Working Paper No. 2018002

The Effect of Technological Imitation on Corporate Innovation: Evidence from US Patent Data

Hyun Joong Im

Janghoon Shon

The Effect of Technological Imitation on Corporate Innovation: Evidence from US Patent Data

Hyun Joong Im[a,*], Janghoon Shon[b]

[a] *Peking University HSBC Business School (PHBS), Peking University, University Town, Nanshan District, Shenzhen, 518055, China*
[b] *HKUST Business School, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

Abstract

Technological imitation may play a crucial role in motivating firms to innovate. However, theoretical predictions and empirical findings on the role of imitation have not yet reached a consensus. One major gap in the previous studies is that the empirical tests are based on samples consisting of only one industry over a short period of time. This study uses a novel measure of industry-level technological imitation proxied by quick citations by competitors to examine the relationship between imitation and innovation. Using US patent data for the period 1977–2005, we find that there are inverted U-shaped relationships between the degree of industry-level technological imitation and industry-level innovation activities and between the degree of industry-level technological imitation and the value of firm-level innovation. Our results suggest that positive externalities from the interactions among firms during the innovation process outweigh the negative effects of free-riding concerns on firms' innovation activities and incentives to innovate up to a high degree of technological imitation, while free-riding concerns outweigh the positive externalities when the level of technological imitation is extremely high. The sector-by-sector analyses show that the relationship between technological imitation and the quantity and market value of innovation are not very different across Pavitt sectors. A comparative analysis on the role of imitation between agglomerated and non-agglomerated industries suggests that the positive effect of a moderate level of imitation and the negative effect of an excessive level of imitation are more pronounced for agglomerated industries. The results suggest that creating innovation clusters, such as Silicon Valley in the United

*Corresponding author. Tel: +86 755 2603 3627; E-mail address: hyun.im@phbs.pku.edu.cn

States and Shenzhen City in China, and allowing different innovators to cooperate, imitate and compete with each other would be very effective in promoting corporate innovation. However, an excessively high level of technological imitation is more detrimental for firms in innovation clusters because it lowers those firms' incentives to innovate more radically.

Keywords: Corporate innovation, Technological imitation, Value of innovation, Innovation Cluster, Agglomeration

---

1. Introduction

Corporate innovation is crucial in that it improves total factor productivity and allows firms to achieve higher potential output with lower manufacturing costs in a more efficient and environmentally friendly way (Schoonhoven et al., 1990); innovation also brings new growth engines into different industries, thus increasing demand in most developed economies (Brozen, 1951; Huang and Rozelle, 1996; Grossman and Helpman, 1991). Although corporate innovation is very important to firms and economies as a whole, it is extremely costly in that it requires massive fixed investments at the early stage and may require substantial support for long-term capital and human resources from companies themselves or from national institutions. Therefore, various determinants of corporate innovation, such as hostile takeovers (Atanassov, 2013), stock liquidity (Fang et al., 2014), corporate taxes (Mukherjee et al., 2017), policy uncertainty (Bhattacharya et al., 2017), and product market competition (Aghion et al., 2005; Greenhalgh and Rogers, 2006; Im et al., 2015) have been studied in the literature. In this paper, we investigate how the degree of industry-level technological imitation influences industry-level innovation activities and firms' motivation to innovate.

The relationship between technological imitation and corporate innovation has been studied by several scholars, but their theoretical predictions and empirical findings have not yet reached consensus. The first view is that technological imitation has a *positive effect* on corporate innovation due to the positive externalities in the process of innovation. Among others, Bessen and Maskin (2009) argue that if innovation is sequential (such that each successive innovation is made based on its predecessors' earlier innovations) and complementary (such that each potential innovator takes a different research line), technological imitation will enhance an inventor's prospective profits. In this case, patent protection (an obstacle to imitation) may not be useful for encouraging corporate innovation. The second view is that technological imitation has a *negative effect* on corporate innovation due to free-riding problems. For example, Zeng (2001) found that an increase in subsidies to technological imitation would increase investment in technological imitation and decrease investment in technological innovation. Given the assumption that innovation is independent, unlike the assumptions made by Bessen and Maskin (2009), technological imitation will decrease the value of a firm's innovation outcomes, thereby reducing its incentives to innovate. The third view predicts an *inverted U-shaped relationship* between technological imitation and corporate innovation. Positive externalities from the interactions among firms during the process of innovation outweigh the negative effects of free-riding concerns on firms' innovation activities and incentives to innovate up to a high degree of technological imitation, while free-riding concerns outweigh the positive externalities when the level of technological imitation is extremely high. In this spirit, Aghion et al. (2001) argued that a small amount of imitation almost always contributes to growth because it promotes more frequent close competition, whereas extremely high imitation unambiguously reduces growth due to free-riding problems.

In this study, we empirically investigate whether the degree of industry-level technological imitation increases or decreases firms' innovation activities and their incentives to innovate by utilizing firm-level patent data for US firms between 1977 and 2005. First, we perform an industry-level analysis as in Aghion et al. (2001) by regressing an industry-average innovation measure (i.e., number of patents and number of citations) on a competitor-quick-citation ratio for each industry-year as a measure of technological imitation. This study finds that the increase in technological imitation leads to an increase in the quantity of innovation up to the 89th percentile of technological imitation, but the effect becomes negative after that point. This result implies that the positive externalities from the interactions among firms during the process of innovation outweigh the negative effects of free-riding concerns on firms' innovation activities up to a high degree of technological imitation, while free-riding concerns outweigh the positive externalities when the level of technological imitation is extremely high. In addition, we repeat the analysis for each Pavitt technological sector to investigate whether the relationship between the degree of technological imitation and the quantity of industry-average innovation is heterogeneous across sectors. In general, all Pavitt sectors have peak points at similar imitation levels (between the 83rd and 90th percentile), although Pavitt 4 has a peak point at an imitation level significantly lower than Pavitt sector 2 at the 5% level. The results imply that regardless of Pavitt sectors, the positive externalities from the interactions among firms during the innovation process dominate the negative effects of free-riding concerns on firms' innovation activities up to a rather high degree of technological imitation, whereas free-riding concerns dominate the positive externalities in the case of extremely high levels of technological imitation.

We then investigate the impact of technological imitation on the value of firm-level innovation using the approach of Im et al. (2015), Faulkender and Wang (2006), and Dittmar and Mahrt-Smith (2007). We find that an increase in technological imitation leads to an increase in the value of innovation up to the 82nd to 85th percentile of technological imitation, but the effect becomes negative after that point. This finding implies that the positive externalities from the interactions among firms during the innovation process outweigh the negative effects of free-riding concerns on firms' incentives to innovate up to a high degree of technological imitation; however, free-riding concerns outweigh the positive externalities when technological imitation is at extremely high levels. To further examine whether the relationship between technological imitation and the value of innovation is heterogeneous across sectors, we repeat the analysis for each Pavitt technological sector, finding that the relationships between technological imitation and the market value of firm-level innovation are not very different across Pavitt technology sectors.

Finally, we further investigate how the relationship between imitation and innovation differs between agglomerated and non-agglomerated industries. This study finds that the impacts of technological imitation on both the quantity and the market value of innovation are stronger for agglomerated industries than for non-agglomerated industries; thus, the positive effect of a moderate level of imitation and the negative effect of an excessive level of imitation are more pronounced for agglomerated industries. The results suggest that creating innovation clusters, such as Silicon Valley in the United States and Shenzhen City in China, and allowing different innovators to cooperate, imitate and compete with each other would be very effective in promoting corporate innovation.[1] However, an excessively high level of technological imitation is more detrimental for firms in innovation clusters because it lowers those firms' incentives to innovate more radically.

The main contribution of this study is twofold. First, we propose a novel technological imitation measure to empirically test the impact of imitation on the quantity and value of innovation. Although we are not the first to investigate the imitation-innovation relationship, previous studies are mainly based on a small sample limited to a specific industry and/or short horizon of time (Baptista and Swann, 1998; Casadesus-Masanell and Zhu, 2013; Czarnitzki and Kraft, 2011; Kim, 1997). Our imitation measure can be applied to a large sample across different industries. Our findings support the inverted U-shaped relationship between imitation and the quantity and value of innovation. It is also notable that our findings suggest that even a developed country such as the United States could benefit

[1] An article in the South China Morning Post on 28 September 2016 introduced the success of Shenzhen City in promoting corporate innovation: "*Beginning in 2013, Shenzhen funnelled more than 4 per cent of its annual GDP into research and development, putting it on par with South Korea and Israel. The city now accounts for almost half of the mainland's international patent filings—about 13,300 last year, even outpacing the UK or France. In the first six months of this year, Shenzhen filed 9,002 patent applications under the international patent system, 50 per cent up year on year, according to the municipal government.*"

from a certain allowance of imitation. Second, the large sample used in this study makes it possible to examine heterogeneity across industries. Our findings suggest that the impact of imitation across industries differs and that the industry's level of agglomeration moderates the relationship between imitation and the quantity and value of innovation.

The remainder of the paper is organized as follows. In Section 2, we first review the literature regarding the relationship between technological imitation and the quantity and value of innovation. We then review the literature regarding the moderating effect of agglomeration on those relationships. Section 3 describes the sample, the measurement of variables, and the descriptive statistics. In Section 4, we present our empirical models and results. Section 5 concludes the paper.

2. Related literature

Innovation is a key competitive advantage for a firm (Porter, 1980). Therefore, the mechanism of protecting it from being imitated by competitors is crucial for a firm's motivation to innovate (Rumelt and Lamb, 1984). Various studies suggest that such protection mechanisms against imitation generate incentives for firms to perform active research and development (R&D). In addressing the mechanism, the complexity of the innovation (Ethiraj et al., 2008; Rivkin, 2000) and legislative aspects, such as intellectual property rights (Arrow, 1972), are the major passive and active components, respectively. One of the main findings in the literature is that the complexity of the technology naturally acts as a protection mechanism, while strong patents directly impede imitation. As a result, patent protection encourages firms to innovate; however, it could spur monopoly pricing by patent owners (Arrow, 1972; Tirole, 1988; Grossman and Helpman, 1991; Aghion and Howitt, 1992). Amara et al. (2008) also suggest that the selection of protection mechanisms against imitation is particularly important for knowledge-intensive businesses, as the imitation from competitors will directly impinge a firm's profitability and diminish its value of knowledge. In this spirit, Zeng (2001) argues that technological imitation would decrease the value of a firm's innovation activities due to free-riding concerns. In an industry where imitation activities are prevalent, the value of a firm's innovation outcomes will be diminished due to the risk of being copied and the possibility of being surpassed by competitors.

On the other hand, a recent stream of studies started investigating the potential implication of imitation on corporate innovation, focusing on the externalities of innovation and technological learning. Czarnitzki and Kraft (2011) provide solid empirical evidence for the existence of knowledge spillovers among competitors, thereby stimulating firms to learn and imitate competitors' leaked knowledge. Roper et al. (2013) introduce the notion of externalities of openness in innovation. Their findings suggest that openness enables enhanced knowledge diffusion and that externalities positively influence innovation outputs by intensifying market competition. Furthermore, another strand of research establishes the importance of the learning effect due to the sequential characteristic of innovation. Bessen and Maskin (2009) argue that if innovation is sequential (such that each successive innovation is made based on its predecessors' earlier innovations) and complementary (such that each potential innovator takes a different research line), technological imitation will enhance a firm's incentives to innovate and thus increase the quantity of innovation. Given that innovation is sequential and complementary in an industry, imitation activities will provide learning opportunities to newcomers. Once innovation is achieved by imitating a predecessor, it will trigger other firms to deliver more creative and valuable ideas by using the precedent outcome as a basis (i.e., other firms will apply for new patents quickly after citing a predecessor's patent and implement further development). This is also in line with the findings by Hopenhayn et al. (2006).

However, it is likely that both the positive and negative sides of imitation (i.e., technological free riding and positive externalities) coexist. Aghion et al. (2001) argue that a small amount of imitation almost always enhances growth, as it promotes more frequent close competition, whereas a high degree of imitation evidently slows growth due to free-riding problems. When the level of imitation is moderately low, positive externalities from the interactions among firms outweigh the negative effect of free-riding concerns on firms' innovation activities. When the imitation level in a certain industry is extremely high, however, the negative effect of free-riding concerns outweighs the positive effect of technological externalities. In this case, a firm cannot enjoy the monopolistic advantage from innovation, resulting in strong discouragement towards further development. An innovation outcome can be instantly discovered or duplicated by competitors; therefore, the value of patent or the invention of new products or services will be diminished, which will hinder firms from facilitating enhanced R&D activities. Thus, a newly granted patent is much less valuable to companies in an industry with a more intensive degree of imitation. In turn, the quantity of innovation decreases with the degree of technological imitation prevalent in the industry.

Meanwhile, imitation is one of the major drivers of developing countries' economic development (Dobson and Safarian, 2008; Kim, 1997; Kim and Nelson, 2000). Developing countries generally tolerate imitation by imposing relatively weak intellectual property right protection in their early stage. Once they reach a certain stage of economic development, however, those countries showed a tendency to impose strong intellectual property right protection to support their own non-tangible assets (Hwang et al., 2016). Similarly, in a certain industry where firms are in their initial stage of development, a certain level of imitation provides a better environment for innovation. In such cases, both quality (measured by the increase in the market value of equity driven by a one-unit increase in innovation) and quantity (measured by the number of patents or citations) of innovation in the industry is anticipated to increase relative to the degree of imitation.

Because the main drivers of the effect of imitation are learning and free-riding, we expect that agglomeration would play an important role. Firms in a cluster would be able to learn and imitate their competitors' innovation much easier than those firms not in a cluster. For instance, information sharing has become a hallmark of Silicon Valley firms. Various high-tech companies, such as Facebook, Google, and Twitter, have designed their work spaces to enhance the interaction and sharing of ideas among workers. This knowledge spillover effect is also well supported by several recent studies. Carlino and Kerr (2015) note that knowledge spillovers are one of the main theoretical linkages between innovation and agglomeration, while Jaffe et al. (1993) find that citation-linked patent pairs tend to be in closer spatial proximity than the control patent pairs, which implies strong evidence that knowledge spillovers are geographically concentrated.

Such findings are further supported by Jaffe et al. (2000) and Thompson (2006). Along with the results that spillovers dissipate rapidly with distance (Conley et al., 2003; Moretti, 2004; Adams and Jaffe, 1996), we expect the impact of imitation to be more significant in agglomerated industries. As a byproduct, our results also contribute to the discussion on the policy implications of agglomeration. Chatterji et al. (2014) suggest that our understanding of which policy supports entrepreneurial clusters and how it works is quite limited. Therefore, the existence of an optimal technological imitation level and heterogeneity among industries implies the complexity of the optimal policy on agglomeration. This statement is also in line with the argument that there is no single ideal model for innovation policy because innovation activities differ significantly between central, peripheral and old industrial areas (Tödtling and Trippl, 2005).

Various theoretical models based on different assumptions and model settings do not demonstrate unanimous predictions regarding the relationship between imitation and innovation. Therefore, this study empirically investigates whether technological imitation and the quantity and value of innovation have upward-sloping, downward-sloping, or inverted U-shaped relationships. Furthermore, Greenhalgh and Rogers (2006) and Im et al. (2015) find that the relationship between the market value of innovation and product market competition differs across industries. The innate characteristics of industries are the main determinants for the costs and benefits of imitation. The benefits from learning would be higher for science-based industries but lower for information-based industries. Furthermore, legislative backgrounds and innovation types also differ across industries. For instance, the pharmaceutical industry and software industry are subject to very different regulations and have very different innovation outputs. Thus, the impact of imitation on innovation could also be very different across industries. We investigate whether the impacts of imitation on the quantity and value of innovation are heterogeneous across sectors. Finally, we revisit the question raised by Baptista and Swann (1998) and investigate whether the impact of imitation is greater for agglomerated industries.

3. Sample selection and variable construction

3.1. Sample selection

Our key dataset is the latest version of the National Bureau of Economic Research (NBER) US Patent Citations Data File, which contains firms' patent-related information, including the patent identifier, citing patent identifier, patent assignee names, number of citations received by each patent, and each patent's application year over the period 1976–2006. Truncation issues in our patent dataset are handled by implementing the method of Hall et al. (2001, 2005). We exclude observations before 1977 and after 2005 to further mitigate concerns arising from truncations. Thus, our patent dataset covers all patents applied during the 1977–2005 period.

We use data from Compustat North America to construct industry-average and firm-level variables based on the information contained in financial statements. We also use data for returns to individual firms' stocks from the Center for Research in Security Prices (CRSP) and data for returns to the 25 portfolios formed based on size and book-to-market (5×5) from Kenneth French's data library (or industry-average stock returns) to calculate excess stock returns. We exclude firms in the utilities and financial service sectors and restrict the sample to firms whose common shares are publicly traded on the three major US stock exchanges (NYSE, NASDAQ, and AMEX).

We match our patent dataset with Compustat/CRSP data using a match table that contains a firm identifier (i.e., GVKEY) as well as patent assignee and patent identifier data. When we calculate firm-level patent and citation numbers, we assume that firms without any information in our patent dataset have no patents. Therefore, our sample is not constrained by the NBER database. Our sample covers new firms that are listed in the stock market and firms that are delisted from the stock market or that go out of business, as long as they are covered by Compustat/CRSP. Our final sample is an unbalanced panel of 9,064 firms among 296 four-digit Standard Industrial Classification (SIC) industries over the 29-year period of 1978–2006.[2]

3.2. Construction of Variables

As measures for firm-level innovation activities, we use *i)* the number of patents that firm i applied for in year t ($COUNT_{i,t}$) and *ii)* the number of citations of the patents that firm i applied for in year t ($CITE_{i,t}$). Similarly, to measure industry-average innovation activities, we use *i)* the industry-average number of patents that firms in industry j applied for in year t ($\overline{COUNT}_{j,t}$) and *ii)* the industry-average number of citations of the patents that firms in industry j applied for in year t ($\overline{CITE}_{j,t}$). As both firm-level and industry-average measures are skewed to the right, the natural logarithm of one plus each of the original measures is used in the industry-level regressions reported in Subsection 4.1 and the firm-level regressions reported in Subsection 4.2.

To measure the intensity of technological imitation in industry j in year t, $IMI^{Raw}_{j,t}$, we use the *industry-average competitor-quick-citation ratio*, where the

[2]Note that we use the lagged value of our imitation measure constructed based on our patent dataset.

competitor-quick-citation ratio is defined as the ratio of *Number of competitors' within-five-year citations for the patents that any firms in industry j applied for in year t* to the *Number of all citations for the patents that any firms in industry j applied for in year t*, where competitors are defined as all peers with the same four-digit SIC industry code.[3] For example, $IMI_{j,t}^{Raw} = 0$ means that no patents applied for in year t by any firms in industry j were cited by any competitors within five years after the patent application, implying that the degree of imitation in industry j is extremely low in year t. By contrast, $IMI_{j,t}^{Raw} = 0.5$ means that the patents applied for in year t by any firms in industry j have been heavily cited by competitors within five years after the patent application, implying that the degree of technological imitation in industry j is quite high in year t. In this sense, we believe that the degree of imitation in a certain industry should be positively correlated with the industry-average competitor-quick-citation ratio. We use the standardized technological imitation measure ($IMI_{j,t}$) in regression models.

However, our measure has some potential limitations. First, an imitation could take place without patent citations. For example, competitors could adopt similar functions or designs without citing patents. This could occur quite often in the

[3]We are the first to use the industry-average competitor-quick-citation ratio as an indicator of industry-level technological imitation. We tried to find alternative indicators of imitation in the prior literature, but we could not find any promising alternative measures.

case of production innovation. However, our focus is more on a common type of technological innovation that can be protected only by applying for patents. When they apply for new patents, firms are required to cite relevant patents. Otherwise, their applications may not be successful. Moreover, imitating without citing and stealing others' technologies are extremely costly for firms.[4] Therefore, we assume that the majority of technological imitation often involves the citation of competitors' patents. Second, there might be some "Type 1" errors in our measure. Consider a situation in which an inventor who was not aware of the existing patent might have invented a technology independently without imitation, but the patent examiner requested that the inventor make the citation. In this case, our measure should identify the independent work as an imitation. However, we expect those cases to be few because researchers in a firm tend to be aware of the patents in the same industry and consistently get updated of the new patents. Third, patterns of patent citations may differ across industries and types of technology. Therefore, it is very important to ensure that our imitation measure does not capture the variation driven by the heterogeneity across industries. To obtain

[4]In the United States, patent infringement is normally treated as a civil offense, and the National Intellectual Property Rights Coordination Center (IPR Center) stands at the forefront of the government's response to global intellectual property theft and enforcement of its international trade laws. In some countries (e.g., Argentina, China, France, Japan, Russia, and South Korea), the cases are handled in criminal laws. Thus, those uncaptured imitation using our proxy might exist, but it could be extremely costly for firms.

the imitation measure to be used in the regression models, we standardize the industry-average competitor-quick-citation ratio using the within-industry mean and within-industry standard deviation. In addition, we conduct sector-by-sector analyses.

Overall, our imitation measure based on patent citations would be an effective proxy for the level of imitation of patents or technology given a firm's strong incentive of being protected from patent infringement under the strong legal enforcement system in the United States. However, the measure should be treated with caution when used for developing countries with low litigation costs or when measuring non-patent innovation, such as innovation in business strategies or product designs.

All variables are winsorized at the 1st and 99th percentiles, and their definitions are reported in Appendices A and B.[5] Table 1 reports the summary statistics for those variables. Panel A is related to the industry-level analysis concerning the effect of imitation on the quantity of innovation (Subsection 4.1), and Panel B is related to the firm-level analysis regarding the effect of imitation on the market value of innovation (Subsection 4.2).

[5] Firm age is winsorized at 37 years following Hadlock and Pierce (2010).

Table 1: Summary statistics

Panel A. Industry-level variables

Variable	Obs	Mean	S.D.	Min	Q1	Median	Q3	Max
$\overline{COUNT}_{j,t}$	6,400	7.578	16.128	0.000	0.333	1.531	6.167	88.394
$\overline{CITE}_{j,t}$	6,400	88.837	187.668	0.000	2.864	16.246	72.231	1009.677
$ln(1+\overline{COUNT}_{j,t})$	6,400	3.110	1.964	0.000	1.609	2.890	4.537	7.460
$ln(1+\overline{CITE}_{j,t})$	6,400	5.061	2.640	0.000	3.495	5.203	6.877	10.114
$\overline{Size}_{j,t-1}$	6,400	4.772	1.433	1.580	3.750	4.556	5.556	9.075
$\overline{ROA}_{j,t-1}$	6,400	-0.053	0.555	-4.230	-0.013	0.085	0.135	0.284
$\overline{R\&D}_{j,t-1}$	6,400	0.041	0.067	0.000	0.004	0.014	0.045	0.364
$\overline{PPE}_{j,t-1}$	6,400	0.301	0.136	0.056	0.203	0.273	0.371	0.770
$\overline{Lev}_{j,t-1}$	6,400	0.284	0.133	0.008	0.187	0.271	0.368	0.781
$\overline{Capex}_{j,t-1}$	6,400	0.066	0.036	0.008	0.043	0.059	0.080	0.262
$\overline{MB}_{j,t-1}$	6,400	2.511	5.174	0.455	0.930	1.304	2.026	42.707
$\overline{Age}_{j,t-1}$	6,400	13.425	5.601	2.000	9.348	12.619	16.567	34.000
$\overline{KZ}_{j,t-1}$	6,400	2.613	8.629	-32.735	0.422	1.596	3.239	61.104
$IMI_{j,t-1}$	6,400	-0.006	0.949	-1.173	-0.588	-0.353	0.247	3.507
$IMI^2_{j,t-1}$	6,400	0.905	1.948	0.000	0.112	0.288	0.656	12.302

Note: This table shows summary statistics for the industry-level variables used in Table 2, Table 4, and Table 8.

Panel B. Firm-level variables

Variable	Obs	Mean	S.D.	Min	Q1	Median	Q3	Max
$r_{i,t}$	67,537	0.164	0.697	-0.856	-0.263	0.040	0.390	3.292
$r_{i,t}-R_{p,t}$	67,537	0.002	0.682	-1.089	-0.409	-0.109	0.230	3.046
$r_{i,t}-R_{j,t}$	67,537	-0.017	0.619	-1.302	-0.369	-0.084	0.213	2.612
$INN1_{i,t-1}$	67,537	0.614	1.111	0.000	0.000	0.000	0.701	4.615
$INN2_{i,t-1}$	67,537	1.236	2.078	0.000	0.000	0.000	2.504	7.153
$\Delta Earnings_{i,t}$	67,537	0.025	0.235	-0.985	-0.036	0.010	0.057	1.905
$\Delta Assets_{i,t}$	67,537	0.065	0.637	-4.186	-0.057	0.054	0.195	3.529
$\Delta R\&D_{i,t}$	67,537	0.000	0.031	-0.184	0.000	0.000	0.005	0.117
$\Delta Dividends_{i,t}$	67,537	0.001	0.013	-0.093	0.000	0.000	0.000	0.082
$LnTA_{i,t-1}$	67,537	4.526	2.070	-1.952	3.048	4.354	5.830	10.141
$Leverage_{i,t-1}$	67,537	0.572	1.292	0.000	0.024	0.186	0.578	15.524
$MB_{i,t-1}$	67,537	1.783	2.290	0.240	0.752	1.108	1.882	30.731
$Financing_{i,t}$	67,537	0.052	0.300	-1.224	-0.028	0.002	0.078	2.057
$\Delta Interests_{i,t}$	67,537	0.002	0.043	-0.386	-0.002	0.000	0.006	0.242
$Age_{i,t-1}$	67,537	12.598	9.313	2.000	5.000	10.000	18.000	37.000

Note: This table shows summary statistics for the variables used in Table 5, Table 7, and Table 9.

4. Empirical models and results

4.1. Effects of technological imitation on the quantity of innovation: An industry-level analysis

4.1.1. Full-sample analyses

To examine the relationship between the degree of technological imitation and industry-level innovation activities, we estimate the following regression models:

$$\begin{aligned} y_{j,t} &= \beta_0 + \beta_1 IMI_{j,t-1} + \beta_2 IMI^2_{j,t-1} + \beta'_{CONTROLS}\mathbf{CONTROLS} \\ &\quad + Industry\ FE + Year\ FE + \varepsilon_{j,t}, \end{aligned} \tag{1}$$

where $y_{j,t}$ is an industry-level innovation measure for industry j in year t, and $IMI_{j,t-1}$ is the (standardized) industry-year-average competitor-quick-citation ratio for industry j in year $t-1$. The control variables include industry-average values for the following measures: size, profitability, R&D intensity, asset tangibility, leverage, investment, market-to-book ratio, firm age, and a financial constraint measure. We also add year dummies to capture unobserved heterogeneity across years.

Table 2 presents the regression results. We first use fixed-effects regression models as in Fang et al. (2014). Two industry-average innovation measures, i.e.,

Table 2: Effects of technological imitation on the quantity of innovation: An industry-level analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable	$ln(1+\overline{COUNT}_{j,t})$		$\overline{COUNT}_{j,t}$		$ln(1+\overline{CITE}_{j,t})$		$\overline{CITE}_{j,t}$		
Regression model	FE	Poisson	NB	ZiNB	FE	Poisson	NB	ZiNB	VIF
$IMI_{j,t-1}$	0.475***	0.357***	0.396***	0.392***	0.658***	0.433***	0.462***	0.388***	2.18
	(0.038)	(0.052)	(0.046)	(0.022)	(0.052)	(0.059)	(0.041)	(0.025)	
$IMI^2_{j,t-1}$	-0.184***	-0.128***	-0.156***	-0.154***	-0.255***	-0.162***	-0.177***	-0.147***	2.38
	(0.015)	(0.022)	(0.017)	(0.011)	(0.022)	(0.025)	(0.016)	(0.012)	
$\overline{Size}_{j,t-1}$	0.101***	0.082**	0.132***	0.119***	0.131***	0.052	0.145***	0.136***	1.39
	(0.034)	(0.036)	(0.038)	(0.017)	(0.048)	(0.051)	(0.044)	(0.021)	
$\overline{ROA}_{j,t-1}$	-0.037	0.102*	0.041	0.049	-0.013	0.115*	0.024	0.032	2.35
	(0.044)	(0.055)	(0.050)	(0.037)	(0.076)	(0.063)	(0.073)	(0.046)	
$\overline{R\&D}_{j,t-1}$	2.019***	0.024	0.364	0.292	2.262***	0.173	0.718	0.534*	1.52
	(0.580)	(0.586)	(0.572)	(0.270)	(0.712)	(0.548)	(0.548)	(0.307)	
$\overline{PPE}_{j,t-1}$	-0.828**	-0.506	0.027	-0.028	-1.121*	-0.062	-0.095	-0.212	2.19
	(0.417)	(0.555)	(0.496)	(0.249)	(0.625)	(0.552)	(0.474)	(0.288)	
$\overline{Lev}_{j,t-1}$	-0.496**	-0.465*	-0.404*	-0.418***	-0.819***	-0.427**	-0.472**	-0.463***	1.52
	(0.206)	(0.256)	(0.222)	(0.122)	(0.305)	(0.201)	(0.226)	(0.167)	
$\overline{Capex}_{j,t-1}$	-1.110	-0.763	-1.287*	-0.949*	-0.648	-1.194	-0.293	0.121	1.96
	(0.751)	(0.805)	(0.740)	(0.486)	(1.217)	(0.852)	(0.700)	(0.592)	
$\overline{MB}_{j,t-1}$	0.001	0.006	0.001	0.001	0.001	-0.001	0.000	0.002	2.13
	(0.004)	(0.004)	(0.003)	(0.003)	(0.006)	(0.005)	(0.005)	(0.004)	
$\overline{Age}_{j,t-1}$	-0.042	0.009	0.018**	0.296***	-0.044	0.005	0.017**	0.266***	2.41
	(0.070)	(0.006)	(0.007)	(0.038)	(0.103)	(0.006)	(0.008)	(0.048)	
$\overline{KZ}_{j,t-1}$	0.000	-0.001	-0.000	-0.000	0.001	-0.001	0.000	0.001	1.13
	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	
Constant	3.602***	1.481***	1.103***	0.656***	5.875***	3.438***	3.022***	-7.303***	
	(0.228)	(0.251)	(0.256)	(0.212)	(0.331)	(0.313)	(0.258)	(1.528)	

Table 2 (Continued): Effects of technological imitation on the quantity of innovation: An industry-level analysis

Dependent variable	(1) $ln(1+\overline{COUNT}_{j,t})$	(2) $\overline{COUNT}_{j,t}$	(3)	(4)	(5) $ln(1+\overline{CITE}_{j,t})$	(6) $\overline{CITE}_{j,t}$	(7)	(8)	(9)
Regression model	FE	Poisson	NB	ZiNB	FE	Poisson	NB	ZiNB	VIF
Logistic model predicting the membership in the "certain zero" group									
Prop. of zero R&D firms				21.628***				2.761***	
				(7.270)				(0.786)	
$IMI_{j,t-1}$				0.068				-1.693***	
				(0.774)				(0.270)	
$IMI^2_{j,t-1}$				-0.052				0.730***	
				(0.217)				(0.119)	
$\overline{Size}_{j,t-1}$				-0.507				0.054	
				(0.348)				(0.173)	
$\overline{ROA}_{j,t-1}$				7.797*				-0.102	
				(4.026)				(0.403)	
$\overline{PPE}_{j,t-1}$				-3.147				-5.160***	
				(4.071)				(1.983)	
$\overline{Lev}_{j,t-1}$				-2.402				-0.091	
				(3.076)				(2.540)	
$\overline{Capex}_{j,t-1}$				7.761				4.503	
				(11.686)				(8.086)	
$\overline{MB}_{j,t-1}$				-1.009*				0.041	
				(0.594)				(0.033)	
$\overline{Age}_{j,t-1}$				-4.003***				0.732	
				(1.188)				(0.821)	
$\overline{KZ}_{j,t-1}$				0.052**				0.008	
				(0.026)				(0.020)	
Constant				-13.551**				-0.753***	
				(6.651)				(0.029)	
LR test of $\alpha = 0$			3098.66***				1.2E+05***		
p-value			0.000				0.000		
Vuong z-stat				-1.90				8.29***	
p-value				0.971				0.000	
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	6,400	6,400	6,400	6,400	6,400	6,400	6,400	6,400	
R-squared	0.383				0.490				
Number of industries	296	296	296	296	296	296	296	296	

Note: This table reports the results of regressions designed to estimate the impact of technological imitation on the quantity of industry-level innovation. The results of four different models are reported in Columns (1) through (8). "FE" stands for a fixed-effects regression model, "Poisson" stands for a Poisson regression model, "NB" stands for a negative binomial regression model, and "ZiNB" stands for a zero-inflated negative binomial regression model. In all regression models, we include industry and year dummies to capture unobserved heterogeneity across industries and across years, respectively. In Column (9), variance inflation factors (VIFs) are reported. "LR test of $\alpha = 0$" is the test statistic of the likelihood-ratio χ^2 test that the dispersion parameter (α) is equal to zero. "Vuong z-stat" is the test statistic for the Vuong test that compares the zero-inflated negative binomial model to the standard negative binomial model. Standard errors clustered by industry are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

$\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$, are skewed to the right, so we transform the variables by adding one and then taking the natural logarithm (i.e., $ln(1+\overline{COUNT}_{j,t})$ or $ln(1+\overline{CITE}_{j,t})$) as in Fang et al. (2014). The results reported in Columns (1) and (5) suggest that there is an inverted U-shaped relationship between technological imitation and two log-transformed industry-average innovation measures. As we include a quadratic term, $IMI^2_{j,t-1}$, we test for multicollinearity using variance inflation factors (VIFs) based on an ordinary-least-squares (OLS) regression model. The maximum VIF for the set of independent variables is only 2.41 (i.e., much smaller than 10), so multicollinearity does not appear to be a serious issue.[6] In addition, our main finding is robust to *i)* using three-digit SIC codes to classify industries; *ii)* defining the degree of imitation as an industry-average competitor-citation ratio without the five-year restriction; *iii)* restricting the sample to the industry-years with at least 30 patents; and *iv)* controlling for product market competition as measured by (1-Lerner's index). We also find very similar results when using firm-level variables instead of industry-average variables.

However, it is often reported that log transformations perform poorly compared to Poisson and negative binomial models, except when the dispersion is

[6]A maximum VIF greater than 10 is believed to signal serious multicollinearity (Marquaridt, 1970).

small and the mean counts are large (e.g., O'hara and Kotze (2010)). Thus, we employ three types of count data regression models (i.e., Poisson regression model, negative binomial regression model, and zero-inflated negative binomial regression model) in which the dependent variable is one of the two original industry-average innovation measures (i.e., $\overline{COUNT}_{j,t}$ or $\overline{CITE}_{j,t}$) as in Aghion et al. (2005). Columns (2) and (6) present the results from Poisson regression models. Regardless of the choice of the dependent variable, we find evidence that there is an inverted U-shaped relationship between technological imitation and industry-average innovation.[7] However, as the summary statistics in Table 1 Panel A indicate, the standard deviation of $\overline{COUNT}_{j,t}$ is 2.13 times its mean, and the standard deviation of $\overline{CITE}_{j,t}$ is 2.11 times its mean; thus, there is a strong possibility that these variables are over-dispersed. In such a case, negative binomial models would be more appropriate.

Columns (3) and (7) present the results from negative binomial regression models. We first test whether the dispersion parameter α is equal to zero using the likelihood-ratio χ^2 test. The test statistic in Column (3) (Column (7)) is negative two times the difference of the log-likelihood from the Poisson model and the negative binomial model, 3,099 (120,000) with an associated p-value of 0.000

[7]We conduct further tests to conclude that there is an inverted U-shaped relationship.

(0.000). The high test statistics suggest that both $\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$ are over-dispersed and are not sufficiently described by the simpler Poisson distribution. Again, we find evidence that there is an inverted U-shaped relationship between technological imitation and industry-average innovation measured by $\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$.

However, the low median values of $\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$ in Table 1 Panel A also suggest that there may be many zeros for these variables. According to our investigation, $\overline{COUNT}_{j,t}$ ($\overline{CITE}_{j,t}$) has 470 (669) zeros among 6,400 industry-year observations. If there are many zeros, a zero-inflated negative binomial model that explicitly models excess zeros (or certain zeros) would be appropriate. Columns (4) and (8) present the results from zero-inflated negative binomial regressions. To predict membership in the "certain zero" group, first-stage logistic models include various predictors such as *Prop. of zero R&D firms*, imitation, imitation squared, firm size, return on assets, asset tangibility, leverage, capital expenditures, market-to-book ratio, firm age, and a financial constraint measure.[8] Only *Prop. of zero R&D firms* has significant coefficients in both first-stage logistic models reported

[8]We do not include an agglomeration measure in this regression model because sample size drops significantly in this case. The concentration data are available only for the manufacturing sector, and the state and area employment, hours, and earnings database is publicly available only for the period between 1990 and 2002. Nevertheless, we have tested whether our main results still hold when agglomeration is controlled for, finding that our results with and without the agglomeration measure are quite similar.

in Columns (4) and (8). Significantly positive coefficients suggest that the higher the proportion of zero R&D firms, the more likely it is that the industry has a certain zero.

A Vuong test is often used to compare a zero-inflated negative binomial model with a corresponding standard negative binomial model. In Column (4), the Vuong test statistic is not significantly different from zero at the conventional level of significance (p-value=0.971), suggesting that the zero-inflated negative binomial model reported in Column (4) does not have a better fit than the corresponding standard negative binomial model reported in Column (3). Columns (3) and (4) both provide evidence that there is an inverted U-shaped relationship between technological imitation and the quantity of innovation measured by $\overline{COUNT}_{j,t}$. However, in Column (8), the Vuong test statistic is significant at the 1% level (p-value=0.000), suggesting that the zero-inflated negative binomial model reported in Column (8) has a better fit than the corresponding standard negative binomial model reported in Column (7). Once again, we find evidence that there is an inverted U-shaped relationship between technological imitation and quantity of innovation measured by $\overline{CITE}_{j,t}$.

To thoroughly examine whether there is an inverted U-shaped relationship between technological imitation and the quantity of innovation, we follow the procedures suggested by Haans et al. (2016).[9] First, we conduct Lind and Mehlum's (2010) test to examine whether the slopes at the lower and upper bounds are significantly positive and negative, respectively. The Lind and Mehlum's test results based on zero-inflated negative binomial models reported in Column (4) (Column (8)) suggest that the slope at the lower bound is 0.745 [t-stat=8.88] (0.738 [t-stat=9.57]) and the slope at the upper bound is -0.674 [t-stat=-7.81] (-0.664 [t-stat=-7.51]), and the overall test of an inverted U-shaped relationship rejects the null hypothesis of a monotone or U-shaped relationship at the 1% level of significance with a t-statistic of 7.81 (7.51). Thus, Lind and Mehlum's tests confirm that slopes at the lower and upper bounds are significantly positive and negative at the 1% level. Second, we test whether the turning point and its 95% confidence interval are within the data range of technological imitation. The estimation results based on zero-inflated negative binomial models suggest that the industry-average number of patents ($\overline{COUNT}_{j,t}$) peaks when $IMI_{j,t-1}$ has a value of 1.272 with the 95% confidence interval of [1.139, 1.405], while the industry-average number of citations ($\overline{CITE}_{j,t}$) peaks when $IMI_{j,t-1}$ has a value of 1.317 with the 95% confidence interval of [1.162, 1.471].[10] An investigation of the distribution of our

[9]We thank the anonymous referee for suggesting the procedures.

[10]The standardized imitation level corresponding to each of the peak points is estimated as $-\beta^{IMI_{j,t-1}}/2\beta^{IMI^2_{j,t-1}}$, where $\beta^{IMI_{j,t-1}}$ is the regression coefficient of $IMI_{j,t-1}$ and $\beta^{IMI^2_{j,t-1}}$ is the

imitation measure suggests that turning points correspond to approximately the 89th percentile with their 95% confidence intervals being [the 87th percentile, the 90th percentile],[11] suggesting that the turning points and their 95% confidence intervals are within the data range of technological imitation. Thus, an increase in technological imitation leads to an increase in the quantity of innovation up to the 89th percentile of technological imitation, but after that point, the effect becomes negative. This result implies that the positive externalities from the interactions among firms during the process of innovation outweigh the negative effects of free-riding concerns on firms' innovation activities up to a high degree of technological imitation. By contrast, free-riding concerns outweigh the positive externalities when the level of technological imitation is extremely high.

4.1.2. Sector-by-sector analyses

To further examine whether the relationship between technological imitation and the quantity of innovation is heterogeneous across sectors, we repeat the analysis for each sector, where the sector is defined following Greenhalgh and Rogers' (2006) classification of six technology sectors. Greenhalgh and Rogers (2006) expanded Pavitt's (1984) classification of technology sectors. Pavitt (1984) origi-

regression coefficient of $IMI^2_{j,t-1}$.

[11] The summary statistics for $IMI_{j,t-1}$ reported in Table 1 suggest that the sample mean (median) is -0.006 (-0.353),and the minimum (maximum) is -1.173 (3.507).

Table 3: Pavitt technology sectors

Pavitt sector	Description	SIC	Relative balance between product and process innovation	Average proportion of R&D firms	Average proportion of firms with patents	Average R&D expenses to total assets	Average imitation level before standardization
Pavitt 1: Supplier dominated	Traditional manufacturing. Generally small firms with weak in-house R&D and engineering capabilities. Innovations come from suppliers of equipment or materials	12, 13, 15, 16, 22, 23, 24, 25, 26, 30, 31, 7300, 7312, 7336, 7361	Process	25.70%	18.77%	0.71%	9.39%
Pavitt 2: Production intensive, Scale intensive	Large firms producing standard materials or durable goods, including cars	20, 21, 32, 33, 34, 37	Process	49.19%	36.10%	1.59%	8.79%
Pavitt 3: Production intensive, Specialised suppliers	Machinery and instruments. Tend to be smaller firms which are technologically specialised	35, 38, 39	Product	86.93%	41.30%	8.60%	11.09%
Pavitt 4: Science based	Electronics, electrical and chemicals. Often large firms. Technology from in-house R&D but based on basic science from elsewhere	28, 29, 36	Mixed	83.34%	38.80%	12.49%	12.71%
Pavitt 5: Information intensive	Includes finance (not in the sample), retail, communications and publishing industries. In-house software or systems development, plus IT hardware and software purchases	27, 48, 50–67, 7313 and 7383 (media)	Mixed	11.69%	8.22%	0.69%	5.20%
Pavitt 6: Software-related firms	Computer software and services firms	All SIC 73 sub-codes not shown above	Product	68.32%	11.91%	10.63%	9.25%

Note: We closely follow Greenhalgh and Rogers' (2006) classification of technology sectors. Pavitt (1984) originally introduced four industrial classifications based on technological trajectories: "supplier dominated" (Pavitt 1), "production intensive" (scale intensive) (Pavitt 2), "production intensive (specialist suppliers)" (Pavitt 3), and "science based" (Pavitt 4). Tidd et al. (2005) included a new sector called "information intensive" (Pavitt 5), which includes firms in finance, retail and publishing. Greenhalgh and Rogers (2006) allocated "software-related firms" (Pavitt 6) to a separate sector. The first four columns of this table are from Greenhalgh and Rogers (2006), and the last four columns are based on our sample.

nally introduced four industrial classifications based on technological trajectories: "supplier dominated" (Pavitt 1), "production intensive (scale intensive)" (Pavitt 2), "production intensive (specialist suppliers)" (Pavitt 3), and "science based" (Pavitt 4). Tidd et al. (2005) included a new sector called "information intensive" (Pavitt 5), which includes firms in finance, retail and publishing. Greenhalgh and Rogers (2006) allocated "software-related firms" (Pavitt 6) to a separate sector.

Table 3 describes the Pavitt technology sectors and provides some summary statistics. This table shows that innovation patterns are very heterogeneous across the Pavitt sectors. Pavitt sectors 3 and 4 have significantly higher levels of innovation: larger proportions of R&D firms, higher proportions of firms with patents, and higher R&D-to-total-assets ratios. Software industries (Pavitt 6) have a larger proportion of R&D firms and a higher R&D-to-total-assets ratio, but have a relatively lower proportion of firms with patents. Pavitt sectors 5 and 1 tend to have significantly lower levels of innovation based on the three measures. This table also shows that imitation levels are very heterogeneous across Pavitt sectors. Pavitt sectors 3 and 4 with significantly higher levels of innovation have significantly higher levels of imitation as well. Pavitt sector 5, with the lowest level of innovation, also has the lowest level of imitation. Pavitt sector 4, with a higher imitation level than Pavitt sector 3, has a lower proportion of R&D firms, a lower proportion of firms with patents, and a higher R&D-to-total-assets ratio. This preliminary analysis of the summary statistics implies that there may be an inverted U-shaped or positive relationship between innovation activities and the degree of industry-level imitation.

Table 4 reports the zero-inflated negative binomial regression results for each sector. The dependent variable is the industry-average number of citations of the patents that any firms in industry j applied for in year t ($\overline{CITE}_{j,t}$). In all six Pavitt sectors, *Prop. of zero R&D firms* has significantly positive signs in the logistic models predicting membership in the "certain zero" group, suggesting that the higher the proportion of zero R&D firms, the more likely the industry is to have a certain zero. In the six sectors, the test statistics for the Vuong tests are significant at the 1% (10%) level for Pavitt sectors 1, 2, 3, 4, and 5 (Pavitt sector 6), suggesting that in all Pavitt sectors, the zero-inflated negative binomial model is a better fit than the standard negative binomial model. We find a clear inverted U-shaped relationship between technological imitation and the quantity of innovation, regardless of the Pavitt sector.

The estimation results based on the zero-inflated negative binomial regressions suggest that $\overline{CITE}_{j,t}$ has peaks at the 89th, 90th, 89th, 83rd, 86th, and 88th percentiles in Pavitt sectors 1, 2, 3, 4, 5, and 6, respectively. The 95% confidence intervals of the imitation levels corresponding to the peak points are as follows: (1) Pavitt sector 1: [85th percentile, 93rd percentile]; (2) Pavitt sector 2: [88th percentile, 93rd percentile]; (3) Pavitt sector 3: [86th percentile, 91st percentile]; (4) Pavitt sector 4: [77th percentile, 87th percentile]; (5) Pavitt sector 5: [84th percentile, 88th percentile]; and (6) Pavitt sector 6: [83rd percentile, 92nd percentile]. The results suggest that the peak points and their 95% confidence intervals are within the data range. In general, all Pavitt sectors have peak points at similar imitation levels, although Pavitt sector 4 has a peak point at a significantly lower imitation level compared to Pavitt sector 2 at the 5% level.[12] The differences across Pavitt sectors are not economically meaningful given that all 6 sectors have peaks between 83rd percentile and 90th percentile. The results imply that regardless of the Pavitt sector, the positive externalities from the interactions among firms during the innovation process dominate the negative effects of free-riding concerns on firms' innovation activities up to a rather high degree of technological imitation, whereas free-riding concerns dominate the positive externalities only when technological imitation is at extremely high levels.

[12] Note that the 95% confidence intervals for Pavitt sectors 2 and 4 do not overlap with each other.

Table 4: Effects of technological imitation on the quantity of innovation: By Pavitt technology sectors

Pavitt sector	(1) Pavitt 1	(2) Pavitt 2	(3) Pavitt 3	(4) Pavitt 4	(5) Pavitt 5	(6) Pavitt 6
$IMI_{j,t-1}$	0.345***	0.489***	0.464***	0.195***	0.493***	0.670***
	(0.051)	(0.050)	(0.056)	(0.059)	(0.082)	(0.170)
$IMI^2_{j,t-1}$	-0.132***	-0.166***	-0.177***	-0.137***	-0.242***	-0.275***
	(0.027)	(0.023)	(0.024)	(0.037)	(0.044)	(0.079)
$\overline{Size}_{j,t-1}$	0.161***	0.168***	0.130***	0.182***	0.138	-0.060
	(0.048)	(0.041)	(0.040)	(0.044)	(0.089)	(0.159)
$\overline{ROA}_{j,t-1}$	-0.177	0.138	-0.132	0.086	-0.048	0.393
	(0.120)	(0.089)	(0.098)	(0.091)	(0.127)	(0.278)
$\overline{R\&D}_{j,t-1}$	0.012	1.232	0.403	-0.116	6.308*	2.425
	(1.377)	(0.886)	(0.447)	(0.515)	(3.616)	(2.134)
$\overline{PPE}_{j,t-1}$	-2.053***	-1.616***	-1.180*	0.288	0.865	0.033
	(0.547)	(0.546)	(0.701)	(0.723)	(1.027)	(1.701)
$\overline{Lev}_{j,t-1}$	-1.190***	0.411	0.138	-0.453	-0.666	0.630
	(0.276)	(0.319)	(0.312)	(0.380)	(0.598)	(1.212)
$\overline{Capex}_{j,t-1}$	-0.026	-0.487	-0.509	1.489	-3.452	0.335
	(1.059)	(1.176)	(1.133)	(1.333)	(2.464)	(4.930)
$\overline{MB}_{j,t-1}$	0.009	-0.004	-0.008	0.011	0.014	-0.007
	(0.015)	(0.010)	(0.006)	(0.009)	(0.009)	(0.020)
$\overline{Age}_{j,t-1}$	0.195*	0.045	0.118	0.432***	0.140	-0.516
	(0.110)	(0.092)	(0.099)	(0.126)	(0.189)	(0.533)
$\overline{KZ}_{j,t-1}$	0.002	-0.004	0.009***	-0.003	-0.005	0.007
	(0.003)	(0.004)	(0.003)	(0.004)	(0.005)	(0.006)
Constant	2.735***	2.573***	4.836***	3.901***	1.361**	2.254*
	(0.752)	(0.382)	(0.351)	(0.511)	(0.617)	(1.276)
Vuong z-stat	3.37	4.87	3.54	2.94	3.32	1.37
p-value	0.000	0.000	0.000	0.007	0.000	0.091
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,001	1,343	1,427	1,064	683	239

Note: This table reports the second-stage results of sector-by-sector zero-inflated negative binomial (ZiNB) regressions designed to estimate the impact of technological imitation on the quantity of industry-level innovation. The results are reported for the following sector: (1) Pavitt 1: Supplier dominated; (2) Pavitt 2: Production intensive (scale intensive); (3) Pavitt 3: Production intensive (specialized suppliers); (4) Pavitt 4: Science based; (5) Pavitt 5: Information intensive; (6) Pavitt 6: Software-related firms. Note that we exclude firms in the utilities and financial service sectors. The dependent variable is the industry-average number of citations of the patents applied for in year t by any firms in industry j ($\overline{CITE}_{j,t}$). We include year and industry dummies to capture unobserved heterogeneity across years and industries, respectively. Standard errors clustered by industry are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

4.2. Effects of technological imitation on the market value of innovation: A firm-level analysis

4.2.1. Full-sample analyses

To further investigate the impact of technological imitation on the value of firm-level innovation, we follow the approach used by Im et al. (2015).[13] To measure the market value of firm-level innovation, we estimate the sensitivity of raw (excess) stock returns to a firm-level innovation measure controlling for various factors affecting stock returns. Specifically, we estimate the coefficient of a firm-level innovation measure in a regression model in which the dependent variable is raw (excess) stock returns. In this study, we model the regression coefficient as a quadratic function of technological imitation to investigate the effect of the degree of technological imitation on a firm's incentive to innovate as measured by the market value of firm-level innovation.

The model is specified as follows:

$$\begin{aligned} r_{i,t}\ (\text{or } r_{i,t} - R_{B,t}) = \beta_0 + \beta_1 INN_{i,t-1} + \beta'_{CONTROLS}\mathbf{CONTROLS} \\ + \mathit{Firm\ FE} + \mathit{Year\ FE} + \varepsilon_{i,t}, \end{aligned} \quad (2)$$

[13] Im et al. (2015) employed the approach used by Faulkender and Wang (2006) and Dittmar and Mahrt-Smith (2007) to measure the market value of cash holdings.

Table 5: Effects of technological imitation on the market value of firm-level innovation: Annual raw or excess stock returns as the dependent variable

Dependent variable	(1) $r_{i,t}$	(2) $r_{i,t}-R_{p,t}$	(3) $r_{i,t}-R_{j,t}$	(4) VIF	(5) $r_{i,t}$	(6) $r_{i,t}-R_{p,t}$	(7) $r_{i,t}-R_{j,t}$	(8) VIF
$INN1_{i,t-1}$	0.032***	0.032***	0.020***	1.82				
	(0.006)	(0.006)	(0.005)					
$INN1_{i,t-1} \times IMI_{j,t-1}$	0.012***	0.017***	0.009***	2.92				
	(0.004)	(0.004)	(0.003)					
$INN1_{i,t-1} \times IMI^2_{j,t-1}$	-0.009***	-0.008***	-0.006***	3.26				
	(0.002)	(0.002)	(0.001)					
$INN2_{i,t-1}$					0.017***	0.016***	0.011***	1.77
					(0.002)	(0.002)	(0.002)	
$INN2_{i,t-1} \times IMI_{j,t-1}$					0.011***	0.013***	0.007***	2.53
					(0.002)	(0.002)	(0.002)	
$INN2_{i,t-1} \times IMI^2_{j,t-1}$					-0.006***	-0.005***	-0.003***	2.78
					(0.001)	(0.001)	(0.001)	
$IMI_{j,t-1}$	0.011***	0.007*	0.003	1.99	0.009**	0.004	0.001	1.96
	(0.004)	(0.004)	(0.004)		(0.004)	(0.004)	(0.004)	
$\Delta Earnings_{i,t}$	0.569***	0.576***	0.467***	1.11	0.569***	0.577***	0.468***	1.11
	(0.020)	(0.020)	(0.018)		(0.020)	(0.020)	(0.018)	
$\Delta Assets_{i,t}$	0.231***	0.223***	0.194***	1.70	0.231***	0.223***	0.194***	1.70
	(0.009)	(0.009)	(0.008)		(0.009)	(0.009)	(0.008)	
$\Delta R\&D_{i,t}$	0.763***	0.791***	0.676***	1.10	0.764***	0.790***	0.676***	1.10
	(0.126)	(0.127)	(0.119)		(0.126)	(0.127)	(0.119)	
$\Delta Dividends_{i,t}$	1.359***	1.416***	1.112***	1.01	1.358***	1.416***	1.112***	1.01
	(0.281)	(0.289)	(0.257)		(0.281)	(0.289)	(0.257)	
$LnTA_{i,t-1}$	-0.236***	-0.236***	-0.196***	1.83	-0.236***	-0.235***	-0.196***	1.74
	(0.007)	(0.007)	(0.006)		(0.007)	(0.007)	(0.006)	
$Leverage_{i,t-1}$	0.090***	0.089***	0.077***	1.21	0.090***	0.089***	0.077***	1.21
	(0.005)	(0.005)	(0.005)		(0.005)	(0.005)	(0.005)	
$MB_{i,t-1}$	-0.076***	-0.068***	-0.068***	1.15	-0.077***	-0.069***	-0.068***	1.15
	(0.003)	(0.003)	(0.003)		(0.003)	(0.003)	(0.003)	
$Financing_{i,t}$	0.110***	0.120***	0.102***	1.42	0.110***	0.120***	0.102***	1.42
	(0.018)	(0.018)	(0.016)		(0.018)	(0.018)	(0.016)	
$\Delta Interests_{i,t}$	-1.423***	-1.395***	-1.163***	1.26	-1.422***	-1.394***	-1.163***	1.26
	(0.106)	(0.107)	(0.095)		(0.106)	(0.107)	(0.095)	
$Age_{i,t-1}$	0.008	-0.002	0.009**	1.43	0.011**	-0.001	0.011**	1.43
	(0.005)	(0.005)	(0.004)		(0.005)	(0.005)	(0.004)	
Constant	1.123***	0.731***	0.652***		1.127***	0.733***	0.653***	
	(0.030)	(0.030)	(0.026)		(0.030)	(0.030)	(0.026)	
Year fixed effects	Yes	Yes	Yes		Yes	Yes	Yes	
Firm fixed effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	67,537	67,537	67,537		67,537	67,537	67,537	
Adjusted R-squared	0.261	0.191	0.158		0.262	0.191	0.158	

Note: This table reports the results of regressions designed to estimate the impact of technological imitation on the value of firm-level innovation. Raw stock returns ($r_{i,t}$) or excess returns ($r_{i,t}-R_{p,t}$ or $r_{i,t}-R_{j,t}$) are used as the dependent variable. The regression models reported in Columns (1) through (3) and Columns (5) through (7) are estimated using the within-groups (i.e., fixed-effects) estimator. Standard errors clustered by firm are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. In Columns (4) and (8), variance inflation factors (VIFs) are reported.

where

$$\beta_1 = \gamma_0 + \gamma_1 IMI_{j,t-1} + \gamma_2 IMI_{j,t-1}^2. \tag{3}$$

$r_{i,t}$ is the annualized stock return of firm i in year t, and $R_{B,t}$ is the annualized return of the benchmark portfolio in year t. The benchmark portfolios are Fama and French's 5×5 portfolios of size and book-to-market ratio and industry portfolio, and the corresponding portfolio returns are denoted $R_{p,t}$ and $R_{j,t}$, respectively. $IMI_{j,t-1}$ is the lagged technological imitation measure, and $INN_{i,t-1}$ is the lagged value of a firm-level innovation measure ($INN1_{i,t-1}$ or $INN2_{i,t-1}$). Both measures are defined in Appendix B. The control variables include the ratio of the change in earnings to market equity, the ratio of the change in total assets to market equity, the ratio of the change in R&D expenses to market equity, the ratio of the change in dividends to market equity, the ratio of the change in interest expenses to market equity, the ratio of new financing to market equity, the lagged leverage ratio, the lagged natural logarithm of total assets, the lagged market-to-book ratio, firm age, and the linear term of imitation ($IMI_{j,t-1}$).[14] We employ within-groups (i.e., fixed-effects) estimators to capture unobserved heterogeneity across firms.[15] We

[14]In addition to the control variables used in Im et al. (2015), we include firm age and the linear term of imitation because firm age and the linear term may have an impact on stock returns. We do not include the quadratic term of imitation because the quadratic term is not significant in most regression models.

[15]Stock returns may vary with industry affiliation, so it is necessary to control for industry

also include year dummies to capture unobserved heterogeneity across years. The definitions of the control variables are provided in Appendix B.

Table 5 reports the regression results for the model specified in Equations (2) and (3). We use two different measures for firm-level innovation and three different regression specifications to measure the market value of innovation. Columns (1) through (3) are based on $INN1_{i,t-1}$ as a firm-level innovation measure, while Columns (5) through (7) are based on $INN2_{i,t-1}$. We use three dependent variables: $r_{i,t}$ is the raw return, $r_{i,t} - R_{p,t}$ is the excess return based on Fama and French's 5×5 portfolios, and $r_{i,t} - R_{j,t}$ is the excess return based on industry portfolios. Regardless of the model specifications, the regression coefficient of $INN_{i,t-1} \times IMI^2_{j,t-1}$ is significantly negative at the 1% level and the regression coefficient of $INN_{i,t-1} \times IMI_{j,t-1}$ is significantly positive at the 1% level, suggesting that the relationship between technological imitation and the market value of innovation has an inverted U-shaped relationship. That is, a firm's incentive to innovate increases with the intensity of technological imitation up to a certain point, beyond which it starts to decrease with the intensity of technological imitation.

The estimation results with raw returns as the dependent variables suggest that

effects. However, we do not include industry dummies because the industry fixed effects are subsumed by firm fixed effects.

the value of innovation measured by the regression coefficient of $INN1_{i,t-1}$ peaks when $IMI_{j,t-1}$ has a value of 0.672, while the value of innovation measured by the regression coefficient of $INN2_{i,t-1}$ peaks when $IMI_{j,t-1}$ has a value of 0.920.[16] The values correspond to approximately the 82nd and 85th percentiles, respectively. The 95% confidence intervals for the peak points are [77th percentile, 85th percentile] and [81st percentile, 88th percentile], respectively. The results suggest that the peak points and their confidence intervals are within the data range. Thus, an increase in technological imitation leads to an increase in the market value of innovation up to the 82nd to 85th percentile of technological imitation, but the effect then becomes negative after that point. This finding implies that the positive externalities from the interactions among firms during the innovation process outweigh the negative effects of free-riding concerns on firms' incentives to innovate up to a high degree of technological imitation, while free-riding concerns outweigh the positive externalities when there is a very high level of technological imitation.

As we include interaction terms, we test for multicollinearity using VIFs based on an OLS regression model. The maximum VIF for the first (second) set of

[16]The standardized imitation level corresponding to each of the peak points is estimated as $-\beta^{INN_{i,t-1} \times IMI_{j,t-1}} / 2\beta^{INN_{i,t-1} \times IMI^2_{j,t-1}}$, where $\beta^{INN_{i,t-1} \times IMI_{j,t-1}}$ is the regression coefficient of $INN_{i,t-1} \times IMI_{j,t-1}$ and $\beta^{INN_{i,t-1} \times IMI^2_{j,t-1}}$ is the regression coefficient of $INN_{i,t-1} \times IMI^2_{j,t-1}$.

independent variables is only 3.26 (2.78), so multicollinearity does not seem to be a serious issue. Our main finding is robust to *i)* using three-digit SIC codes to classify industries; *ii)* restricting the sample to the industry-years with at least 30 patents; *iii)* including firm fixed effects; and *iv)* controlling for the effect of product market competition.

In the firm-level model specified in Equations (2) and (3), there might be an endogeneity problem concerning the relationship between firm innovation ($INN_{i,t-1}$) and technological imitation ($IMI_{j,t-1}$). The causal relationship between $INN_{i,t-1}$ and $IMI_{j,t-1}$ is actually the rationale underlying the industry-level model in Equation (1). A possible solution to the endogeneity concern is to construct a new dependent variable, such as the stock return divided by $INN_{i,t-1}$, which indicates the average annual return of firm innovations.[17] By directly including $IMI_{j,t-1}$ and $IMI_{j,t-1}^2$ as explanatory variables, we can test for the curvilinear effects of technological imitation on the average annual return of firm innovations. We use the same control variables as those used in Table 5. We employ within-groups (i.e., fixed-effects) estimators to capture unobserved heterogeneity across firms.[18] We

[17]We are grateful to an anonymous referee for pointing out the potential endogeneity problem and proposing the solution.

[18]The average annual return of firm innovations may vary with industry affiliation, so it is necessary to control for industry effects. However, we do not include industry dummies because the industry fixed effects are subsumed by firm fixed effects.

Table 6: Effects of technological imitation on the market value of firm-level innovation: Annual raw or excess stock returns per innovation as the dependent variable

Dependent variable	(1) $\frac{r_{i,t}}{INN1_{i,t-1}}$	(2) $\frac{r_{i,t}-R_{p,t}}{INN1_{i,t-1}}$	(3) $\frac{r_{i,t}-R_{j,t}}{INN1_{i,t-1}}$	(4) VIF	(5) $\frac{r_{i,t}}{INN2_{i,t-1}}$	(6) $\frac{r_{i,t}-R_{p,t}}{INN2_{i,t-1}}$	(7) $\frac{r_{i,t}-R_{j,t}}{INN2_{i,t-1}}$	(8) VIF
$IMI_{j,t-1}$	0.032***	0.032***	0.019**	5.32	0.016***	0.016***	0.010***	4.88
	(0.010)	(0.010)	(0.009)		(0.004)	(0.004)	(0.004)	
$IMI^2_{j,t-1}$	-0.011***	-0.012***	-0.008**	3.35	-0.005***	-0.005***	-0.003**	2.94
	(0.004)	(0.004)	(0.004)		(0.002)	(0.002)	(0.002)	
$\Delta Earnings_{i,t}$	0.742***	0.740***	0.625***	1.13	0.259***	0.261***	0.223***	1.13
	(0.055)	(0.055)	(0.049)		(0.020)	(0.020)	(0.018)	
$\Delta Assets_{i,t}$	0.221***	0.217***	0.184***	1.76	0.086***	0.085***	0.077***	1.79
	(0.020)	(0.020)	(0.018)		(0.007)	(0.007)	(0.007)	
$\Delta R\&D_{i,t}$	0.570***	0.619***	0.470**	1.16	0.204**	0.211**	0.148*	1.15
	(0.218)	(0.217)	(0.201)		(0.084)	(0.084)	(0.078)	
$\Delta Dividends_{i,t}$	0.462	0.713	0.132	1.03	0.188	0.328*	0.076	1.03
	(0.498)	(0.505)	(0.453)		(0.204)	(0.198)	(0.186)	
$LnTA_{i,t-1}$	-0.226***	-0.205***	-0.165***	1.33	-0.073***	-0.067***	-0.057***	1.35
	(0.012)	(0.012)	(0.011)		(0.005)	(0.005)	(0.004)	
$Leverage_{i,t-1}$	0.078***	0.076***	0.068***	1.18	0.026***	0.024***	0.024***	1.18
	(0.013)	(0.013)	(0.011)		(0.005)	(0.005)	(0.004)	
$MB_{i,t-1}$	-0.057***	-0.050***	-0.051***	1.21	-0.019***	-0.016***	-0.016***	1.21
	(0.004)	(0.004)	(0.004)		(0.001)	(0.001)	(0.001)	
$Financing_{i,t}$	0.143***	0.162***	0.134***	1.50	0.032**	0.037***	0.030**	1.50
	(0.040)	(0.040)	(0.036)		(0.014)	(0.014)	(0.013)	
$\Delta Interests_{i,t}$	-1.914***	-1.891***	-1.428***	1.22	-0.731***	-0.716***	-0.586***	1.23
	(0.262)	(0.260)	(0.226)		(0.100)	(0.101)	(0.094)	
$Age_{i,t-1}$	0.004	-0.002	0.004	1.63	0.005***	0.001	0.002	1.67
	(0.004)	(0.004)	(0.004)		(0.002)	(0.002)	(0.002)	
Constant	1.228***	0.870***	0.695***		0.406***	0.277***	0.235***	
	(0.060)	(0.060)	(0.053)		(0.024)	(0.023)	(0.021)	
Year fixed effects	Yes	Yes	Yes		Yes	Yes	Yes	
Firm fixed effects	Yes	Yes	Yes		Yes	Yes	Yes	
Observations	22,274	22,274	22,274		20,388	20,388	20,388	
Adjusted R-squared	0.235	0.173	0.141		0.230	0.168	0.140	

Note: This table reports the results of regressions designed to estimate the impact of technological imitation on the value of firm-level innovation. Raw stock returns per innovation or excess stock returns per innovation are used as the dependent variable. The regression models reported in Columns (1) through (3) and Columns (5) through (7) are estimated using the within-groups (i.e., fixed-effects) estimator. Standard errors clustered by firm are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. In Columns (4) and (8), variance inflation factors (VIFs) are reported.

also include year dummies to capture unobserved heterogeneity across years.

Table 6 shows the regression results for the alternative model. We use two different measures for firm-level innovation and three different specifications. Columns (1) through (3) are based on the annual raw or excess returns divided by $INN1_{i,t-1}$ as the dependent variable, while Columns (5) through (7) are based on the annual raw or excess returns divided by $INN2_{i,t-1}$ as the dependent variable. To thoroughly examine whether there is an inverted U-shaped relationship between technological imitation and the value of innovation, we check if the peak points and their confidence intervals are within the data range. For example, the estimation results reported in Columns (1) and (5) suggest that $r_{i,t}/INN1_{i,t-1}$ peaks when $IMI_{j,t-1}$ has a value of 1.479, while $r_{i,t}/INN2_{i,t-1}$ peaks when $IMI_{j,t-1}$ has a value of 1.742. The values correspond to approximately the 90th and 93rd percentiles, respectively. The 95% confidence intervals for the peak points are [84th percentile, 95th percentile] and [85th percentile, 97th percentile], respectively. These analyses suggest that all the peak points and their confidence intervals are within the data range.

Regardless of the specifications, the relationship between technological imitation and the market value of firm-level innovation has an inverted U-shaped relationship, suggesting that a firm's incentive to innovate increases with the intensity of technological imitation up to a certain point, beyond which it starts to decrease with the intensity of technological imitation. Therefore, we conclude that the results reported in Table 5 are not driven by the endogeneity problem mentioned above. As we include a quadratic term, we test for multicollinearity using VIFs based on an OLS regression model. The maximum VIF in Column (4) (in Column (8)) is only 5.32 (4.88), so it appears that multicollinearity is not a serious issue.

4.2.2. Sector-by-sector analyses

To further examine whether the relationship between technological imitation and the market value of firm-level innovation is heterogeneous across sectors, we repeat the analysis specified in Equations (2) and (3) for each Pavitt sector. Table 7 reports the results of the sector-by-sector regressions. Using raw stock returns as the dependent variable, we find that the regression coefficient of $INN2_{i,t-1} \times IMI^2_{j,t-1}$ is negative regardless of Pavitt sectors. Note that the regression coefficient is statistically significant at the 10% level for the first three Pavitt sectors and is statistically significant at the 1% level for the fourth Pavitt sector but is not statistically significant for the fifth and sixth Pavitt sectors.[19] The value of

[19]The lack of significance of the coefficient estimates could be due to relatively smaller sample sizes.

Table 7: Effects of technological imitation on the market value of firm-level innovation: By Pavitt technology sectors

Pavitt sector	(1) Pavitt 1	(2) Pavitt 2	(3) Pavitt 3	(4) Pavitt 4	(5) Pavitt 5	(6) Pavitt 6
$INN2_{i,t-1}$	0.011*	0.008	0.020***	0.020***	0.023	0.031***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.017)	(0.011)
$INN2_{i,t-1} \times IMI_{j,t-1}$	-0.005	0.005	0.012**	0.025***	0.009	0.002
	(0.005)	(0.004)	(0.005)	(0.004)	(0.012)	(0.014)
$INN2_{i,t-1} \times IMI^2_{j,t-1}$	-0.008*	-0.002*	-0.007*	-0.015***	-0.010	-0.005
	(0.005)	(0.001)	(0.003)	(0.003)	(0.017)	(0.010)
$IMI_{j,t-1}$	0.024***	0.006	0.013	0.016	-0.017	-0.030**
	(0.009)	(0.009)	(0.010)	(0.010)	(0.011)	(0.012)
$\Delta Earnings_{i,t}$	0.514***	0.341***	0.872***	1.032***	0.699***	0.778***
	(0.039)	(0.052)	(0.049)	(0.059)	(0.069)	(0.073)
$\Delta Assets_{i,t}$	0.261***	0.027**	0.437***	0.464***	0.249***	0.367***
	(0.020)	(0.012)	(0.028)	(0.032)	(0.028)	(0.040)
$\Delta R\&D_{i,t}$	-0.071	0.739**	1.191***	1.023***	4.887***	1.294***
	(1.233)	(0.343)	(0.239)	(0.230)	(1.852)	(0.288)
$\Delta Dividends_{i,t}$	1.733**	0.098**	2.661***	1.215	0.854	1.056
	(0.804)	(0.046)	(0.794)	(0.810)	(1.126)	(1.142)
$LnTA_{i,t-1}$	-0.138***	-0.180***	-0.236***	-0.234***	-0.207***	-0.297***
	(0.015)	(0.017)	(0.014)	(0.013)	(0.038)	(0.020)
$Leverage_{i,t-1}$	0.075***	0.033***	0.164***	0.167***	0.124***	0.357***
	(0.010)	(0.010)	(0.017)	(0.018)	(0.016)	(0.040)
$MB_{i,t-1}$	-0.182***	-0.017**	-0.132***	-0.105***	-0.204***	-0.125***
	(0.012)	(0.007)	(0.007)	(0.005)	(0.017)	(0.007)
$Financing_{i,t}$	0.005	0.027	0.093*	0.373***	0.097*	0.446***
	(0.035)	(0.021)	(0.051)	(0.052)	(0.051)	(0.091)
$\Delta Interests_{i,t}$	-1.774***	-0.537***	-3.495***	-3.434***	-2.230***	-2.857***
	(0.240)	(0.195)	(0.337)	(0.371)	(0.339)	(0.612)
$Age_{i,t-1}$	0.014	0.027***	0.011	0.011	0.016	-0.011
	(0.012)	(0.009)	(0.011)	(0.010)	(0.020)	(0.021)
$Constant$	0.971***	0.910***	1.049***	1.109***	1.005***	0.835***
	(0.078)	(0.090)	(0.060)	(0.061)	(0.154)	(0.217)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	9,542	7,449	14,650	15,848	6,234	7,294
Adjusted R-squared	0.275	0.202	0.319	0.347	0.295	0.384

Note: This table reports the results of sector-by-sector regressions designed to estimate the impact of technological imitation on the market value of firm-level innovation. Each column represents the following sector: (1) Pavitt 1: supplier dominated; (2) Pavitt 2: Production intensive (scale intensive); (3) Pavitt 3: Production intensive (specialized suppliers); (4) Pavitt 4: Science based; (5) Pavitt 5: Information intensive; (6) Pavitt 6: Software-related firms. Note that we exclude firms in the utilities and financial service sectors. The dependent variable is raw stock returns ($r_{i,t}$). Standard errors clustered by firm are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

innovation measured by the regression coefficient of the innovation measure peaks at the 55th, 87th, 85th, 84th, 79th, and 73rd percentiles of our imitation measure in Pavitt sectors 1, 2, 3, 4, 5, and 6, respectively. The 95% confidence intervals of the imitation levels corresponding to the peak points are as follows: (1) Pavitt sector 1: [1st percentile, 79th percentile]; (2) Pavitt sector 2: [60th percentile, 97th percentile]; (3) Pavitt sector 3: [73rd percentile, 92nd percentile]; (4) Pavitt sector 4: [81st percentile, 87th percentile]; (5) Pavitt sector 5: [2nd percentile, 95th percentile]; and (6) Pavitt sector 6: [Below minimum, 97th percentile]. The results suggest that all the peak points and their 95% confidence intervals are within the data range, with the exception of the 95% lower bound for Pavitt sector 6. In general, all Pavitt sectors have peak points at similar imitation levels, although Pavitt sector 1 has a peak point at a significantly lower imitation level compared to Pavitt sector 4 at the 5% level.[20] Thus, technological imitation and the market value of firm-level innovation, regardless of Pavitt sectors, have an inverted U-shaped relationship. This result suggests that regardless of technological sectors, a firm's incentive to innovate increases with the intensity of technological imitation up to a certain point, beyond which it starts to decrease with the intensity of technological

[20]Note that the 95% confidence intervals for Pavitt sectors 1 and 4 do not overlap with each other.

imitation.

4.3. Effects of agglomeration on the relationship between technological imitation and the quantity and market value of innovation

4.3.1. Defining agglomerated and non-agglomerated industries

To derive implications for innovation cluster policies, we further investigate whether the relationships between imitation and the quantity and value of innovation differ between agglomerated and non-agglomerated industries. To measure the degree of agglomeration for industry j in year t, we closely follow Ellison and Glaeser (1997). The Ellison-Glaeser index ($\gamma^{EG}_{j,t}$) is defined as follows:

$$\gamma^{EG}_{j,t} = \frac{\sum_{i=1}^{S}(s_{i,j,t} - x_{i,t})^2 - (1 - \sum_{i=1}^{S} x_{i,t}^2) HHI_{j,t}}{(1 - \sum_{i=1}^{S} x_{i,t}^2)(1 - HHI_{j,t})}, \tag{4}$$

where $s_{i,j,t}$ is the share of industry j's employment in state i measured in year t, $x_{i,t}$ is the share of total employment in state i measured in year t, and $HHI_{j,t}$ is Herfindahl-Hirschman index (HHI) for industry j measured in year t. We calculate $s_{i,j,t}$ and $x_{i,t}$ using the state and area employment, hours, and earnings database provided by the US Bureau of Labor Statistics, and we calculate the HHI using the economic census database provided by the US Census Bureau.[21] We first calculate the Ellison-Glaeser index for each of 2-digit SIC industries in the manufacturing sector because the concentration data are available only for the manufacturing sector and the employment data are available only for each of 2-digit SIC industries. We then categorize industry-year observations into two groups (i.e., agglomerated and non-agglomerated industries) based on the sample median of the Ellison-Glaeser index. A dummy variable indicating agglomerated industries ($HighEG_{j,t-1}$) has a value of one if the Ellison-Glaeser index for industry j in year $t-1$ is greater than or equal to the sample median of the Ellison-Glaeser index and a value of zero otherwise.

4.3.2. Does agglomeration moderate the relationship between imitation and the quantity of innovation?

To investigate whether the moderating effect of agglomeration on the imitation-innovation relationship exists, we estimate zero-inflated negative binomial models with two additional interaction terms (i.e., $IMI_{j,t-1} \times HighEG_{j,t-1}$ and $IMI^2_{j,t-1} \times HighEG_{j,t-1}$) and an additional control variable (i.e., $HighEG_{j,t-1}$). Table 8 re-

[21]The economic census database contains HHI data for each of 4-digit SIC industries. We obtain the HHI for each of 2-digit SIC industries by value-weighting the HHIs for constituent 4-digit SIC industries. The database is updated every 5 years, so we use linear interpolation to obtain HHIs for the years between two census years. The state and area employment, hours, and earnings database is publicly available only for the period between 1990 and 2002.

Table 8: Moderating effect of agglomeration on the relationship between imitation and the quantity of innovation

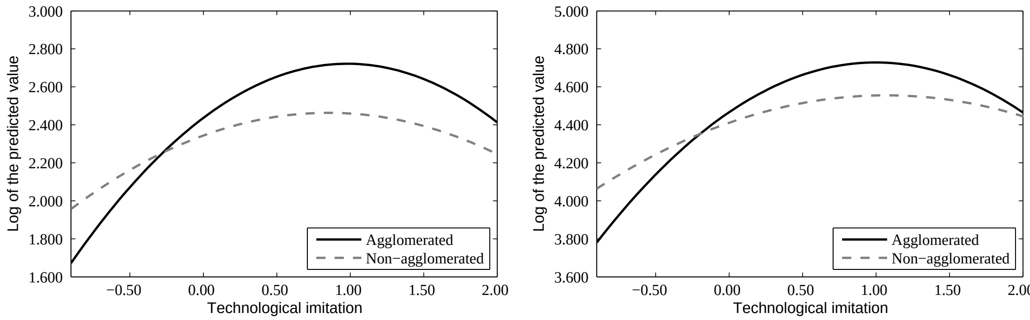
Dependent variable	(1) $\overline{COUNT}_{j,t}$	(2) $\overline{CITE}_{j,t}$
$IMI_{j,t-1}$	0.281*	0.271*
	(0.168)	(0.160)
$IMI^2_{j,t-1}$	-0.165**	-0.127*
	(0.071)	(0.069)
$IMI_{j,t-1} \times HighEG_{j,t-1}$	0.301*	0.254*
	(0.175)	(0.153)
$IMI^2_{j,t-1} \times HighEG_{j,t-1}$	-0.132*	-0.136*
	(0.077)	(0.073)
$HighEG_{j,t-1}$	-0.092	0.060
	(0.225)	(0.205)
$\overline{Size}_{j,t-1}$	0.412***	0.445***
	(0.090)	(0.077)
$\overline{ROA}_{j,t-1}$	0.425**	0.260
	(0.183)	(0.191)
$\overline{R\&D}_{j,t-1}$	8.070***	6.754***
	(1.809)	(2.057)
$\overline{PPE}_{j,t-1}$	-2.148**	-1.721
	(1.025)	(1.063)
$\overline{Lev}_{j,t-1}$	-0.314	-0.135
	(0.675)	(0.588)
$\overline{Capex}_{j,t-1}$	4.365	1.105
	(2.865)	(2.655)
$\overline{MB}_{j,t-1}$	0.005	0.010
	(0.021)	(0.020)
$\overline{Age}_{j,t-1}$	0.054***	0.065***
	(0.020)	(0.021)
$\overline{KZ}_{j,t-1}$	-0.008	-0.008
	(0.007)	(0.006)
Constant	1.947***	-1.080**
	(0.433)	(0.450)
Vuong z-stat	2.88***	1.31*
p-value	0.002	0.095
Year fixed effects	Yes	Yes
Industry fixed effects	Yes	Yes
Observations	1,408	1,408
Number of industries	140	140

Note: This table reports the second-stage results of zero-inflated negative binomial (ZiNB) regression analyses designed to estimate the moderating effect of agglomeration on the relationship between imitation and the quantity of innovation. "Vuong z-stat" is the test statistic for the Vuong test that compares the zero-inflated negative binomial model with industry dummies to the standard negative binomial model with industry dummies. Standard errors clustered by industry are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

ports the second-stage regression results of zero-inflated negative binomial regression analyses. The dependent variables are $\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$ in Columns (1) and (2), respectively. Note that Vuong test statistics are significantly greater than zero in both columns, suggesting that zero inflated negative binomial models have better fits than the corresponding standard negative binomial models. First, we find evidence that there are inverted-U relationships between imitation and innovation for both agglomerated and non-agglomerated industries.[22] Second, we find evidence that the moderating effect of agglomeration on the imitation-innovation relationship exists. The estimated regression coefficients for the additional interaction terms suggest that the impact of imitation on the quantity of innovation is significantly stronger for agglomerated industries than for non-agglomerated industries.

Figure 1 visually shows the relationship between the natural logarithm of the quantity of innovation and the degree of imitation for agglomerated and non-

[22]Note that turning points and their 95% confidence intervals for both agglomerated and non-agglomerated industries are within the data range. For agglomerated industries, $\overline{COUNT}_{j,t}$ peaks when $IMI_{j,t-1}$ has a value of 0.982 (86th percentile) with a 95% confidence interval of [0.708, 1.256] ([83rd percentile, 88th percentile]), while the $\overline{CITE}_{j,t}$ peaks when $IMI_{j,t-1}$ has a value of 0.995 (86th percentile) with a 95% confidence interval of [0.691, 1.299] ([82nd percentile, 89th percentile]). For non-agglomerated industries, $\overline{COUNT}_{j,t}$ peaks when $IMI_{j,t-1}$ has a value of 0.853 (85th percentile) with a 95% confidence interval of [0.327, 1.380] ([76th percentile, 90th percentile]), while $\overline{CITE}_{j,t}$ peaks when $IMI_{j,t-1}$ has a value of 1.065 (87th percentile) with a 95% confidence interval of [0.420, 1.709] ([78th percentile, 92nd percentile]). Note also that locations of turning points for agglomerated and non-agglomerated industries are quite similar.



(a) Using $\overline{COUNT}_{j,t}$ as the dependent variable (b) Using $\overline{CITE}_{j,t}$ as the dependent variable

Figure 1. The impact of technological imitation on the quantity of innovation: agglomerated industries *vs.* non-agglomerated industries
This figure plots the natural logarithm of the quantity of innovation against the degree of imitation for agglomerated and non-agglomerated industries. The variables presented on the vertical axis in Panels (a) and (b) are $ln(\overline{COUNT}_{j,t})$ and $ln(\overline{CITE}_{j,t})$, respectively. The vertical-axis values are the natural logarithms of the predicted values obtained using estimated regression coefficients of the zero-inflated negative binomial models presented in Table 8 and sample mean values of the control variables. The solid curves show the imitation-innovation relationship for agglomerated industries, while the dotted curves show the relationship for non-agglomerated industries. The figure is drawn for the range between the 5th and 95th percentiles of the imitation measure in the manufacturing sector sample.

agglomerated industries. The variables presented on the vertical axis in Panels (a) and (b) are $ln(\overline{COUNT}_{j,t})$ and $ln(\overline{CITE}_{j,t})$, respectively. The solid curves show the imitation-innovation relationship for agglomeration industries, while the dotted curves show the relationship for non-agglomeration industries. The solid curves are steeper on both sides of the peaks than dotted curves in both panels, suggesting that the impact of imitation on the quantity of innovation is stronger for agglomerated industries than for non-agglomerated industries. Easier access to competitors, suppliers, customers, and R&D collaborators in agglomerated industries enhances both the positive and negative impacts of technological imitation on the quantity of innovation.

4.3.3. Does agglomeration moderate the relationship between imitation and the market value of innovation?

To investigate whether the moderating effect of agglomeration on the relationship between imitation and the value of innovation exists, we estimate the fixed-effects regression models described by Equations (2) and (3) with three additional interaction terms (i.e., $INN_{i,t-1} \times HighEG_{j,t-1}$, $INN_{i,t-1} \times IMI_{j,t-1} \times HighEG_{j,t-1}$ and $INN_{i,t-1} \times IMI^2_{j,t-1} \times HighEG_{j,t-1}$) and an additional control variable (i.e., $HighEG_{j,t-1}$). Table 9 reports the fixed-effects regression results. We use three different dependent variables, as shown in Table 5. Columns (1) through (3) are based on $INN1_{i,t-1}$ as a firm-level innovation measure, while Columns (4) through (6) are based on $INN2_{i,t-1}$. First, we find evidence that there are inverted U-shaped relationships between imitation and the value of innovation for both agglomerated and non-agglomerated industries.[23] Second, we find that the moderating effect of agglomeration on the relationship between im-

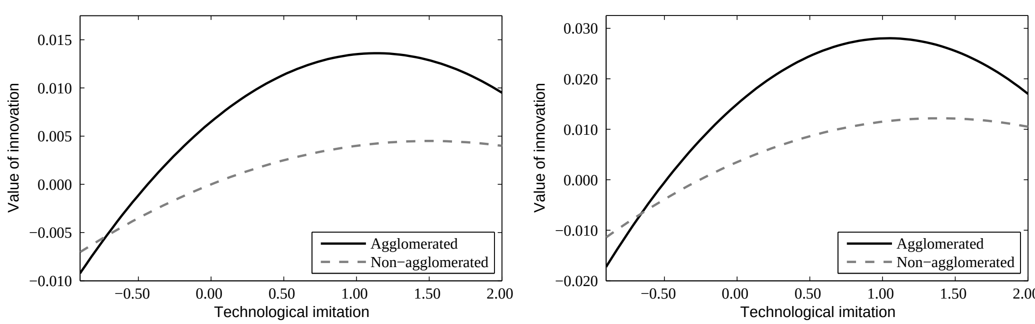
[23]Note that, regardless of the choice of the dependent variable and the innovation measure, all turning points and their 95% confidence intervals (except one upper bound only) are within the data range. Note also that the locations of turning points for agglomerated and non-agglomerated industries are quite similar.

Table 9: Moderating effect of agglomeration on the relationship between imitation and the value of innovation

Dependent variable	(1) $r_{i,t}$	(2) $r_{i,t} - R_{p,t}$	(3) $r_{i,t} - R_{j,t}$	(4) $r_{i,t}$	(5) $r_{i,t} - R_{p,t}$	(6) $r_{i,t} - R_{j,t}$
$INN1_{i,t-1}$	0.027***	0.025**	0.015			
	(0.009)	(0.010)	(0.009)			
$INN1_{i,t-1} \times IMI_{j,t-1}$	0.025***	0.026***	0.021***			
	(0.008)	(0.008)	(0.007)			
$INN1_{i,t-1} \times IMI^2_{j,t-1}$	-0.009***	-0.007**	-0.007**			
	(0.003)	(0.003)	(0.003)			
$INN1_{i,t-1} \times HighEG_{j,t-1}$	0.023***	0.013*	0.020***			
	(0.007)	(0.007)	(0.007)			
$INN1_{i,t-1} \times IMI_{j,t-1} \times HighEG_{j,t-1}$	0.025***	0.030***	0.019**			
	(0.009)	(0.010)	(0.009)			
$INN1_{i,t-1} \times IMI^2_{j,t-1} \times HighEG_{j,t-1}$	-0.015***	-0.014***	-0.008**			
	(0.004)	(0.005)	(0.004)			
$INN2_{i,t-1}$				0.010**	0.008*	0.006
				(0.004)	(0.004)	(0.004)
$INN2_{i,t-1} \times IMI_{j,t-1}$				0.13***	0.014***	0.12***
				(0.004)	(0.004)	(0.004)
$INN2_{i,t-1} \times IMI^2_{j,t-1}$				-0.004**	-0.003*	-0.004**
				(0.002)	(0.002)	(0.002)
$INN2_{i,t-1} \times HighEG_{j,t-1}$				0.012***	0.008**	0.009***
				(0.003)	(0.004)	(0.003)
$INN2_{i,t-1} \times IMI_{j,t-1} \times HighEG_{j,t-1}$				0.015***	0.016***	0.011**
				(0.005)	(0.006)	(0.005)
$INN2_{i,t-1} \times IMI^2_{j,t-1} \times HighEG_{j,t-1}$				-0.007***	-0.006**	-0.004
				(0.003)	(0.003)	(0.002)
$HighEG_{j,t-1}$	0.016	0.023	-0.034*	0.008	0.017	-0.034*
	(0.022)	(0.022)	(0.020)	(0.022)	(0.022)	(0.020)
$IMI_{j,t-1}$	0.020**	0.015*	-0.005	0.018**	0.013	-0.006
	(0.008)	(0.009)	(0.007)	(0.009)	(0.009)	(0.008)
Constant	1.218***	1.633***	0.902***	1.207***	1.641***	0.920***
	(0.103)	(0.105)	(0.091)	(0.102)	(0.104)	(0.090)
Firm-level control variables	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22,458	22,458	22,458	22,458	22,458	22,458
Adjusted R-squared	0.281	0.220	0.189	0.281	0.220	0.188

Note: This table reports the results of regression analyses designed to estimate the moderating effect of agglomeration on the relationship between imitation and the value of innovation. Raw stock returns ($r_{i,t}$) or excess returns ($r_{i,t} - R_{p,t}$ or $r_{i,t} - R_{j,t}$) are used as the dependent variable. All the regression models are estimated using the within-groups (i.e., fixed-effects) estimator. Standard errors clustered by firm are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

itation and the value of innovation exists. Note that in Columns (1) through (5), the coefficients of $INN_{i,t-1} \times IMI_{j,t-1} \times HighEG_{j,t-1}$ and $INN_{i,t-1} \times IMI^2_{j,t-1} \times HighEG_{j,t-1}$ are significantly positive and negative at the 5% or 1% level, respectively. In Column (6), the coefficient of $INN_{i,t-1} \times IMI_{j,t-1} \times HighEG_{j,t-1}$ is significantly positive at the 5% level but the coefficient of $INN_{i,t-1} \times IMI^2_{j,t-1} \times HighEG_{j,t-1}$ is only marginally insignificant at the 10% level (t-stat=-1.64; p-value=0.102). Thus, the results suggest that the impact of imitation on the value of innovation is significantly stronger for agglomerated industries than for non-agglomerated industries.



(a) Using $INN1_{i,t-1}$ as an innovation measure (b) Using $INN2_{i,t-1}$ as an innovation measure

Figure 2. The impact of technological imitation on the value of innovation: agglomerated industries *vs.* non-agglomerated industries

This figure plots the value of innovation against the degree of imitation for agglomerated and non-agglomerated industries. The value of innovation is measured as the sensitivity of raw stock returns to an innovation measure ($INN1_{i,t-1}$ or $INN2_{i,t-1}$). The innovation measures used in Panels (a) and (b) are $INN1_{i,t-1}$ and $INN2_{i,t-1}$, respectively. The solid curves show the relationship for agglomerated industries, while the dotted curves show the relationship for non-agglomerated industries. The figure is drawn for the range between the 5th and 95th percentiles of the imitation measure in the manufacturing sector sample.

Figure 2 shows the relationship between the value of innovation (β_1) and the degree of imitation ($IMI_{j,t-1}$) for agglomerated and non-agglomerated industries. The solid curves show the relationship for agglomerated industries, while the dotted curves show the relationship for non-agglomerated industries. The solid curves are steeper on both sides of the peaks than the dotted curves in both panels, suggesting that the impact of imitation on the value of innovation is stronger for agglomerated industries than for non-agglomerated industries. Easier access to competitors, suppliers, customers, and R&D collaborators in agglomerated industries enhances both the positive and negative impacts of technological imitation on the value of innovation.

Overall, these results suggest that creating innovation clusters and allowing different innovators to cooperate, imitate and compete with each other would be very effective in promoting corporate innovation. Firms in an innovation cluster are located close to the supplier, customer, and R&D collaborator. Thus, firms in a cluster would be able to learn from and imitate their competitors' innovations much easier than those firms not in a cluster. However, a level of technological imitation that is too high is more detrimental for firms in an innovation cluster, as it lowers their incentives to innovate more significantly.

5. Conclusion

This study examines the relationship between technological imitation and firms' innovation activities and their incentives to innovate using US firm-level patent data for the period 1977–2005. The findings reveal inverted U-shaped relationships between technological imitation and industry-average innovation activities and between technological imitation and the market value of firm-level innovation. The results are driven by the trade-off of two different effects. The first effect is the positive externalities of the interactions among firms during the process of technological innovation. Particularly when innovation is sequential and complementary, interactions among innovative firms can enhance the firms' innovation activities and incentives to innovate. The second effect is the negative effect of free-riding problems on firms' innovation activities and their incentives to innovate. This effect may be quite significant when innovation outcomes can be easily extended or imitated by competing firms and imitators can extract significant parts of the benefits that would have been enjoyed by the original innovators. Our results suggest that the first effect outweighs the second effect up to a high level of technological imitation, while the second effect outweighs the first effect when the level of technological imitation is extremely high. The positive effect of a moderate level of technological imitation and the negative effect of an excessive level of technological imitation are more pronounced for agglomerated industries. This finding suggests that creating innovation clusters such as Silicon Valley in the United States and Shenzhen City in China and allowing different innovators to cooperate, imitate and compete with each other would be very effective in promoting corporate innovation. However, an excessively high level of technological imitation is more detrimental for firms in innovation clusters because it lowers firms' incentives for technological innovation more radically. Because our imitation measure is computed based on patent data, this study is limited to the role of patent-based imitation. Future studies could investigate the role of imitation in the innovation of product designs and business strategies.

Acknowledgements

The authors are grateful to anonymous referees for their insightful comments on the earlier version of the manuscript. We are grateful to Bronwyn Hall, Adam Jaffe, and Manuel Trajtenberg for making the NBER US patent citation data file available on the NBER website. We also thank Gege Zhao and Yang Liu for their research assistance. The research has been partially supported by the Peking University HSBC Business School Annual Research Fund and the Bridge Trust Asset Management Research Fund.

References

Adams, J. D., Jaffe, A. B., 1996. Bounding the effects of R&D: an investigation using matched establishment-firm data. Tech. rep., National Bureau of Economic Research.

Aghion, P., Bloom, N., Blundell, R., Griffith, R., Howitt, P., 2005. Competition and innovation: An inverted-U relationship. The Quarterly Journal of Economics 120 (2), 701–728.

Aghion, P., Harris, C., Howitt, P., Vickers, J., 2001. Competition, imitation and growth with step-by-step innovation. The Review of Economic Studies 68 (3), 467–492.

Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. Econometrica (1986-1998) 60 (2), 323.

Amara, N., Landry, R., Traoré, N., 2008. Managing the protection of innovations in knowledge-intensive business services. Research Policy 37 (9), 1530–1547.

Arrow, K. J., 1972. Economic welfare and the allocation of resources for invention. In: Readings in Industrial Economics. Springer, pp. 219–236.

Atanassov, J., 2013. Do hostile takeovers stifle innovation? Evidence from antitakeover legislation and corporate patenting. The Journal of Finance 68 (3), 1097–1131.

Baptista, R., Swann, P., 1998. Do firms in clusters innovate more? Research Policy 27 (5), 525–540.

Bessen, J., Maskin, E., 2009. Sequential innovation, patents, and imitation. The RAND Journal of Economics 40 (4), 611–635.

Bhattacharya, U., Hsu, P.-H., Tian, X., Xu, Y., 2017. What affects innovation more: Policy or policy uncertainty? Journal of Financial and Quantitative Analysis 52 (5), 1869–1901.

Brozen, Y., 1951. Invention, innovation, and imitation. The American Economic Review 41 (2), 239–257.

Carlino, G., Kerr, W. R., 2015. Agglomeration and innovation. In: Handbook of Regional and Urban Economics. Vol. 5. Elsevier, pp. 349–404.

Casadesus-Masanell, R., Zhu, F., 2013. Business model innovation and competitive imitation: The case of sponsor-based business models. Strategic Management Journal 34 (4), 464–482.

Chatterji, A., Glaeser, E., Kerr, W., 2014. Clusters of entrepreneurship and innovation. Innovation Policy and the Economy 14 (1), 129–166.

Conley, T. G., Flyer, F., Tsiang, G. R., 2003. Spillovers from local market human capital and the spatial distribution of productivity in Malaysia. Advances in Economic Analysis & Policy 3 (1).

Czarnitzki, D., Kraft, K., 2011. Spillovers of innovation activities and their profitability. Oxford Economic Papers 64 (2), 302–322.

Dittmar, A., Mahrt-Smith, J., 2007. Corporate governance and the value of cash holdings. Journal of Financial Economics 83 (3), 599–634.

Dobson, W., Safarian, A. E., 2008. The transition from imitation to innovation: An enquiry into China's evolving institutions and firm capabilities. Journal of Asian Economics 19 (4), 301–311.

Ethiraj, S. K., Levinthal, D., Roy, R. R., 2008. The dual role of modularity: innovation and imitation. Management Science 54 (5), 939–955.

Fang, V. W., Tian, X., Tice, S., 2014. Does stock liquidity enhance or impede firm innovation? The Journal of Finance 69 (5), 2085–2125.

Faulkender, M., Wang, R., 2006. Corporate financial policy and the value of cash. The Journal of Finance 61 (4), 1957–1990.

Greenhalgh, C., Rogers, M., 2006. The value of innovation: The interaction of competition, R&D and IP. Research Policy 35 (4), 562–580.

Grossman, G., Helpman, E., 1991. Innovation and growth in the world economy.

Haans, R. F., Pieters, C., He, Z.-L., 2016. Thinking about U: Theorizing and testing U-and inverted U-shaped relationships in strategy research. Strategic Management Journal 37 (7), 1177–1195.

Hadlock, C. J., Pierce, J. R., 2010. New evidence on measuring financial constraints: Moving beyond the KZ index. The Review of Financial Studies 23 (5), 1909–1940.

Hall, B. H., Jaffe, A., Trajtenberg, M., 2005. Market value and patent citations. RAND Journal of Economics, 16–38.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2001. The NBER patent citation data file: Lessons, insights and methodological tools. Tech. rep., National Bureau of Economic Research.

Hopenhayn, H., Llobet, G., Mitchell, M., 2006. Rewarding sequential innovators: Prizes, patents, and buyouts. Journal of Political Economy 114 (6), 1041–1068.

Huang, J., Rozelle, S., 1996. Technological change: Rediscovering the engine of productivity growth in China's rural economy. Journal of Development Economics 49 (2), 337–369.

Hwang, H., Wu, J. Z., Yu, E. S., 2016. Innovation, imitation and intellectual property rights in developing countries. Review of Development Economics 20 (1), 138–151.

Im, H. J., Park, Y. J., Shon, J., 2015. Product market competition and the value of innovation: Evidence from US patent data. Economics Letters 137, 78–82.

Jaffe, A. B., Trajtenberg, M., Fogarty, M. S., 2000. Knowledge spillovers and patent citations: Evidence from a survey of inventors. American Economic Review 90 (2), 215–218.

Jaffe, A. B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics 108 (3), 577–598.

Kaplan, S. N., Zingales, L., 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? The Quarterly Journal of Economics 112 (1), 169–215.

Kim, L., 1997. Immitation to innovation; The dynamics of Korea's technological learning. Harvard Business School Press.

Kim, L., Nelson, R. R., 2000. Technology, learning, and innovation: Experiences of newly industrializing economies. Cambridge University Press.

Lind, J. T., Mehlum, H., 2010. With or without U? The appropriate test for a U-shaped relationship. Oxford Bulletin of Economics and Statistics 72 (1), 109–118.

Marquaridt, D. W., 1970. Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. Technometrics 12 (3), 591–612.

Moretti, E., 2004. Human capital externalities in cities. In: Handbook of Regional and Urban Economics. Vol. 4. Elsevier, pp. 2243–2291.

Mukherjee, A., Singh, M., Žaldokas, A., 2017. Do corporate taxes hinder innovation? Journal of Financial Economics 124 (1), 195–221.

O'hara, R. B., Kotze, D. J., 2010. Do not log-transform count data. Methods in Ecology and Evolution 1 (2), 118–122.

Pavitt, K., 1984. Sectoral patterns of technical change: Towards a taxonomy and a theory. Research Policy 13 (6), 343–373.

Porter, M., 1980. Corporate strategy. New York. New York, NY.

Rivkin, J. W., 2000. Imitation of complex strategies. Management Science 46 (6), 824–844.

Roper, S., Vahter, P., Love, J. H., 2013. Externalities of openness in innovation. Research Policy 42 (9), 1544–1554.

Rumelt, R. P., Lamb, R., 1984. Competitive strategic management. Toward a Strategic Theory of the Firm, 556–570.

Schoonhoven, C. B., Eisenhardt, K. M., Lyman, K., 1990. Speeding products to market: Waiting time to first product introduction in new firms. Administrative Science Quarterly, 177–207.

Thompson, P., 2006. Patent citations and the geography of knowledge spillovers: evidence from inventor-and examiner-added citations. The Review of Economics and Statistics 88 (2), 383–388.

Tidd, J., Bessant, J., Pavitt, K., 2005. Managing innovation integrating technological, market and organizational change. John Wiley and Sons Ltd.

Tirole, J., 1988. The theory of industrial organization. MIT Press.

Tödtling, F., Trippl, M., 2005. One size fits all?: Towards a differentiated regional innovation policy approach. Research Policy 34 (8), 1203–1219.

Zeng, J., 2001. Innovative vs. imitative R&D and economic growth. Journal of Development Economics 64 (2), 499–528.

Appendix A. Definition of industry-level variables

The following table shows the definitions of the industry-level variables used in Table 2, Table 4, and Table 8. The italicized codes in brackets([]) represent item codes in CRSP/Compustat Merged Database. All control variables are measured in year $t-1$.

Variable	Definition
Dependent variables	
$\overline{COUNT}_{j,t}$	Industry-average number of the patents applied for in year t by any firms in industry j
$\overline{CITE}_{j,t}$	Industry-average number of citations of the patents applied for in year t by any firms in industry j
$ln(1+\overline{COUNT}_{j,t})$	The natural logarithm of 1 plus $\overline{COUNT}_{j,t}$
$ln(1+\overline{CITE}_{j,t})$	The natural logarithm of 1 plus $\overline{CITE}_{j,t}$
Imitation-related variables	
$IMI_{j,t-1}$	Technological imitation for industry j in year $t-1$, defined as the average ratio of the citations made by industry peers within five years after the application of the patents to the number of citations of the patents that any firms in industry j applied for in year $t-1$
$IMI^2_{j,t-1}$	The square of the technological imitation measure
Control variables	
$\overline{Size}_{j,t-1}$	Industry-average value of firm size (*Size*) where *Size* is measured as the natural logarithm of market value of total assets ($[prcc\_f] \times [cshpri] + [pstkl] + [dlc] + [dltt] - [txditc]$)
$\overline{ROA}_{j,t-1}$	Industry-average value of return on assets (*ROA*) where *ROA* is measured as the ratio of operating income before depreciation ($[oibdp]$) to book value of the total assets ($[at]$)
$\overline{R\&D}_{j,t-1}$	Industry-average value of R&D intensity (*R&D*) where *R&D* is measured as the ratio of R&D expenditures ($[xrd]$) to book value of total assets ($[at]$)
$\overline{PPE}_{j,t-1}$	Industry-average value of asset tangibility (*PPE*) where *PPE* is measured as the ratio of net property, plant and equipment ($[ppent]$) to book value of total assets($[at]$)
$\overline{Lev}_{j,t-1}$	Industry-average value of market leverage ratio (*Lev*) where *Lev* is measured as the ratio of total debt ($[dlc]+[dltt]$) to market value of total assets ($[prcc\_f] \times [cshpri] + [pstkl] + [dlc] + [dltt] - [txditc]$)
$\overline{Capex}_{j,t-1}$	Industry-average value of investment rate (*Capex*) where *Capex* is measured as the ratio of capital expenditures ($[capx]$) to book value of total assets ($[at]$)
$\overline{MB}_{j,t-1}$	Industry-average value of market-to-book ratio (*MB*) where *MB* is measured as the ratio of market value of total assets ($[prcc\_f] \times [cshpri] + [pstkl] + [dlc] + [dltt] - [txditc]$) to book value of total assets ($[at]$)
$\overline{Age}_{j,t-1}$	Industry-average value of firm age (*Age*) where *Age* is defined as the number of years preceding the observation year that the firm has a non-missing stock price on the Compustat file and winsorized at 37 years following Hadlock and Pierce (2010)
$\overline{KZ}_{j,t-1}$	Industry-average value of Kaplan-Zingales (*KZ*) Index where *KZ* index is defined as $-1.002 \times CashFlow + 0.283 \times Q + 3.139 \times Leverage - 39.368 \times Dividends - 1.315 \times CashHoldings$, where each component is defined in line with Kaplan and Zingales (1997)

Appendix B. Definition of firm-level variables

The following table shows the definitions of the firm-level variables used in Table 5, Table 7, and Table 9. The italicized codes in brackets([]) represent item codes in CRSP/Compustat Merged Database. In line with Im et al. (2015), Faulkender and Wang (2006), and Dittmar and Mahrt-Smith (2007), some control variables are measured in year $t-1$ and other control variables in year t.

Variable	Definition
Dependent variables	
$r_{i,t}$	Firm i's annual stock returns in year t
$r_{i,t} - R_{p,t}$	Firm i's annual stock returns in year t in excess of annual returns to the 5×5 Fama and French portfolios formed on "Size" and "Book-to-Market"
$r_{i,t} - R_{j,t}$	Firm i's annual stock returns in year t in excess of industry j's annual stock returns, where firm i belongs to industry j in year t
Firm-level innovation measures	
$INN1_{i,t-1}$	Natural logarithm of 1 plus the number of patents that firm i applied for in year $t-1$
$INN2_{i,t-1}$	Natural logarithm of 1 plus the number of citations of the patents that firm i applied for in year $t-1$
Control variables	
$\Delta Earnings_{i,t}$	Ratio of change in earnings ([*ebit*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$\Delta Assets_{i,t}$	Ratio of change in total assets ([*at*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$\Delta R\&D_{i,t}$	Ratio of change in R&D expenditures ([*xrd*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$\Delta Dividends_{i,t}$	Ratio of change in dividends ([*dvc*] + [*dvp*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$LnTA_{i,t-1}$	Natural logarithm of book total assets ([*at*])
$Leverage_{i,t-1}$	Ratio of total debt ([*dlc*] + [*dltt*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$MB_{i,t-1}$	Ratio of market value of total assets ([*prcc_f*] $\times$ [*cshpri*] + [*pstkl*] + [*dlc*] + [*dltt*] $-$ [*txditc*]) to book value of the total assets ([*at*])
$Financing_{i,t}$	Ratio of new financing ([*dltis*] $-$ [*dltr*] + [*sstk*] $-$ [*prstkc*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$\Delta Interests_{i,t}$	Ratio of interest expenditures ([*xint*]) to market capitalization ([*prcc_f*] $\times$ [*cshpri*]) at the previous fiscal end
$Age_{i,t}$	Firm age which is defined as the number of years preceding the observation year that the firm has a non-missing stock price on the Compustat file and winsorized at 37 years following Hadlock and Pierce (2010)