' innovation policies influence the focal firm's corporate innovation beyond other firm-specific characteristics.

Following Leary and Roberts (2014), and Adhikari and Agrawal (2018), we use peer firms' idiosyncratic stock returns ($\overline{EShock}$) and risks ($\overline{ERisk}$) as our instruments for the endogenous variable (peer average R&D). The instruments are relevant for the following reasons. First, idiosyncratic stock returns and risks are unique to a specific firm, and as such, are less likely to affect the innovation decisions of the focal firm directly. Second, stock returns are widely available for all firms and are not easy to manipulate, unlike accounting-based measures of performance.[13] Third, several studies find a significant relationship between innovation and stock returns (Hirshleifer et al., 2013; Gu, 2016; Hirshleifer et al., 2018). Hence, we estimate the extent of peer influence on R&D using the peer firms' idiosyncratic stock returns and risks as instruments.

We estimate the idiosyncratic stock returns and risks using the augmented Carhart

[13]This reduces concerns about the external validity and reliability of our empirical tests. While we do not claim that the instruments are perfect, we contend that the extant literature supports the instruments and, as such, their predictive power for corporate innovation has been widely demonstrated.

(1997) model with four factors (size, book-to-market, and momentum) as follows:[14]

$$
\begin{aligned}
R_{ijt} =& \alpha_{ijt} + \beta^{M}_{ijt}(RM_t - RF_t) + \beta^{SMB}_{ijt} \times SMB_t + \beta^{HML}_{ijt} \times HML_t \\
&+ \beta^{MOM}_{ijt} \times MOM_t + \beta^{IND}_{ijt}(\overline{R}_{-ijt} - RF_t) + \eta_{ijt}
\end{aligned} \tag{3}
$$

where R_{ijt} is the total stock return for firm i in industry j over the month t, $RM_t - RF_t$ is the excess market return, SML_t is the size factor, HML_t is the book-to-market factor, MOM_t is the momentum factor, $\overline{R}_{-ijt} - RF_t$ is the excess return on an equally weighted industry portfolio (where the industry is defined at the three-digit SIC code), excluding firm i's return, and η_{ijt} is the error term.

We estimate Equation (3) on a rolling annual basis using monthly stock returns. At a minimum, we require that each firm should have at least 24 months of historical returns data and use up to 60 months of data for the estimations. Using the estimated coefficients from Equation (3) for the previous year (*t-1*) and the monthly factors returns for the current year (t), we use Equation (3) to compute the expected return and Equation (4) for the idiosyncratic return and risk as follows:-

$$
\begin{aligned}
\textit{Expected Return} \equiv \widehat{R}_{ijt} = & \widehat{\alpha}_{ijt} + \widehat{\beta}^{M}_{ijt}(RM_t - RF_t) + \widehat{\beta}^{SMB}_{ijt} \times SMB_t \\
&+ \widehat{\beta}^{HML}_{ijt} \times HML_t + \widehat{\beta}^{MOM}_{ijt} \times MOM_t \\
&+ \widehat{\beta}^{IND}_{ijt}(\overline{R}_{-ijt} - RF_t)
\end{aligned} \tag{4}
$$

$$
\textit{Idiosyncratic Return} \equiv \widehat{\eta}_{ijt} = R_{ijt} - \widehat{R}_{ijt} \tag{5}
$$

We compute the annual idiosyncratic return as the geometric average of the monthly idiosyncratic returns from Equation (5). Consistent with Adhikari and Agrawal (2018), the annual idiosyncratic risk is the standard deviation of the idiosyncratic returns from Equation (5). Our instruments, the average peer idiosyncratic return ($\overline{EShock}$) and aver-

[14]Our approach is in line with Adhikari and Agrawal (2018), who examine peer effects on dividends pay-out policies.

age peer idiosyncratic risk ($\overline{ERisk}$), are calculated analogously as discussed above for the peer-firm averages in Equation (1).

PLEASE INSERT TABLE 2 HERE

Table 2 presents the summary statistics of the factor loadings estimated using Equation (3). The average (median) number of months for the regressions is 57 (60), and the R^2 is 0.343 (0.334). Consistent with Adhikari and Agrawal (2018), our factors load positively on the market (β_{ijt}^{M}), size (β_{ijt}^{SMB}), book-to-market (β_{ijt}^{HML}) and industry (β_{ijt}^{IND}), while they load negatively on momentum (β_{ijt}^{MOM}). The mean (median) β_{ijt}^{SMB} for our sample is 0.414 (0.359), which is relatively lower than 0.945 (0.922) for Adhikari and Agrawal (2018), but similar to Leary and Roberts (2014), who report 0.399 (0.422). All the other factor loadings are consistent with the literature.

We first discuss the models' diagnostic tests as our results largely depend on the validity of our instruments. To assess the appropriateness and validity of our instruments, we report the First-Stage F-Statistic (weak instruments test — the relevance of our instruments) and Hansen J-Statistic of over-identification restrictions (which tests the null hypothesis that our instrumental variables are jointly exogenous). A small Hansen J-Statistic indicates that the instruments are valid and appropriate (Sargan, 1958; Baum et al., 2007; Barth et al., 2013). For all our models, the First-Stage F-Statistic is greater than Staiger and Stock (1997)'s rule of thumb (a minimum critical value of 10), suggesting no apparent weak instrument problem (supporting the relevance of our instrument). In addition, the Wald Test of Exogeneity is significant in all cases, indicating sufficient evidence to reject the null hypothesis of exogeneity. In our case, the significant Wald Test of Exogeneity suggests that peer firm averages are not exogenous, hence, the need to address potential endogeneity issues *via* instrumental variables (IV) estimations (see Baum et al., 2003, 2007; Huang et al., 2018; Xing et al., 2020). Moreover, the Hansen J-Statistic is small in all our models, suggesting that the null is not rejected and over-identification restrictions are valid (see Sargan, 1958; Baum et al., 2007; Roodman, 2006). In summary, the models' diagnostic tests provide sufficient confidence that our instruments address potential endogeneity issues and significantly predict endogenous variable

— peer average R&D.[15] Hence, the validity of our instruments is confirmed, and the instrumental variable regressions (IV-Tobit, 2SLS and IV-Probit) are robust estimations for our empirical analyses.

5. Empirical Results

5.1. Peer influence on corporate innovation

In this section, we test Hypothesis 1 (H1a & H1b), which examines whether or not peer firms influence the focal firm's R&D. Table 3 presents the Tobit (Column (1)), OLS (Column (2)), 2SLS (Columns (3) and (4)), and IV-Tobit (Columns (5) and (6)) regression estimations. For all the estimations, the dependent variable is R&D intensity measured using R&D expenditure to total assets (RD/TA). Using peer firms' idiosyncratic stock returns ($\overline{EShock}$) and risks ($\overline{ERisk}$) as instruments, the 2SLS and IV-Tobit estimations correct for endogeneity concerns associated with identifying peer averages of R&D. We control for both peer and firm-specific characteristics, such as cash ratio, market-to-book ratio, debt, firm size, and return on assets in all regressions. All models include industry and year-fixed effects (but not reported for brevity of presentation).

PLEASE INSERT TABLE 3 HERE

Table 3 presents the results estimating the Tobit, OLS, 2SLS, and IV-Tobit regressions. We are, however, cognisant that using OLS and 2SLS might lead to biased inferences as our dependent variable, R&D, is censored at zero and one. We nonetheless present results based on OLS and 2SLS as a form of robustness and to facilitate comparisons with prior studies that use similar approaches (e.g., Chen and Chang, 2012; Leary and Roberts, 2014; Adhikari and Agrawal, 2018). To address problems associated with estimating a model with a censored regressor, we use Tobit and IV-Tobit regressions for our main analyses (see Adhikari and Agrawal, 2018). Using IV-Tobit also enables us to

[15] Appendix B shows that the correlations between firm-specific factors and both contemporaneous and one-period lead peer equity shocks and risks are quite low and not significant. This indicates that the peer equity shocks and risks that we use as instruments are quite reliable as they contain less information about the firm's current or near-future observable R&D determinants.

address endogeneity concerns associated with the study of peer effects. Our results show that the coefficient of peer average R&D ($\overline{RD/TA}$) is consistently positive and significant at 1% level for the Tobit (Column (1)) and OLS (Column (2)) estimations. Similarly, we find a positive and significant relationship between peer average R&D and a firm's R&D investments in the second stage of the 2SLS and IV-Tobit models as reported in Columns (3) and (5), respectively. The results show that, on average, firms increase R&D expenditure by between 3% and 4% for a one standard deviation increase in peer R&D.

These significant and positive peer effects are in line with Hypothesis 1b (H1b) and corroborate studies documenting significant peer influence on other corporate decisions (e.g. Chen and Chang, 2012; Francis et al., 2016; Bustamante and Frésard, 2017; Park et al., 2017; Adhikari and Agrawal, 2018). The peer effects that we document are particularly important due to the unique features of R&D - high irreversibility, information asymmetry and asset substitution concerns, long investment horizons, and low-collateral values - that make it difficult or less desirable for firms to mimic.[16] The effects of these unique features lead to Hypothesis 1a (H1a), which is not supported in our case. Therefore, as we find significant peer effects despite the aforementioned militating factors, we contend that our study provides more rigorous tests and persuasive empirical evidence of peer influence on corporate decisions.

We also find that other peer average factors, which we include to control for other channels through which peer firms might influence innovation, mostly have a low or insignificant effect on R&D intensity. The few significant coefficients of the peer averages tend to have an opposite sign to that of the corresponding firm-specific factors. For the firm-specific factors, we find that the coefficients of cash, market-to-book ratio, and size are consistently significant and positively correlated with R&D. The positive effect of cash and market-to-book ratio on RD/TA is in line with Brown et al. (2009, 2012) and Borisova and Brown (2013), who find that most R&D is attributable to cash-rich and high-growth firms. Similarly, Brown and Petersen (2011) find that firms tend to hoard

[16]These unique or special features of R&D investments are well documented in the literature (see Brown et al., 2009; Brown and Petersen, 2011; Borisova and Brown, 2013).

cash in order to smooth out R&D. Recently, He and Wintoki (2016) also link the increase in cash holdings to the surge in R&D. These prior studies collectively show that cash and R&D are positively correlated in line with Table 3. For debt and profitability, we find that they are negatively correlated with RD/TA as is consistent with prior literature (e.g. Hall, 2002; Hall et al., 2009; Aghion et al., 2004; Hall and Lerner, 2010). As our results for firm-specific factors are consistent with the literature, we, therefore, do not further discuss them for brevity.

To summarise, we find significant peer influence on innovation. This finding suggests that peer effects documented in the extant literature on other corporate decisions also influence innovation beyond other firm-specific and macroeconomic determinants.

5.2. Competition and peer firm effects on corporate innovation

In this section, we test Hypothesis 2 (H2), which posits that peer influence on R&D increases with product market competition. Hypothesis 2 (H2) is premised on Hart (1983) and Aghion et al. (1999), who contend that competition drives corporate innovation by acting as a disciplinary device that spurs managerial action. According to Aghion et al. (2001), firms innovate to keep abreast of competition and enhance their growth prospects in highly competitive product markets. This implies that peer effects increase with product market competition, which is in contrast to the traditional Schumpeterian model predicting a negative effect of competition on innovation and monopoly rents (Schumpeter, 1947; Hart, 1983; Grossman and Helpman, 1991; Aghion et al., 1999). Following the traditional Schumpeterian model, peer influence could decrease rather than increase with product market competition. Against this background, it is, therefore, not *a priori* clear how competitive pressures moderate peer effects on corporate innovation.

To examine the above opposing predictions, we create a dummy variable to indicate whether a firm faces high (low) product market competition based on whether the firm is below (above) the median Herfindahl–Hirschman Index (HHI-Assets and HHI-Sales), Concentration Index for top 4 firms (CR4-Assets and CR4-Sales). Giroud and Mueller (2011) and Adhikari and Agrawal (2018) use similar proxies and approaches to examine the effects of product market competition on equity prices and peer effects on dividends, respectively. Using similar approaches enable us to test Hypothesis 2 (H2), which is premised on the rivalry theory. The rivalry theory posits that firms respond by matching or exceeding their peers in order to maintain or enhance competitive positions in the product markets (Lieberman and Asaba, 2006). Therefore, the independent variables of interest, in Equation (2), are the interaction terms between peer average R&D and dummies for capturing the level of competition that the focal firm faces. The dummy variable Low ($High$) denote a firm that has below (above) median HHI-Assets, HHI-Sales, CR4-Assets, and CR4-Sales, which implies the firm faces high (low) product market competition. Accordingly, the interaction terms $\overline{RD/TA} \times Low$ and $\overline{RD/TA} \times High$ are for firms facing high and low product market competition, respectively. Table 4 summarises the results.

PLEASE INSERT TABLE 4 HERE

Column (1), which uses the Herfindahl Index of assets (HHI-Assets) to measure product market competition, show that peer effects on R&D are more pronounced and significant when a firm operates in highly competitive product markets ($\overline{RD/TA} \times Low$) relative to those in less competitive product markets ($\overline{RD/TA} \times High$). Specifically, we find that the average firm in very (less) competitive industries increases R&D by 4.4% (1.8%) for a one standard deviation increase in peer firms' R&D. This effect is statistically and economically significant at 1%. We find a similar response based on the Herfindahl Index of sales (HHI-Sales) in Column (2), with firms in more competitive industries increasing R&D by 4.5% for a one standard deviation increase in peer firms' R&D compared to 1.4% for those operating in less competitive industries. The χ^2 statistic for the difference in the coefficients for low *versus* high HH–Assets and HHI–Sales are 46.12 and 71.04, respectively. This indicates that even though the coefficients of peer firms' R&D are positive and significant when firms face high and low product market competition, they are statistically different. The differences suggest that the incentive to mimic peer firms is more pronounced for firms in highly competitive product markets. This result collaborates Aghion et al. (2001) and Aghion et al. (2005), who find a positive link between product market competition and corporate innovation. Our results build on and extend these two prior studies by showing that product market competition not only motivates firms to innovate but also to follow peer firms in order to remain competitive.

Next, we use alternative measures of product market competition, such as the Concentration Index based on assets (CR4-Assets) and sales (CR4-Sales) of the top 4 firms in the industry to test the robustness of our results. Further comparisons using the Concentration Index (CR4-Assets in Column (3) and CR4-Sales in Column (4)) show similar differences (as those based on the Herfindahl Index), with peer effects being more pronounced for firms in more competitive industries relative to those in less-competitive industries. Specifically, the coefficients of peer average R&D is positive and significant at the 1% level in Columns (3) and (4) for high-product market competition ($\overline{RD/TA} \times Low$) and low-product market competition ($\overline{RD/TA} \times High$). Again, the χ^2 statistic of 41.69 and 45.62 indicates that the coefficients for low and high product market competition are statistically different. Based on these analyses, we conclude that our results are robust to how we measure or define product market competition.

Overall, our results suggest that peer effects increase (decrease) with product market competition (concentration). This provides new empirical support on the predictions of the rivalry theory (Lieberman and Asaba, 2006), which posits that mimicking increases with product market competition. For R&D, matching or exceeding peer firms is important for two reasons. First, technological laggards need to catch up with the leaders and enhance their profitability (Aghion et al., 2001, 2005). Second, innovation reduces production costs and enhances the survival of firms operating in highly competitive industries (Nickell, 1996; Blundell et al., 1999). Thus, innovation enables firms to reduce competitive pressures and generate superior and sustainable profitability. The increase in peer effects with product market competition that we find is in line with Hypothesis 2 (H2) and shows that firms are more responsive to their rivals when they face high product market competition.

5.3. Which firms are mimicked?

In this section, we test Hypothesis 3 (H3), which predicts that firms subject to information asymmetry mimic those perceived to have superior information. This hypothesis is based on the information theory and implies significant heterogeneity in how firms respond to their peers. To test this hypothesis, we use four firm-specific variables (profitability, firm size, sales, and analyst followings) that are widely used to proxy for information asymmetry (see Leary and Roberts, 2014; Bustamante and Frésard, 2017; Adhikari and Agrawal, 2018). Thus, we create a dummy variable, *Low*, if the firm is less profitable, small, generate low sales and have few analysts followings. Similarly, *High* is a dummy variable for firms that are more profitable, large, have high sales, and high analyst followings. We define low (high) profit margins, small (large) firm size, low (high) sales, and low (high) analyst followings as firms reporting below (above) the median profit margin, firm size, sales, and analyst followings in each year, respectively. Following the information theory, we expect *Low* and *High* firms, which can be defined as *Followers* and *Leaders*, respectively, to be more (less) responsive to their peers. Table 5 summarises the results estimating Equation (2), where the independent variables are the interaction terms between peer average R&D and dummies for *Followers* ($\overline{RD/TA} \times Low$) and *Leaders* ($\overline{RD/TA} \times High$).

PLEASE INSERT TABLE 5 HERE

Table 5 shows significant heterogeneity in peer influence. The coefficients of $\overline{RD/TA} \times Low$ are positive and significant at the 1% level, as shown in Columns (1)–(4). Our findings suggest that *Followers* increase R&D, on average, by 4.5% for a one standard deviation increase in peer R&D. Similarly, we find also that $\overline{RD/TA} \times High$ is positively correlated with the focal firm's R&D, with an average increase of about 2% for a one standard deviation increase in peer firms' R&D. While these results suggest that both *Followers* and *Leaders* tend to mimic peer firms' R&D, the magnitude of the coefficients of $\overline{RD/TA} \times Low$ is larger than $\overline{RD/TA} \times High$. This is confirmed by the statistically significant χ^2 statistics for the test of difference between the coefficients of $\overline{RD/TA} \times Low$ and $\overline{RD/TA} \times High$ (Diff — *High vs Low*). Our results hold regardless of the measure or approach used to identify follower or leader firms. We further examine, in untabulated results, cross-cohort mimicking behaviour and find that *Followers* (*Leaders* tend to mimic other followers (leaders). Interestingly, *Followers* mimic *Leaders* and *vice versa*, with the mimicking by *Followers* being more pronounced relative to that by *Leaders*. Our results, therefore, suggest that peer effects matter for both followers and leaders, with this effect being more important for follower firms.

The above asymmetric and bidirectional cross-cohort peer effects, which are, on average, higher for *Followers* (low-profitability, small size, low sales, and fewer analyst followings) relative to *Leaders* (high-profitability, large size, high sales, and more analyst followings), are in line with the information-based theory as they show that firms with less informational advantages mimic those perceived to have more informational advantages. This form of mimicking appears to be beneficial as it enables a firm to play it safe or learn new information from its peers without expending resources to find or determine the optional R&D policy. The significant bidirectional peer effects that we document also show that the feedback theory partly explains how firms respond to their peers as leaders copy policies of their followers to undercut or drive them out of business. This is akin to a predatory response to followers that undermines any strategic inroads or serve to dampen competitive pressures in competitive product markets (Bolton and Scharfstein, 1990). This empirical evidence is new to the literature on peer effects and shows important strategic interactions that might potentially amplify both positive and negative firm-specific shocks within and across industries.

To summarise, our results show that peer effects are bidirectional and asymmetric between cohorts or subgroups. The incentive to innovate is higher for followers who are, on average, still seeking to establish themselves within the product markets. As our results show, peer influence significantly drives innovation in addition to other factors so far examined in the literature.

5.4. The implications of mimicking corporate innovation

While the previous sections show significant peer effects, it remains an empirical question as to whether mimicking innovation is beneficial or not. Accordingly, we test Hypothesis 4 (H4), which posits that mimicking peer firms has positive implications on corporate outcomes. Understanding the implications of mimicking is important as peer effects amplify firm-specific shocks both across and within industries. To tackle this under-researched, but yet important question, we follow Nason and Patel (2016) and Gyimah et al. (2020), and estimate the following model:

$$y_{ijt+1,t+5} = \gamma_0 + \gamma_1 \overline{RD/TA}_{-ijt} + \boldsymbol{\theta} \boldsymbol{Z}_{ijt} + \mu_j + \mu_t + \xi_{ijt} \tag{6}$$

where $y_{ijt+1,t+5}$ is the corporate outcome (firm value and performance, patents counts and value, citations, product similarity and fluidity) for firm i in industry j over the window from time $t+1$ to $t+5$, γ_0 is a constant. γ_1, γ_2 and $\boldsymbol{\theta}$ are the vectors of coefficients to be estimated. $\overline{RD/TA}_{-ijt}$ are the peer effects calculated as the industrial average excluding the focal firm. As discussed previously in Section 4, we use the peer firms' idiosyncratic stock returns ($\overline{EShock}$) and risks ($\overline{ERisk}$) as our instruments for the endogenous peer effects. $\boldsymbol{Z}_{ijt}$ is a vector of sales growth (*Sales Growth*), cash (*Cash*), debt (*Debt*), property, plant and equipment (*PPE*), firm-size (*Size*) and the logarithm of firm-age (*LogAge*). μ_k and μ_t are industry and year-fixed effects, respectively. Finally, ξ_{ijt} is the error term.

Table 6 presents the results estimating Equation (6) for the implications of mimicking on stock market valuation using Tobin's q, innovation outputs, such as patent counts ($LOGPATS$) and patent value ($LOGTCW$ and TSM/TA), and novelty of products using product similarity (SIM) and product fluidity ($FLUIDITY$). Output data for innovation, where patent counts and forward citations are measured using the grant-year, is from Kogan et al. (2017).[17] $LOGPATS$ is the logarithm of the number of patent counts, $LOGTCW$ is the citation-weighted value of patents, and TSM/TA is the stock market value of patents. Product novelty SIM and $FLUIDITY$ is the measure of product similarity (Hoberg and Phillips, 2016) and product fluidity (Hoberg et al., 2014), respectively.[18] We recognise the uncertain process of translating R&D inputs into

[17]This patents data is available from https://iu.app.box.com/v/patents and the NBER website — https://data.nber.org/patents/.

[18]Product similarity, according to Hoberg and Phillips (2016), is a firm-by-firm pairwise measure of product similarity based on the product descriptions from firms' 10-K annual reports. Product fluidity, on the other hand, represents a firm-level measure based on each firm's unique product-market vocabulary, assessing how intensely product market changes around a firm (Hoberg et al., 2014). We extract this data from the authors' website on `http://hobergphillips.tuck.dartmouth.edu/`.

outputs, which might take considerable time to materialise. At the same time, a firm can only improve its product novelty once it has achieved some advances in technology and successfully commercialises these advancements. Therefore, we follow Chan et al. (2001), Hirshleifer et al. (2013), Bena and Li (2014) and Hsu et al. (2018) in measuring our dependent variables, innovation outputs and product market performance, over a five-year window.

PLEASE INSERT TABLE 6 HERE

Table 6 summarises the results estimating Equation (6). Columns (1)–(6) show that mimicking R&D has a significant positive effect on corporate outcomes, such as stock market value (Tobin's q), innovation outputs ($LOGPATS$, $LOGTCW$ and TSM/TA) and product novelty (SIM and $FLUIDITY$). This result underscores the notion that current managerial decisions regarding R&D can enhance or hurt the viability, growth, and competitiveness of an organisation in the long-run (see Morbey, 1988; Ehie and Olibe, 2010). This positive effect on corporate outcomes is in line with Kallunki et al. (2009), who report that R&D has significant positive effects on stock valuation following technology-oriented M&As. Our findings corroborate this study by showing that mimicking R&D increases firm value, patent counts, patent value and product novelty. This finding is important to corporate strategy as these factors directly affect firm survival and growth in increasingly competitive product markets (see Harford, 1999; Cockburn and MacGarvie, 2011; Kogan et al., 2017).

To ensure the robustness of our results, we also use two other alternative approaches to identify mimicking firms. First, we use a modified version of Equation (1) that includes a firm dummy (IDD) and the interaction term, $\overline{RD/TA}_{-ijt} \times IDD$, to identify mimicking firms. Specifically, we categorise firms as mimickers if they are in the upper quartile of the distribution of the $\overline{RD/TA}_{-ijt} \times IDD$ coefficient in each year. We then substitute $\overline{RD/TA}_{-ijt}$ with the mimicking dummy, $Mimicking$, in Equation (6). Based on this first alternative approach, we find qualitatively similar results to those in Table 6, which suggests that mimicking R&D has real and long-term implications on innovation outputs and product market performance. Second, we follow Fairhurst and Nam (2018) and use the regression diagnostic statistic (DFBETA) to identify mimicking firms. The DFBETA test statistics are traditionally used to quantify the impact of deleting an observation on the regression coefficients. Accordingly, the DFBETAs could be used, as in our case, to identify firms that have the highest impact on the peer R&D coefficient, which for the purposes of this study, we categorise as mimicking firms. To obtain the DFBETAs, we run an OLS regression of Equation (1) and then categorise firms in the upper quartile of the distribution of DFBETAs in each year as mimicking firms. Specifically, we create a mimicking dummy, $Mimicking$, that takes the value of one if a firm is categorised in the upper quartile of the DFBETAs distribution in each year and zero otherwise. Using this approach yields qualitatively similar results as those reported in Table 6. These additional robustness checks suggest that our findings are not driven by the approach or method we use to identify mimicking firms.[19]

In summary, our analyses show that mimicking peer firms' R&D is beneficial as it reduces search costs incurred when a firm wants to optimise its decisions. These findings provide new and important empirical insights on a relatively-unexplored research question of whether mimicking peer firms is beneficial or not.

6. Robustness

In this section, we implement a battery of robustness checks. First, we re-estimate our main results based on several broad industrial sector categorisations. We categorise firms into three broad industries, namely; Others (all other industries except mining and manufacturing), mining ($1000 \leq$ SIC code ≤ 1499) and manufacturing ($2000 \leq$ SIC code ≤ 3999). We also further partition firms in the manufacturing sector into durables ($2400 \leq$ SIC code ≤ 2500 and $3200 \leq$ SIC code ≤ 3800) and non-durables ($2000 \leq$ SIC code ≤ 2300 and $2600 \leq$ SIC code ≤ 3100). Table 7 summarises the results estimating Equation (1) for the sub-samples.

PLEASE INSERT TABLE 7 HERE

[19]The results from these additional robustness checks are available upon request. We thank an anonymous reviewer for highlighting this insightful extension to our paper.

Columns (1)–(5) of Table 7 show that peer effects vary significantly across the sub-samples. We find, in Columns (1)–(3), that peer influence on R&D is significant across the three broad industries.[20] For firms in the manufacturing sectors, in Columns (4) and (5), we further find that peer effects are significant for firms in durables, but not the non-durables sector, which highlights the central role of innovation in corporate strategy and how the interplay of industry dynamics influence firm's innovation policies. Overall, the results for the sub-samples are qualitatively similar to our main findings.

PLEASE INSERT TABLE 8 HERE

To address concerns that our measure of corporate innovation might drive our results, we re-estimate our baseline model using five other definitions of innovation, namely: the R&D dummy (*RDD*), R&D to net assets (*RD/NA*), change in R&D to total assets (Δ*RD/TA*), the sum of R&D and selling, general and administrative expenses (SG&A) to total assets (*RDSGA/TA*), and R&D *plus* selling, general and administrative expenses to sales (*RDSGA/Sales*). Using *RDSGA/TA* and *RDSGA/Sales* address concerns that some firms might attempt to strategically manage R&D expenses by selectively reporting them as part of selling, general and administrative expenses (see Banker et al., 2011; Lévesque et al., 2012). These comprehensive measures of innovation also address the selective or strategic reporting of R&D (see Koh and Reeb, 2015), which could significantly bias our inferences. Consistent with Table 3, Columns (1)–(5) of Table 8 show that peer effects are significant across our five alternative measures of innovation. These findings suggest that our results are not driven by how we define or identify innovation.

PLEASE INSERT TABLE 9 HERE

In addition to using the three-digit SIC code as the primary definition of our industries, we also conduct several robustness tests using other five alternative industry classifications. Specifically, we define our alternative industries using the one-digit SIC

[20] Our unreported results that plot the average R&D across the different industry groupings, we find that firms in the manufacturing (durables) sectors report significantly higher R&D than those in the non-manufacturing (non-durables) sectors.

code (SIC1), two-digit SIC code (SIC2), two-digit North America Industry Classification System (NAICS2) and 10-K-based Fixed Industry Classifications (FIC200 and FIC300) (see Hoberg and Phillips, 2010, 2016). Table 9, which tabulates the results for our alternative industry classifications, shows that changing the level or way in which we define the industries does not significantly affect our main findings. Appendix C and D similarly show that our results are robust to using different estimation techniques, excluding and including industry and year-fixed effects, and excluding firms that change industries or have multiple business segments. Finally, Appendix E shows that our results are robust to controlling for several factors linked to changes in credit markets and macroeconomic conditions.

Based on the tests above, we conclude that peer firms significantly influence innovation and that results are robust to a host of factors that could potentially bias our inferences.

7. Conclusion

We extend the literature on peer effects on corporate decisions by examining whether peer firms influence corporate innovation in the US over the past five decades. This period is interesting as it is marked by a surge in R&D to levels that exceed physical or tangible investments. To the best of our knowledge, we make the first attempt at testing whether peer effects matter for corporate innovation. By focusing on corporate innovation, we provide sharper tests on peer effects as mimicking is difficult or less desirable given that corporate innovation is susceptible to high-asymmetric information problems, asset substitution issues, low-collateral values, irreversibility and long-investment horizons.

Our results, which are robust to a battery of tests, suggest that peer effects matter for corporate innovation. Specifically, we find that firms are more responsive to peers in highly competitive product markets and more importantly, firms learn from one another in defining their corporate policy choices. The information theory underpins such a learning process by which firms follow peers that are perceived to have superior information about the market. Even though we find bidirectional peer effects, the analyses emphasise that *leaders* more significantly influence their *followers* relative to the effects of *followers* on *leaders*.

Next, we explore the implications of mimicking corporate innovation on innovation outputs and product market performance. We find evidence suggesting that adopting innovation policies of rivals is beneficial as it reduces information search costs incurred when attempting to optimise corporate decisions. Our results confirm this hypothesis, indicating the beneficial effects of mimicking on both long-term innovation outputs and product market performance.

Overall, peer effects emerge from our findings as an important determinant of corporate innovation in addition to other factors so far considered in the literature. Our findings of significant peer effects on R&D have both academic and policy implications. Academic researchers should consider peer effects as one of the new and important determinants of corporate innovation. Policymakers, on the other hand, should have a keen interest in peer effects as they could potentially amplify the impact of firm-specific shocks both within and across industries. This has real economic and welfare implications, and at the same time, represents an exciting future research area that could enhance our understanding of industry dynamics and corporate decisions.

References

Acemoglu, D., Aghion, P., Zilibotti, F., 2006. Distance to frontier, selection, and economic growth. Journal of the European Economic Association 4, 37–74.

Acharya, V., Xu, Z., 2017. Financial dependence and innovation: The case of public versus private firms. Journal of Financial Economics 124, 223–243.

Adhikari, B.K., Agrawal, A., 2018. Peer influence on payout policies. Journal of Corporate Finance 48, 615–637.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-U relationship. The Quarterly Journal of Economics 120, 701–728.

Aghion, P., Bond, S., Klemm, A., Marinescu, I., 2004. Technology and financial structure: Are innovative firms different? Journal of the European Economic Association 2, 277–288.

Aghion, P., Dewatripont, M., Rey, P., 1999. Competition, financial discipline and growth. The Review of Economic Studies 66, 825–852.

Aghion, P., Harris, C., Howitt, P., Vickers, J., 2001. Competition, imitation and growth with step-by-step innovation. The Review of Economic Studies 68, 467–492.

Aghion, P., Harris, C., Vickers, J., 1997. Competition and growth with step-by-step innovation: An example. European Economic Review 41, 771–782.

Aghion, P., Van Reenen, J., Zingales, L., 2013. Innovation and institutional ownership. American Economic Review 103, 277–304.

Atanassov, J., 2015. Arms length financing and innovation: Evidence from publicly traded firms. Management Science 62, 128–155.

Atkeson, A., Burstein, A., 2019. Aggregate implications of innovation policy. Journal of Political Economy 127, 2625–2683.

Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. The Journal of Finance 61, 1645–1680.

Banbury, C.M., Mitchell, W., 1995. The effect of introducing important incremental innovations on market share and business survival. Strategic Management Journal 16, 161–182.

Banker, R.D., Huang, R., Natarajan, R., 2011. Equity incentives and long-term value created by SG&A expenditure. Contemporary Accounting Research 28, 794–830.

Barth, J.R., Lin, C., Ma, Y., Seade, J., Song, F.M., 2013. Do bank regulation, supervision and monitoring enhance or impede bank efficiency? Journal of Banking & Finance 37, 2879–2892.

Bates, T.W., Kahle, K.M., Stulz, R.M., 2009. Why do U.S. firms hold so much more cash than they used to? The Journal of Finance 64, 1985–2021.

Baum, C.F., Schaffer, M.E., Stillman, S., 2003. Instrumental variables and GMM: Estimation and testing. The Stata Journal 3, 1–31.

Baum, C.F., Schaffer, M.E., Stillman, S., 2007. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. The Stata Journal 7, 465–506.

Bena, J., Li, K., 2014. Corporate innovations and mergers and acquisitions. The Journal of Finance 69, 1923–1960.

Bird, A., Edwards, A., Ruchti, T.G., 2018. Taxes and peer effects. The Accounting Review 93, 97–117.

Blundell, R., Griffith, R., Van Reenen, J., 1999. Market share, market value and innovation in a panel of British manufacturing firms. The Review of Economic Studies 66, 529–554.

Bolton, P., Scharfstein, D.S., 1990. A theory of predation based on agency problems in financial contracting. The American Economic Review 80, 93–106.

Borisova, G., Brown, J.R., 2013. R&D sensitivity to asset sale proceeds: New evidence on financing constraints and intangible investment. Journal of Banking & Finance 37, 159–173.

Brander, J.A., Lewis, T.R., 1986. Oligopoly and financial structure: The limited liability effect. The American Economic Review 76, 956–970.

Brennan, M.J., Subrahmanyam, A., 1995. Investment analysis and price formation in securities markets. Journal of Financial Economics 38, 361–381.

Brown, J.R., Fazzari, S.M., Petersen, B.C., 2009. Financing innovation and growth: Cash flow, external equity, and the 1990s R&D boom. The Journal of Finance 64, 151–185.

Brown, J.R., Martinsson, G., Petersen, B.C., 2012. Do financing constraints matter for R&D? European Economic Review 56, 1512–1529.

Brown, J.R., Petersen, B.C., 2011. Cash holdings and R&D smoothing. Journal of Corporate Finance 17, 694–709.

Brown, J.R., Petersen, B.C., 2015. Which investments do firms protect? Liquidity management and real adjustments when access to finance falls sharply. Journal of Financial Intermediation 24, 441–465.

Bui, D.G., Chen, Y., Lin, C.Y., Lin, T.C., 2019. R&D expenditure as a response to peer influence. SSRN Scholarly Paper ID 3412770. Social Science Research Network. Rochester, NY, USA.

Bustamante, M.C., Frésard, L., 2017. Does firm investment respond to peers' investment? SSRN Scholarly Paper ID 2827803. Social Science Research Network. Rochester, NY, USA.

Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Chan, L.K., Lakonishok, J., Sougiannis, T., 2001. The stock market valuation of research and development expenditures. The Journal of Finance 56, 2431–2456.

Chang, X., Dasgupta, S., Hilary, G., 2006. Analyst coverage and financing decisions. The Journal of Finance 61, 3009–3048.

Chen, i.W., Chang, Y., 2012. Peer effects on corporate cash holdings. Working Paper. National Chengchi University. Taiwan.

Cockburn, I.M., MacGarvie, M.J., 2011. Entry and patenting in the software industry. Management Science 57, 915–933.

Cyert, R.M., March, J.G., 1963. A behavioral theory of the firm. Englewood Cliffs, NJ: Prentice-Hall.

Ehie, I.C., Olibe, K., 2010. The effect of R&D investment on firm value: An examination of US manufacturing and service industries. International Journal of Production Economics 128, 127–135.

Fairhurst, D.D., Nam, Y., 2018. Corporate governance and financial peer effects. Financial Management , 1–29

Falato, A., Sim, J., 2014. Why do innovative firms hold so much cash? Evidence from changes in state R&D tax credits. SSRN Scholarly Paper ID 2503457. Social Science Research Network. Rochester, NY, USA.

Foucault, T., Frésard, L., 2014. Learning from peers' stock prices and corporate investment. Journal of Financial Economics 111, 554–577.

Francis, B.B., Hasan, I., Kostova, G.L., 2016. When do peers matter?: A cross-country perspective. Journal of International Money and Finance 69, 364–389.

Frésard, L., Valta, P., 2016. How does corporate investment respond to increased entry threat? The Review of Corporate Finance Studies 5, 1–35.

Frydman, C., 2015. What drives peer effects in financial decision-making? Neural and behavioral evidence. SSRN Scholarly Paper ID 2561083. Social Science Research Network. Rochester, NY, USA.

Geroski, P., Machin, S., Van Reenen, J., 1993. The profitability of innovating firms. The RAND Journal of Economics , 198–211.

Giroud, X., Mueller, H.M., 2011. Corporate governance, product market competition, and equity prices. The Journal of Finance 66, 563–600.

Gorodnichenko, Y., Svejnar, J., Terrell, K., 2010. Globalization and innovation in emerging markets. American Economic Journal: Macroeconomics 2, 194–226.

Grennan, J.P., 2019. Dividend payments as a response to peer influence. Journal of Financial Economics, 131, 549–570.

Grossman, G.M., Helpman, E., 1991. Quality ladders in the theory of growth. The Review of Economic Studies 58, 43–61.

Gu, L., 2016. Product market competition, R&D investment, and stock returns. Journal of Financial Economics 119, 441–455.

Gyimah, D., Machokoto, M., Sikochi, A.S., 2020. Peer influence on trade credit. Journal of Corporate Finance 64, 101685.

Hall, B.H., 2002. The financing of research and development. Oxford Review of Economic Policy 18, 35–51.

Hall, B.H., 2009. The financing of innovative firms. European Investment Bank Papers 14, 8–28.

Hall, B.H., Lerner, J., 2010. The financing of R&D and innovation, in: Rosenberg, B.H.H.a.N. (Ed.), Handbook of the Economics of Innovation, Volume 1. North-Holland, pp. 609–639.

Hall, B.H., Mairesse, J., Mohnen, P., 2009. Measuring the returns to R&D. Working Paper 15622. National Bureau of Economic Research.

Harford, J., 1999. Corporate cash reserves and acquisitions. The Journal of Finance 54, 1969–1997.

Hart, O.D., 1983. The market mechanism as an incentive scheme. The Bell Journal of Economics , 366–382.

Hartmann, G.C., Myers, M.B., Rosenbloom, R.S., 2006. Planning your firm's R&D investment. Research-Technology Management 49, 25–36.

He, Z., Wintoki, M.B., 2016. The cost of innovation: R&D and high cash holdings in U.S. firms. Journal of Corporate Finance 41, 280–303.

Hirschey, M., Skiba, H., Wintoki, M.B., 2012. The size, concentration and evolution of corporate R&D spending in U.S. firms from 1976 to 2010: Evidence and implications. Journal of Corporate Finance 18, 496–518.

Hirshleifer, D., Hsu, P.H., Li, D., 2013. Innovative efficiency and stock returns. Journal of Financial Economics 107, 632–654.

Hirshleifer, D., Hsu, P.H., Li, D., 2018. Innovative originality, profitability, and stock returns. The Review of Financial Studies 31, 2553–2605.

Hoberg, G., Maksimovic, V., 2015. Redefining financial constraints: A text-based analysis. The Review of Financial Studies 28, 1312–1352.

Hoberg, G., Phillips, G., 2010. Real and financial industry booms and busts. The Journal of Finance 65, 45–86.

Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. Journal of Political Economy 124, 1423–1465.

Hoberg, G., Phillips, G., Prabhala, N., 2014. Product market threats, payouts, and financial flexibility. The Journal of Finance 69, 293–324.

Hong, H., Lim, T., Stein, J.C., 2000. Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. The Journal of Finance 55, 265–295.

Hsu, P.H., Lee, H.H., Peng, S.C., Yi, L., 2018. Natural disasters, technology diversity, and operating performance. Review of Economics and Statistics 100, 619–630.

Huang, J., Su, C., Joseph, N.L., Gilder, D., 2018. Monitoring mechanisms, managerial incentives, investment distortion costs, and derivatives usage. British Accounting Review 50, 93–141.

Kallunki, J.P., Pyykkö, E., Laamanen, T., 2009. Stock market valuation, profitability and R&D spending of the firm: The effect of technology mergers and acquisitions. Journal of Business Finance & Accounting 36, 838–862.

Kaustia, M., Rantala, V., 2015. Social learning and corporate peer effects. Journal of Financial Economics 117, 653–669.

Kogan, L., Papanikolaou, D., Seru, A., Stoffman, N., 2017. Technological innovation, resource allocation, and growth. The Quarterly Journal of Economics 132, 665–712.

Koh, P.S., Reeb, D.M., 2015. Missing R&D. Journal of Accounting and Economics 60, 73–94.

Kubick, T.R., Lynch, D.P., Mayberry, M.A., Omer, T.C., 2015. Product market power and tax avoidance: Market leaders, mimicking strategies, and stock returns. The Accounting Review 90, 675–702.

Leary, M.T., Roberts, M.R., 2014. Do peer firms affect corporate financial policy? The Journal of Finance 69, 139–178.

Lévesque, M., Joglekar, N., Davies, J., 2012. A comparison of revenue growth at recent-IPO and established firms: The influence of SG&A, R&D and COGS. Journal of Business Venturing 27, 47–61.

Lieberman, M.B., Asaba, S., 2006. Why do firms imitate each other? Academy of Management Review 31, 366–385.

Machlup, F., 1967. Theories of the firm: Marginalist, behavioral, managerial. The American Economic Review 57, 1–33.

Mairesse, J., Mohnen, P., Kremp, E., Kremp, E., 2005. The importance of R&D and innovation for productivity: A reexamination in light of the French innovation survey. Annales d'Economie et de Statistique , 487–527.

Mama, H.B., 2018. Innovative efficiency and stock returns: Should we care about non-linearity? Finance Research Letters 24, 81–89.

Manski, C.F., 1993. Identification of endogenous social effects: The reflection problem. The Review of Economic Studies 60, 531–542.

Morbey, G.K., 1988. R&D: Its relationship to company performance. Journal of Product Innovation Management 5, 191–200.

Mudambi, R., Swift, T., Hannigan, T.J., 2015. Sometimes cutting R&D spending can yield more innovation. Harvard Business Review URL: https://hbr.org/2015/01/sometimes-cutting-rd-spending-can-yield-more-innovation.

Nason, R.S., Patel, P.C., 2016. Is cash king? Market performance and cash during a recession. Journal of Business Research 69, 4242–4248.

Nickell, S.J., 1996. Competition and corporate performance. Journal of Political Economy 104, 724–746.

Park, K., Yang, I., Yang, T., 2017. The peer-firm effect on firms investment decisions. The North American Journal of Economics and Finance 40, 178–199.

Roberts, P.W., 1999. Product innovation, product–market competition and persistent profitability in the US pharmaceutical industry. Strategic Management Journal 20, 655–670.

Roodman, D., 2006. How to do xtabond2. The Stata Journal 9, 86–136.

Sargan, J.D., 1958. The estimation of economic relationships using instrumental variables. Econometrica: Journal of the Econometric Society 26, 393–415.

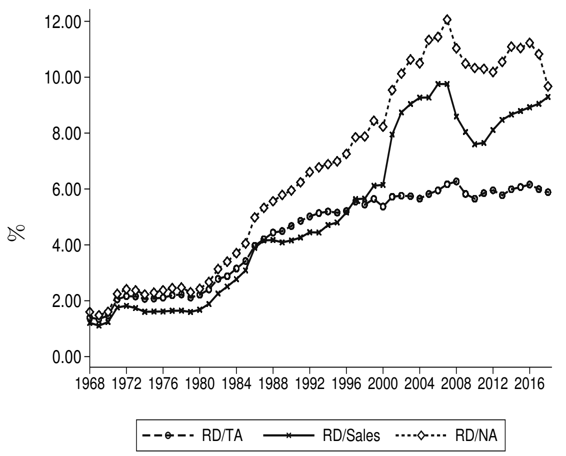
Scharfstein, D., 1988. Product-market competition and managerial slack. The RAND Journal of Economics 19, 147–155.

Scharfstein, D.S., Stein, J.C., 1990. Herd behavior and investment. The American Economic Review 80, 465–479.

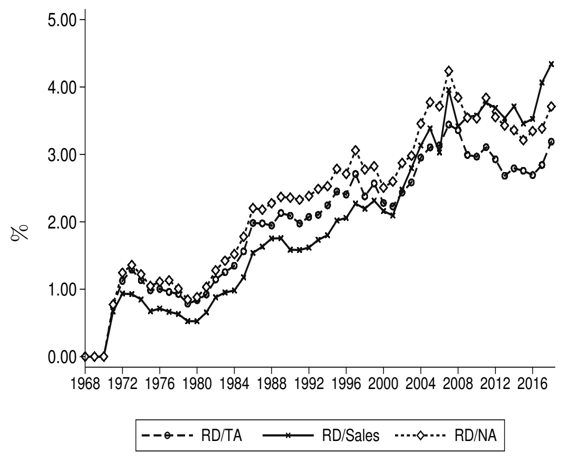
Schumpeter, J.A., 1947. Capitalism, socialism and democracy. London: Allen & Unwin.

Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557.

Xing, L., Gonzalez, A., Sila, V., 2020. Does cooperation among women enhance or impede firm performance? British Accounting Review, 100936.



(a) Mean



(b) Median

Figure 1 Time variations in corporate innovation

The figure plots the mean *RD/TA*, *RD/Sales* and *RD/NA* over the sample period. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period from 1968 to 2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles.

Table 1 Basic statistics

The table presents the summary statistics for all the variables used. The firm-specific characteristics are defined as follows: RD/TA is research and development to total assets, RDD is the research and development dummy, RD/NA is research and development to net assets, $\Delta RD/TA$ is change in research and development to total assets, $RDSGA/TA$ is research and development plus selling, general and administrative expenses to total assets, and $RDSGA/Sales$ is research and development plus selling, general and administrative expenses to total sales. *Cash* is cash and cash equivalent, Q is market to book ratio, *Debt* is total debt, *Size* is logarithm of total assets, ROA is return on assets, *EShock* is equity shock, and *ERisk* is equity risk. The peer firms' average characteristics (e.g., $\overline{RD/TA}$) are calculated as the average of all firms within an industry-year, excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles.

Variables	Mean	SD	Min	p25	Median	p75	Max
Firm-specific factors							
RD/TA	0.045	0.065	0.000	0.019	0.000	0.067	0.529
RDD	0.647	0.478	0.000	1.000	0.000	1.000	1.000
RD/NA	0.069	0.132	0.000	0.021	0.000	0.084	1.580
$\Delta RD/TA$	0.001	0.020	-0.135	0.000	-0.001	0.002	0.167
$RDSGA/TA$	0.355	0.264	0.000	0.301	0.166	0.476	1.623
$RDSGA/Sales$	0.325	0.275	0.000	0.252	0.145	0.433	2.966
Cash	0.152	0.167	0.000	0.086	0.030	0.219	0.822
Q	1.685	1.128	0.413	1.329	1.005	1.931	12.485
Debt	0.193	0.161	0.000	0.177	0.044	0.301	0.729
Size	5.620	2.006	0.952	5.460	4.142	6.918	11.716
ROA	0.079	0.102	-0.641	0.093	0.051	0.133	0.324
EShock	-0.010	0.044	-0.247	-0.008	-0.033	0.015	0.166
ERisk	0.122	0.071	0.025	0.104	0.073	0.150	0.668
Peer firms' average characteristics							
$\overline{RD/TA}$	0.045	0.040	0.000	0.030	0.007	0.081	0.139
$\overline{RDD}$	0.647	0.323	0.000	0.759	0.357	0.922	1.000
$\overline{RD/NA}$	0.069	0.071	0.000	0.036	0.008	0.124	0.366
$\overline{\Delta RD/TA}$	0.001	0.004	-0.020	0.000	-0.001	0.002	0.021
$\overline{\Delta RDSGA/TA}$	0.355	0.158	0.017	0.373	0.235	0.464	0.861
$\overline{\Delta RDSGA/Sales}$	0.325	0.173	0.021	0.282	0.180	0.464	0.919
$\overline{Cash}$	0.152	0.091	0.012	0.122	0.078	0.213	0.384
$\overline{Q}$	1.679	0.543	0.562	1.577	1.277	2.006	4.456
$\overline{Debt}$	0.194	0.073	0.030	0.192	0.138	0.243	0.522
$\overline{Size}$	5.590	0.914	3.401	5.523	4.920	6.198	9.037
$\overline{ROA}$	0.077	0.035	-0.140	0.079	0.055	0.101	0.194
$\overline{EShock}$	-0.010	0.010	-0.069	-0.010	-0.016	-0.004	0.038
$\overline{ERisk}$	0.122	0.034	0.050	0.115	0.098	0.141	0.278
N	51,990						
SIC3 Industries	74						
Firms	4,545						
$RD/TA > 0$ N	33,663						
$RD/TA > 0$ Firms	3,027						
$RD/TA = 0$ N	18,327						
$RD/TA = 0$ Firms	1,518						

Table 2 Stock return regressions

The table presents the summary statistics of the estimation results of the following model:-

$$
\begin{aligned}
R_{ijt} =& \alpha_{ijt} + \beta^{M}_{ijt}(RM_t - RF_t) + \beta^{SMB}_{ijt} \times SMB_t + \beta^{HML}_{ijt} \times HML_t \\
&+ \beta^{MOM}_{ijt} \times MOM_t + \beta^{IND}_{ijt}(\overline{R}_{-ijt} - RF_t) + \eta_{ijt}
\end{aligned} \tag{3}
$$

where R_{ijt} is the total stock return for firm i in industry j over the month t, $RM_t - RF_t$ is the excess market return, SML_t is the size factor, HML_t is the book-to-market factor, MOM_t is the momentum factor, $\overline{R}_{-ijt} - RF_t$ is the excess return on an equally weighted industry portfolio (where the industry is defined at the three-digit SIC code), excluding firm i's return, and η_{ijt} is the error term.

Equation (3) is estimated on a rolling annual basis using monthly returns. At a minimum, each firm is required to have at least 24 months of historical returns data and up to 60 months of data is used for the estimations. Equation (4) is then used to compute the idiosyncratic return and risk using the estimated coefficients from Equation (3) for the previous year (*t-1*) and the monthly factors returns for the current year (t) as follows:-

$$
\begin{aligned}
\textit{Expected Return} \equiv \widehat{R}_{ijt} = & \widehat{\alpha}_{ijt} + \widehat{\beta}^{M}_{ijt}(RM_t - RF_t) + \widehat{\beta}^{SMB}_{ijt} \times SMB_t \\
&+ \widehat{\beta}^{HML}_{ijt} \times HML_t + \widehat{\beta}^{MOM}_{ijt} \times MOM_t \\
&+ \widehat{\beta}^{IND}_{ijt}(\overline{R}_{-ijt} - RF_t)
\end{aligned} \tag{4}
$$

$$
\textit{Idiosyncratic Return} \equiv \widehat{\eta}_{ijt} = R_{ijt} - \widehat{R}_{ijt} \tag{5}
$$

The annual idiosyncratic return is calculated by taking the geometric average of the monthly idiosyncratic returns from Equation (5). Consistent with Adhikari and Agrawal (2018), the annual idiosyncratic risk is calculated as the standard deviation of the idiosyncratic returns from Equation (5). The instruments which are the average peer idiosyncratic return and risk, are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period from 1968 to 2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles.

Variables	Mean	Median	SD
$\widehat{\alpha}_{ijt}$	0.007	0.006	0.019
$\widehat{\beta}^{M}_{ijt}$	0.476	0.515	1.055
$\widehat{\beta}^{SMB}_{ijt}$	0.414	0.359	1.295
$\widehat{\beta}^{HML}_{ijt}$	0.047	0.051	1.076
$\widehat{\beta}^{MOM}_{ijt}$	-0.048	-0.032	0.734
$\widehat{\beta}^{IND}_{ijt}$	0.510	0.431	0.845
Obs per Regression	57	60	8
R^2	0.343	0.334	0.166
Monthly Return	0.049	0.023	0.180
Expected Monthly Return	0.016	0.015	0.036
Idiosyncratic Monthly Return	-0.001	-0.002	0.048

Table 3 The effect of peer firms on corporate innovation

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics. There are four models presented in this table: Tobit (Column (1)), OLS (Column (2)), 2SLS (Columns (3)&(4)), and IV-Tobit (Columns (5)&(6)). The dependent variable, *RD/TA*, is the firm's R&D to total assets. The independent variable is the peer firms' average firm characteristics defined as follows: $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The other peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent to total assets, *Q* is lagged market-to-book ratio, *Debt* is lagged debt to total assets, *Size* is lagged size (logarithm of total assets), *ROA* is lagged profitability (earnings before interest and tax to total assets), *EShock* is the lagged idiosyncratic stock returns, and *ERisk* is the standard deviation of the idiosyncratic stock returns. The instruments used in the first stage of the 2SLS and IV-Tobit models are the lagged peer average idiosyncratic stock returns ($\overline{EShock}$) and the average standard deviation of the peer idiosyncratic stock returns ($\overline{ERisk}$). The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

Estimations	Tobit	OLS	2SLS		IV-Tobit	
			2^{nd} Stage	1^{st} Stage	2^{nd} Stage	1^{st} Stage
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{RD/TA}$	1.024***	0.744***	0.712***		0.914***	
	(0.025)	(0.018)	(0.064)		(0.100)	
$\overline{Cash}$	-0.017*	-0.006	0.000	0.208***	0.004	0.208***
	(0.010)	(0.007)	(0.015)	(0.002)	(0.023)	(0.002)
$\overline{Q}$	0.001	-0.006***	-0.006***	0.006***	0.002	0.006***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.002)	(0.000)
$\overline{Debt}$	0.024**	0.021***	0.024***	-0.030***	0.026**	-0.030***
	(0.010)	(0.005)	(0.005)	(0.002)	(0.011)	(0.002)
$\overline{Size}$	0.003***	-0.001	-0.001*	0.003***	0.002***	0.003***
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
$\overline{ROA}$	0.106***	0.094***	0.095***	-0.100***	0.095***	-0.100***
	(0.016)	(0.010)	(0.014)	(0.003)	(0.021)	(0.003)
Cash	0.072***	0.063***	0.063***	0.007***	0.074***	0.007***
	(0.003)	(0.002)	(0.002)	(0.000)	(0.002)	(0.000)
Q	0.010***	0.009***	0.009***	0.000	0.010***	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Debt	-0.066***	-0.039***	-0.043***	-0.000	-0.070***	-0.000
	(0.002)	(0.002)	(0.002)	(0.000)	(0.002)	(0.000)
Size	0.003***	0.000***	0.002***	0.000***	0.004***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ROA	-0.184***	-0.155***	-0.145***	0.001	-0.172***	0.001
	(0.005)	(0.004)	(0.004)	(0.001)	(0.003)	(0.001)
$\overline{EShock}$				-0.037***		-0.037***
				(0.007)		(0.007)
$\overline{ERisk}$				0.222***		0.222***
				(0.004)		(0.004)
EShock			-0.007	-0.000	-0.006	-0.000
			(0.007)	(0.001)	(0.007)	(0.001)
ERisk			0.068***	0.007***	0.090***	0.007***
			(0.005)	(0.001)	(0.006)	(0.001)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	47,445	47,445	47,445	47,445	47,445	47,445
Adj.R^2		0.450		0.453		0.890
First-Stage F-Statistic	656.00***	641.10***	406.50***		406.50***	
Wald Test of Exogeneity			122.60***		84.26***	
Hansen J-Statistic			0.02		1.59	

Table 4 Product market competition and peer effects on corporate innovation

The table presents estimation results of Equation (2), which relate R&D to firm-specific and peer firms' average characteristics. The dependent variable, RD/TA, is the firm's R&D to total assets. *Low* indicates that a firm faces high competition with above median HHI-Assets, HHI-Sales, CR4-Assets, and CR4-Sales whereas *High* is the dummy variable when a firm faces low product market competition based on below median of the measures of product market competition. Thus, the independent variables are the interactions between the peer average R&D and each of the dummy variables *Low* ($\overline{RD/TA} \times Low$) and *High* ($\overline{RD/TA} \times High$). $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The other peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent to total assets, Q is lagged market-to-book ratio, *Debt* is lagged debt to total assets, *Size* is lagged size (logarithm of total assets), *ROA* is lagged profitability (earnings before interest and tax to total assets). The instruments for the IV-Tobit regression models are the interactions between each of the instruments and the dummy variables *Low* ($\overline{EShock} \times Low$ and $\overline{ERisk} \times Low$) and *High* ($\overline{EShock} \times High$ and $\overline{ERisk} \times High$). $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{ERisk}$ is the standard deviation of the peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

Product competition	HHI-Assets	HHI-Sales	CR4-Assets	CR4-Sales
Variables	(1)	(2)	(3)	(4)
$\overline{RD/TA} \times Low$	1.147***	1.182***	1.202***	1.204***
	(0.107)	(0.106)	(0.106)	(0.106)
$\overline{RD/TA} \times High$	0.543***	0.415***	0.647***	0.621***
	(0.112)	(0.116)	(0.110)	(0.111)
Diff (*High vs Low*)	46.12***	71.04***	41.69***	45.62***
First stage regressions				
$\overline{EShock} \times Low$	-0.120***	-0.113***	-0.277***	-0.285***
	(0.011)	(0.011)	(0.014)	(0.014)
$\overline{ERisk} \times Low$	0.275***	0.279***	0.292***	0.293***
	(0.006)	(0.006)	(0.006)	(0.006)
$\overline{EShock} \times High$	-0.439***	-0.456***	-0.288***	-0.284***
	(0.017)	(0.017)	(0.014)	(0.014)
$\overline{ERisk} \times High$	0.374***	0.367***	0.273***	0.273***
	(0.005)	(0.005)	(0.006)	(0.006)
Firm factors	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	45,656	45,653	47,427	47,419
R^2	0.799	0.798	0.732	0.732
First-Stage F-Statistic	168.00***	191.20***	66.85***	62.20***
Wald Test of Exogeneity	124.70***	150.70***	134.10***	137.10***
Hansen J-Statistic	3.09	2.84	1.66	1.78

Table 5 Do firms mimic leaders or followers?

The table presents estimation results of Equation (2), which relate change in R&D to firm-specific and peer firms' average characteristics. The dependent variable, *RD/TA*, is the firm's R&D to total assets. *Low* indicates that a firm generates below median profitability, sales, size, and analyst followings. Similarly, *High* is indicator variable for firms with above median profitability, sales, size, and analyst followings. *Low* and *High* are defined as *Followers* and *Leaders*, respectively. Thus, the independent variables are the interactions between the peer average R&D and each of the dummy variables *Low* ($\overline{RD/TA} \times Low$) and *High* ($\overline{RD/TA} \times High$). $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The other peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent to total assets, *Q* is lagged market-to-book ratio, *Debt* is lagged debt to total assets, *Size* is lagged size (logarithm of total assets), *ROA* is lagged profitability (earnings before interest and tax to total assets). The instruments for the IV-Tobit regression models are the interactions between each of the instruments and the dummy variables *Low* ($\overline{EShock} \times Low$ and $\overline{ERisk} \times Low$) and *High* ($\overline{EShock} \times High$ and $\overline{ERisk} \times High$). $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{ERisk}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	Profitability	Size	LogSales	Analyst Followings
Variables	(1)	(2)	(3)	(4)
$\overline{RD/TA} \times Low$	1.023***	1.292***	1.215***	1.296***
	(0.105)	(0.106)	(0.105)	(0.183)
$\overline{RD/TA} \times High$	0.772***	0.506***	0.587***	0.485**
	(0.104)	(0.107)	(0.107)	(0.197)
Diff (*High vs Low*)	8.87***	84.37***	55.05***	27.14***
First stage regressions				
$\overline{EShock} \times Low$	-0.255***	-0.283***	-0.284***	-0.067*
	(0.014)	(0.015)	(0.015)	(0.035)
$\overline{ERisk} \times Low$	0.287***	0.287***	0.289***	0.404***
	(0.006)	(0.006)	(0.006)	(0.014)
$\overline{EShock} \times High$	-0.297***	-0.293***	-0.288***	-0.024
	(0.015)	(0.015)	(0.015)	(0.037)
$\overline{ERisk} \times High$	0.261***	0.281***	0.277***	0.385***
	(0.006)	(0.006)	(0.006)	(0.014)
Firm factors	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	44,288	44,288	44,272	11,713
R^2	0.739	0.736	0.737	0.815
First-stage F-statistic	91.22***	56.47***	56.04***	9.39***
Wald Test of Exogeneity	96.51***	173.20***	144.90***	58.93***
Hansen J-Statistic	1.32	2.61	2.15	0.55

Table 6 Peer effects and corporate outcomes

The table presents estimation results of Equation (6), which relate long-term innovation outputs and product market performance to peers' R&D to corporate outcomes, given firm-specific characteristics. The dependent variables are the measures of long-term innovation outputs and product market performance and are computed over a 5-year window (Period $[t+1, t+5]$): Tobin's q is market value of equity *plus* total debt to total assets (Q), log patent counts ($LOGPATS$), logarithm of citation-weighted value of patents ($LOGTCW$), market value of patents to total assets (TSM/TA), product similarity (SIM), and product fluidity ($FLUIDITY$). The independent variable is $\overline{RD/TA}$, which is the ratio of average peer firms' R&D to total assets. The firm-specific characteristics are defined as follows: *Sales Growth* change in sales, *Cash* is lagged cash and cash equivalent, *Debt* is lagged total debt, *PPE* is lagged logarithm of property, plant and equipment, *Size* is lagged logarithm of total assets, and *LogAge* is the logarithm of firm age. The instruments for the 2SLS regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

Long-term $[t+1, t+5]$	Tobin's q	LOGPATS	LOGTCW	TSM/TA	SIM	FLUIDITY
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{RD/TA}$	2.318***	3.344*	5.818***	0.994***	15.488***	29.808***
	(0.758)	(1.950)	(2.231)	(0.233)	(3.349)	(2.797)
First stage regressions						
$\overline{EShock}$	-0.143***	-0.070***	-0.070***	-0.070***	-0.160***	-0.160***
	(0.012)	(0.020)	(0.020)	(0.020)	(0.023)	(0.023)
$\overline{ERisk}$	0.376***	0.380***	0.380***	0.380***	0.338***	0.338***
	(0.006)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	28,571	6,045	6,045	6,045	10,021	10,021
R^2	0.81	0.88	0.88	0.88	0.81	0.81
First-Stage F-Statistic	251.50***	176.60***	176.60***	176.60***	40.85***	40.85***
Wald Test of Exogeneity	9.35***	2.94*	6.80***	18.14***	21.39***	113.60***
Hansen J-Statistic	0.13	0.10	0.22	1.62	0.76	0.31

Table 7 Peer effects on corporate innovation across industries

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics across different industries. The dependent variable, RD/TA, is the firm's R&D to total assets. The independent variable is the peer firms' average firm characteristics defined as follows: $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent, Q is lagged market to book ratio, *Debt* is lagged total debt, *Size* is lagged logarithm of total assets, *ROA* is lagged return on assets, and *EShock* is lagged equity shock. The instruments for the IV-Tobit regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. Firms are classified into five industrial categorisations as follows: Others (industries except mining and manufacturing), mining (1000 $\leq$ SIC code $\leq$ 1499), manufacturing (2000 $\leq$ SIC code $\leq$ 3999), durables (2400 $\leq$ SIC code $\leq$ 2500 and 3200 $\leq$ SIC code $\leq$ 3800) and non-durables (2000 $\leq$ SIC code $\leq$ 2300 and 2600 $\leq$ SIC code $\leq$ 3100). The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	Others	Mining	Manufacturing	Non-Durables	Durables
Variables	(1)	(2)	(3)	(4)	(5)
$\overline{RD/TA}$	0.955**	7.939***	0.728***	0.190	0.587***
	(0.407)	(2.270)	(0.148)	(2.086)	(0.151)
First stage regressions					
$\overline{EShock}$	0.103***	-0.024*	-0.058***	-0.028***	-0.048***
	(0.019)	(0.013)	(0.007)	(0.008)	(0.009)
$\overline{ERisk}$	0.220***	-0.068***	0.158***	-0.018***	0.196***
	(0.010)	(0.008)	(0.004)	(0.005)	(0.005)
Firm factors	Yes	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes	Yes
N	11,331	2,689	33,425	8,729	24,140
R^2	0.87	0.61	0.93	0.98	0.92
First-Stage F-Statistic	198.00***	27.08***	132.50***	20.36***	174.90***
Wald Test of Exogeneity	1.42	8.46***	1.87	0.09	0.22
Hansen J-Statistic	1.81	1.36	0.05	0.28	0.12

Table 8 Alternative proxies of corporate innovation

The table presents estimation results of Equation (1), which relate alternative proxies of R&D to firm-specific and peer firms' average characteristics. The alternative dependent variables are as follows: *RDD* is a dummy variable which is equal to 1 if a firm has R&D and 0 otherwise (Column (1)), *RD/NA* is research and development to net assets (Columns (2)), $\Delta RD/TA$ is change in research and development to total assets (Columns (3)), *RDSGA/TA* is research and development plus selling, general and administrative expenses to total assets (Columns (4)), and *RDSGA/Sales* is research and development plus selling, general and administrative expenses to total sales (Columns (5)). The independent variables are the peer-averages of the measures of R&D: $\overline{RDD}$, $\overline{RD/NA}$, $\overline{\Delta RD/TA}$, $\overline{RDSGA/TA}$, and $\overline{RDSGA/Sales}$. The peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific variables are defined as follows: *Cash* is lagged cash and cash equivalent, *Q* is lagged market to book ratio, *Debt* is lagged total debt, *Size* is lagged logarithm of total assets, *ROA* is lagged return on assets, *EShock* is lagged equity shock, and *ERisk* is lagged equity risk. The instruments for the IV-Probit (Column (1)) and IV-Tobit (Columns (2)–(5)) regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	RDD	RD/NA	ΔRD/TA	RDSGA/TA	RDSGA/Sales
Variables	(1)	(2)	(3)	(4)	(5)
$\overline{RDD}$	3.179***				
	(0.543)				
$\overline{RD/NA}$		1.025***			
		(0.137)			
$\overline{\Delta RD/TA}$			2.614***		
			(0.381)		
$\overline{RDSGA/TA}$				0.581***	
				(0.089)	
$\overline{RDSGA/Sales}$					0.369***
					(0.132)
First stage regressions					
$\overline{EShock}$	0.356***	-0.102***	0.016***	0.129***	-0.116***
	(0.066)	(0.013)	(0.002)	(0.033)	(0.029)
$\overline{ERisk}$	0.931***	0.285***	0.024***	0.690***	0.380***
	(0.037)	(0.007)	(0.001)	(0.019)	(0.016)
Method	IV-Probit	IV-Tobit	IV-Tobit	IV-Tobit	IV-Tobit
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	47,445	47,445	46,206	47,445	47,389
R^2	0.87	0.90	0.55	0.87	0.91
First-Stage F-Statistic	256.90***	135.30***	275.70***	409.60***	55.10***
Wald Test of Exogeneity	34.23***	55.73***	47.15***	42.61***	7.80***
Hansen J-Statistic	0.10	1.67	2.27	1.17	1.39

Table 9 Alternative industrial definitions

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics using alternative industrial classifications. The dependent variable, RD/TA, is the firm's R&D to total assets. The independent variable is the peer firms' average firm characteristics defined as follows: $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent, Q is lagged market to book ratio, *Debt* is lagged total debt, *Size* is lagged logarithm of total assets, *ROA* is lagged return on assets, and *EShock* is lagged equity shock. The instruments for the IV-Tobit regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the one-digit SIC code (SIC1), two-digit SIC code (SIC2), the two-digit North America Industry Classification System (NAICS2) and the 10-K-based Fixed Industry Classifications (FIC200, and FIC300) (see Hoberg and Phillips, 2010, 2016). The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	SIC1	SIC2	NAICS2	FIC200	FIC300
Variables	(1)	(2)	(3)	(4)	(5)
$\overline{RD/TA}$	0.740***	0.936***	0.486***	0.202**	0.233***
	(0.145)	(0.114)	(0.167)	(0.081)	(0.073)
First stage regressions					
$\overline{EShock}$	0.136***	-0.034***	0.042***	-0.231***	-0.208***
	(0.007)	(0.007)	(0.008)	(0.013)	(0.013)
$\overline{ERisk}$	0.337***	0.270***	0.254***	0.282***	0.323***
	(0.004)	(0.004)	(0.005)	(0.007)	(0.007)
Controls	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	44,323	44,227	43,479	34,180	33,090
R^2	0.94	0.94	0.91	0.83	0.83
First-Stage F-Statistic	2,668.00***	821.00***	818.90***	9.63***	50.33***
Wald Test of Exogeneity	26.17***	67.96***	8.48***	6.20**	10.12***
Hansen J-Statistic	1.56	0.37	2.19	0.26	0.03

Appendix A Variable definitions

The table lists the definitions of all variables used and the account items obtained from Compustat databases.

Variable	Definition
RDD	Research and development (*XRD*) dummy.
RD/TA	Research and development (*XRD*) to total assets (*AT*).
RD/NA	Research and development (*XRD*) to net asset. Net assets are total assets (*AT*) *less* cash and cash equivalent (*CHE*)
Δ*RD/TA*	Change in research and development to total assets (*AT*).
RDSGA/TA	Research and development (*XRD*) plus selling, general and administrative expenses (*XSGA*) to total assets (*AT*).
RDSGA/Sales	Research and development (*XRD*) plus selling, general and administrative expenses (*XSGA*) to total sales (*SALE*).
Cash	Cash and cash equivalent (*CHE*) to total assets (*TA*).
Tobin's q (Q)	Market value of equity (*PRCC_F* $\times$ *CSHO*) *plus* total debt (*DLC* + *DLTT*) to total assets (*AT*).
Debt	Total debt (*DLC* + *DLTT*) to total assets (*AT*).
Size	Logarithm of total assets (*AT*).
ROA	Earnings before interest and tax (*EBIT*) plus depreciation (*DP*) to total assets (*AT*).
EShock	The annual idiosyncratic returns is the geometric mean of the monthly idiosyncratic returns from Equation (5).
ERisk	The annual idiosyncratic risk is the standard deviation of the idiosyncratic returns calculated from the monthly idiosyncratic returns from Equation (5).
Sales Growth	The change in sales (*SALE less* lagged *SALE* to lagged *SALE*).
LTDA	Long-term debt (*DLTT*) to total assets (*AT*).
STDA	Short-term debt (*DLC*) to total assets (*AT*).
PPE	Property, plant and equipment to total assets (*AT*).
LOGPATS	A measure of innovation output defined as the logarithm of the number of patents granted to a firm. It is computed over a 5-year window following investments in R&D. This data is extracted from the website of Kogan et al. (2017) at https://iu.app.box.com/v/patents.
LOGTCW	It is computed as the logarithm of the citation-weighted value of patents using the grant year and averaged over a 5-year window. This data is provided by Kogan et al. (2017) and available at https://iu.app.box.com/v/patents.
TSM/TA	This is the dollar stock market value of patents over a 5-year window granted to a firm scaled by total assets (*AT*). It is provided by Kogan et al. (2017) and available at https://iu.app.box.com/v/patents.
SIM	Firm-level product similarity measure by Hoberg and Phillips (2016) based on the product descriptions from firms 10-K annual reports. This measure s computed over a 5-year window following R&D investments. This data is available at http://hobergphillips.tuck.dartmouth.edu
FLUIDITY	A firm-level measure of product fluidity, indicating how intensely product market changes around a firm This measure s computed over a 5-year window following R&D investments. (Hoberg et al., 2014). The data is available at http://hobergphillips.tuck.dartmouth.edu.
GDP growth	Annual GDP growth rate.
Inflation	Annual inflation rate.
UMCSENT	University of Michigan Consumer Sentiment Index available at https://fred.stlouisfed.org/series/UMCSENT.
SENT	Baker and Wurgler's (2006) Investor Sentiment Index available at http://people.stern.nyu.edu/jwurgler/.

Appendix B Peer equity residuals

The table presents estimation results of the following equation:-

$$\overline{Instrument_{ijt}} = \vartheta + \theta' X_{ijt-1} + \varphi' \overline{X}_{-ijt-1} + \pi' \zeta_t + e_{ijt} \quad (7)$$

where $\overline{Instrument_{ijt}}$ is the peer equity shock ($\overline{EShock_{ijt}}$) or risk ($\overline{ERisk_{ijt}}$) for firm i in industry j at time t, ϑ is a constant, θ' and φ' are the vectors of coefficients to be estimated, $\overline{X}_{-ijt-1}$ and X are vectors of peer firm averages excluding firm i and firm-specific characteristics, respectively. ζ_t and e_{ijt} are year-fixed effects and firm-year specific error term, respectively. The vector of firm-specific factors, X_{-ijt-1}, include lagged cash and cash equivalent (*Cash*), lagged market to book ratio (Q_{ijt}), lagged total debt ($Debt_{ijt}$), lagged logarithm of total assets ($Size_{ijt}$), lagged return on assets (ROA_{ijt}) and lagged equity shock ($EShock_{ijt}$) or lagged equity risk ($ERisk_{ijt}$). The peer firms' average characteristics, $\overline{X}_{-ijt-1}$, are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period from 1968 to 2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses.

	Panel A: Contemporaneous independent variables		Panel B: Lagged independent variables	
	$\overline{EShock}$	$\overline{ERisk}$	$\overline{EShock}$	$\overline{ERisk}$
Variables	(1)	(2)	(3)	(4)
Cash	-0.001	0.001	-0.000	0.002*
	(0.000)	(0.001)	(0.000)	(0.001)
Q	0.000***	-0.000	0.000***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Debt	0.000	-0.005***	0.001*	-0.004***
	(0.000)	(0.001)	(0.000)	(0.001)
PPE	-0.000	0.001***	-0.000	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
ROA	-0.001***	0.004***	-0.000	0.005***
	(0.000)	(0.001)	(0.000)	(0.001)
Firm Equity Shock	Yes	No	Yes	No
Firm Equity Risk	No	Yes	No	Yes
Peer averages	Yes	Yes	Yes	Yes
N	47,445	47,445	47,445	47,445
Firms	4,545	4,545	4,545	4,545
R^2	0.35	0.79	0.34	0.77

Appendix C Alternative estimations of peer effects on corporate innovation

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics. Columns (1)–(3) presents the results of the IV-GMM (generalised method of moments) whereas the results in Columns (4)–(6) are the IV-LIML (limited-information maximum likelihood) estimations. The dependent variable, RD/TA, is the firm's R&D to total assets. The independent variable is the peer firms' average R&D, $\overline{RD/TA}$, which is defined as the average peer firms' R&D to total assets. The other peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: $Cash$ is lagged cash and cash equivalent to total assets, Q is lagged market-to-book ratio, $Debt$ is lagged debt to total assets, $Size$ is lagged size (logarithm of total assets), ROA is lagged profitability (earnings before interest and tax to total assets), and $EShock_{ijt}$ is the lagged idiosyncratic stock returns. The instruments for the IV-GMM and IV-LIML regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	IV-GMM			IV-LIML		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{RD/TA}$	0.664***	0.774***	0.713***	0.657***	0.774***	0.712***
	(0.052)	(0.045)	(0.064)	(0.053)	(0.045)	(0.064)
First stage regressions						
$\overline{EShock}$	-0.037***	-0.036***	-0.037***	-0.037***	-0.036***	-0.037***
	(0.007)	(0.011)	(0.007)	(0.007)	(0.011)	(0.007)
$\overline{ERisk}$	0.222***	0.294***	0.222***	0.222***	0.294***	0.222***
	(0.004)	(0.006)	(0.004)	(0.004)	(0.006)	(0.004)
Firm factors	Yes	Yes	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	No	No	Yes	No	No	Yes
Year FE	No	Yes	Yes	No	Yes	Yes
N	47,445	47,445	47,445	47,445	47,445	47,445
R^2	0.89	0.76	0.89	0.89	0.76	0.89
First-Stage F-Statistic	406.50***	369.60***	406.50***	406.50***	369.60***	406.50***
Wald Test of Exogeneity	161.40***	300.30***	122.80***	155.60***	299.30***	122.60***
Hansen J-Statistic	1.42	0.03	0.02	1.42	0.03	0.02

Appendix D Peer effects on corporate innovation excluding firms that change industries or with multiple segments

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics. Columns (1)–(3) presents the results when excluding firms that change industries, and Columns (4)–(6) provide the estimation results after excluding firms with multiple segments. The dependent variable, *RD/TA*, is the firm's R&D to total assets. The independent variable is the peer firms' average firm characteristics defined as follows: $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent, *Q* is lagged market to book ratio, *Debt* is lagged total debt, *Size* is lagged logarithm of total assets, *ROA* is lagged return on assets, and *EShock* is lagged equity shock. The instruments for the IV-Tobit regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

	Excluding firms that change industries			Excluding firms with multiple segments		
Variables	(1)	(2)	(3)	(4)	(5)	(6)
$\overline{RD/TA}$	1.116***	0.816***	1.014***	1.351***	1.061***	1.328***
	(0.059)	(0.059)	(0.111)	(0.078)	(0.077)	(0.159)
First stage regressions						
$\overline{EShock}$	-0.128***	-0.128***	-0.051***	-0.101***	-0.125***	-0.074***
	(0.010)	(0.010)	(0.008)	(0.015)	(0.016)	(0.012)
$\overline{ERisk}$	0.385***	0.368***	0.233***	0.450***	0.429***	0.244***
	(0.005)	(0.005)	(0.005)	(0.007)	(0.008)	(0.007)
Firm factors	No	Yes	Yes	No	Yes	Yes
Peer averages	No	No	Yes	No	No	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	37,959	34,453	34,453	22,632	20,680	20,680
R^2	0.83	0.84	0.90	0.83	0.84	0.91
First-Stage F-Statistic	457.90***	380.40***	336.30***	362.70***	260.10***	131.70***
Wald Test of Exogeneity	360.10***	188.60***	83.74***	299.20***	187.70***	69.88***
Hansen J-Statistic	0.05	0.36	1.86	0.01	0.17	0.09

Appendix E Macroeconomic conditions, market sentiment and peer effects on corporate innovation

The table presents estimation results of Equation (1), which relate R&D to firm-specific and peer firms' average characteristics. The dependent variable, RD/TA, is the firm's R&D to total assets. The independent variable is the peer firms' average firm characteristics defined as follows: $\overline{RD/TA}$, is the average peer firms' R&D to total assets. The peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent, Q is lagged market to book ratio, *Debt* is lagged total debt, *Size* is lagged logarithm of total assets, ROA is lagged return on assets, and *EShock* is lagged equity shock. $GDPgrowth$, $Inflation$ $UMCSENT$, and $SENT$ are the measures of macroeconomic conditions and stock market sentiment. The instruments for the IV-Tobit regression models are $\overline{EShock}$ and $\overline{ERisk}$. $\overline{EShock}$ is the lagged average peer idiosyncratic stock returns and $\overline{EShock}$ is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All Variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

Variables	(1)	(2)	(3)	(4)	(5)
$\overline{RD/TA}$	0.914***	0.914***	0.914***	0.914***	0.914***
	(0.100)	(0.100)	(0.100)	(0.100)	(0.100)
GDP growth	-0.046				-0.053
	(0.038)				(0.040)
Inflation		-0.030			-0.062
		(0.084)			(0.087)
UMCSENT			-0.002		0.000
			(0.003)		(0.001)
SENT				-0.044	-0.041
				(0.059)	(0.080)
First stage regressions					
$\overline{EShock}$	-0.037***	-0.037***	-0.037***	-0.037***	-0.037***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$\overline{ERisk}$	0.222***	0.222***	0.222***	0.222***	0.222***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Firm factors	Yes	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
N	47,445	47,445	47,445	47,445	47,445
R^2	0.89	0.89	0.89	0.89	0.89
First-Stage F-Statistic	406.50***	405.90***	406.50***	406.50***	405.90***
Wald Test of Exogeneity	84.25***	84.25***	84.26***	84.26***	84.24***
Hansen J-Statistic	1.58	1.60	1.59	1.59	1.59

Online Appendices

Not For Publication

Appendix A.1 Correlations

Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively.

#	Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	RD/TA	1												
(2)	RDD	0.515***	1											
(3)	$\Delta RD/TA$	0.228***	0.041***	1										
(4)	$RDSGA/TA$	0.574***	0.252***	0.138***	1									
(5)	$RDSGA/Sales$	0.728***	0.395***	0.133***	0.612***	1								
(6)	RD/NA	0.875***	0.387***	0.184***	0.463***	0.740***	1							
(7)	$Cash$	0.591***	0.509***	0.035***	0.343***	0.539***	0.498***	1						
(8)	Q	0.464***	0.635***	0.025***	0.243***	0.395***	0.357***	0.774***	1					
(9)	$Debt$	0.130***	0.101***	0.074***	0.079***	0.087***	0.096***	0.244***	0.155***	1				
(10)	$Size$	0.351***	0.272***	0.023***	0.568***	0.409***	0.279***	0.590***	0.413***	0.154***	1			
(11)	ROA	0.520***	0.420***	0.022***	0.387***	0.610***	0.472***	0.873***	0.634***	0.150***	0.669***	1		
(12)	$EShock$	0.562***	0.443***	0.029***	0.309***	0.558***	0.514***	0.953***	0.671***	0.200***	0.530***	0.906***	1	
(13)	$ERisk$	0.429***	0.217***	0.060***	0.216***	0.524***	0.566***	0.424***	0.258***	0.055***	0.265***	0.457***	0.464***	1
(14)	$\overline{RD/TA}$	0.282***	0.164***	-0.002	0.165***	0.300***	0.282***	0.279***	0.170***	0.052***	0.175***	0.341***	0.309***	0.350***
(15)	$\overline{RDD}$	-0.294***	-0.210***	-0.017***	-0.253***	-0.293***	-0.284***	-0.304***	-0.230***	-0.040***	-0.248***	-0.312***	-0.303***	-0.477***
(16)	$\overline{\Delta RD/TA}$	-0.145***	0.019***	-0.009*	-0.335***	-0.186***	-0.130***	-0.144***	-0.130***	-0.037***	-0.210***	-0.120***	-0.084***	-0.134***
(17)	$\overline{\Delta RDSGA/TA}$	-0.272***	-0.069***	0.068***	-0.218***	-0.339***	-0.303***	-0.167***	-0.111***	-0.005	-0.108***	-0.178***	-0.180***	-0.102***
(18)	$\overline{\Delta RDSGA/Sales}$	-0.061***	-0.019***	-0.078***	-0.045***	-0.082***	-0.054***	-0.065***	-0.033***	0.008*	-0.038***	-0.055***	-0.061***	-0.069***
(19)	$\overline{RD/NA}$	0.221***	0.034***	0.013***	0.254***	0.222***	0.186***	0.216***	0.118***	0.038***	0.156***	0.176***	0.188***	0.082***
(20)	$\overline{Cash}$	0.472***	0.314***	0.017***	0.291***	0.531***	0.459***	0.782***	0.473***	0.112***	0.486***	0.851***	0.868***	0.507***
(21)	$\overline{Q}$	0.348***	0.235***	0.016***	0.217***	0.443***	0.338***	0.581***	0.356***	0.104***	0.363***	0.708***	0.644***	0.348***
(22)	$\overline{Debt}$	-0.412***	-0.341***	-0.014***	-0.329***	-0.438***	-0.364***	-0.679***	-0.516***	-0.089***	-0.557***	-0.699***	-0.678***	-0.399***
(23)	$\overline{Size}$	-0.200***	-0.194***	-0.013***	-0.278***	-0.172***	-0.109***	-0.327***	-0.275***	-0.079***	-0.486***	-0.284***	-0.201***	-0.055***
(24)	$\overline{ROA}$	-0.320***	-0.226***	-0.002	-0.199***	-0.349***	-0.304***	-0.540***	-0.340***	-0.005	-0.344***	-0.574***	-0.586***	-0.316***
(25)	$\overline{EShock}$	-0.186***	-0.111***	0.007	-0.107***	-0.166***	-0.156***	-0.308***	-0.164***	-0.005	-0.183***	-0.278***	-0.293***	-0.138***
(26)	$\overline{ERisk}$	0.277***	0.169***	0.010**	0.200***	0.229***	0.212***	0.457***	0.247***	0.073***	0.343***	0.380***	0.400***	0.150***

#	Variables	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)
(14)	$\overline{RD/TA}$	1												
(15)	$\overline{RDD}$	-0.263***	1											
(16)	$\overline{\Delta RD/TA}$	0.058***	0.175***	1										
(17)	$\overline{\Delta RDSGA/TA}$	0.174***	-0.111***	0.214***	1									
(18)	$\overline{\Delta RDSGA/Sales}$	-0.186***	0.056***	0.049***	-0.049***	1								
(19)	$\overline{RD/NA}$	-0.003	0.020***	-0.459***	-0.325***	-0.015***	1							
(20)	$\overline{Cash}$	0.314***	-0.329***	-0.037***	-0.186***	-0.054***	0.133***	1						
(21)	$\overline{Q}$	0.440***	-0.238***	-0.002	-0.088***	-0.042***	0.126***	0.662***	1					
(22)	$\overline{Debt}$	-0.259***	0.380***	0.116***	0.132***	0.055***	-0.134***	-0.771***	-0.553***	1				
(23)	$\overline{Size}$	-0.009*	0.123***	0.392***	0.043***	0.038***	-0.189***	-0.109***	-0.006	0.283***	1			
(24)	$\overline{ROA}$	-0.141***	0.188***	0.057***	0.255***	0.045***	-0.138***	-0.573***	-0.245***	0.363***	0.164***	1		
(25)	$\overline{EShock}$	-0.102***	0.115***	0.080***	0.061***	0.096***	-0.233***	-0.256***	-0.245***	0.259***	0.191***	0.178***	1	
(26)	$\overline{ERisk}$	0.131***	-0.126***	-0.179***	-0.084***	-0.106***	0.425***	0.271***	0.256***	-0.259***	-0.450***	-0.321***	-0.479***	1

Appendix A.2 Who mimics who? Further analysis

The table presents estimation results of Equation (1), which relate change in R&D to firm-specific and peer firms' average characteristics. Columns (1)–(4) in Panel A estimate the regression when *Followers* mimic other *Followers* whereas Columns (5)–(8) show the results when *Leaders* mimic other *Leaders*. In Panel B, Columns (1)–(4) give results for *Followers* mimicking *Leaders*, and Columns (5)–(8) produce the output for *Leaders* mimicking *Followers*. *Followers* (*Leaders*) are firms with below (above) median profitability, firm size, sales, and analyst followings. The dependent variable, RD/TA, is the firm's R&D to total assets. The independent variable is the $\overline{RD/TA}$, which is the average peer firms' R&D to total assets. The other peer control variables are as follows: $\overline{Cash}$ is lagged peer cash and cash equivalent to total assets, $\overline{Q}$ is lagged peer market-to-book ratio, $\overline{Debt}$ is lagged peer debt to total assets, $\overline{Size}$ is lagged peer size (logarithm of total assets), and $\overline{ROA}$ is lagged peer profitability (earnings before interest and tax to total assets). The firm-specific characteristics are defined as follows: *Cash* is lagged cash and cash equivalent to total assets, Q is lagged market-to-book ratio, *Debt* is lagged debt to total assets, *Size* is lagged size (logarithm of total assets), *ROA* is lagged profitability (earnings before interest and tax to total assets). The instruments for the IV-Tobit regression models are $\overline{EShock}$ defined as the lagged average peer idiosyncratic stock returns and $\overline{ERisk}$, which is the standard deviation of the average peer idiosyncratic stock returns. The peer firms' average characteristics are calculated as the average of all firms within an industry-year excluding the i^{th} observations. Industries are defined at the three-digit SIC code. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects (but are not reported). ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses. The Wald test of the exogeneity of the instrumented variables, First-Stage F-statistic for the instruments, and Hansen J-Statistic of over-identification restrictions are presented.

Panel A: Within peer group mimicking

	Followers mimicking followers				Leaders mimicking leaders			
	Profitability	Size	LogSales	Analyst Followings	Profitability	Size	LogSales	Analyst Followings
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\overline{RD/TA}$	0.664***	1.109***	1.086***	0.956***	0.832***	0.675***	0.821***	0.519**
	(0.153)	(0.209)	(0.210)	(0.361)	(0.150)	(0.093)	(0.089)	(0.222)
First stage regressions								
$\overline{EShock}$	-0.043***	-0.046***	-0.062***	-0.056***	-0.024**	-0.006	0.003	0.078***
	(0.010)	(0.009)	(0.009)	(0.015)	(0.012)	(0.011)	(0.011)	(0.016)
$\overline{ERisk}$	0.178***	0.149***	0.147***	0.105***	0.219***	0.297***	0.346***	0.217***
	(0.005)	(0.005)	(0.005)	(0.010)	(0.008)	(0.008)	(0.008)	(0.013)
Firm factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	13,195	13,004	12,954	3,349	13,226	13,404	13,475	3,514
R^2	0.88	0.88	0.88	0.89	0.86	0.88	0.88	0.89
First-Stage F-Statistic	153.00***	96.83***	67.21***	7.74***	173.20***	394.60***	576.40***	166.00***
Wald Test of Exogeneity	18.78***	28.06***	26.77***	7.01***	30.80***	52.32***	85.31***	5.49**
Hansen J-Statistic	0.73	0.17	0.00	0.21	0.76	0.00	0.02	0.02

Appendix A.2 Who mimics who? Further analysis (continued)

Panel B: Cross peer group mimicking

	Followers mimicking leaders				Leaders mimicking followers			
	Profitability	Size	LogSales	Analyst Followings	Profitability	Size	LogSales	Analyst Followings
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\overline{RD/TA}$	1.139***	0.780***	0.892***	0.551**	1.236***	0.527***	0.578***	0.294
	(0.201)	(0.147)	(0.137)	(0.277)	(0.122)	(0.128)	(0.132)	(0.273)
First stage regressions								
$\overline{EShock}$	-0.081***	-0.024**	-0.003	0.020	-0.057***	-0.055***	-0.048***	0.010
	(0.012)	(0.011)	(0.011)	(0.021)	(0.013)	(0.012)	(0.012)	(0.021)
$\overline{ERisk}$	0.208***	0.372***	0.393***	0.356***	0.214***	0.178***	0.174***	0.202***
	(0.008)	(0.008)	(0.008)	(0.018)	(0.007)	(0.007)	(0.007)	(0.013)
Firm factors	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Peer averages	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	13,120	12,920	12,852	3,223	13,218	13,409	13,479	3,501
R^2	0.83	0.87	0.86	0.86	0.88	0.87	0.87	0.82
First-stage F-statistic	71.00***	546.70***	695.60***	151.70***	110.80***	86.11***	87.88***	79.26***
Wald Test of Exogeneity	32.01***	28.12***	42.43***	3.95**	102.20***	16.83***	19.31***	1.16
Hansen J-Statistic	1.59	0.15	0.08	0.31	2.01	1.85	3.06*	2.76*

Appendix A.3 Peer effects and corporate outcomes: Alternative approaches

The table presents estimation results of the modified version of Equation (6) relating long-term innovation outputs and product market performance to the Mimicking dummy (*Mimicking*) and control variables. The dependent variables are the measures of long-term innovation outputs and product market performance and are computed over a 5-year window (Period $[t+1, t+5]$): Tobin's q is market value of equity *plus* total debt to total assets (Q), log patent counts ($LOGPATS$), logarithm of citation-weighted value of patents ($LOGTCW$), market value of patents to total assets (TSM/TA), product similarity (SIM), and product fluidity ($FLUIDITY$). For the purposes of the analyses in this table, $\overline{RD/TA}$ in the Equation (6) is replaced with the Mimicking dummy (*Mimicking*). The main independent variable, *Mimicking*, is constructed as follows: *Mimicking* in Columns (1)–(6) is constructed from the rankings of firms based on the coefficient from a modified version of Equation (1) that includes a firm dummy (IDD) and the interaction term ($\overline{RD/TA} \times IDD$). In this case, the coefficient of $\overline{RD/TA} \times IDD$ is used to rank firms into quartiles and those in the upper quartile of the distribution in each year are categorised as mimickers ($Mimicking = 1$). For Columns (7)–(12), the Mimicking dummy (*Mimicking*) is constructed following Fairhurst and Nam (2018). In this second case, the regression diagnostic statistic (DFBETA) from running an OLS regression of Equation (1) is used identify mimicking firms. Firms ranked in the upper quartiles of the DFBETAs distribution in each year are categorised as mimicking firms (*Mimicking*). The control variables (*Controls*) in the modified version of Equation (6) are defined as follows: *Sales Growth* change in sales, *Size* is logarithm of total assets, *Cash* is cash and cash equivalent, *LTDA* is long-term debt, *STDA* is short-term debt and *PPE* is lagged logarithm of property, plant and equipment. The sample consists of listed non-utility and non-financial firms in the US drawn from *Compustat* over the period 1968–2018. All variables used are defined in Appendix A, and are winsorised at the lower and upper one percentiles. All regression models are estimated with a constant and include both industry and year-fixed effects but are not reported. ***, **, * indicate significance at the 1%, 5%, and 10% levels, respectively, based on robust standard errors. Standard errors are reported in parentheses.

	Mimicking dummy based on the rankings of the coefficients of $\overline{RD/TA} \times IDD$						Mimicking dummy based on the rankings of DFBETAs					
Long-term $[t+1, t+5]$	Tobin's q	LOGPATS	LOGTCW	TSM/TA	SIM	FLUIDITY	Tobin's q	LOGPATS	LOGTCW	TSM/TA	SIM	FLUIDITY
Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Mimicking	0.082***	0.103***	0.138***	0.005	0.640***	0.387***	0.073***	0.195***	0.248***	0.018***	0.527***	0.158***
	(0.013)	(0.029)	(0.032)	(0.005)	(0.072)	(0.049)	(0.013)	(0.029)	(0.033)	(0.005)	(0.081)	(0.052)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	25,060	4,186	4,186	4,186	8,099	8,099	25,060	4,186	4,186	4,186	8,099	8,099
R^2	0.30	0.70	0.67	0.42	0.22	0.39	0.30	0.70	0.68	0.42	0.22	0.38