Working Paper No. 2017002

Peking University HSBC Business School

PHBS

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

Hyun Joong Im

Yang Liu

Janghoon Shon

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

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

[a] *HSBC Business School, 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

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 dominate the negative effects of free-riding concerns on firms' innovation activities and incentives to innovate up to quite a high degree of technological imitation, while free-riding concerns dominate the positive externalities when the level of technological imitation is extremely high. Thus, creating innovation clusters and allowing different innovators to cooperate, imitate and compete with each other would be very effective in promoting corporate innovation.

Keywords: Corporate innovation, Technological imitation, Value of innovation, Clustering

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, Eisenhardt, and

*Corresponding author. Tel: +852 6756 7630; E-mail address: jshon@connect.ust.hk

Lyman, 1990), as well as bringing 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, Tian, and Tice, 2014), corporate taxes (Mukherjee, Singh, and Zaldokas, 2016), policy uncertainty (Bhattacharya, Hsu, Tian, and Xu, 2015), and product market competition (Aghion, Bloom, Blundell, Griffith, and Howitt, 2005; Greenhalgh and Rogers, 2006; Im, Park, and Shon, 2015) have been studied in the literature. In this paper, we investigate whether the degree of industry-level technological imitation decreases or increases 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 against 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 dominate 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 dominate the positive externalities when the level of technological imitation is extremely high. In this spirit, Aghion, Harris, Howitt, and Vickers (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, Harris, Howitt, and Vickers (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 upto the 85th 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 dominate the negative effects of free-riding concerns on firms' innovation activities up to quite a high degree of technological imitation, while free-riding concerns dominate the positive externalities when the level of technological imitation is extremely high. In addition, we repeat the analysis for each Pavitt technological sector in order 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 (around the 85th percentile), although Pavitt 4 has a peak point at a slightly lower imitation level (at the 79th percentile). 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 by Im, Park, and Shon (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 upto the 81st to 83rd 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 dominate the negative effects of free-riding concerns on firms' incentives to innovate up to quite high degrees of technological imitation; however, free-riding concerns dominate 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 the agglomeration and non-agglomeration industries. This study finds that the impacts of 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

[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*

high level of technological imitation is more detrimental for firms in innovation clusters because it lowers those firms' incentives to innovate more radically.

The remainder of the paper is organized as follows. In Section 2, we first derive predictions regarding the effects of technological imitation on the quantity and value of innovation based on the existing theoretical and empirical literature. Section 3 describes the sample, the measurement of variables, and the descriptive statistics. In Section 4, we present our main results. Section 5 concludes the paper.

2. Predictions regarding the effects of technological imitation on the quantity and value of innovation

Some prior studies have examined the relationship between technological imitation and corporate innovation, but their theoretical predictions and empirical findings have not yet reached consensus. The first view is that technological imitation has a *positive effect* on the quantity and value of innovation due to the positive externalities during the process of innovation. Among others, Bessen and Maskin (2009) argued 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

up year on year, according to the municipal government."

imitation will ironically enhance firm incentives to innovate and thus increase the quantity of innovation due to the positive externalities in the innovation process. When innovation is sequential and complementary in one industry, imitation activities bring learning opportunities to followers. After an innovation outcome is achieved by a predecessor (i.e., a patent is granted), imitation activities will allow other firms to turn more creative (and more valuable) ideas related to the earlier innovation into successive innovations (i.e., other firms will apply for new patents after quickly citing the predecessor's patent and developing it further). In this way, the quantity and market value of future innovations will increase.

The second view is that technological imitation has a *negative effect* on the value and the quantity of corporate innovation due to free-riding concerns. Namely, Zeng (2001) argued that under the assumption of independently researched innovations, technological imitation may decrease the value of a firm's innovation activities and in turn the quantity of innovation because of the threat of being imitated by followers (i.e., free-riding concerns). One could imagine that when imitation activities are prevalent in the industry where firms conduct innovation/invention independently, the value of a firm's innovation outcomes will be reduced because of the risk of being copied and surpassed by followers (i.e., competitors in the same industry). Thus, the motivation to innovate and, in turn, the quantity of innovation will decrease.

However, it is likely that the two effects (i.e., positive externalities and technological free riding) coexist, in which case an *inverted-U-shaped relationship* is predicted. Positive externalities from the interactions among firms during the process of innovation dominate the negative effects of free-riding concerns on firms' innovation activities and incentives to innovate up to a high degree of technological imitation, whereas when the level of technological imitation is extremely high, free-riding concerns dominate the positive externalities. In this spirit, Aghion, Harris, Howitt, and Vickers (2001) argued that a small amount of imitation almost always enhances growth, as it promotes more frequent close competition, whereas extremely high imitation unambiguously slows growth due to free-riding problems.

In certain industries that are in the early stage of technological imitation, a relatively low degree of imitation actually provides a better environment for innovation. It makes it easier for companies in this industry to learn from each other's innovation achievements and create new innovation outcomes. In this case, the quality or value of innovation (measured by the increase in the market value of equity driven by a one-unit increase in innovation) in the industry increases as the degree of imitation rises, since the positive effect from technological externalities is greater than the negative effect from free riding concerns. In turn, the quantity of innovation (measured by the number of patents or citations) in the industry also increases with the degree of technological imitation in the industry.

When the imitation level in a certain industry is extremely high, however, the negative effect from free riding concerns dominates the positive effect from technological externalities, and thus, the total effect may be negative. An innovation outcome can be quickly learnt or copied by competitors in such an industry, so companies have less motivation to engage in innovation activities. Thus, a newly granted patent is much less valuable to companies in the industry than to those in an industry with a less intensive degree of imitation. In turn, the quantity of innovation decreases with the degree of technological imitation in the industry.

Different theoretical models based on different assumptions and model settings have different and sometimes conflicting 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.

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, Jaffe, and Trajtenberg (2001, 2005). We exclude observations before 1977 and in 2006 to further mitigate concerns arising from truncations. Thus, our patent dataset covers all patents applied for during the period 1977-2005.

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. Variable construction

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

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

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, $IMI_{j,t}$, we use the *industry-average competitor–quick citation ratio* standardized based on the within-industry mean and within-industry standard deviation, where the industry-average competitor–quick citation ratio is defined as the industry average of the ratio of *Competitors' citations received within 5 years for the patents that any firms in industry j applied for in year t* to the *Total number of 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, $IMI_{j,t} = 0$ 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, $IMI_{j,t} = 0.5$ means that the patents applied for in year t by any

[3]We are the first to use the (standardized) 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.

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.

However, our measure might have some potential problems. First, an imitation could take place without patent citations. For example, competitors could adopt similar functions or designs without citing patents. This could happen quite often in the case of production innovation. However, our focus is more on technological innovation, which 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. Therefore, we assume that technological imitation often involves the citation of competitors' patents. Second, 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 our 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.

All variables are winsorized at the 1st and 99th percentiles, and their definitions are reported in Appendices A and B. Table 1 reports the summary statistics for those variables. Panel A is related to the industry-level analysis concerning the effect of imitation on corporate 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).

4. Empirical models and results

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

4.1.1. Full-sample analyses

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

$$\begin{aligned} y_{j,t} &= \beta_0 + \beta_1 IMI_{j,t-1} + \beta_2 IMI^2_{j,t-1} + \beta_{Controls} Controls \\ &\quad + Industry\ FE + Year\ FE + \varepsilon_{j,t}, \end{aligned} \tag{1}$$

Table 1: Summary statistics

Panel A. Industry-average 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	2.191	0.384	1.113	1.925	2.173	2.434	3.383
$\overline{KZ}_{j,t-1}$	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-average variables used in Table 2.

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

Note: This table shows summary statistics for the variables used in Table 5.

where $y_{j,t}$ is an industry-average 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, assets tangibility, leverage, investment, market-to-book ratio, 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, Tian, and Tice (2014). Two industry-average innovation measures, i.e., $\overline{COUNT}_{j,t}$ 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, Tian, and Tice (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 corporate innovation measures. As we include a squared 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 2.38 (i.e., much smaller than 10), so multicollinearity does not appear to be a serious issue.[4] 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 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 corporate innovation measures (i.e., $\overline{COUNT}_{j,t}$ or $\overline{CITE}_{j,t}$) as in Aghion, Bloom, Blundell, Griffith, and Howitt (2005). Columns (2) and (6) present the results from Poisson regressions with fixed effects. Regardless of the choice of the dependent variable, we find an inverted-U-shaped relationship be-

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

tween technological imitation and industry-average corporate innovation. 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 regressions with fixed effects. 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, 54,000 (830,000) with an associated p-value of 0.000 (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 an inverted-U-shaped relationship between technological imitation and industry-average corporate 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. Our investigation confirms that $\overline{COUNT}_{j,t}$ ($\overline{CITE}_{j,t}$) has 470 (669) zeros among 6,400 industry-year observations. In such a case, 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. *Prop. of R&D firms* has significantly negative signs in the logistic models predicting membership in the "certain zero" group, suggesting that the higher the proportion of R&D firms, the less likely it is that the industry has a certain zero. A Vuong test compares a zero-inflated negative binomial model to a corresponding standard negative binomial model. Because the z-values are significant at the 1% level, Vuong tests suggest that the zero-inflated negative binomial models have better fits than the corresponding standard negative binomial models. Once again, we find an inverted-U-shaped relationship between technological imitation and the quantity of corporate innovation measured by $\overline{COUNT}_{j,t}$ and $\overline{CITE}_{j,t}$.

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 0.957, while the industry-average number of citations ($\overline{CITE}_{j,t}$) peaks when $IMI_{j,t-1}$ has a value of 0.881.[5] An investigation of the distribution of our imitation measure suggests that they correspond to approximately the 85th

[5]The standardized imitation level corresponding to the peak points are 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 regression coefficient of $IMI^2_{j,t-1}$.

percentile.[6] Thus, an increase in technological imitation leads to an increase in the quantity of innovation up to the 85th 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 dominate the negative effects of free-riding concerns on firms' innovation activities up to a high degree of technological imitation. By contrast, free-riding concerns dominate 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 corporate 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) 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

[6] The summary statistics for $IMI_{j,t-1}$ reported in Table 1 suggest that the sample mean (median) is -0.006 (-0.353) and that the first (third) quartile is -0.588 (0.247).

Table 2: Effects of technological imitation on the quantity of corporate 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
$IMI_{j,t-1}$	0.471***	0.364***	0.373***	0.337***	0.654***	0.443***	0.393***	0.365 ***	2.18
	(0.038)	(0.052)	(0.020)	(0.046)	(0.052)	(0.058)	(0.019)	(0.061)	
$IMI^2_{j,t-1}$	-0.183***	-0.131***	-0.148***	-0.176***	-0.254***	-0.166***	-0.162***	-0.211 ***	2.38
	(0.015)	(0.022)	(0.009)	(0.022)	(0.022)	(0.025)	(0.010)	(0.028)	
$\overline{Size}_{j,t-1}$	0.108***	0.088***	0.119***	0.641***	0.139***	0.052	0.111***	0.561 ***	1.39
	(0.032)	(0.033)	(0.012)	(0.018)	(0.045)	(0.048)	(0.011)	(0.023)	
$\overline{ROA}_{j,t-1}$	-0.030	0.097*	0.035	0.206***	-0.005	0.114*	0.088**	0.245 **	2.35
	(0.043)	(0.054)	(0.031)	(0.078)	(0.076)	(0.062)	(0.035)	(0.103)	
$\overline{R\&D}_{j,t-1}$	1.935***	-0.019	0.223	9.467***	2.169***	0.187	1.671***	10.743 ***	1.52
	(0.571)	(0.586)	(0.205)	(0.561)	(0.702)	(0.563)	(0.205)	(0.694)	
$\overline{PPE}_{j,t-1}$	-0.680	-0.571	0.012	-2.439***	-0.956	-0.149	0.122	-2.569 ***	2.19
	(0.414)	(0.536)	(0.195)	(0.262)	(0.622)	(0.510)	(0.165)	(0.321)	
$\overline{Lev}_{j,t-1}$	-0.464**	-0.475*	-0.427***	-0.543***	-0.783**	-0.449**	-0.494***	-0.236	1.52
	(0.205)	(0.247)	(0.106)	(0.238)	(0.303)	(0.193)	(0.104)	(0.323)	
$\overline{Capex}_{j,t-1}$	-1.363*	-0.532	-0.877**	1.581	-0.931	-0.869	-0.635	4.208 ***	1.96
	(0.762)	(0.777)	(0.438)	(1.021)	(1.227)	(0.829)	(0.423)	(1.274)	
$\overline{MB}_{j,t-1}$	0.001	0.005	0.002	0.000	0.001	-0.001	0.003	-0.006	2.13
	(0.003)	(0.004)	(0.003)	(0.006)	(0.006)	(0.005)	(0.003)	(0.008)	
$\overline{Age}_{j,t-1}$	-0.267***	0.266***	0.365***	0.900***	-0.295**	0.248**	0.318***	0.946 ***	2.41
	(0.088)	(0.083)	(0.042)	(0.100)	(0.134)	(0.097)	(0.042)	(0.118)	
$\overline{KZ}_{j,t-1}$	0.000	-0.001	0.000	0.001	0.001	-0.000	0.001	0.002	1.13
	(0.001)	(0.002)	(0.001)	(0.003)	(0.002)	(0.002)	(0.001)	(0.004)	
Constant	3.795***		1.247***	-1.585***	6.092***		0.104	0.856***	
	(0.224)		(0.107)	(0.192)	(0.327)		(0.095)	(0.243)	
Logistic model predicting the membership in the "certain zero" group									
Prop. of R&D firms				-1.258***				-14.90***	
				(0.124)				(1.970)	
Constant				-27.32***				-1.312***	
				(0.080)				(0.180)	

Table 2: (Continued)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
LR test of $\alpha = 0$			5.4E+04				8.3E+05		
p-value			0.000				0.000		
Vuong z-stat				3.88				11.37	
p-value				0.000				0.000	
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	
Industry fixed effects	Yes	Yes	Yes	No	Yes	Yes	Yes	No	
Observations	6,400	6,388	6,388	6,400	6,400	6,378	6,378	6,400	
R-squared	0.386				0.491				
Number of industries	296	291	291	296	296	289	289	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). "Poisson" stands for a Poisson regression model, "NB" stands for a negative binomial regression model, "ZiNB" stands for a zero-inflated negative binomial regression model, and "FE" stands for a fixed-effects regression model. 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 a 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.

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
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%
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%
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%
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%
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%
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%

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, Bessant, and Pavitt (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 three columns are based our sample.

"science based" (Pavitt 4). Tidd, Bessant, and Pavitt (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 Pavitt technology sectors and provides some summary statistics. The summary statistics of for the three innovation measures show that innovation patterns are very heterogenous across 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 1 and 5 tend to have significantly lower levels of innovation based on the three measures.

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 the five Pavitt sectors, except for Pavitt sector 2, *Prop. of R&D firms* has significantly negative signs in the logistic models predicting membership in the "certain zero" group, suggesting that the higher the proportion of R&D firms, the less likely the industry is to have a certain zero. In the five sectors, the test statistics for the

Vuong tests are significant at the 1% (10%) level for Pavitt sectors 1, 3, 4, and 5 (Pavitt sector 6), suggesting that the zero-inflated negative binomial model is a better fit than the standard negative binomial model. However, the test statistics for the Vuong tests are not significant at all for Pavitt sector 1, suggesting that the zero-inflated negative binomial model is not a better fit than the standard negative binomial model. Thus, we also report the negative binomial regression results for this sector. We find a clear inverted-U-shaped relationship between technological imitation and the quantity of corporate innovation, regardless of Pavitt sectors.

The estimation results based on the zero-inflated negative binomial regressions suggest that $\overline{CITE}_{j,t}$ has peaks at the 85th, 89th, 88th, 79th, 84th, and 87th percentiles in Pavitt sectors 1, 2, 3, 4, 5, and 6, respectively. Similarly, the estimation results based on the negative binomial regressions suggest that $\overline{CITE}_{j,t}$ has peaks at the 90th percentile in Pavitt sector 2. In general, all Pavitt sectors have peak points at similar imitation levels (around the 85th percentile), although Pavitt 4 has a peak point at a slightly lower imitation level (at the 79th percentile). 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 when technological imitation is at extremely high levels.

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, Park, and Shon (2015).[7] To measure the market value of firm-level innovation, we estimate the sensitivity of a firm-level innovation measure to raw (excess) 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 value of firm-level innovation.

The model is specified as follows:

$$\begin{aligned} r_{i,t} \text{ (or } r_{i,t} - R_{B,t}) = \beta_0 + \beta_1 INN_{i,t-1} + \beta_{Controls} Controls \\ + Firm\ FE + Year\ FE + \varepsilon_{i,t}, \end{aligned} \tag{2}$$

[7]Im, Park, and Shon (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 4: Effects of technological imitation on the quantity of corporate innovation: By Pavitt technology sectors

Pavitt sector	(1) Pavitt 1	(2) Pavitt 2	(3) Pavitt 2	(4) Pavitt 3	(5) Pavitt 4	(6) Pavitt 5	(7) Pavitt 6
Regression model	ZiNB	ZiNB	NB	ZiNB	ZiNB	ZiNB	ZiNB
$IMI_{j,t-1}$	0.267***	0.564***	0.470***	0.484***	0.135	0.396***	1.059**
	(0.098)	(0.083)	(0.038)	(0.124)	(0.098)	(0.110)	(0.458)
$IMI^2_{j,t-1}$	-0.153***	-0.210***	-0.173***	-0.195***	-0.149***	-0.254***	-0.621***
	(0.048)	(0.038)	(0.019)	(0.054)	(0.051)	(0.053)	(0.185)
$\overline{Size}_{j,t-1}$	0.608***	0.480***	0.070***	0.703***	0.567***	0.754***	-0.205
	(0.055)	(0.033)	(0.022)	(0.040)	(0.031)	(0.052)	(0.172)
$\overline{ROA}_{j,t-1}$	0.520**	0.403***	0.073	0.164	0.404***	0.210	-0.107
	(0.248)	(0.127)	(0.094)	(0.125)	(0.115)	(0.174)	(0.383)
$\overline{R\&D}_{j,t-1}$	25.310**	13.420***	1.365**	6.696***	4.303***	20.320***	-3.720
	(10.223)	(2.849)	(0.640)	(0.824)	(0.708)	(5.142)	(3.277)
$\overline{PPE}_{j,t-1}$	0.277	-5.719***	1.123***	0.123	-1.777***	2.229***	-12.625***
	(0.486)	(0.540)	(0.373)	(1.032)	(0.554)	(0.658)	(3.136)
$\overline{Lev}_{j,t-1}$	-2.790***	3.386***	-0.154	0.097	0.562	0.263	3.461
	(0.383)	(0.435)	(0.217)	(0.469)	(0.479)	(0.738)	(2.894)
$\overline{Capex}_{j,t-1}$	2.156	3.079	0.324	8.408***	5.138***	-9.861***	0.959
	(1.570)	(2.113)	(0.987)	(2.412)	(1.709)	(3.452)	(7.774)
$\overline{MB}_{j,t-1}$	-0.034*	-0.031*	-0.006	-0.008	0.027**	-0.004	-0.016
	(0.019)	(0.016)	(0.012)	(0.009)	(0.013)	(0.011)	(0.033)
$\overline{Age}_{j,t-1}$	0.505**	0.604***	0.223***	0.152	0.650***	0.431	-0.590
	(0.200)	(0.190)	(0.086)	(0.192)	(0.174)	(0.264)	(0.935)
$\overline{KZ}_{j,t-1}$	-0.012**	-0.002	-0.000	0.021***	-0.007	0.021*	-0.007
	(0.005)	(0.007)	(0.003)	(0.006)	(0.006)	(0.012)	(0.011)
Constant	0.940**	1.310***	-0.147	0.832**	1.260***	-1.429**	10.628***
	(0.428)	(0.391)	(0.216)	(0.368)	(0.348)	(0.590)	(1.932)
Logistic model predicting the membership in the "certain zero" group							
Prop. of R&D firms	-71.446***	5.162***		-10.574***	-12.430***	-34.113***	-143.834***
	(25.789)	(1.804)		(1.529)	(3.594)	(8.852)	(44.974)
Constant	1.739	-28.585***		0.128	1.518*	-2.518***	9.592***
	(1.242)	(1.297)		(0.659)	(0.923)	(0.951)	(3.270)
Vuong z-stat	3.53	-12.13		2.46	2.64	3.07	1.55
p-value	0.000	1.000		0.007	0.004	0.001	0.061
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry fixed effects	No	No	Yes	No	No	No	No
Observations	1,001	1,343	1,342	1,427	1,064	683	239

Note: This table reports the results of sector-by-sector 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 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 ($\overline{CITE}_{j,t}$). "NB" stands for a negative binomial regression model, and "ZiNB" stands for a zero-inflated negative binomial regression model. "Vuong z-stat" is the test statistic for the Vuong test that compares the zero-inflated negative binomial model to a 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.

where

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

$r_{i,t}$ 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 ($R_{p,t}$) and industry portfolio ($R_{j,t}$). $IMI_{j,t-1}$ 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, and the lagged market-to-book ratio. We employ within-groups (i.e., fixed-effects) estimators to capture unobserved heterogeneity across firms. We also include year dummies to capture unobserved heterogeneity across years.

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 specifications to measure the value of innovation. Columns (1) through

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***	0.033***	0.020***	1.76				
	(0.006)	(0.006)	(0.005)					
$INN1_{i,t-1} \times IMI_{j,t-1}$	0.010***	0.016***	0.006**	1.97				
	(0.003)	(0.003)	(0.003)					
$INN1_{i,t-1} \times IMI^2_{j,t-1}$	-0.008***	-0.008***	-0.005***	2.22				
	(0.001)	(0.001)	(0.001)					
$INN2_{i,t-1}$					0.018***	0.018***	0.012***	1.75
					(0.002)	(0.003)	(0.002)	
$INN2_{i,t-1} \times IMI_{j,t-1}$					0.009***	0.012***	0.005***	1.69
					(0.002)	(0.002)	(0.002)	
$INN2_{i,t-1} \times IMI^2_{j,t-1}$					-0.006***	-0.005***	-0.003***	1.87
					(0.001)	(0.001)	(0.001)	
$\Delta Earnings_{i,t}$	0.568***	0.576***	0.467***	1.11	0.568***	0.576***	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***	1.69	0.231***	0.223***	0.194***	1.69
	(0.009)	(0.009)	(0.008)		(0.009)	(0.009)	(0.008)	
$\Delta R\&D_{i,t}$	0.762***	0.793***	0.675***	1.1	0.763***	0.792***	0.675***	1.10
	(0.126)	(0.127)	(0.119)		(0.126)	(0.127)	(0.119)	
$\Delta Dividends_{i,t}$	1.365***	1.418***	1.116***	1.01	1.365***	1.418***	1.117***	1.01
	(0.281)	(0.289)	(0.257)		(0.281)	(0.289)	(0.257)	
$LnTA_{i,t-1}$	-0.236***	-0.236***	-0.195***	1.55	-0.236***	-0.236***	-0.196***	1.44
	(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***	1.13	-0.077***	-0.069***	-0.068***	1.13
	(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)	
$\Delta Interests_{i,t}$	-1.425***	-1.395***	-1.164***	1.26	-1.424***	-1.394***	-1.164***	1.26
	(0.106)	(0.107)	(0.095)		(0.106)	(0.107)	(0.095)	
Constant	1.109***	0.732***	0.640***		1.112***	0.734***	0.641***	
	(0.029)	(0.029)	(0.025)		(0.029)	(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.

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
$IMI_{j,t-1}$	0.034***	0.034***	0.019**	2.93	0.023***	0.022***	0.011***	3.28
	(0.009)	(0.009)	(0.008)		(0.004)	(0.004)	(0.004)	
$IMI^2_{j,t-1}$	-0.014***	-0.015***	-0.009***	2.24	-0.006***	-0.007***	-0.002**	1.85
	(0.003)	(0.004)	(0.003)		(0.001)	(0.002)	(0.001)	
$\Delta Earnings_{i,t}$	0.742***	0.738***	0.625***	1.11	0.159***	0.161***	0.138***	1.23
	(0.055)	(0.055)	(0.049)		(0.053)	(0.054)	(0.045)	
$\Delta Assets_{i,t}$	0.221***	0.217***	0.183***	1.69	0.059***	0.061***	0.054***	1.29
	(0.020)	(0.020)	(0.018)		(0.015)	(0.016)	(0.014)	
$\Delta R\&D_{i,t}$	0.574***	0.629***	0.471**	1.1	-0.192	-0.179	-0.214	1.08
	(0.218)	(0.217)	(0.201)		(0.199)	(0.196)	(0.195)	
$\Delta Dividends_{i,t}$	0.469	0.716	0.138	1.01	-0.000	0.002	-0.007	1.01
	(0.497)	(0.505)	(0.453)		(0.041)	(0.041)	(0.041)	
$LnTA_{i,t-1}$	-0.227***	-0.208***	-0.165***	1.13	-0.100***	-0.095***	-0.077***	1.11
	(0.012)	(0.012)	(0.011)		(0.007)	(0.007)	(0.006)	
$Leverage_{i,t-1}$	0.078***	0.076***	0.068***	1.19	0.023***	0.021***	0.022***	1.22
	(0.013)	(0.013)	(0.011)		(0.007)	(0.007)	(0.006)	
$MB_{i,t-1}$	-0.057***	-0.051***	-0.051***	1.11	-0.022***	-0.020***	-0.020***	1.15
	(0.004)	(0.004)	(0.004)		(0.002)	(0.002)	(0.002)	
$Financing_{i,t}$	0.144***	0.162***	0.134***	1.41	0.045*	0.046*	0.037	1.21
	(0.040)	(0.040)	(0.036)		(0.027)	(0.027)	(0.026)	
$\Delta Interests_{i,t}$	-1.913***	-1.889***	-1.428***	1.25	0.009	0.008	0.038	1.28
	(0.262)	(0.260)	(0.227)		(0.109)	(0.111)	(0.091)	
Constant	1.255***	0.883***	0.713***		0.575***	0.436***	0.352***	
	(0.060)	(0.059)	(0.053)		(0.034)	(0.034)	(0.032)	
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	
Adj. R-squared	0.235	0.173	0.141		0.169	0.125	0.098	

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.

(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 specifications, the relationship between technological imitation and the market value of 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.

The estimation results with raw returns as the dependent variables suggest that the value of innovation measured by the regression coefficient of $INN1_{i,t-1}$ peaks when $IMI_{j,t-1}$ has a value of 0.625, 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 0.750.[8] The values correspond to approximately the 81st and 83rd percentiles, respectively. Thus, an increase in technological imitation leads to an increase in the market value of innovation up to the 81st to 83rd percentile of technological imitation, but the effect then becomes negative after that point. This finding im-

[8] The standardized imitation level corresponding to the peak points are 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 $INN_{i,t-1} \times IMI^2_{j,t-1}$.

plies that the positive externalities from the interactions among firms during the innovation process dominate the negative effects of free-riding concerns on firms' incentives to innovate up to a high degree of technological imitation, while free-riding concerns dominate 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 independent variables is only 2.22 (1.87), 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.[9] By directly including $IMI_{j,t-1}$ and $IMI^2_{j,t-1}$ as explanatory variables, we can test the curvilinear effects. 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. 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 squared term, we test for multicollinearity using VIFs based on an OLS regression model. The maximum VIF in Column (4) (in Column (8)) is only 2.93 (3.28), so it appears that multicollinearity is not a serious issue.

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

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 coefficient is statistically significant at the 5% 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.[10] Thus, the relationship between technological imitation and the market value of firm-level innovation, regardless of Pavitt sectors, has 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 imitation.

[10]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
$INN2_{i,t-1}$	0.015**	0.011**	0.022***	0.025***	0.018	0.032***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.018)	(0.010)
$INN2_{i,t-1} \times IMI_{j,t-1}$	0.001	0.006	0.011**	0.017***	-0.008	0.002
	(0.005)	(0.004)	(0.004)	(0.003)	(0.009)	(0.009)
$INN2_{i,t-1} \times IMI^2_{j,t-1}$	-0.011**	-0.006**	-0.008**	-0.021***	-0.001	-0.010
	(0.004)	(0.003)	(0.004)	(0.003)	(0.010)	(0.008)
$\Delta Earnings_{i,t}$	0.094*	0.370***	0.594***	0.284***	0.322***	0.422***
	(0.055)	(0.058)	(0.051)	(0.110)	(0.082)	(0.063)
$\Delta Assets_{i,t}$	0.057***	0.030**	0.128***	0.068*	0.029	0.079***
	(0.016)	(0.014)	(0.038)	(0.040)	(0.030)	(0.029)
$\Delta R\&D_{i,t}$	0.886	0.834**	0.761***	0.187	0.183	0.601***
	(0.545)	(0.417)	(0.170)	(0.148)	(0.329)	(0.185)
$\Delta Dividends_{i,t}$	0.144**	0.088**	0.254*	0.006	-0.010	0.296**
	(0.070)	(0.043)	(0.142)	(0.094)	(0.103)	(0.121)
$LnTA_{i,t-1}$	-0.195***	-0.191***	-0.259***	-0.296***	-0.228***	-0.306***
	(0.018)	(0.019)	(0.014)	(0.015)	(0.037)	(0.021)
$Leverage_{i,t-1}$	0.042**	0.037***	0.063***	0.080***	0.053***	0.157***
	(0.018)	(0.011)	(0.014)	(0.017)	(0.009)	(0.043)
$MB_{i,t-1}$	-0.093***	-0.018**	-0.063***	-0.047***	-0.088***	-0.050***
	(0.010)	(0.008)	(0.009)	(0.005)	(0.019)	(0.012)
$Financing_{i,t}$	0.041**	0.022	0.075	0.275***	0.072*	0.298***
	(0.018)	(0.031)	(0.048)	(0.036)	(0.038)	(0.074)
$\Delta Interests_{i,t}$	-0.051	-0.578***	-0.606**	-0.191	-0.385**	-0.338
	(0.067)	(0.222)	(0.296)	(0.323)	(0.161)	(0.226)
Constant	1.219***	1.117***	1.092***	1.367***	1.135***	0.902***
	(0.087)	(0.091)	(0.059)	(0.067)	(0.162)	(0.111)
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
R-squared	0.185	0.201	0.246	0.257	0.222	0.315

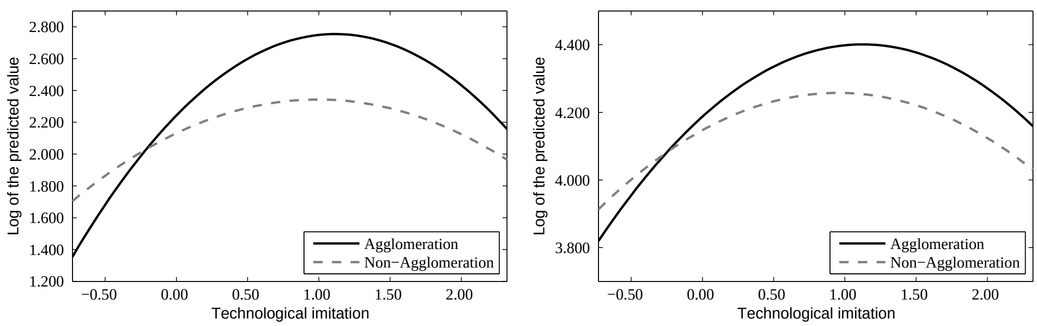
Note: This table reports the results of sector-by-sector regressions designed to estimate the impact of technological imitation on market value of firm-level innovation. Each column represents the following sector: (1) Pavitt 1: supplier dominated; (2) Pavitt 2: Production intensive (scale intensive); (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 utilities and financial service sectors. Raw stock returns are used as the dependent variable. Standard errors clustered by firm are reported in brackets. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

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

To derive implications for innovation cluster policies, we further investigate how the relationship between imitation and innovation differs between the agglomeration and non-agglomeration industries. We do so because the most important factor in innovation clusters is the agglomeration or proximity to the supplier, customer, and R&D collaborator. To measure the level of agglomeration for the industry j in year t, we closely follow Ellison and Glaeser (1997). We calculate $EG-index$ ($\gamma^{EG}_{j,t}$) as follows:

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

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 the Herfindahl index 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 Herfindahl index using the economic census database provided by the US Census Bureau. The $EG-index$ is calculated for each of the 2-digit SIC industries in the manufacturing sector since the concentration data are available only for the manufacturing sector and the employment data are available only for each of the 2-digit SIC industries. We then categorize industry-year observations into two groups (i.e., agglomeration and non-agglomeration industries) based on the sample mean of the $EG-index$.[11]



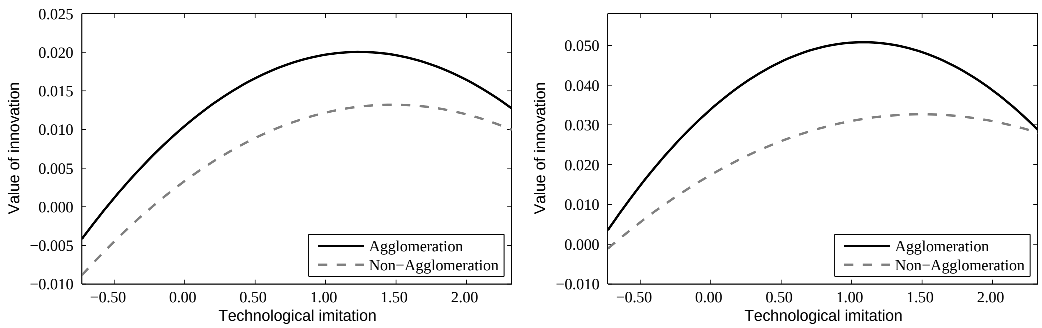
(a) Using $\overline{COUNT}_{j,t}$ as the dependent variable (b) Using $\overline{CITE}_{j,t}$ as the dependent variable

Figure 1. The impact of imitation on the quantity of innovation: agglomeration industries vs. non-agglomeration industries. This figure plots the natural logarithm of the quantity of innovation against the degree of imitation separately for agglomeration and non-agglomeration 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 and subsample mean values of the control variables. The solid curves are based on the subsample of agglomeration industries, while the dotted curves are based on the subsample of non-agglomeration industries. The figure is drawn for a range between the 1st and 99th percentiles of the imitation measure in the manufacturing sample.

To investigate how the relationship between imitation and innovation differs between the agglomeration and non-agglomeration industries, we estimate the zero-inflated negative binomial models as in Table 2 and the fixed-effects regres-

[11]An industry-year observation is classified as an agglomerated industry if the $EG-index$ is greater than or equal to its sample mean and as a non-agglomerated industry otherwise.

sion models as in Table 5, separately for the two subsamples (i.e., agglomeration and non-agglomeration industries). Figures 1 and 2 summarize the estimation results. In Figure 1, we plot the natural logarithm of the quantity of innovation against the degree of imitation separately for the two subsamples. 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 are based on the subsample of agglomeration industries, while the dotted curves are based on the subsample of non-agglomeration industries. The solid curves are steeper 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. In Figure 2, we plot the value of innovation against the degree of imitation separately for agglomeration and non-agglomeration 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 are based on the subsample of agglomeration industries, while the dotted curves are based on the subsample of non-agglomeration industries. Again, the solid curves are steeper than the dotted curves in both panels, suggesting a stronger impact of imitation on the market value of innovation for agglomerated industries than for non-agglomerated industries. Overall, 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.



(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 imitation on the value of innovation: agglomeration industries vs. non-agglomeration industries. This figure plots the value of innovation against the degree of imitation separately for agglomeration and non-agglomeration 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 are based on the subsample of agglomeration industries, while the dotted curves are based on the subsample of non-agglomeration industries. The figure is drawn for a range between the 1st and 99th percentiles of the imitation measure in the manufacturing sample.

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. However, a level of technological imitation that is too high is more detrimental for firms in innovation clusters, as it lowers their incentives to innovate more radically.

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 positive externalities from the interactions among firms during the process of technological innovation. Particularly when innovation is sequential and complementary, interactions among innovative firms can enhance firms' innovation activities and incentives to innovate. The second effect is the negative effect of free-riding problems on firms' innovation activities and 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 dominates the second effect up to a high level of technological imitation, while the second effect dominates 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.

Acknowledgements

The authors are grateful to two anonymous referees for insightful comments on the earlier version. 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. The research has been partially supported by the Peking University HSBC Business School Annual Research Funds and Bridge Trust Asset Management Research Funds.

Appendix A. Definition of industry-average variables

The following table shows the definitions of the industry-average variables used in Table 2 and Table 4. 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 the 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 calculated as the time elapsed since a firm appears in the data for the first time.
$\overline{KZ}_{j,t-1}$	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. The italicized codes in brackets([]) represent item codes in CRSP/Compustat Merged Database. In line with Im, Park, and Shon (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

References

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

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

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

Bessen, J., and Maskin, E., 2009, "Sequential innovation, patents, and imitation," *RAND Journal of Economics*, 40, 611–635.

Bhattacharya, U., Hsu, P.H., Tian, X., and Xu, Y., 2015, "What Affects Innovation More: Policy or Policy Uncertainty?" *Kelley School of Business Research Paper*.

Brozen, Y., 1951, "Invention, innovation, and imitation," *American Economic Review*, 41, 239–257.

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

Ellison, G., and Glaeser, E.L., 1997, "Geographic concentration in US manufacturing industries: a dartboard approach," *Journal of Political Economy*, 105, pp. 889–927.

Fang, V.W., Tian, X., and Tice, S., 2014, "Does stock liquidity enhance or impede firm innovation?" *Journal of Finance*, 69, 2085–2125.

Faulkender, M., and Wang, R., 2006, "Corporate financial policy and the value of cash," *Journal of Finance*, 61, 1957–1990.

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

Grossman, G., and Helpman, E., 1991, "Innovation and growth in the world economy", Cambridge, MA: MIT Press.

Hall, B.H., Jaffe, A., and Trajtenberg, M., 2001, "The NBER patent citation data file: Lessons, insights and methodological tools," No. w8498, National Bureau of Economic Research.

Hall, B.H., Jaffe, A., and Trajtenberg, M., 2005, "Market value and patent citations," *RAND Journal of Economics*, 36, 16–38.

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

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

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

Mukherjee, A., Singh, M., and Zaldokas, A., 2016, "Do Corporate Taxes Hinder Innovation?" *Journal of Financial Economics*, Forthcoming.

O'Hara, R.B. and Kotze, D.J., 2010, "Do not log-transform count data," *Methods in Ecology and Evolution*, 1, 118–122.

Pavitt, K., 1984, "Sectoral patterns of technical change: towards a taxonomy and a theory," *Research policy*, 13, 343–373.

Schoonhoven, C.B., Eisenhardt, K.M., and Lyman, K., 1990, "Speeding products to market: Waiting time to first product introduction in new firms," *Administrative Science Quarterly*, 35, 177–207.

Tidd, J., Bessant, J. and Pavitt, K., 2005, Managing Innovation: Integrating Technological, Market and Organizational Change. John Wiley and Sons Ltd.

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