Congruence and Firm Innovation: Evidence from "Chinese NASDAQ"

Yong Wang[*], Xiuping Hua[†], Junjie Xia[‡], and Haochen Zhang[§]

Abstract

This paper studies the determinants of firm innovation through the lens of congruence, measuring the distance between the local endowment structure and the industry factor inputs structure. Using comprehensive data for firms listed in the National Equities Exchange and Quotations (NEEQ)—the counterpart of NASDAQ in China—we identify the mechanisms through which congruence facilitates firm innovation. We highlight that congruence is more pronounced for firms with tighter financial constraint and firms without venture capital support, and it is more salient for firms in regions where the productivity levels are further away from the technology frontier. Further tests suggest that congruence associated with inter-industry network and the placed-based industrial policy drives heterogeneous effects. Finally, we find that firms with higher congruence are more likely to graduate to the Chinese main stock markets.

Keywords: congruence; network; innovation; R&D; patents

JEL Classification: D80, G30, O30, R10

[*] Institute of New Structural Economics (INSE), Peking University, Room 503, Langrun Garden, 5 Yiheyuan Road, Haidian District, Beijing, 100871. email: yongwang@nsd.pku.edu.cn;

[†] Corresponding author, Business School, UNNC-NFTZ Blockchain Laboratory, University of Nottingham Ningbo China, Yinzhou District, Ningbo, China, 315100, email: xiuping.hua@hottingham.edu.cn;

[‡] Central University of Finance and Economics and Peking University, email: junjiexia@nsd.pku.edu.cn;

[§] Institute of New Structural Economics (INSE), Peking University, email: zhanghc@pku.edu.cn.

1. Introduction

Innovation boosts productivity and drives economic growth. For a growing economy like China, which has been gradually losing its comparative advantage in labor-intensive sectors, innovation becomes an increasingly important driving force for industrial upgrading and economic transformation. Over the years, the Chinese government has implemented various policies to encourage firm innovation. One example is the establishment of National Equities Exchange and Quotations (NEEQ), the counterpart of NASDAQ in China,[5] which aims to work as an alternative investment market to promote innovation and entrepreneurship for small and medium-sized firms (SMEs).[6] SMEs account for the majority of economic activities. According to the China Ministry of Industry and Information Technology, SMEs contributes up to 60% of GDP and 70% of technological innovation in 2020. Therefore, our investigation on SMEs provides important implications to the subject of innovation.

In this paper, we emphasize on the role of endowment structure in explaining firm innovation performance, which is distinct from the existing studies that focus on institutions, culture, and market development (e.g., Chen et al., 2014; Zhu et al., 2020).[7] In particular, we construct a measure of congruence between the local endowment structure and the industry factor inputs structure, which has been used in the development literature (e.g., Lin, 2009; Ju et al., 2015). Our paper is the first attempt to explore empirically how congruence affects firm innovation. We find that, after controlling other factors, firms in industries with higher congruence invest more in research and development (R&D) and produce more patents.

To examine how congruence affects firm innovation, we use a fixed-effects identification approach which captures both cross-sectional and time-series dynamics between endowment structure and innovation and we allow for reliable statistical inferences. In addition, to enhance the empirical analysis, we follow the literature on agglomeration and innovation by

[5] See media coverage at the People's Daily on May 29, 2013 with the Chinese title "The establishment of NEEQ – Chinese Nasdaq Launching", discussing how NEEQ could develop to be the Chinese Nasdaq.
[6] Source: the official website of NEEQ (http://www.neeq.com.cn/en/about_neeq/introduction.html).
[7] He and Tian (2021) provide a comprehensive survey for recent studies on institutions and innovation.

incorporating network spillover effects into our empirical specifications. In particular, we explore the network effects through three channels: knowledge spillover, labor pooling, and production network. While the first two channels have been studied in the literature,[8] few studies have explored how innovation is affected by production network.[9]

We investigate two mechanisms motivated by the theory on the endowment structure and growth. One is the channel of financial support, and the other is the channel of technology selection. First, in terms of financial support, theoretically, if the factor input structure of a firm is less consistent with the endowment structure of the economy, the new products generated by innovation activities will be less cost efficient and hence less competitive in the market for violating comparative advantages. Consequently, the market value of new patents is lower and firms' incentives to invest in R&D are weaker. Therefore, firms that are more financially constrained would be even more reluctant to invest in R&D. As a result, financially constrained firms are expected to be more affected by the degree of congruence in their R&D performance. In contrast, firms with financial slack are less sensitive to the congruence impact as their R&D investment decisions are less resource constrained by the profitability of new products/technologies due to R&D. Similarly, the VC institutional support at the firm level is another effective influencing mechanism (Hsu, 2006; Hochberg et al., 2007). If an innovative firm lacks of innovation support like VC support, it faces higher uncertainty in innovation commercialization, and thus cost efficiency through the congruence effect matters more for them, while a VC-backed firm is less sensitive to the congruence effect.

Second, in connection with the channel of technology selection, firms that are further away from the technological frontier are more likely to use mature and publicly available technologies, so cost efficiency means more to them in order to maintain market competitiveness and to generate revenues available to finance R&D expenditure. Furthermore, such firms would generally find it more effective to achieve technological progress by imitation rather than innovation (Acemoglu, Aghion and Zilibotti, 2006), so they would be more cautious

[8] Carlino and Kerr (2015) review the literature on agglomeration and innovation and argue that there is more empirical evidence supporting the channel of knowledge spillover.

[9] A growing macro and trade literature has studied on the topic of production network (e.g., Acemoglu et al., 2012; Liu, 2019; and Bigio and La'O, 2020).

and selective when making R&D decisions. Since more congruence implies higher cost efficiency and higher market reward for a given innovation, firms further away from the frontier are more likely to restrict themselves to R&Ds that are more sensitive to the congruence effect. In addition, we assert that this effect is particularly relevant for developing countries such as China, as a large fraction of the economy is still away from the world technological frontier.

To test our hypotheses, we utilize several different sources to obtain data for firms listed on NEEQ. We first manually collect firm balance sheet information from the China Stock Market and Accounting Research Database (CSMAR) and Wind Database. We then match them with the patent data from the China National Intellectual Property Administration and Incopat database. To construct a host of empirical measures, we also resort to five other data sources: 1) enterprise income tax records from the Chinese State Administration of Tax, 2) the China City Statistical Yearbook, 3) the input-output table from the China National Bureau of Statistics, 4) the China Population Census, and 5) venture capital (VC) data from Zero2IPO and CVSource databases.

Using this comprehensive set of data, we document that firms with higher congruence invest more in R&D and produce more patents. On average, an increase in congruence by one standard deviation would raise R&D intensity by 13.5% and the number of patents by 2.1%. Furthermore, we provide confirming evidence on the two main mechanisms. First, the congruence effect is more pronounced for firms that are more financially constrained, and firms without VC investment depend more on congruence than those with VC investment. These findings also contribute to the literature on the financing of innovation by illuminating the role of congruence effect.[10] Second, we show that the congruence effect is more salient for firms in regions where the productivity levels of the corresponding industries are further away from the technology frontier.

We also conduct additional heterogeneity tests and report the following two empirical findings: First, firms located outside high-tech special economic zones (SEZ) are more affected

[10] See Hall and Lerner (2010) and Kerr and Nanda (2015) for a detailed survey of the literature on the financing of R&D expenditures and innovation.

by congruence than firms located inside. Our explanation is that firms lack of place-based industrial policy support are similar to those firms with financially constrained or lack of VC support, the congruence effect on innovation is more pronounced. Second, firms that are more separated from an agglomerative economy are more sensitive to the congruence effect, because those firms face tougher external conditions.

Finally, we find that firms with higher congruence are more likely to be listed successfully in the main stock market, such as Shanghai Stock Exchange Market and Shenzhen Stock Exchange Market, which could provide policy implications for the new reform of NEEQ – the establishment of Beijing Stock Market.

Our paper contributes to several strands of literature. First, our work provides new insights to the literature on how the external environment influences firm innovation. Some studies examine the impact of culture on innovation activities. Chen et al. (2014) present evidence that firms located in gambling-prone areas tend to invest more in R&D and produce more innovation output. Bénabou et al. (2015) find that religiosity is negatively correlated with innovation. Another strand of literature investigates the impact of market development on innovation. Hsu et al. (2014) show that countries with better developed equity markets generate more innovation. Zhu et al. (2020) suggest that countries with a higher level of financial development have slower innovation growth, while Xiao and Zhao (2012) find that banking sector development enhances firm innovation for countries with lower government ownership of banks. Our work contributes to the literature by exploring the role of the local endowment structure and provides empirical evidence that congruence is an important determinant for firm innovation in a developing country.

In addition, our paper is closely related to the growth literature on technology and endowment structure. Jones (2005) and Caselli and Coleman (2006) study how the properties of the endogenous aggregate production function for developed countries (on the world technology frontier) are affected by technology choices that optimally respond to factor endowment structure (also see Leon-Ledesma and Satchi (2019)). Boldrin and Levine (2002) demonstrate how rising wages drive innovation for new vintages of labor-saving capital. Basu and Weil (1998) highlight that the appropriate technologies for developing countries should be consistent with factor endowment structure. Lin (2009) argues that macroeconomic performance of an economy is largely affected by the congruence of industrial structure with the comparative advantages determined by the endowment structure. Ju et al. (2015) develop a theory of endowment-driven structural change in explaining shifts in industrial structures, life-cycle industry dynamics and aggregate economic growth, and they find that industries that are more congruent with endowment structures tend to have a larger value-added share (and also employment share) in the economy. Lin et al. (2021) present empirical evidence that congruence is an important determinant for the success of industrial policy. Our paper adds to the literature by showing how congruence of industrial input structure with the local endowment structure affects firm innovation in China, which has important implications for China's productivity growth.

Lastly, our paper also sheds light on how urban agglomeration affects innovation. Empirical studies have shown that knowledge spillover is crucial to promoting innovation through spatial concentration (e.g., Jaffe et al., 1993; Thompson and Fox-Kean, 2005; and Muarta et al., 2014) and through universities and research units (e.g., Audretsch and Stephan, 1996; Anselin et.al., 1997; and Carlino et al., 2007). Some articles suggest that labor pooling or thick labor markets improve the quality of matches and thus enhance innovation (e.g., Moretti, 2004; Rosenthal and Strange, 2008).[11] In addition to the knowledge spillover and the labor pooling, we also consider production network. Thus, our paper complements the literature by investigating how agglomeration may affect innovation through three different channels, and we show that agglomeration does affect innovation, and the impact depends on congruence.

The remainder of the paper is organized as follows. In Section 2, we explore the theoretical and empirical evidence of congruence effect on innovation in the existing literature, and put forward three testable hypotheses. Section 3 introduces the institutional background and data. Section 4 presents our empirical specification, baseline results and possible mechanisms. Section 5 discusses the heterogeneity and extension. Section 6 concludes.

[11] For further reference, see Glaeser (1999), Berliant et al. (2006), and Fallick et al. (2006).

2. Related Literature and Hypothesis Development

2.1 Congruence and firm innovation

Academic studies about congruence can be traced back to Lin (2009), who argues that a wide spectrum of economic development issues, including firm performance, industrial upgrading, structural change, development strategies, financial institutions, industrial policies and economic growth, can all be better understood through the lens of the congruence of the industrial structure and technological structure with the economy's endowment structure and its change. In this literature, (factor) endowment structure refers to the composition of production factors, such as labor, human capital, physical capital, land and other natural resources. The core argument goes as follows: the economic performance of a firm, an industry or an economy as a whole would be better, *ceteris paribus*, if the factor intensities of the embodied technologies are more congruent with the factor endowment structure, which is given at a time and changing over time, because more congruence implies that production costs are lower when productions utilize more abundant and hence cheaper factors. As a result, more congruence means higher cost efficiency and more competitiveness, as comparative advantage is followed. As a flip side of the coin, many distortions in policies and institutions observed in reality are endogenous consequences of deviations from congruence, which are in turn perhaps due to hasty catch-up political motives, like the "Great Leap Forward" movement in China in the 1960s (Lin, 2009).

Using the NBER-CES data of the US and the cross-country UNIDO data sets, Ju et al. (2015) document and also develop a dynamic general equilibrium model to explain a "congruence fact" in steadily growing economies such as the US, namely, the further the deviation of an industry's capital-labor ratio from the aggregate capital-labor ratio (endowment structure) of the economy, the smaller is the employment (and value added) share of this industry. It is consistent with the finding of Lin (2009) that documents a huge amount of cross-country evidence showing that countries would exhibit undesirable macroeconomic performance when their industries are not congruent with factor endowment.

However, it remains unexplored how congruence would impact firm innovation in the literature. Theoretically speaking, firms in industries that are congruent to factor endowment structure of the local economy would in general have low production cost, so the newly innovated products and services are also cost efficient and hence competitive in the market, and the market value of the patents would be high, so firms would have stronger incentives to invest more in R&D and their innovation output would be also high. In contrast, industries with too low or too high capital intensities are inconsistent with the comparative advantages of the (local) economy, and thus the potential market value of a patent in such industries is lower. Therefore, both the share of R&D investment and that of patents will be lower.

The congruence impact is particularly important in developing countries like China. It is because, unlike developed countries, developing countries are more plagued with market frictions and distorting policies, so a significant fraction of firms is nonviable as they are in industries that are incongruent with factor endowment. For example, many developing countries want to prematurely promote capital-intensive industries even though capital is scarce at those early stages of development, resulting in that firms in those industries tend to use more capital and less labor, so the output costs are too high and firms have no competitive advantage. Such firms are easily defeated by the firms from developed countries because capital-intensive industries are consistent with their comparative advantages, so they have to rely on government subsidies and protection in order to survive. Firms, no matter state-owned or private, would be nonviable if they are in industries incongruent with factor endowment structure. Such industries would not be able to operate in free market, but they might survive in countries like China because of government support such as subsidies or entry restrictions on external firms. In contrast, in developed countries, firms are mostly viable and industries are generally congruent with factor endowment structures, because otherwise they would not be able to survive the market competition. Note that markets are mature and more efficient in developed countries and governments generally do not pursue catching-up development strategies.

Besides, in China there exist a wide array of industries that are heterogeneous in both capital intensities and distances from the world technological frontier. Therefore, China offers good opportunities to exploit these rich variations when studying how endowment structure affects firm innovation in heterogeneous industries. We hypothesize that the congruence effect on innovation varies with technology development levels of industries. “Leading” industries are close to or at the technology frontier of the world and consistent with comparative advantage. They usually rely on their own R&D to achieve technological progress, and the shares of R&D expenditure and patents in such industries are increasing over time. The “catching-up” industries are sufficiently away from the frontier but also consistent with comparative advantage. They benefit from technology spillovers and account for smaller shares in the R&D expenditure. As firms in such industries gradually approach the technology frontier, innovation activities become increasingly necessary and vital for their survival. Thus, more R&D investment and patents are expected. In contrast, the “exiting” industries are gradually losing their comparative advantages as their capital labor ratio deviates from the endowment structure, so new patents are less valuable. Consequently, the incentives for innovation are weak, and the R&D expenditure shares are getting smaller and smaller, which result in less subsequent patenting behavior.

To sum up, we contemplate that innovative firms with higher congruence, such as those in “leading” and “catch-up” industries, invest more in research and development (R&D) and produce more patents, than those with lower congruence and in “exiting” industries that lack of incentives to innovate. It thus drives our first hypothesis.

Hypothesis 1 *The congruence is positively associated with innovation performance of SMEs in China.*

2.2. Influencing mechanisms: The channel of financial support

Previous work has not explored how congruence impacts innovation at the firm level either. Theoretically, firm innovation of heterogeneous industries can be influenced by congruence through different mechanisms. Based on these theoretical predictions, we propose the two channels through which congruence exerts an influence on firm innovation performance, namely the channel of financial support and the channel of technology selection. First, we discuss how the financial support plays an important role in transmitting the effects of congruence on firm innovation through two perspectives: financial constraints and VC institutions support.

It is well noted in the prior research that financially constrained and unconstrained firms behave heterogeneously in their innovation investment activities, and a financially constrained R&D-intensive firm is more likely to suspend or discontinue its R&D projects. Empirical evidence generally indicates that financial constraints discourage investment in R&D. Garmaise (2008) models the contrasting capital-labor decisions of financially constrained and unconstrained firms and indicate that financially restricted firms use relatively more labor than physical capital.[12] Guariglia and Liu (2014) investigate the extent to which financing constraints affect the innovation activities of unlisted Chinese firms over the period 2000-2007, and find that Chinese firms' innovation activities are constrained by the availability of internal finance and specifically, private firms suffer the most.

We conjecture that the innovative activities of SMEs in China face significant adverse selection and moral hazard problems, and thus financing constraints are more likely to drive R&D investment below the privately optimal level. As aforementioned, congruence of the capital intensity of a firm with the endowment structure of the aggregate economy is able to mitigate financial constraints due to cost efficiency and business viability. Thus, the congruence is able to help firms to bring in more investment in R&D and patent applications. In particular, this impact is more salient for firms with a high level of financial constraints. More financially constrained companies are facing more severe challenges in accessing to both internal and external finance, and thus enhance the link between congruence and firm innovation. On the contrary, this congruence impact can be mitigated by financial slack. This leads to our second hypothesis.

Hypothesis 2A *The congruence effect is stronger on the SMEs that have a higher level of financial constraint.*

---

[12] In the case of China, Cull et al. (2015) find that the sensitivity of investment to internal cash flows is higher for Chinese firms that report higher financial constraint and greater obstacles to obtaining external funds.

In a similar logic, the VC institutional support can be another possible influencing mechanism in terms of financial support. Innovation is a very risky endeavor. Firms with VC institutional support are more willing to take extra risk incurred by innovative activities. As an important form of financial support, VCs' contribution to innovations is well recorded in the literature. Recent research shows that VCs play significant roles in promoting firm innovation by not only providing risk money but also supporting and monitoring the management and operation of portfolio companies by finding key managerial personnel, lining up quality suppliers, improving customer relationships and providing consultancy (Kortum and Lerner, 2001; Chemmanur, Krishnan, and Nandy, 2011; Bernstein, Giroud, and Townsend, 2016). Kortum and Lerner (2001) find consistent evidence using the data of U.S. manufacturing firms that the firms that receive venture capital financing file more patents. Van Den Berghe and Levrau (2002) focus on the role of VCs as a monitor of high-tech VC-backed companies and state that VCs play significant monitoring roles that are different from other types of shareholders and provide value-added consultancy through board activities. Chemmanur et al. (2011) indicate that VC monitoring improves efficiency in VC-backed firms after investment.

In the past two decades, China has experienced rapid growth in venture capital (VC) market (Guo and Jiang, 2013). Ayyagari et al. (2011) find that the externally financed proportion of a firm's investment expenditures is positively related to firm innovation. Our conjecture is that if a firm is able to obtain the VC support, it will not only mitigate the firm' s constraint on financing innovation, but also counter affect the adverse impact of deviation from aggregate endowment structure. Thus, VC investment will reduce firms' dependence on internal profitability and business viability that is exerted by the congruence to invest in R&D. Those innovative SMEs that cannot get access to VC investment shall be more likely to be affected by the congruence. This conjecture lies with and expands the literature on examining how VC affects the process and outcomes of innovation (e.g., Manso, 2011; Chemmanur et al., 2014; Tian and Wang, 2014). This leads us to the following hypothesis.

Hypothesis 2B *The congruence effect on innovation is weaker on the SMEs that are backed by VC institutions.*

2.3. Influencing mechanisms: The channel of technology selection

If firms are further away from the technological frontier, they are more likely to use mature and publicly available technologies. In the case, cost efficiency is more important to them in order to maintain market competitiveness and to generate revenues available to finance R&D expenditure. Thus, the congruence effect is more salient for firms in regions where the productivity levels of the corresponding industries are further away from the technology frontier.

Furthermore, such firms would generally find it more effective to achieve technological progress by imitation rather than innovation (Acemoglu, Aghion and Zilibotti, 2006), and thus they would be more cautious and selective when making R&D decisions. Because a higher degree of congruence implies higher cost efficiency and higher market reward for a given innovation, firms that are further away from the frontier are more likely to restrict themselves to R&Ds that are more sensitive to the congruence effect. In addition, we assert that this effect is particularly relevant for developing countries such as China, as a large fraction of the economy is still away from the world technological frontier. Therefore, we derive this following hypothesis.

Hypothesis 3 *The congruence effect on innovation is more pronounced for firms in regions where the productivity levels of the corresponding industries are further away from the technology frontier.*

3. Institutional Background and Data

We start by introducing the relevant institutional background, in particular, the historical facts of innovation and related policies in China. We then describe the data used in our paper.

3.1. Innovation in China

Thanks to market-oriented policy reforms and a favorable demographic structure, China has enjoyed rapid economic growth in the past four decades. However, as it gradually loses its comparative advantage in labor-intensive sectors and gets closer to the world technology frontier, economic growth has been slowing down since 2010. The average annual GDP growth rate from the 1990s to 2010 is more than 10%, while it drops to less than 7% after 2010.[13]

In order to upgrade its industrial structure, China has implemented various policies to promote innovation and indeed has made remarkable progress in innovation. R&D spending rose from $722 billion in 2000 to $2.2 trillion in 2017, with more than a 17% annual growth rate, compared to the U.S.'s average of 4.3%.[14] In addition, in 2019, China surpassed the U.S. in the number of patent applications for the first time. In 2020, China filed 68,720 patent applications while the U.S. filed 59,230.[15] It is important to understand the fundamental determinants of firm innovation in China.

3.2. National Equities Exchange and Quotations

The National Equities Exchange and Quotations (NEEQ), the counterpart of NASDAQ in China and also known as the New Third Board Market, was officially established in 2013 and under the supervision of the China Securities Regulatory Commission. The NEEQ aims to serve SMEs to enhance innovation and entrepreneurship and energize new drivers of economic growth. The Chinese government has conducted several rounds of reforms to the NEEQ. For example, on September 3rd, 2021, Chinese President Xi Jinping announced a new reform—the formation of a Beijing Stock Exchange to better and more effectively steer investment into innovation.[16]

There are many rules and criteria of entering the NEEQ. In brief, firms that are successfully listed at the NEEQ are considered to have well-organized corporate governance, lawful and regulated operations and well-defined shareholding structure. However, unlike initial public offerings at Shanghai Stock Exchange and Shenzhen Stock Exchange, there is no particular requirement on financial indicators for entering the NEEQ. In other words, the NEEQ is inclined to provide better funding opportunities to SMEs.

---

[13] Source: China National Bureau of Statistics.

[14] Data is based on the biennial Science and Engineering Indicator report, released by the U.S. National Science Foundation. See also media coverage at Nature (Viglione, 2020) and Forbes (McCarthy, 2020).

[15] Source: The World Intellectual Property Organization

[16] See media coverage at the *Wall Street Journal* (Areddy, 2021).

In sum, the NEEQ is one of the main policy instruments to promote alternative investments on firm innovation in China. Studying innovation behaviors for firms listed at the NEEQ is useful to improve the understanding of innovation in China, especially for SMEs in all sectors, and also provide strong policy implications.

3.3. Data

Our data sample consists of yearly data on all the listed firms available from NEEQ during the period of 2013-2019. Firm balance sheet information is manually collected from two professional Chinese enterprise databases: CSMAR and Wind.[17] Patent data are mainly collected from the official website of China National Intellectual Property Administration (CNIPA) and Incopat Database. There are three types of patents in China: invention, utility model and design. To construct other measures in our empirical analysis, our paper also relies on four other sources: 1) enterprises income tax records from the Chinese State Administration of Tax, 2) the China City Statistical Yearbook, 3) the input-output table from the China National Bureau of Statistics, 4) the China Population Census, and 5) venture capital (VC) data from Zero2IPO and CVSource databases. Our final dataset contains 88 two-digit level industries and covers most manufacturing and service sectors.[18]

Following most studies, we use two ways to measure innovation: (1) one is by the inputs used in the innovation process – R&D intensity, measured by the ratio of R&D expenditures to total assets; (2) one is by the intermediate outputs of innovation efforts – patenting, measured by granted patent applications. Carlino and Kerr (2015) provide a discussion about the strengths and limitations of the above two measures for innovation. Following previous literature like Chuluun et al (2017), we use both measures in our study.

---

[17] Table A1 in the appendix reports the number of observations across years. Note that firms report their previous three years balance sheet information before they are listed on NEEQ, thus the data structure is an unbalanced panel.

[18] There are in total 97 two-digit industries in China. Table A2 in the appendix lists firm distribution in the top 20 industries.

Table 1 reports the statistical summary of all variables that we use in our paper.[19] In terms of innovation, the average R&D expenditures is 5.09% of total assets.[20] The average number of patent applications per firm is 3.76. In terms of firm performance, the average ROA, ROE and sales growth are 4.79%, 4.81, and 22.54%, respectively.

[Insert Table 1 about here]

Figure 1 depicts the change of firm innovation performance before and after being listed on NEEQ, suggesting that on average, firms spend more on R&D and generate more patent applications in post-listing period.

[Insert Figure 1 about here]

Table 2 presents the changes of firm level characteristics before and after being listed on NEEQ. The result shows that after being listed on NEEQ, firms obtain more subsidies; attract more equity financing, and invest more on R&D, suggesting that NEEQ facilitates firms' external financing and motivates firms' investment on innovation.

[Insert Table 2 about here]

4. Empirical analysis

We first describe our empirical specification, and then discuss how we construct the key variables. Finally, we present the results from our baseline regression.

4.1. Empirical specifications

To examine how the external economic environment affects firm innovation, we estimate the following equation:

$$y_{isct} = \beta Congruence_{sc} + \gamma Network_{sc} + \rho X_{isct} + \varphi_{ct} + \theta_{st} + \varepsilon_{isct}, \tag{1}$$

where i indexes firms; s indexes industry; c indexes city; t indexes time (i.e., year); y_{isct} represents the dependent variables of interest (e.g., R&D expenditures and patent applications);

[19] To lessen the influence of outliers, we winsorize all variables at the 1st and 99th percentiles.

[20] For the comparison, public traded firms, which are considered to be larger, have lower R&D intensity. The average R&D expenditures is 1.59% of total assets.

$Network_{sc}$ refers to the network spillover effect, including the knowledge spillover, the labor pooling and the input-output linkage (i.e., production network);[21] X_{isct} represents firm level control variables, including age, size and leverage ratio; φ_{ct} and θ_{st} are city-time and industry-time fixed effects, respectively; and ε_{isct} is the error term.

Following Ju et al. (2015) and Lin et al. (2021), we construct a measure of the congruence between the local factor endowment structure and the firm factor inputs structure. The mathematical expression is as follows:

$$congruence_{sc} = -\left[\left|\log\left(\frac{K_{sc}/L_{sc}}{K_s/L_s}\right) - \log\left(\frac{\overline{K_c}/\overline{L_c}}{\overline{K}/\overline{L}}\right)\right|\right] \tag{2}$$

where $K_s \equiv \sum_c K_{\mathrm{sc}}$, $L_s \equiv \sum_c L_{\mathrm{sc}}$, $\overline{K} \equiv \sum_c \overline{K_c}$, and $\overline{L} \equiv \sum_c \overline{L_c}$.

In equation (2), K_{sc} and L_{sc} are total fixed assets and employment of industry *s* at city *c*, respectively. Both variables are calculated by using the enterprises income tax records from the Chinese State Administration of Tax. $\frac{K_{sc}}{L_{sc}}$ measures factor inputs structure (or technology choice) of industry *s* at city *c*. K_s and L_s are total fixed assets and employment in industry *s* at the national aggregate level, so K_s/L_s measures the national average level of factor input structure (technology choice) of industry *s*. $\frac{K_{sc}/L_{sc}}{K_s/L_s}$ measures the capital intensity of industry *s* at city *c* relative to the national average. As we know, within the same industry there still exist heterogeneous sub-industries, products and tasks, depending on the disaggregated level, and their capital intensities may well be different. For example, in capital abundant cities such as Shanghai there are still labor-intensive industries such as apparels and shoes, but firms may choose more capital-intensive technologies or specialize on more capital-intensive products/tasks than firms in the same industry but in capital scarce cities such as Changchun.

$\overline{K_c}$ refers to the total fixed assets of city *c* and $\overline{L_c}$ refers to the total employment of city *c*, so $\overline{K_c}/\overline{L_c}$ measures the factor endowment structure of city *c*. Likewise, $\overline{K}/\overline{L}$ measures the factor

[21] The construction of network index is consistent with that in the literature on agglomeration spillover (e.g., Greenstone et al., 2010 and Ellison et al., 2010). The online appendix provides the technical explanation in details.

endowment structure at the national level. Thus, $\frac{\overline{K_c}/\overline{L_c}}{\bar{K}/\bar{L}}$ represents the endowment structure of city c relative to the national average. Therefore, $\left|\log\left(\frac{K_{sc}/L_{sc}}{K_s/L_s}\right) - \log\left(\frac{\overline{K_c}/\overline{L_c}}{\bar{K}/\bar{L}}\right)\right|$ captures how congruent the relative technological choice (capital intensities) of industry s at city c is with the relative endowment structure of city c. If the absolute value of the difference is larger, it means less congruent, so we add a negative sign before the absolute value for the congruence index. The higher the congruence index, the more congruent of industry s in city c with its endowment structure.[22]

The above fixed effects identification approach captures both cross-sectional and time-series variations between congruence and firm innovation. The city-year fixed effect absorbs time varying city characteristics, e.g., local government policies, city-wide reforms, and economic differences. Industry-year fixed effect absorbs the effects of industrial variation. This interacted fixed effects allows us to control for a wide array of omitted variables (See a similar approach used in Rajan and Zingales (1998) and Hsu et al. (2014)).

Moreover, because changes in a firm's factor input structure may be related to unobserved changes in its production decisions, we purge our specifications of this variation by using two years prior to the establishment of NEEQ. Specifically, we use the data in 2011 to calculate the measure of congruence.[23] This can also mitigate the reverse causality issue. Finally, we cluster standard errors by city and by industry. Overall, we focus on the sign and significance level of β when we interpret our empirical results.

4.2. Results

Table 3 presents the estimates of R&D intensity from our baseline regression. We only keep the estimated coefficients of congruence which is of primary interest. All regressions include city-year fixed effect and industry-year fixed effect. Column (1) of Table 3 estimates the basic impact of congruence on firm innovation, giving a positive and significant correlation

[22] In the regressions, congruence is standardized with a mean of zero and a standard deviation of 1.

[23] This identification framework is similar to an instrumental variables approach (see also Duchin et al., 2010).

coefficient. Column (2) includes the network spillover effects in three ways: knowledge spillover, production network and labor pooling as controls, and shows that the sign and magnitude of the coefficient of congruence do not change.

Column (3) of Table 3 presents our complete baseline regression in which we additionally control firm level characteristics, including age, size, ROA and leverage ratio. The coefficient for congruence is 0.135 and is statistically significant and better than the 1% level, suggesting that increasing congruence by one standard deviation would increase R&D intensity by 13.5%. The magnitude is still close to that in Columns (1) and (2).

[Insert Table 3 about here]

Columns (4), (5) and (6) of Table 3 report the estimates of innovation output—patent applications—from our baseline regression. The coefficient for congruence is positive and significant in all the three specifications. Focusing on Column (6) in which we include firm-level controls, inter-industry network controls and fixed effects, we find that increasing one standard deviation of congruence would increase patent applications by 2.1%.

We also present evidence that firms with higher congruence would generate better performance. Table 4 reports the estimates of firm performance, including profitability, productivity and sales performance, from our baseline regression. In particular, if congruence increases by one standard derivation, ROA, ROE, labor productivity and sales growth would increase by 32.9%, 75%, 2.5%, and 94.8%, respectively. Overall, the results are consistent with our assertion that higher congruence implies higher cost efficiency and hence better firm performance.

[Insert Table 4 about here]

As the robustness check, we conduct two more tests. First, we show that our results are still valid when we employ the propensity score matching (PSM) approach to mitigate the concern of selection biases (Table 5). Second, our results continue to hold when we repeat our baseline regression by using a firm-level measure of congruence, which may allow more time-series variations to our empirical analysis (Table 6).

[Insert Table 5 and Table 6 about here]

Based on the empirical findings from Table 3, we conclude that firms in industries with higher congruence indexes invest more in R&D and generate more patents applications. We then explore the mechanism in explaining this finding. Our hypothesis is that for a developing country with underdeveloped financial markets, obtaining external financing for R&D investment is more challenging, and thus, internal and fundamental factors, such as profitability and business viability, become even more important for firms to finance innovation activities. Also, the role of congruence is more prominent for firms with less favorable internal and external conditions, including financial constraint, external financing by venture capital, government policy support and regional agglomeration economies.

4.3. Influencing channels

First, we hypothesize that higher levels of congruence can help firms alleviate the financial constraints, thus improving firms' innovation. To test this hypothesis, we separate our sample into two groups based on the size-age (SA) index measuring the firms' financial constraint (Hadlock & Pierce, 2010). The SA index is defined as SA = $-0.737 \times Size + 0.043 \times Size^2 - 0.040 \times Age$, with higher values implying higher levels of financial constraint. We separate the sample into two parts according to the within-year median of the SA index and conduct the baseline regression specifications in the two subsamples. The results are presented in Table 7.

[Insert Table 7 about here]

In Table 7, we examine two dependent variables, R&D intensity and patent applications. The regression results for the more financially constrained firms are presented in Columns (1) and (3) in Table 7, and Columns (2) and (4) contain results for less financially constrained firms. We find that the coefficient of congruence is significantly positive for the firms with higher financial constraint, and the magnitude is larger than that in the baseline regressions. In particular, increasing congruence by one standard derivation would increase R&D intensity by 23.5% and patents applications by 3.4%. For the subsample with lower levels of financial constraint, however, we find no significant effects of congruence on firm innovation. Therefore, the results support our hypothesis that the congruence is more necessary for the financially constrained firms.

Secondly, we examine the role of VC support as a form of external financing and value-added inputs in the effects of congruence on firm innovation. We hypothesize that the firms with external institutional support are less sensitive to congruence than those without it. To examine this, we separate our sample into two groups, firm with venture capital investment after listing on NEEQ, and those without venture capital investment. We conduct the same specification as the baseline regressions, and the results are reported in Table 8.

[Insert Table 8 about here]

In Table 8, the dependent variables are still firms' R&D intensity and patent applications. We find that the coefficient of congruence is only positively significant for the firms without VC investment, while it is not significant for the subsample with VC investment. In particular, for firms without VC support, increasing congruence by one standard derivation would increase R&D intensity by 15.8% and patents applications by 2.8%. Therefore, we conclude that the association between congruence and firm innovation outcome is more prominent in firms with less external financing. Such finding is also consistent with the analysis on firms' financial constraint, implying congruence can play a significant role in addressing the financial bottleneck for firms to conduct innovation.

We then test our final hypothesis that the congruence effect is more salient for firms in regions where the productivity levels of the corresponding industries are further away from the technology frontier. Following Acemoglu et al. (2006), we first use labor productivity to calculate the measure for the distance to the technology frontier, and then we separate our samples into two groups based on the median value of the distance. Table 9 presents the result, indicating that the effect of congruence is more pronounced for the group of firms that locate in an industry with a larger distance to the technology frontier. Specifically, increasing congruence by one standard derivation would increase R&D intensity by 18.2% and increase patent applications by 2.2%. This is consistent with the prediction in Hypothesis 3.

[Insert Table 9 about here]

5. Heterogeneity and extension

5.1. Heterogeneous effects

We conduct two additional tests on the heterogeneous effects. We find that the effects of congruence are more prominent for firms with a lower level of local inter-industry agglomeration and firms without support from government placed-based innovation policy.

First, we consider the heterogeneous impacts of inter-industry network. Specifically, we explore the inter-industry network through three perspectives: knowledge spillover, production network, and labor pooling. To investigate the heterogeneity in terms of the inter-industry network, we first separate our sample into two groups based on the median of the inter-industry network variables and then repeat our baseline regression to each group.

[Insert Table 10 about here]

The results are presented in Table 10 in which Panel A, B, and C refer to the categories of knowledge spillover, production network and labor pooling, respectively. The dependent variables are R&D intensity in the first two columns and patent applications in the last two columns. We find consistent results from these three panels. The effect of congruence is only significantly positive in the subsample of firms with lower levels of inter-industry agglomerations. In the subsample of high agglomeration, the coefficient of congruence is not significant. Such findings provide support for our hypothesis that the role of congruence is more prominent for firms with less favorable external environment in terms of the inter-industry agglomeration, and the role of congruence is substitutable rather than complementary to the role of external networks.

Second, we examine the heterogeneity in terms of the government place-based industrial policy support. In particular, we identify whether a firm is located in a high-tech special economic zone according to the firms' address information. In our sample, about 7% of the firm-year observations are in the high-tech SEZs. In such high-tech SEZs, firms are considered to enjoy government policies that support firm innovation in form of subsidies, tax credit, lower land price, etc., and hence firms may benefit from the geographical agglomeration of other high-tech firms in the SEZs. In this case, our conjecture is that the role of congruence is more prominent when firms face less favorable external economic environment, and thus more prominent for firms outside the high-tech SEZs. We conduct the subsample regressions according to whether firms are located in high-tech SEZs.

[Insert Table 11 about here]

Table 11 presents the result. From Columns (1) and (3), we find that for the firms outside high-tech SEZs, the level of congruence is positively associated with firms' R&D intensity and patent applications, and the magnitudes of the coefficients are close to those in the baseline regressions. In contrast, for the firms in the high-tech SEZs, we find no significant effect of congruence on firm innovation. Such results, together with the heterogeneity results of the inter-industry network variables, support our hypothesis that firms with worse external economic conditions, in terms of agglomeration or government policy, rely more on congruence to conduct innovation.[24]

5.2. Extension

In our last exercise, we extend our empirical analysis to investigate how congruence affects the likelihood of firms being listed successfully in the stock market, such as Shanghai Exchange Stock and Shenzhen Exchange Stock. This provides policy implication for the recent NEEQ reform. On September 3rd, 2021, Chinese President Xi Jinping announced a new reform—the formation of a Beijing Stock Exchange (BSE) to steer investment into innovation, making BSE become the third main stock exchange in China. A group of firms from the NEEQ will be select to IPO in BSE.

Although our sample does not include the period of BSE, we collect firms that were original listed on the NEEQ and have eventually graduated to the Shanghai, Shenzhen or Hong Kong

[24] We implement an additional test for the geographic heterogeneity based on the regional difference, and find that the role of congruence is more significant in central areas. The results are reported in Table A4 in our Appendix.

Stock Exchanges. Among all firms in our data sample, there are 251 firms that graduate successfully from NEEQ to the main stock markets by the end of 2020. We then repeat our baseline regression to examine the relation between congruence and the probability of main market IPOs. Table 12 reports the corresponding result and suggests that firms with higher congruence have higher chance to be public listed in the main stock markets. The statistical difference between Non-VC backed and VC-backed firms provides consistent evidence on influencing mechanisms of congruence.

[Insert Table 12 about here]

6. Conclusion

This paper examines the relationship between congruence and firm innovation. Our paper sheds light on innovation literature by proffering a new determinant: the congruence between the industry factor inputs structure and the regional endowment structure. Using a comprehensive dataset, we have found that congruence plays an important role in explaining firm innovation. The empirical analysis suggests that firms with higher congruence are more willing to invest in innovation and file more patent applications. We find that firms with financial constraints or without venture capital (VC) investment rely more on congruence. Moreover, the congruence effect is more salient for firms in regions where the productivity levels of the corresponding industries are further away from the technology frontier. Further tests suggest congruence associated with inter-industry network and government support drives heterogeneous effects. Lastly, we find that firms with higher congruence are more likely to IPO successfully in the stock market.

In sum, our empirical findings provide strong policy implications. It is true that technological advancement is crucial for a country losing its comparative advantage in labor-intensive sectors. However, if this country does not have a fully developed financial market, financing innovation through external sources become challenging. Therefore, policies may generate better results for supporting firms with higher congruence, because these firms have better business viability and stronger incentives to pursue innovation. For the future study, it is also promising to develop a theoretical framework linking congruence to innovation as well as interpreting its influencing mechanisms.

Acknowledgments

This work was supported by the National Office for Philosophy and Social Science of China (Project No. 20AJL017).

References

Acemoglu, D., Carvalho, V. M., Ozdaglar, A., & Tahbaz-Salehi, A. (2012). The network origins of aggregate fluctuations. *Econometrica*, *80*(5), 1977-2016.

Acemoglu, D., Aghion, P., & Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, *4*(1), 37-74.

Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, *42*(3), 422-448.

Areddy, James T. (2021). China to launch Beijing Stock Exchange to steer investment into innovation. *The Wall Street Journal*.

Audretsch, D. B., & Stephan, P. E. (1996). Company-scientist locational links: The case of biotechnology. *The American Economic Review*, *86*(3), 641-652.

Ayyagari, M., Demirgüç-Kunt, A., & Maksimovic, V. (2011). Firm innovation in emerging markets: The role of finance, governance, and competition. *Journal of Financial and Quantitative Analysis*, *46*(6), 1545-1580.

Basu, S., & Weil, D. N. (1998). Appropriate technology and growth. *The Quarterly Journal of Economics*, *113*(4), 1025-1054.

Berliant, M., Reed III, R. R., & Wang, P. (2006). Knowledge exchange, matching, and agglomeration. *Journal of Urban Economics*, *60*(1), 69-95.

Bernstein, S., Giroud, X., & Townsend, R. R. (2016). The impact of venture capital monitoring. *The Journal of Finance*, *71*(4), 1591-1622.

Bénabou, R., Ticchi, D., & Vindigni, A. (2015). Religion and innovation. *American Economic Review*, *105*(5), 346-51.

Bigio, S., & La'o, J. (2020). Distortions in production networks. *The Quarterly Journal of Economics*, *135*(4), 2187-2253.

Boldrin, M., & Levine, D. (2002). The case against intellectual property. *American Economic Review*, *92*(2), 209-212.

Carlino, G. A., Chatterjee, S., & Hunt, R. M. (2007). Urban density and the rate of invention. *Journal of Urban Economics*, *61*(3), 389-419.

Carlino, G., & Kerr, W. R. (2015). Agglomeration and innovation. *Handbook of Regional and Urban Economics*, *5*, 349-404.

Caselli, F., & Coleman II, W. J. (2006). The world technology frontier. *American Economic Review*, *96*(3), 499-522.

Chemmanur, T. J., Krishnan, K., & Nandy, D. K. (2011). How does venture capital financing improve efficiency in private firms? A look beneath the surface. *The Review of Financial Studies*, *24*(12), 4037-4090.

Chemmanur, T. J., Loutskina, E., & Tian, X. (2014). Corporate venture capital, value creation, and innovation. *The Review of Financial Studies*, *27*(8), 2434-2473.

Chen, Y., Podolski, E. J., Rhee, S. G., & Veeraraghavan, M. (2014). Local gambling preferences and corporate innovative success. *Journal of Financial and Quantitative Analysis*, *49*(1), 77-106.

Chuluun, T., Prevost, A., & Upadhyay, A. (2017). Firm network structure and innovation. *Journal of Corporate Finance*, *44*, 193-214.

Cull, R., Li, W., Sun, B., & Xu, L. C. (2015). Government connections and financial constraints: Evidence from a large representative sample of Chinese firms. *Journal of Corporate Finance*, *32*, 271-294.

Duchin, R., Ozbas, O., & Sensoy, B. A. (2010). Costly external finance, corporate investment, and the subprime mortgage credit crisis. *Journal of Financial Economics*, *97*, 418-435.

Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2006). Job-hopping in Silicon Valley: Some evidence concerning the microfoundations of a high-technology cluster. *The Review of Economics and Statistics*, *88*(3), 472-481.

Garmaise, M. J. (2008). Production in entrepreneurial firms: The effects of financial constraints on labor and capital. *The Review of Financial Studies*, *21*(2), 543-577.

Glaeser, E. L. (1999). Learning in cities. *Journal of Urban Economics*, *46*(2), 254-277.

Guariglia, A., & Liu, P. (2014). To what extent do financing constraints affect Chinese firms' innovation activities? *International Review of Financial Analysis*, *36*, 223-240.

Guo, D., & Jiang, K. (2013). Venture capital investment and the performance of entrepreneurial firms: Evidence from China. *Journal of Corporate Finance*, *22*, 375-395.

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., & Lerner, J. (2010). The financing of R&D and innovation. In *Handbook of the Economics of Innovation* (Vol. 1, pp. 609-639). North-Holland.

He, J., & Tian, X. (2021) Institutions and innovation: A review of recent literature. *Annual Review of Financial Economics*, *12*(2), 377-398.

Hochberg, Y. V., Ljungqvist, A., & Lu, Y. (2007). Whom you know matters: Venture capital networks and investment performance. *The Journal of Finance*, *62*(1), 251-301.

Hsu, D. H. (2006). Venture capitalists and cooperative start-up commercialization strategy. *Management Science*, *52*(2), 204-219.

Hsu, P. H., Tian, X., & Xu, Y. (2014). Financial development and innovation: Cross-country evidence. *Journal of Financial Economics*, *112*(1), 116-135.

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.

Jones, C. I. (2005). The shape of production functions and the direction of technical change. *The Quarterly Journal of Economics*, *120*(2), 517-549.

Ju, J., Lin, J. Y., & Wang, Y. (2015). Endowment structures, industrial dynamics, and economic growth. *Journal of Monetary Economics*, *76*, 244-263.

Kerr, W. R., & Nanda, R. (2015). Financing innovation. *Annual Review of Financial Economics*, *7*, 445-462.

Kortum, S., & Lerner, J. (2001). *Does venture capital spur innovation?* Emerald Group Publishing Limited.

León-Ledesma, M. A., & Satchi, M. (2019). Appropriate technology and balanced growth. *The Review of Economic Studies*, *86*(2), 807-835.

Lin, J. Y. (2009). *Marshall Lectures: Economic Development and Transition: Thought, Strategy, and Viability.* Cambridge University Press.

Lin. J. Y., Wang, F., Xia, J., and Zhang, H. (2021). What determines the success of industrial policy: evidence from China's special economic zones. Peking University working paper.

Lin, J. Y., & Wang, Y. (2018). Remodeling structural change. In *The Oxford Handbook of Structural Transformation*. New York: Oxford University Press.

Liu, E. (2019). Industrial policies in production networks. *The Quarterly Journal of Economics*, *134*(4), 1883-1948.

Manso, G. (2011). Motivating innovation. *The Journal of Finance*, *66*(5), 1823-1860.

McCarthy, N. (2020). China Is Closing the Gap with the US in R&D Expenditure. *Forbes*.

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*, *94*(3), 656-690.

Murata, Y., Nakajima, R., Okamoto, R., & Tamura, R. (2014). Localized knowledge spillovers and patent citations: A distance-based approach. *Review of Economics and Statistics*, *96*(5), 967-985.

Rajan, R., & Zingales, L. (1998). Financial dependence and growth. *American Economic Review*, 88, 559-586.

Rosenthal, S. S., & Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, *64*(2), 373-389.

Thompson, P., & Fox-Kean, M. (2005). Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*, *95*(1), 450-460.

Tian, X., & Wang, T. Y. (2014). Tolerance for failure and corporate innovation. *The Review of Financial Studies*, *27*(1), 211-255.

Van Den Berghe, L. A., & Levrau, A. (2002). The role of the venture capitalist as monitor of the company: A corporate governance perspective. *Corporate Governance: An International Review*, *10*(3), 124-135.

Viglione, G. (2020). China is closing gap with United States on research spending. *Nature*.

Xiao, S., & Zhao, S. (2012). Financial development, government ownership of banks and firm innovation. *Journal of International Money and Finance*, *31*(4), 880-906.

Zhu, X., Asimakopoulos, S., & Kim, J. (2020). Financial development and innovation-led growth: Is too much finance better? *Journal of International Money and Finance*, *100*, 102083.

Table 1: Summary Statistics of the Main Variables

Variable	Obs	Mean	Std. Dev.	Min.	Max.
Age	52,408	11.76	5.15	1	54
ROA (%)	43,234	4.79	16.17	-63.881	50.448
ROE (%)	42,741	4.81	29.64	-143.356	94.247
Sales growth (%)	43,198	22.54	70.59	-74.534	611.343
Number of patent applications	52,408	3.67	12.99	0	1892
Number of patent applications - invention	52,408	0.73	3.82	0	226
Number of patent applications - utility model	52,408	2.42	8.16	0	1225
Number of patent applications - design	52,408	0.52	4.85	0	577
Subsidy / Assets (%)	34,589	1.87	2.65	0.002	15.466
Subsidy / Sales (%)	34,597	3.21	5.86	0.001	39.419
R&D Intensity (R&D expenditure / Assets) (%)	43,150	5.09	6.26	0	35.452
Debt / Asset (%)	43,026	42.21	22.73	3.085	142.192

Note: The observations are restricted to the firms listed on the NEEQ since 2013, and the sample only includes observations of firms after being listed on the NEEQ.

Table 2: Within-Firm Changes before and after Being Listed on NEEQ

VARIABLES	(1) Subsidy/Assets	(2) Subsidy/Sales	(3) Equity Financing	(4) R&D Intensity
After	0.931***	1.446***	1.700***	0.565***
	(0.030)	(0.060)	(0.238)	(0.069)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Listing-year FE	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	83,363	83,363	73,739	73,989
R-squared	0.456	0.441	0.671	0.737

Note: The sample are restricted to the firms listed on NEEQ since 2013, and on the observations from 3 years before to 5 years after the year of listing. Equity financing refers to the ratio of equity to total liability. R&D intensity refers to the ratio of R&D expenditure to total assets. The dummy after indicates the observations since the year of listing for each firm. All regressions include firm fixed effects, year fixed effects, listing-year fixed effects, city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 3: Congruence and Innovation

Dependent variable	R&D intensity			Patent applications		
	(1)	(2)	(3)	(4)	(5)	(6)
Congruence	0.112**	0.132**	0.135***	0.025***	0.023***	0.021**
	(0.052)	(0.052)	(0.052)	(0.009)	(0.009)	(0.009)
Network controls	No	Yes	Yes	No	Yes	Yes
Firm-level controls	No	No	Yes	No	No	Yes
City-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	39,140	38,789	38,651	47,598	47,149	42,679
R-squared	0.458	0.456	0.460	0.281	0.279	0.299

Note: The dependent variable in Column (1), (2) and (3) is R&D intensity that is the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Column (4), (5) and (6) is patent applications measured by the logarithm of one plus the total number of patent applications, which are eventually granted and filed by a firm in a year. Congruence is standardized with a mean of 0 and a standard deviation of 1. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 4: Congruence and Firm Performance

	(1) ROA	(2) ROE	(3) Sales/worker	(4) Sales Growth
Congruence	0.329**	0.750***	0.025**	0.948**
	(0.154)	(0.276)	(0.011)	(0.465)
Network controls	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	38,960	38,516	38,957	38,891
R-squared	0.100	0.093	0.286	0.102

Note: The dependent variable is patent applications measured by the logarithm of one plus the total number of patent applications, which are eventually granted and filed by a firm in a year. Congruence is standardized with a mean of 0 and a standard deviation of 1. ROA is the ratio of firm profit to total assets. ROE is the ratio of firm profit to equity. Sales/work is the ratio of total sales to total employment. Sales growth is the growth of sales from year t-1 to year t. The three network indexes are all in logarithm. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 5: Robustness check – PSM

	R&D intensity			Patent applications		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
High congruence	0.187*	0.195**	0.195**	0.044**	0.043**	0.043**
	(0.098)	(0.096)	(0.095)	(0.018)	(0.018)	(0.017)
Network controls	No	Yes	Yes	No	Yes	Yes
Firm-level controls	No	No	Yes	No	No	Yes
City-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32,920	32,920	32,920	36,303	36,303	36,303
R-squared	0.466	0.467	0.472	0.284	0.284	0.300

Note: The dependent variable in Column (1), (2) and (3) is R&D intensity that is the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Column (1), (2) and (3) is patent applications measured by the logarithm of one plus the total number of patent applications, which are eventually granted and filed by a firm in a year. The treatment variable "high congruence dummy" is a dummy variable indicating the firm has a level of congruence higher than the median in the sample. The variables used for the propensity score matching include the network controls, the firm-level controls, and the city dummies as well as 1-digit industry dummies. The three network indexes are all in logarithm. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 6: Robustness Check – Firm-level congruence

	R&D intensity			Patent applications		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Firm congruence	0.164***	0.171***	0.157***	0.067***	0.068***	0.062***
	(0.045)	(0.046)	(0.045)	(0.007)	(0.007)	(0.007)
Network controls	No	Yes	Yes	No	Yes	Yes
Firm-level controls	No	No	Yes	No	No	Yes
City-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	38,983	38,637	38,547	39,148	38,802	38,609
R-squared	0.471	0.468	0.469	0.297	0.296	0.307

Note: The dependent variable in Column (1), (2) and (3) is R&D intensity that is the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Column (1), (2) and (3) is patent applications measured by the logarithm of one plus the total number of patent applications, which are eventually granted and filed by a firm in a year. Congruence is standardized with a mean of 0 and a standard deviation of 1. The three network indexes are all in logarithm. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

The mathematical expression in constructing firm congruence index is as follows:

$$congruence_{isct} = -\left[\left|\log\left(\frac{K_{isct}/L_{isct}}{K_s/L_s}\right) - \log\left(\frac{\overline{K_c}/\overline{L_c}}{\overline{K}/\overline{L}}\right)\right|\right]$$

Table 7: Mechanism Analysis: Financial Constraint

Dependent variable:	R&D intensity		Patent applications	
	(1)	(2)	(3)	(4)
	More Constrained	Less Constrained	More Constrained	Less Constrained
Congruence	0.235***	0.035	0.034**	0.017
	(0.082)	(0.081)	(0.016)	(0.013)
Network controls	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	12,950	12,502	12,950	12,503
R-squared	0.589	0.451	0.366	0.350

Note: The dependent variable in Columns (1) and (2) is R&D intensity measured by the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Columns (3) and (4) is patent applications measured by the logarithm of one plus the total number of patent applications, eventually granted and filed by a firm in a year. The subsample of firms in Columns (1) and (3) are the firm-year observations with SA index (what is it ?) higher than the cross-sectional median within a year, and those in Columns (2) and (4) are firm-year observations with SA index lower than or equal to the cross-sectional median within a year. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 8: Mechanism Analysis: Venture Capital

Dependent variable:	R&D intensity		Patent applications	
	(1)	(2)	(3)	(4)
	Non-VC-Backed	VC-Backed	Non-VC-Backed	VC-Backed
Congruence	0.158***	0.074	0.028***	0.002
	(0.056)	(0.112)	(0.009)	(0.024)
Network controls	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	30,206	7,778	34,193	7,818
R-squared	0.470	0.521	0.308	0.407

Note: The dependent variable in Columns (1) and (2) is R&D intensity measured by the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Columns (3) and (4) is patent applications measured by the logarithm of one plus the total number of patent applications, eventually granted and filed by a firm in a year. The subsample of firms in Columns (2) and (4) are the firms that receive venture capital (VC) financing, and those in Columns (1) and (3) are firms without VC financing. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls include age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 9: Mechanism Analysis: Technology Selection

Dependent variable:	R&D intensity		Patent applications	
	(1)	(2)	(3)	(4)
	Small	Large	Small	Large
Congruence	0.062	0.182**	0.021	0.022**
	(0.080)	(0.076)	(0.017)	(0.011)
Network controls	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	19,045	18,998	19,045	18,998
R-squared	0.465	0.487	0.465	0.487

Note: The dependent variable in Columns (1) and (2) is R&D intensity that is the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Columns (3) and (4) is patent applications that is the logarithm of one plus the total number of patent applications, eventually granted filed by a firm in a year. The small and large subsamples comprise firms with the measure for the distance to the technology frontier below and above the sample median. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls includes age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 10: Heterogeneity Analysis: Inter-Industry Agglomeration

	R&D intensity		Patent applications	
	(1)	(2)	(3)	(4)
	High	Low	High	Low
A. Knowledge spillover				
congruence	0.118	0.116*	0.014	0.029**
	(0.078)	(0.069)	(0.012)	(0.012)
Observations	19,293	18,824	23,889	22,615
R-squared	0.457	0.502	0.324	0.270
B. Production network				
congruence	0.043	0.143**	0.005	0.027**
	(0.085)	(0.070)	(0.012)	(0.013)
Observations	16,630	16,891	18,794	18,778
R-squared	0.471	0.522	0.379	0.320
C. Labor pooling				
congruence	0.066	0.157**	0.008	0.030**
	(0.077)	(0.079)	(0.013)	(0.012)
Observations	17,592	16,013	19,848	17,805
R-squared	0.450	0.534	0.346	0.341
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes

Note: The dependent variable in Columns (1) and (2) is R&D intensity measured by the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Columns (3) and (4) is patent applications measured by the logarithm of one plus the total number of patent applications, eventually granted filed by a firm in a year. We divide the firm-level sample according to the median of three inter-industry agglomeration measures, knowledge spillover in Panel A, production network in Panel B, and labor pooling in Panel C, respectively. The subsample of firms in Columns (1) and (3) are the firms with higher levels of agglomeration, and those in Columns (2) and (4) are firms with lower levels of agglomeration. All regressions include city-by-year fixed effects and industry-by-year fixed effects. Firm-level controls include age, size, ROA and leverage ratio. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Table 11: Heterogeneity Analysis: High-Tech Special Economic Zones

Dependent variable:	R&D intensity		Patent applications	
	(1)	(2)	(3)	(4)
	Outside High-Tech SEZs	In High-Tech SEZs	Outside High-Tech SEZs	In High-Tech SEZs
Congruence	0.151***	-0.022	0.020**	0.004
	(0.055)	(0.164)	(0.009)	(0.044)
Network controls	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes
Observations	35,914	2,357	39,724	2,577
R-squared	0.464	0.490	0.307	0.377

Note: The dependent variable in Columns (1) and (2) is R&D intensity that is the logarithm of one plus a firm's R&D expenditure scaled by total assets. The dependent variable in Columns (3) and (4) is patent applications that is the logarithm of one plus the total number of patent applications, eventually granted filed by a firm in a year. The subsample of firms in Columns (2) and (4) are the firms that are located in high-tech special economic zones (SEZs) according to their address information, and those in Columns (1) and (3) are firms outside high-tech SEZs. Network controls include the measures for knowledge spillover, production network and labor pooling. Firm-level controls includes age, size, ROA and leverage ratio. All regressions include city-by-year fixed effects and industry-by-year fixed effects. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

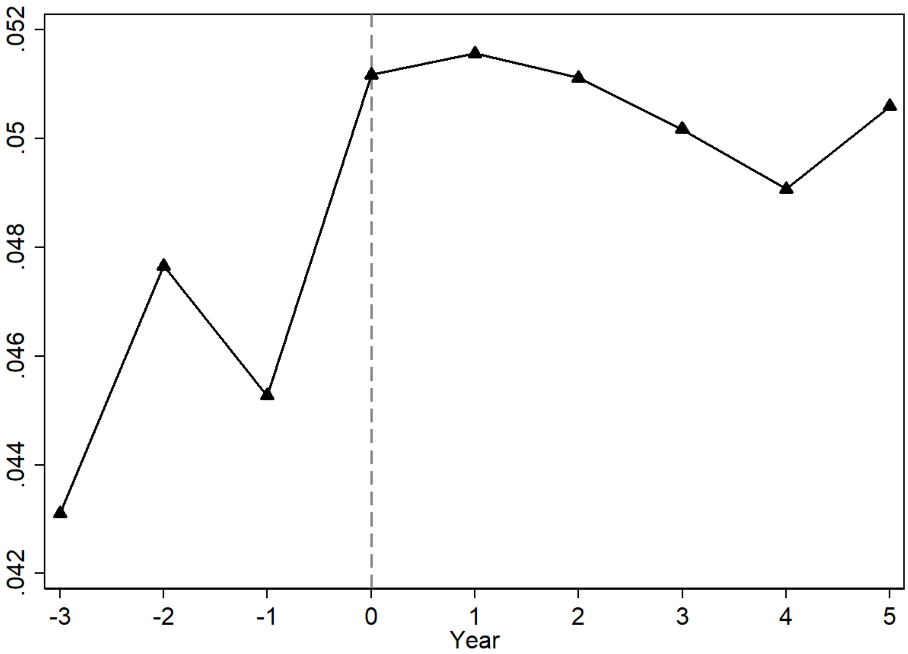
Table 12: Congruence and Main Stock Market IPO

	(1)	(2)	(3)
	Full sample	Non-VC sample	VC sample
Congruence	0.0022**	0.0023*	-0.0005
	(0.0011)	(0.0012)	(0.0021)
Network controls	Yes	Yes	Yes
City FE	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes
Observations	11,362	9,690	1,570
R-squared	0.049	0.055	0.170

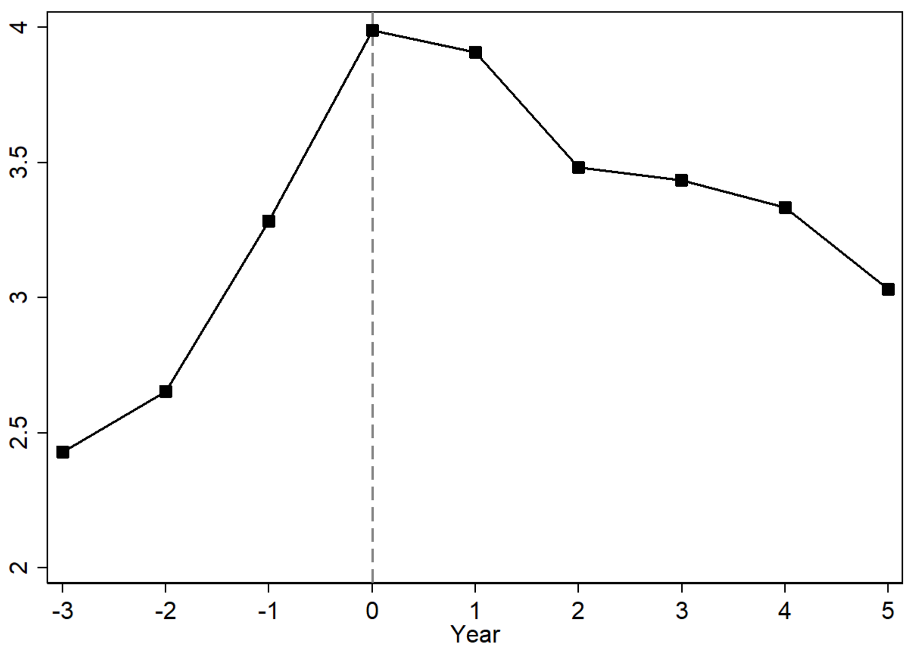
Note: The dependent variable is an indicator, equaling one if a firm has been publicly listed in the Chinese main stock markets by the end of 2020. Network controls include the measures for knowledge spillover, production network and labor pooling. All regressions include city fixed effects and industry fixed effects. Column (1) include all firm observations. Column (2) include the sample in which firms do not receive VC investment. Column (3) include the sample in which firms receive VC investment. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.

Figure 1: Average Patent Applications before and after the Year of Listing on the NEEQ

(a) R&D Intensity



(b) Patent Applications



Note: Figure (a) plots the average R&D intensity (R&D expenditure over total assets) in the years before and after the year being listed on the NEEQ. Figure (b) plots the average number of patent applications the firms file (and eventually granted) in the years before and after the year being listed on the NEEQ.

Appendix

A1. Technical appendix for the construction of network indices

In this section, we describe the way to construct the inter-industry network measures. As described in Section 3.1 in our paper, we calculate $\text{network}_{sc}^{\text{inter}} = \log(\sum_r w_{sr}^{inter} \times share_{rc})$ where $w_{sr}^{inter} = \frac{inter_{sr}}{\sum_r inter_{sr}}$ is the weight. The way to construct the inter-industry connectivity measures, KS_{sr}, LP_{sr}, and IO_{sr}, is as follows:

1) Knowledge spillover (KS_{sr}): we use the inter-industry patent citations to measure knowledge spillover. Specifically, we base our calculation on the sample of the firm tax survey data, match the firms' patent information with the Incopat patent database, which contains the citation information of each pair of patents. We aggregate the inter-firm patent citations to the industry level. Denote $PatentOut_{s\rightarrow r}$ as the share of industry r in the forward citations of patents in the industry s, where $\sum_r PatentOut_{s\rightarrow r} = 1$, and $PatentIn_{s\rightarrow r}$ as the share of industry s in the backward citations of industry r, where $\sum_s PatentIn_{s\rightarrow r} = 1$. Then the inter-industry knowledge spillover connectivity between industry s and industry r is defined as $KS_{sr} = \max\{PatentIn_{s\leftarrow r}, PatentIn_{r\leftarrow s}, PatentOut_{s\rightarrow r}, PatentOut_{r\rightarrow s}\}$.

2) Labor pooling (LP_{sr}): we calculate the correlation between the occupation structures in industry s and industry r. Specifically, based on the 0.35% random sample of the 2010 population census dataset in China (409 occupation codes in total), we calculate the occupation structure of each industry, i.e., the share of each occupation code in each industry's employment. We define $occshare_{so}$ as the fraction of industry s's employment in occupation o, and we measure the similarity of the occupational structure between industries s and r through the correlation of $occshare_{so}$ and $occshare_{ro}$.

3) Input-output (IO_{sr}): we calculate the input-output linkages between industries based on the 135-sector input-output table published by the National Bureau of Statistics (NBS) in China in 2012. We map the 3-digit industry code used in our analysis to the industry code of the

NBS input-output table. We define $Input_{s \leftarrow r}$ as the share of industry s's inputs that come from industry r, and define $Output_{s \rightarrow r}$ as the share of industry s's outputs that are sold to industry r. Then the input-output linkage between the two industries s and r is defined as $IO_{sr} = \max\{Input_{s \leftarrow r}, Input_{r \leftarrow s}, Output_{s \rightarrow r}, Output_{r \rightarrow s}\}$。

A2. Tables and figures

Table A1: Number of Observations

Year	Before Listing on NEEQ	After Listing on NEEQ	Total Number of Observations
2013	7,449	148	7,597
2014	9,882	1,314	11,196
2015	7,469	4,702	12,171
2016	2,902	9,634	12,536
2017	889	11,788	12,677
2018	335	12,361	12,696
2019	86	12,609	12,695

Note: The table reports the number of observations of firms before and after they are listed on the NEEQ by year in our dataset. We restrict the sample to the firms being listed on the NEEQ since 2013.

Table A2: Industry Distribution

Rank	Industry name	Numbers of firms	Percentage (%)
1	Software and information technology	1793	14.22
2	General and special equipment	797	6.32
3	Computer, communication and other electronic equipment	778	6.17
4	Electrical machinery and equipment	693	5.50
5	Business service	638	5.06
6	Chemical raw materials and products	569	4.51
7	Internet services	545	4.32
8	General equipment	515	4.08
9	wholesale	391	3.10
10	Professional scientific and technical service	378	3.00
11	Medical and pharmaceutical products	334	2.65
12	Nonmetallic mineral products	291	2.31
13	Apparatus and instrumentation	264	2.09
14	Rubber and plastics product s	263	2.09
15	Metallic products	259	2.05
16	Auto industry	233	1.85
17	Retailing	220	1.74
18	Agricultural and food processing	211	1.67
19	Environmental protection	180	1.43
20	Civil engineering construction	163	1.29

Table A3: Regional Heterogeneity Analysis

Dependent variable:	R&D Intensity			Patent Applications		
	(1)	(2)	(3)	(4)	(5)	(6)
	Eastern region	Central region	Western region	Eastern region	Central region	Western region
Congruence	0.155***	0.378**	0.088	0.015	0.077***	0.025
	(0.060)	(0.156)	(0.201)	(0.011)	(0.022)	(0.029)
Network controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm-level controls	Yes	Yes	Yes	Yes	Yes	Yes
City-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	32,509	6,360	3,186	32,509	6,360	3,186
R-squared	0.309	0.416	0.469	0.305	0.377	0.371

Note: Congruence is standardized with a mean of 0 and a standard deviation of 1. The three network indexes, including knowledge spillover, production network and labor pooling, are all in logarithm. For each dependent variable, three regressions are conducted based on the subsample of firms in Eastern region, Central region, and Western region, respectively. We control for city-by-year fixed effects and industry-by-year fixed effects, as well as firm-level characteristics including age, size and leverage ratio in all the regressions. The standard errors in parentheses are clustered at the city-by-industry level. ***, **, and * stand for the significance level of 1%, 5% and 10%, respectively.