China Economic Review 67 (2021) 101605

Contents lists available at ScienceDirect

China Economic Review

journal homepage: www.elsevier.com/locate/chieco

Establishing science parks everywhere? Misallocation in R&D and its determinants of science parks in China

Chih-Hai Yang [a,*], Wen-Chieh Lee [b]

[a] *National Central University, Taiwan*
[b] *National Chengchi University, Taiwan*

ARTICLE INFO

JEL classification:
H76
O33
O38
R11
R58

Keywords:
Science Park
R&D
Misallocation

ABSTRACT

The establishment of science parks is a vital strategy to develop high-tech industries and facilitate innovations in China. The success of a science park depends heavily on its supportive environment, suggesting that it is hard to replicate everywhere, while China has established many science parks across regions in the past decade. This study evaluates the degree of misallocation in research and development (R&D) and its determinants across science parks in China. Based on an unbalanced panel data of 145 science parks for the period 2007–2014, we find that the overall R&D efficiency has decreased sharply since 2011 when China began to initiate many new science parks. The newly constructed science parks exhibit a lower R&D efficiency than their incumbent parks, suggesting a considerable misallocation in R&D resource caused by expanding science parks everywhere. We further investigate the determinants of R&D misallocation and find that park characteristics and environmental characteristics matter. Parks which are larger, older, and having a higher human quality experience a lower R&D misallocation. Parks with closer R&D collaboration with universities or research institutes, particularly with universities, exhibit a lower R&D misallocation.

1. Introduction

Inspired by the success of spontaneous industrial clusters in the United States, such as Silicon Valley and Route 128, the idea of science parks is at the center of establishing high-tech industries, upgrading the technological ladder, and promoting regional economic growth in many countries. Consequently, an increased number of science parks (about 1500) were established worldwide up to the mid-2000s (Wainova, 2009), and the government developed most of them.

There are extensive taxonomies of science parks based on their structure and missions over the past four decades (Bigliardi, Dormio, Nosella, & Petroni, 2006, Fig. 1). One consensus, reached after the 1990s, is that the establishment of science parks favors locations near universities and public research institutes. The frequent interaction among the innovation triple helix of industry-university-government relations generates technological externality and agglomeration economies (Fan & Scott, 2003; Westhead & Batstone, 1998), which promotes innovation and regional economic growth. However, Massey, Quintas, and Wield (1992) and Quintas, Wield, and Massey (1992) question the effectiveness of the science park model on facilitating innovation and describe science parks as high-tech fantasies. That is, science parks might have no quantitatively significant impact on their tenants' performance because science parks do not encourage on-park firms to create synergies (Macdonald, 1987), whereas more innovative firms self-select to enter science

* Corresponding author at: Department of Economics, National Central University, 300, Jhongda Road, Jhongli District, Taoyuan 320, Taiwan.
E-mail addresses: chyang@mgt.ncu.edu.tw (C.-H. Yang), jffl1803@nccu.edu.tw (W.-C. Lee).

https://doi.org/10.1016/j.chieco.2021.101605
Received 6 February 2020; Received in revised form 17 June 2020; Accepted 14 February 2021
Available online 20 February 2021
1043-951X/

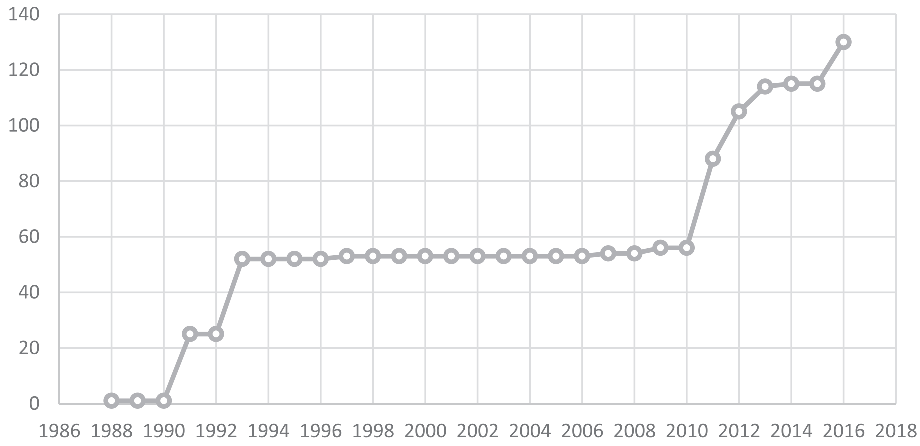

Fig. 1. The Number of STIPs in China, 1988–2015.

parks.

Extant empirical studies also show mixed findings on innovation performance differentials between in-park firms and external counterparts, e.g., Colombo and Delmastro (2002), Diez-Vial and Fernández-Olmos (2017), Fukugawa (2006), Liberati, Marinucci, and Tanzi (2015), Löfsten and Lindelöf (2002), Vásquez-Urriago, Barge-Gil, Rico, and Paraskevopoulou (2014), and Yang, Motohashi, and Chen (2009). Albahari, Barge-Gil, Pérez-Canto, and Modrego-Rico (2018) illustrate science parks' heterogeneity being relevant to their tenants' performance, suggesting that, through services provision and location, science park characteristics can mediate and facilitate the advantages of cluster dynamics; thereby affecting science park performance. The implication is that the effectiveness of science parks depends on their heterogeneity. However, comparative analyses on various aspects of performance across science parks are almost absent: there are only a few science parks in most countries, which prevents rigorous econometric analyses.

Although the evidence of success is not robust, science parks remain a widely-adopted policy tool in emerging economies. From an economic-geographical perspective (where a science park and its surrounding region form an entity), those parks with high-tech firms have an evolving structure of inter-firm linkages and agglomerative effects that help promote R&D productivity and output. Developing technological capability through the cluster of high-tech startups in a superior innovation environment can be an authority's primary policy. Based on this assumption, China has established more than 100 science parks in the past two decades.

If the science park policy is successful, can government transplant this success to other regions? Indeed, China's local governments generally consider science parks as "development catalysts" rather than "innovation catalysts." Both employment and income effects brought about by the establishment of science parks are the key goals of local government, as regional economic growth is the primary concern (Hu, 2007). The widespread setups of science parks thus comprise many seemingly science parks, which lower the functions science parks are expected to play in facilitating innovation.[1] The concept of science parks has become a popular policy tool that enhances innovation and regional development; evaluating the performance of science parks and identifying their determinants are the essential issues of the science parks policy.

However, R&D resource misallocation, an essential aspect of understanding how well a science park operates, is yet to be investigated in the literature. Restuccia and Rogerson (2008) indicated that the allocation of resources across firms may be a significant factor in affecting a country's total factor productivity (TFP). Resource misallocation thus has been identified as a significant source of variation in productivity in emerging economies (Syverson, 2011). In developing countries, R&D is one of the scarce resources for innovation, promoting productivity, and raising the value added of production. Better use of scarce resources can efficiently boost productivity and sustain long-term growth. Directing R&D resource toward government initiated science parks that are located in cities without favorable innovation environments, and well-functioning regional innovation can potentially result in R&D resource misallocation; raising the need for research on the evaluation of different efficiencies in the allocation of resources across science parks. Together, with an understanding of the determinants of R&D misallocation, such research can provide insight into the implications of technology and regional policies.

This study aims to evaluate R&D misallocation and its determinants across science parks in China. It adds three novelties to this line of literature. First, as some countries have established dozens of science parks in the past few decades,[2] studies comparing performance across science parks have emerged recently, but have remained limited. For example, Hu, Han, Yeh, and Lu (2010) and Izadikhah and Saen (2015) adopt the data envelopment analysis (DEA) approach to evaluate the efficiency of science parks in China and Iran, respectively. Unlike previous studies focusing on evaluating performance differentials between in-park and off-park firms, this park-level study helps understand the degree and disparity in R&D misallocation across science parks. A counter-factual welfare analysis is

[1] Luger and Goldstein (1991) argue that there is no pure science or research park, because STPs shoulder multiple missions.

[2] Please see statistics of the United Nations Educational, Scientific, and Cultural Organization (UNESCO) for science parks around the world. http://www.unesco.org/new/en/natural-sciences/science-technology /university-industry-partnerships/science-parks-around-the-world/

also conducted to simulate the loss of R&D outputs due to research input misallocation on the new established parks.

Second, evaluating the performance of science parks is a rather complex undertaking. Instead of using the DEA approach to assess the technical efficiency of science parks in existing studies, this study evaluates the R&D resource misallocations of science parks. Analyzing resource misallocation is an emerging field of literature, while it neglects the misallocation of innovation resource, except for Li, Lee, and Ko (2017). Establishing many science parks in unfavorable innovation environments (regions) might lead to a distortion of the R&D resource usage and an association with a larger R&D misallocation. Building on Hsiech and Klenow's (2009) methodology into misallocation, we calculate the R&D efficiency of science parks. This efficiency measure assumes a larger value if the dispersion of revenue productivity, which is a function of a science park's input and output distortions, is smaller across science parks. This misallocation measure on "revenue productivity" is called TFPR in Foster, Haltiwanger, and Syverson (2008) and Hsiech and Klenow's (2009).

Third, on computing the R&D efficiency measure, we have also obtained individual science park's TFPR, which can be viewed as an inverse measure of distortion (misallocation). We further investigate the determinants of misallocation across science parks. As the science park theory is underpinned by the advantage of agglomeration economies and the network externality of the innovation triple helix, existing studies focus on factors, like university proximity, interaction within the triple helix, and others, for example, Link and Scott (2003), Fritsch and Franke (2004), Yang et al. (2009), and Jongwanich, Kohpaiboon, and Yang (2014). Even though science park characteristics are a crucial dimension of the influential factors affecting performance, they are not well examined. Albahari et al. (2018) indicate that heterogeneity of science parks is relevant to their tenants' performance and affects park performance. However, how science parks' characteristics relate to their performance is less investigated. This study explores how the park's characteristics relate to science parks performance in terms of R&D resource misallocation. Using returns to scale as a conceptual foundation, we also explore whether and how the growth R&D inputs relate to R&D misallocation.

The remainder of this paper is organized as follows. Section 2 briefly introduces the development and importance of science parks in China. Section 3 presents the methodology and demonstrates the data source. Section 4 presents and discusses the R&D misallocation across science parks in China and conduct a counter-factual welfare analysis to simulate the R&D misallocation if considering non-park regions. Section 5 illustrates the empirical results of the determinants of R&D misallocation. Robustness checks are also conducted. The final section summarizes the concluding remarks and policy implications.

2. The development and importance of science parks in China

2.1. Development of science & technology industrial parks in China

China's science park policy was initiated in 1988 when the State Council implemented the Torch Program that aims to promote innovations and develops high-tech industries. One of the foremost strategies is establishing National High-tech Industrial Zones (hereafter, NHIZs or science parks). Like science parks in advanced countries, the NHIZs were established in suitable locations, neighboring renowned universities and research institutes, to provide a catalytic-incubator environment. Through stringent entry criteria, together with liberal policy measures such as tax incentives, R&D grants, and financial assistance to select and support R&D-intensive high-tech startups, NHIZs are tasked with transforming industrial structure by supporting high-tech firm startups and fostering technological diffusion (Hu, 2007).

Zhongguancun Science Park was the first national NHIZ that was launched in Beijing in 1988. At that time, Beijing was not an economic heavyweight of high-tech firms: its selection was attributed to the cluster of research universities and institutions and, probably also the capital bias from a political correctness point. There followed two waves in the establishment of NHIZs in the 1990s and 2010s.

In 1991, the State Council approved the establishment of the first group of 24 NHIZs and then approved another 27 NHIZs in 1993, comprising a total of 52.[3] As depicted in Fig. 1, the number of NHIZs increased slowly to 53 in 1997, and 54 in 2007. The first-wave of the rapid upsurge in NHIZs was in response to Deng Xiaoping's 1992 "South Trip" talk that declared the intent to speed up the pace of the open-door policy. NHIZs are utilized as a policy measure to promote domestic high-tech startups as well as attract the inflow of foreign high-tech firms. The sites of these 54 NHIZs are located either in the largest cities or in business cities to shape a favorable environment.

The period from 1997 to 2007 was a golden decade for China's economic development. The achievement of persistent high economic growth was the most urgent and immediate priority for individual provinces. The building of provincial science and technology industrial parks and specialized industrial bases to accommodate domestic and foreign firms helps provinces to create jobs and increase output, thereby helping to reach provincial GDP growth targets. The State Council thus stopped approving the establishment of new NHIZs and instead focused on reorganizing existing NHIZs, resulting in a stable number of NHIZs.

At the center of China's 11th Five-Year Plan (2006–2010) initialized in 2006, one of the targets was scientific development, aimed at achieving technological self-reliance and becoming an innovation-oriented nation. However, China's economy was also severely affected by the 2008 financial crisis. The government rethought the NHIZs' key role in facilitating high-tech industry development and sustaining economic growth. It upgraded two provincial industrial parks to the national-class science parks in 2009 and then approved the second-wave of NHIZs by improving existing provincial industrial parks or specialized industrial bases. Fig. 1 illustrates that the

[3] The Tiananmen Square protests of 1989 delayed the progress of STIPs project by a few years.

number of NHIZs increased sharply from 56 in 2009 to 83 in 2010, 88 in 2011, 105 in 2012, and 114 in 2013, and finally stabilized at 115 in 2014 and later.

After the second-wave increase, NHIZs are now widespread in every province, except for Tibet. However, their geographical distribution is unequal and is concentrated mainly in coastal regions, followed by central regions. As demonstrated in Table 1, Jiangsu, Shandong, and Guangdong rank the top three provinces by owning 15, 12, and 11 NHIZs respectively, whereas Beijing, Tianjin, Hainan, Guizhou, and Qinghai each have only one NHIZ. Among them, Beijing and Tianjin are municipalities; Guizhou and Qinghai are less developed provinces, and Hainan is an off-shore island.

As indicated in Columns (1) and (2) combined, the number of new NHIZs in the second-wave (91) is much more than the constructed NHIZs in the first-wave (54). Crucially, without a well-developed innovation system, local governments may prioritize science park initiatives to promote economic growth, meaning that the second-wave established or upgraded NHIZs might locate in an innovation environment without enough triple helix interactions that the first-wave NHIZs have due to their location. For example, both Beijing and Qinghai each established only one NHIZ, while these two NHIZs might have diverse performance due to location, park characteristics, and others. It casts the doubt on whether it is necessary to initiate NHIZs everywhere, and to what extent there is R&D resource misallocation.

Along with the increase in the number of NHIZs and the number of firms in each park, NHIZs have played an emerging and critical role in the Chinese economy. Table 2 shows that the output increased approximately 60 times from RMB 310.95 billion in 1997 to RMB 18,601.83 billion in 2015. The corresponding share of GDP increased from 3.90% to 27.14%. Commodities exported from NHIZs accounted for US$ 6.48 billion in 1997, and 4.08% of total exports. By integrating itself into the world trading system after entering to the World Trade Organization (WTO) in 2001, China acts as the primary manufacturing center for the East Asian production network and has gradually become the leading exporter of a variety of products. It has seen a sharp increase in export value to US$ 473.27 billion in 2015, accounting for 20.82% of total exports.

Because the design of NHIZs accommodates R&D-intensive high-tech firms, the tenants naturally devote more efforts to R&D than their outside counterparts. The outlay of R&D rocketed by more than 29 times from an initial RMB 15.52 billion in 2000 to RMB 452.16 billion, in 2015, experiencing an average annual growth rate of 26.29%. The far right column of Table 2 illustrates that NHIZs play a critical role in China's private R&D activity. Their R&D expenditure accounted for 26.30% of total business R&D expenditure in 2000 and, then, increased to an extraordinarily high ratio of 51.66%, in 2007. One surprising phenomenon observed is that although the number of NHIZs has increased since 2009, the corresponding R&D ratio dropped considerably to 32.26% in 2009 and then steadily re-climbed to 42.70% in 2015. The possible causes are three-fold. First, the 11th Five-Year Plan (2006–2010) executed other projects for establishing university technology zones, software zones, and zones for returned talented people to start businesses. The springing up of new parks also accommodated R&D-intensive firms. Second, in the context of becoming a knowledge economy, firms located outside NHIZs began to focus more on R&D activity. Finally, and most interestingly, from an R&D perspective, is resource misallocation. The second-wave upgraded NHIZs might spend less on R&D expenditure because their locations are not so favorable for the innovation network; tenants thus concentrate more on production.

3. Methodology and data

3.1. Measure of R&D misallocation across science parks

The seminal work in Hsieh and Klenow (2009) develops the concept of input distortion in producing outputs to calculate resource misallocation. By adopting the core ideas of Hsieh and Klenow's (2009) methodology, Li et al. (2017) estimate a knowledge production function to evaluate the so-called innovation resource misallocation. Thus, this paper measures the R&D misallocation by estimating a knowledge production function to compute R&D efficiency across science parks in China.

Assuming that there is a competitive innovation system in China where R&D resources are employed to produce a homogeneous product in terms of value-added.[4] Y is the aggregate value-added across M science parks in China and can be specified as follows:

$$Y = \sum_{i=1}^{M} Y_i \tag{1}$$

where Y_i is the value-added of science park i. We assume the NHIZ's innovation production technology is determined by the following decreasing return-to-scale technology[5]:

$$Y_i = A_i \left(L_i^{\alpha} K_i^{1-\alpha} \right)^{\gamma}, \gamma \in (0, 1) \tag{2}$$

where γ governs a park's "operative returns to scale" in the innovation system.

Science parks within the country could be heterogeneous in both their innovative technology A_i and in the distortions associated

[4] In innovation literature, the number of patents is widely adopted as a measure of innovation output, for example, Griliches (1990) and Nagaoka, Motohashi, and Goto (2010), where it is obvious that patents are heterogeneous. Either the R&D outputs are processes or products, it aims to raise the value-added of outputs. Hence, value-added is an adequate measure, which is in concordance with Hsieh and Klenow's (2009) concept.

[5] The innovation production function is assumed to be decreasing returns to scale, as characterized in Jones and Williams (2000) and Weil (2013). This operative return-to-scale is also referred to as the "span-of-control" parameter described in Lucas (1978). Hsieh and Klenow (2009) have proved the isomorphic property between a constant returns to scale model with differentiated goods and a Lucas span of control model formation.

Table 1
Geographical Distribution of STIPs in China, 2015.

Province	-1	-2	Region
	Total Number of NHIZs	Number of NHIZs in 2007	
Jiangsu	15	4	East Coast
Shandong	12	5	East Coast
Guangdong	11	6	East Coast
Zhejiang	8	2	East Coast
Liaoning	8	3	East Coast
Fujian	7	2	East Coast
Hubei	7	2	Center
Henan	7	2	Center
Sichuan	7	2	West
Jiangxi	7	1	Center
Shaanxi	7	3	Center
Hunan	6	2	Center
Hebei	5	2	East Coast
Jilin	5	2	North
Anhui	4	1	Center
Guangxi	4	2	West
Heilongjiang	3	2	North
Xinjiang	3	1	West
Shanghai	2	1	East Coast
Shanxi	2	1	Center
Neimenggu	2	1	North
Gansu	2	1	West
Ningxia	2	0	West
Chongqing	2	1	West
Yunnan	2	1	West
Beijing	1	1	East coast
Tianjin	1	1	East Coast
Hainan	1	1	East Coast
Guizhou	1	1	West
Qinghai	1	0	West
Tibet	0	0	West
Total	145	54	

Source: calculated by the authors.

Table 2
Basic Statistics of STIPs.

Year	Output (RMB billion)	Export (US$ billion)	R&D (RMB billion)	Output to GDP ratio (%)	Export to total exports ratio (%)	R&D to business R&D ratio (%)
1997	310.92	6.48	n.a.	3.9	4.08	n.a.
1998	433.36	8.53	n.a.	5.09	5.22	n.a.
1999	594.4	11.9	n.a.	6.56	6.81	n.a.
2000	794.2	18.58	15.54	7.92	8.31	26.3
2001	1,011.68	22.66	22.18	9.13	9.45	35.22
2002	1,293.71	32.92	31.45	10.63	11.08	39.91
2003	1,725.74	51.02	41.95	12.56	12.65	43.67
2004	2,263.89	82.38	61.38	13.99	14.9	47.53
2005	2,895.76	111.65	80.62	15.46	15.66	49.08
2006	3,589.90	136.1	106.42	16.36	14.86	51.32
2007	4,437.69	172.81	134.88	16.42	14.95	51.66
2008	5,268.47	201.52	166.82	16.49	14.9	50.38
2009	6,115.14	200.72	134.27	17.52	17.63	32.26
2010	8,431.82	264.8	181.25	20.41	23.14	35.8
2011	10,567.96	318.06	226.9	21.6	16.8	35.34
2012	12,860.39	376.04	274.91	23.8	18.4	36.05
2013	15,136.76	413.33	348.88	25.43	18.7	39.48
2014	16,993.69	435.14	399.57	26.39	18.5	40.7
2015	18,601.83	473.27	452.16	27.14	20.82	42.7

Note: n.a. denotes data being unavailable.

with the use of R&D capital and R&D labor. Following the assumptions in Li et al. (2017), NHIZs might experience two types of distortions: output distortion d_{Yi}, which simultaneously affects productivity of R&D capital and R&D labor; R&D capital distortion d_{Ki}, which drives up the productivity of R&D capital relative to that of R&D labor. A science park's innovation payoff is given by

$$\pi_i = (1 - d_{Y_i})P_iY_i - wL_i - (1 + d_{K_i})RK_i \tag{3}$$

Under the assumptions of a competitive R&D input market and a homogeneous output market, the price of a product is as follows, assuming the standard first-order condition on demand holds:

$$P_i = P, \text{for every science parki} \tag{4}$$

To solve the R&D input demand in perfectly competitive factor markets: the derived demand from (4) can be substituted into (3) to solve for both the input demand and the output supply, whose values are determined by the innovation technology A_i and the distortion measures d_{Y_i} and d_{K_i}:

$$L_i = L_i(A_i, d_{Y_i}, d_{K_i}) \tag{5}$$

$$K_i = K_i(A_i, d_{Y_i}, d_{K_i}) \tag{6}$$

Science parks experiencing greater output distortions (d_{Y_i}) and capital distortions (d_{K_i}) will respectively exhibit higher marginal revenue products of R&D labor and R&D capital, because their production pursues profit maximization. We denote the marginal revenue products as follows:

$$MRPL_i = MRPL_i(A_i, L_i, K_i) \approx w\frac{1}{1 - d_{Y_i}} \tag{7}$$

$$MRPK_i = MRPK_i(A_i, L_i, K_i) \approx R\frac{1 + d_{K_i}}{1 - d_{Y_i}} \tag{8}$$

Given the assumption of decreasing returns, highly distorted NHIZs will have an equilibrium scale of production that is smaller than the optimal scale.

Following the concept introduced in Foster et al. (2008) and Hsieh and Klenow (2009), Li et al. (2017) differentiate "physical productivity" from "revenue productivity". The former is denoted as TFPQ and it is NHIZ-specific, whereas the latter is TFPR and it is country-specific if there is no difference in the extent of the distortions across science parks. The reduced forms of TFPQ and TFPR for the science park i can be solved as follows:

$$TFPQ_i \approx w\frac{Y_i}{\left(L_i^{\alpha}K_i^{1-\alpha}\right)^{\gamma}} \tag{9}$$

$$TFPR_i \approx w\frac{PY_i}{L_i^{\alpha}K_i^{1-\alpha}} \tag{10}$$

If NHIZs can be initiated anywhere, that is, establishing science parks in many cities, TFPR will be country-specific, and there is no difference across science parks. However, TFPR could vary across science parks if they have different levels of output and capital distortions. Without the NHIZ-specific output and capital distortions ($TFPQ_i$), $TFPRi$ can be represented by the geometric average of a science park's marginal revenue products of labor and capital, as indicated in Hsieh and Klenow (2009) and Li et al. (2017). Specifically, we use Eqs. (7), (8), and (10) to show that a park's TFPR is, in effect, an indicator of the endured distortions:

$$\begin{aligned} TFPR_i = TFPR_i(d_{Y_i}, d_{K_i}) &\propto \left[\left(\frac{MRPL_i}{w}\right)^{\alpha}\left(\frac{MRPK_i}{R}\right)^{1-\alpha}\right]^{\gamma} \\ &\propto \left[(1 - d_{Y_i})^{\alpha}\left(\frac{1 - d_{Y_i}}{1 + d_{K_i}}\right)^{1-\alpha}\right]^{-\gamma} \end{aligned} \tag{11}$$

Because higher R&D outputs and larger R&D capital distortions raise the marginal products of R&D capital and R&D labor, the science park i will exhibit a smaller scale of output than the efficient scale if it experiences a larger distortion.

Aggregate R&D output (value-added) can be derived by simply aggregating the individual science park's R&D output production, as in eq. (1). Suppose we implicitly define the innovation production efficiency TFP of the country as a whole by:

$$Y = TFP \times L^{\alpha} \times K^{1-\alpha} \tag{12}$$

where $L = \sum_{i=1}^{M} L_i$and $K = \sum_{i=1}^{M} K_i$ represent the aggregate values of the R&D labor and R&D capital devoted to innovation activities, respectively. By simplifying the linear aggregate of the production function in the innovation system, the countrywide innovation production efficiency TFP can be represented by

$$TFP = \frac{Y}{L^{\alpha}K^{1-\alpha}} = \frac{\left[\sum_{i=1}^{M}\left(TFPQ_i\frac{\overline{TFPR}}{TFPR_i}\right)^{\frac{1}{1-\gamma}}\right]^{1-\gamma}}{\left(L_i^{\alpha}K_i^{1-\alpha}\right)^{1-\gamma}} \tag{13}$$

where $\overline{TFPR}$ is a harmonic average of the average marginal revenue product of R&D capital and R&D labor across science parks in

China.

Eq. (13) shows that the countrywide R&D TFP will be that of a CES function aggregated across all of the $TFPQ_i$ if the revenue productivity ($TFPR_i$) is equalized across all science parks in China. In this special case, TFP will be

$$TFP = \overline{A} = \frac{\left[\sum_{i=1}^{M} A_i^{\frac{1}{1-\gamma}}\right]^{1-\gamma}}{\left(L_i^{\alpha} K_i^{1-\alpha}\right)^{1-\gamma}} \tag{14}$$

3.2. Computation of misallocation

To compute R&D efficiency, we adopt the following exogenous parameters: First, we assume a rental rate of R&D capital being $R = 0.23$, following the assumption in Li et al. (2017). The rental rate is a combination of an interest rate (i) of 3% and a depreciation rate (δ) of 20%.[6]

Next, we turn to the choice of Lucas span-of-control parameter γ, and the labor share parameter α.[7] Previous studies on estimating the knowledge production functions, such as, Crépon and Duguet (1997), Ramani, EI-Aroui, and Carrere (2008), and Hu and Jefferson (2009), have shown that innovation production is governed by decreasing returns to scale. Accordingly, in studies estimating firm-level innovation production function, for example, Jefferson, Bai, Guan, and Yu (2006) and Yang (2018), finding the mean innovation elasticity of R&D inputs being approximately 0.8, we thus choose $\gamma = 0.8$.[8] The labor share parameter is assumed to be 0.6 based on the province-level evidence in Li (2009) and Bai (2013) that finds labor share in innovation production is approximately 0.6. Table 3 summarizes our parameter configuration.

Following Li et al. (2017), the idiosyncratic distortions in R&D labor and R&D capital adoption costs and TFPQs across science parks are denoted as follows:

$$\text{Capital distortion}: d_{K_i} = \frac{1-\alpha}{\alpha} \times \frac{wL_i}{RK_i} - 1 \tag{15}$$

$$\text{Output distortion}: 1 - d_{K_i} = \frac{1}{\alpha\gamma} \times \frac{wL_i}{RK_i} \tag{16}$$

In Eq. (15), if the ratio of the labor share to the capital share is greater than $\alpha/(1-\alpha)$, we can infer that R&D capital distortion exists; Eq. (16) illustrates that if the R&D labor share relative to the total output is smaller than $\alpha\gamma$, we have output (value-added) distortion. The $TFPQ_i$ measurement in Eq. (17) is conceptually similar to the TFP in a neoclassical production function. Indeed, these measurements of distortion and science parks productivities are the bases for us to gauge the efficiency loss of the innovation system for science parks.

Li et al. (2017) define the "efficient production" as the output level obtained when there are no idiosyncratic distortions across science parks. Under the optimal scenario, the marginal revenue products of the innovation inputs are equalized across science parks within the innovation system in China; thus,

$$TFPR_i = \overline{TFPR} \tag{17}$$

We can rewrite TFP in Eq. (14) as

$$\overline{A} = \frac{\left[\sum_{i=1}^{M} A_i^{\frac{1}{1-\gamma}}\right]^{1-\gamma}}{\left(L_i^{\alpha} K_i^{1-\alpha}\right)^{1-\gamma}} \tag{18}$$

The ratio of the actual and efficient production levels of innovation output is thus denoted as[9]:

$$Y_R = \frac{Y}{Y_{efficient}} = \left[\sum_{i=1}^{M} \left(\frac{A_i}{\overline{A}} \frac{\overline{TFPR}}{TFPR_i}\right)^{\frac{1}{1-\gamma}}\right]^{1-\gamma} \tag{19}$$

Eq. (19) demonstrates that the ratio Y_R increases as the dispersion of a science park's TFPR decreases. It reaches the maximum value (=1) when all science parks' marginal payoffs of R&D inputs are equalized.[10] Thus, Y_R can be viewed as a measure of innovation

[6] The interest rate is about 3% within the sample period in China. Moreover, the R&D depreciation is estimated to hover between 11% and 36% in various studies (Nadiri & Prucha, 1996).

[7] The span-of-control parameter (γ) records the operative returns to scale by labor (L) and fixed capital (K). This operative return-to-scale parameter is often referred to as the "span-of-control" parameter, as in Lucas (1978), Atkeson and Kehoe (2005) and many other studies. In the current context, our selection of γ can be viewed as replacing the elasticity of the substitution measure in Hsieh and Klenow (2009), and the gains from fewer distortions are increasing in γ.

[8] For robustness, we also consider alternative values of γ: 0.5 and 0.9. Different choices of γ will affect only the numerical values of the measured R&D productivity and not the relative ordering or the trend in productivity.

[9] The subscript "efficient" means the removal of all idiosyncratic barriers or frictions that cause disparities in the marginal products of labor and capital.

[10] This result has been shown in Hsieh and Klenow (2009).

Table 3
Parameters Used in Calibrations.

	α	R = δ + i		γ
		δ	i	
Parameter values	0.60	0.20	0.03	0.8

Note: The parameters are the same as assumed in Li et al. (2017).

efficiency.

3.3. Determinants of misallocation

After calculating the degree of R&D misallocation, this study next investigates the determinants of misallocation in science parks. Considering the cross-sectional feature of misallocation comparison, we specify the empirical models as follows

$$RD\_mis_{it} = \alpha + X_{it}\beta + Z_{it}\gamma + PROV\rho + YEAR\delta + \varepsilon_{it} \tag{20}$$

Here, the dependent variable *RD_mis* is the degree of misallocation, which is measured by $ln\left(TFPR_{it}/\overline{TFPR}_t\right)$, denoting the deviation of park *i*'s TFPR from the mean TFPR in year *t*. A higher value represents a larger distortion (misallocation). Covariates include two vectors of factors, mainly science park characteristics (*X*) and environmental factors (*Z*). *PROV* and *YEAR* are a series of province dummies and year dummies which are used to control for province fixed effect and macroeconomic shocks on R&D efficiency. The term ε denotes the white noise error term.

As indicated in Albahari et al. (2018), heterogeneity in science parks affects tenants' performance: it relates to the park's R&D efficiency. Science park characteristics we consider are discussed next. *SIZE* denotes park size, which is measured by the number of employees in a science park. Larger parks have potential R&D synergy on R&D output and are more efficient in R&D resource usage. Moreover, the cluster of more high-tech firms may generate an R&D spillover effect that has a positive effect on the tenants' R&D efficiency. Therefore, we expect a negative sign associated with the SIZE variable. New_Park is a dummy variable equaling one if a science park was established after 2010. From a managerial viewpoint (e.g., Mccann & Folta, 2011), the services provision and management experience could be relevant to technical efficiency. New_Park is included to capture the managerial experience: more experienced management of park offices can efficiently coordinate tenants and foster R&D activity, leading to a lower misallocation.

Exports also matter to innovations. Exports might facilitate innovation through learning, competition, and customer feedback. Yang (2018) has witnessed a positive relationship between exports and R&D for Chinese firms, implying that exports could stimulate the use of R&D resource more efficiently thereby lower the degree of misallocation. We thus include export intensity (*EXP_ratio*), which is measured by the ratio of exports to commodity sale and expect it is expected to have a negative relationship with R&D misallocation. While Yang (2018) argues that the R&D-enhancing effect of exports depends on the heterogeneity of exports: process exports are found to have a negative association with R&D. The human capital measured by the ratio of university-educated labors to total employees in a science park is *UNIV_ratio*. The success of R&D projects not only depends on the qualified R&D personnel, but also on the coordination between R&D personnel and higher-skilled employees to apply the R&D output. Therefore, human capital is an essential prerequisite for efficiently using the R&D resource and is expected to help lower the R&D misallocation.

An essential issue for provincial market development is regional decentralization that empowers the provincial government to have fiscal autonomy and aligns the interests of local governments with market development (Jin, Qian, & Weingast, 2005). A provincial government facilitates the innovation performance of science parks mainly by establishing a well-functioning regional innovation system because the triple helix of university–industry–government relations is widely known for playing a critical role in regional innovation systems (Etzkowitz & Leydesdorff, 2000). Considering the central role of science parks (industry), we include two variables of the triple helix. The *RES_IND* is the degree of government-industry R&D collaboration in a province that is measured as the ratio of research institutes' expenditure on science and technology activity financed by industry. The other is *UNIV_IND*, the degree of university–industry R&D collaboration. Correspondingly, it is measured by the percentage of university expenditure on science and technology activity financed by industry in a province. Jongwanich et al. (2014) argue that science parks play a crucial role in coordinating R&D collaboration across various R&D performers within the region, and they indirectly contribute to upgrading the regional technological ladder. It suggests that the network externality of innovation should be helpful to lower R&D misallocation of science parks.

As R&D activities out of science parks are probably not neglectable in the spatial sense, it suggests that the difference in regional price (distortion) probably counts for a major part of misallocation calculated for each park. To consider this influence, we adopt two strategies: the first one is a set of province-year dummies and the other is the R&D misallocation of the province (*Prov_distortion*) calculated by Li et al. (2017).[11]

Accordingly, Eq. (20) is rewritten as follows:

[11] Thanks the authors for kindly providing the information regarding R&D misallocation of province they calculated. As the time periods in these two studies are not same, the *Prov_distortion* enters the equation in the one-year lagged form and they are the same in the final two years.

$$\begin{aligned} RD\_mis_{it} = \alpha + \beta_1 lnSIZE_{it} + \beta_2 New\_Park_{it} + \beta_3 EXP\_ratio_{it} \\ + \beta_4 UNIV\_ratio_{it} + \gamma_1 RES\_IND_{it} + \gamma_2 UNIV\_IND_{it} \\ + \beta_5 Prov\_distortion + PROV\rho + YEAR\delta + \varepsilon_{it} \end{aligned} \tag{21}$$

or

$$\begin{aligned} RD\_mis_{it} = \alpha + \beta_1 lnSIZE_{it} + \beta_2 New\_Park_{it} + \beta_3 EXP\_ratio_{it} \\ + \beta_4 UNIV\_ratio_{it} + \gamma_1 RES\_IND_{it} + \gamma_2 UNIV\_IND_{it} \\ + (PROV^*YEAR)\gamma + \varepsilon_{it} \end{aligned} \tag{22}$$

3.4. Data source

In this study, the panel-structure dataset used is drawn from two Chinese databases. The first dataset is China Torch Statistical Yearbook, published by the Ministry of Science and Technology of China. It covers related data of NHIZs, mostly production information, financial information, and innovation statistics, which enabled us to calculate the R&D misallocation across science parks. Moreover, it contains park characteristics that are related to R&D misallocation. As illustrated in Fig. 1, the number of NHIZs shows the second-wave of substantial increases in the 2010s. We adopted an unbalanced panel data of 145 NHIZs for the 2007–2014 period to compare the degree of misallocation between old NHIZs and upgraded new NHIZs.

To further examine the determinants of misallocation, we considered both park characteristics and environmental variables for individual NHIZs. The information of the R&D cooperation between industry and research institute, as well as between industry and university are drawn from various issues of the *China Statistical Yearbook on Science and Technology*. Moreover, we obtained the information of Table 4 to summarize definitions and basic statistics of variables.

4. Empirical results and discussions

Our empirical results will be presented in three parts. First, we compute the annual R&D efficiency of the science park system in China during our sample period. Next, we look at the difference in the annual R&D efficiencies to see if there are any recognizable trends. Finally, we discuss the determinants of R&D misallocation of science parks and try to identify the factors that co-move with the input distortions.

4.1. R&D misallocation across science parks in China

Based on eq. (19), the efficiency measure is the ratio of the actual level of value-added output relative to the "efficient" value-added production level in the denominator. The measure will take on a greater value if the dispersion of TFPRs is smaller across science parks functioning in that year, or equivalently, the extent of distortions is similar across difference science parks. This efficiency measure is computed annually between 2007 and 2014, and the result is presented in Fig. 2.

As depicted in Fig. 2, the efficiency measure increases remarkably during the sample period, starting from 0.42 in 2007 to 0.49 in 2008 and increase even higher at 0.62 in 2009. This increasing pattern of R&D efficiency reveals that although in 2007 about 42% of the efficiency level of output (value-added) of science parks was realized, and for about three years the trend increasing from 42% to 62% of the efficiency level of output. These findings point to a roughly 20% oscillation in the range of input distortions across science parks during those three years.

From Fig. 2, we can clearly observe a significant positive change in the measured R&D efficiency before 2010 (the efficiency measure was highest 0.65), but since 2011 there seems to be a deterioration in efficiency.

4.2. Difference in innovation performance across areas

In Fig. 2, we observe overall large oscillations in the R&D efficiency between 2007 and 2014. Fundamental questions that arise are: what are the differences in the efficiency of science parks and how do these differences evolve?

We use Fig. 3 and Table 5 to tackle this issue. In Fig. 3, we plot the annual "demeaned" TFPRs for all science parks, where a demeaned value is calculated by subtracting a science park's TFPR by the simple average of all TFPRs of science parks in that year. In Fig. 3, we draw three horizontal lines. The middle line represents zero, so if a science park's demeaned TFPR occurred on the line, it implies that the science park's TFPR equals the simple average of all TFPRs in that particular year. The upper and lower lines represent the lines above and below one standard deviation of the demeaned TFPRs of all science parks "across all years," respectively.[12] So, if a science park's demeaned TFPR occurred above the upper line, it implies that the science park's TFPR in that year is more than one (across-years) standard deviation higher than the simple average of all TFPRs in that year; likewise similar interpretation could be made for science parks' demeaned TFPRs that are located below the lower line. For convenience, we will call the area below the lower

[12] To calculate this "across-years" standard deviation of demeaned TFPRs, we first demeaned all science parks TFPRs by the simple average of all science parks' TFPRs in the corresponding years, then calculated the standard deviation of these science parks demeaned TFPRs. The main reason we consider this "across-years" standard deviation instead of the yearly standard deviations is because we hope to highlight the convergence of the science parks' TFPRs throughout the sample period.

Table 4
Variable definitions and basic statistics.

Variable	Definition	Mean (Std. Dev.)
RD_mis	R&D misallocation of a science park, measured by $ln\left(TFPR_{it}/\overline{TFPR_t}\right)$	-3.9e-10 (0.994)
SIZE	Park size: number of employees in a park.	130,323 (174,258)
New_Park	New park dummy: 1 = parks established in 2011 and onward; 0 = others	0.346 (0.476)
EXP	Export of a science park (RMB million)	3685 (6031)
EXP_ratio	The ratio of exports to commodity sales of a park	0.170 (0.185)
SKILL_ratio	Human capital: the ratio of skilled labors to total employees in a park	0.112 (0.062)
UNIV_ratio	Human capital: the ratio of university-educated employees to total employees in a park	0.430 (0.158)
RES_IND	The ratio of research institutes' expenditure on science and technology activity financed by industries (%)	5.287 (2.118)
UNI_IND	The percentage of universities' expenditure on science and technology activity financed by industry	31.389 (10.745)
Prov_dsitortion	Province price distortion: proxized by the dependent variable in Li et al. (2017)	1.030 (0.244)
Province wage	Proxy variable of province price distortion (RMB thousand)	45.358 (13.238)

Note: Mean and standard deviation are calculated using the 2007–2014 data.

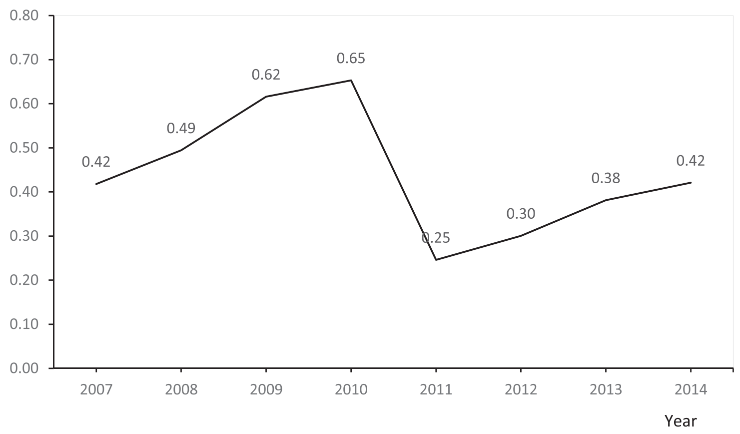

Fig. 2. R&D Efficiency (Y_R), 2007–2014.
Note: The productivity Y_R is computed based on Eq. (20), and utilizing the parameters listed in Table 3 .

line zone 1; the areas within one standard deviation below and above the middle line (which equals to zero) are zones 2 and 3, respectively. The area above the upper line is zone 4.

There are several striking patterns in Fig. 3. First, we see that in earlier years, before 2010, that most science parks' TFPRs lie in zone 2 and zone 3, suggesting only a slight divergence in TFPRs across regions. Besides, we can observe that most of the science parks' TFPRs are in the shallow places (less than 4 in the measure) places into zone 4 for those three years, resulting in a declining trend of input distortion and increasing the trend of production efficiency. Second, we see that after 2010, TFPRs across regions diverge and lie in deeper areas of zone 4 with measurements higher than 5. This divergent pattern is most evident in 2011 and 2012. However, in both 2011 and 2012, the TFPR of one science park lies in a very deep area of zone 4.[13] Third, since 2010, TFPRs across science parks seem to diverge continuously. For example, in 2012 there exists the TFPR located in very deep areas of zone 4.

Are there specific patterns of TFPRs over the years in China? To answer this question, in Table 5, we divide each column year into two sub-columns (except for the first year 2007): old incumbent science parks and new entrant science parks. In this way, we can see the patterns of dispersions of new science parks relative to the old parks for each year. Table 5 shows that three different periods (2007–2008, 2009–2011, 2012–2014) exhibit different patterns for new entrant science parks, and we can see how the TFPR for each science park evolves. In the 2009–2010 period in Table 5, which summarizes the distribution of TFPRs for each year for the entire period, we see that there is a vast difference in the TFPR distribution across these two years.[14] When we compare the TFPRs in different periods, we find fascinating evolution patterns across the science parks. For science parks already existing in the first period (2007–2008) and the third period (2012–2014), their TFPRs are less dispersed relative to the new entering science parks in middle years (2009–2011), in which 50% (2009) and 16% (2011) of the TFPRs of the entrant science parks are in zone 4. However, since 2012, the TFPRs of existing science parks converge toward the mean, and no new science parks are entering outside the middle zones. Similar

[13] However, in 2007–2009, TFPRs located in zones 2 and 3 converge as well, so the measured innovation efficiency is better in that year than in 2019.

[14] The definition of new is to summarize the dispersion of TFPRs for those science parks entering in that specific year.

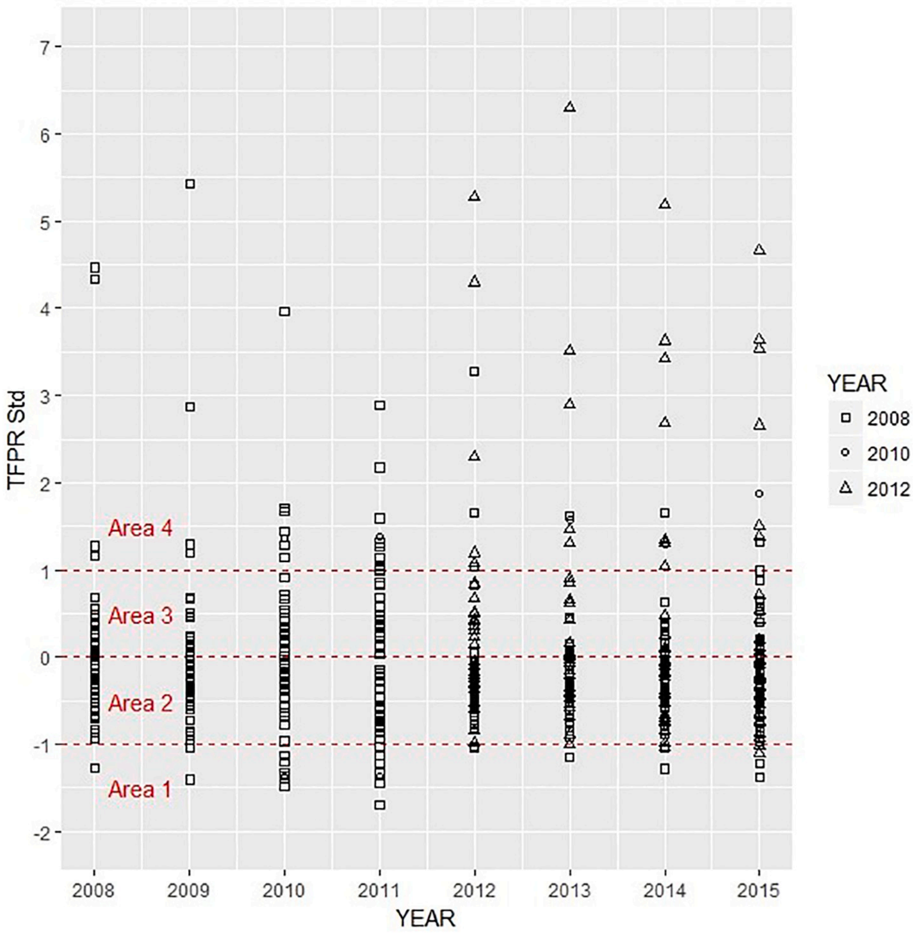

Fig. 3. Distribution of Science Parks' Demeaned TFPRs (2007–2014).

patterns can be observed in the first period (2007–2008). Thus, we can conclude that in 2009–2011, new entrant science parks have more aggregated input distortions than those already existing science parks at any other time.

4.3. A counter-factual welfare analysis

Before we discuss what factors affect resource misallocation across science parks, some theoretical grounds are called for. Here we would like to solidify the theoretical results we have reached in the previous sections. As noted in the misallocation computations, the innovation input distortions have been found to degrade overall R&D efficiency, as shown in Fig. 2. We also observe the establishment of many parks, starting from 2011. Subsequently-established parks recorded more input distortions (see Fig. 3 and Table 5). We therefore doubt that the science parks established after 2011 would have inferior efficiency compared with the existing ones.

We conduct a counter-factual analysis that considers a hypothetical situation to see if post-2011 R&D efficiency (Y_R) maintains the same Y_R as that observed previously. To conduct our hypothetically improved new park setup, the processes are as follows. First, we categorize all the recorded parks into two groups: the old group of parks established before 2011, and the new group of parks established starting from 2011 onward. Each park's TFPQ of the old group in each sample year is calculated based on the Eq. (9). The next step is picking the lowest TFPQ park of the old group for the years post-2011. The TFPQ of the new group of parks is replaced by the lowest TFPQ park of the old group in the years post-2011. Based on the eq. (2), using the benchmarked new parks' TFPQ, we can re-compute each one's R&D output belonging to the new group from each sample year, since each new park is now endowed with hypothetical TFPQ and the original innovation inputs.

Utilizing hypothetically-derived R&D outputs of new park groups and the original innovation inputs, we can again compute the R&D efficiency in the same way as we drew the original Y_R in Fig. 2. Fig. 2H illustrates the counter-factual results of Y_R. As depicted in Fig. 2H, the R&D efficiency improved considerably post-2011, suggesting that even controlling the physical productivity of the new parks, on the lowest level of the old parks, would significantly improve the aggregate research outputs. The key factor to improved R&D outputs embedded within this hypothetical setup lies in eliminating the original innovation input distortions that exist ubiquitously in the new parks. Therefore, we continue, so as to observe the distribution of innovation input distortions following this

Table 5
Distribution of demeaned TFPRs across Science Parks.

Distribution of demeaned TFPRs across different science parks (number)																
Zone	2007		2008		2009		2010		2011		2012		2013		2014	
	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New
Zone1		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Zone2		33	30	0	31	2	34	0	31	27	56	13	63	8	68	0
Zone3		15	18	0	16	0	15	0	14	2	27	1	28	0	27	0
Zone4		6	6	0	7	0	7	0	11	0	2	2	10	0	14	0
Total		54	54	0	54	2	56	0	56	29	85	16	101	8	109	0

Distribution of demeaned TFPRs across different science parks (ratio %)																
Zone	2007		2008		2009		2010		2011		2012		2013		2014	
	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New	Old	New
Zone1		2%	4%		13%	50%	14%		4%	0%	2%		3%		5%	
Zone2		57%	57%		43%	0%	34%		86%	47%	66%		65%		60%	
Zone3		33%	31%		31%	0%	34%		5%	38%	24%		20%		26%	
Zone4		7%	7%		13%	50%	18%		5%	16%	8%		11%		9%	
Total		100%	100%		100%	100%	100%		100%	100%	100%		100%		100%	

Note: The values within each cell in the top panel show the frequency. The numbers in the lower panel show the share of the science parks' demeaned TFPRs located in a particular zone during a specific year.

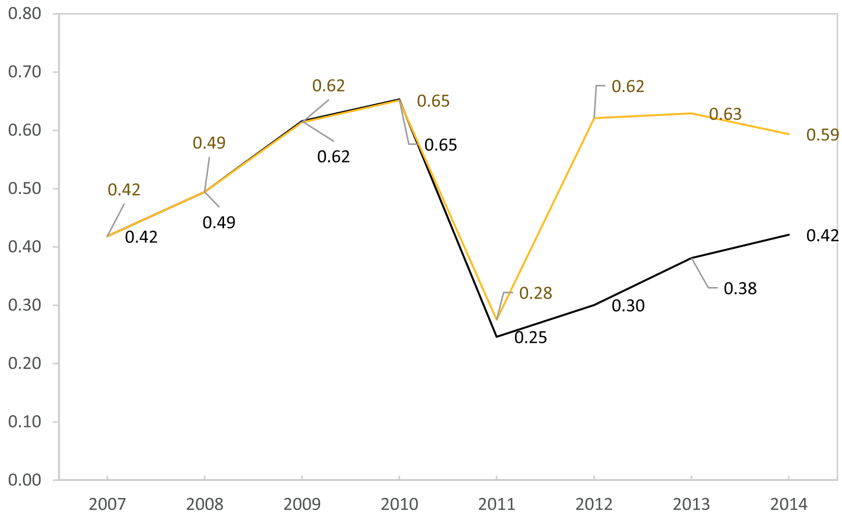

Fig. 2H. R&D efficiency (Y_R), Counter-factual Analysis.

counter-factual analysis.

Correspondingly, we also show the results of innovation input distortions in Fig. 3H and Table 5H, in order to compare with the original results illustrated in Fig. 3 and Table 5.

Fig. 3H and Table 5H also show that the innovation input distortions improved analogously, leading to improved R&D efficiency, as recorded in Fig. 2H. Besides, possessing better environmental factors endowed with new parks should generate greatly optimized innovation input usage.

This section concludes that the worst resource-allocation situation occurs when the least efficient new parks (parks with lower TFPQ compared with old parks) utilize a big chunk of available innovation inputs. That situation results in lower R&D output in the equilibrium. Alternatively, remarkable improvement occurs whenever all of the park-level input distortions have been wiped clean to reflect the optimal allocation based on firm-level TFPQ. Therefore, we determine that better control of new park quality promises higher quality existing parks, leading to improved R&D productivity. Although this conclusion is not limited to the context of science parks, it is that R&D discussion for which we have a ready illustration.

5. Determinants of R&D misallocation and R&D scale

5.1. Determinants of misallocation

Table 6 displays the estimates of various specifications to examine the determinants of R&D misallocation across science parks in China. To begin with, Columns (1) and (2) are a baseline model that includes only park-specific characteristics, as emphasized in Albahari et al. (2018). A series of province-year dummies are included in Column (2) to capture the price distortion across provinces in various years.

How does heterogeneity in science parks affect their R&D efficiency (misallocation)? Results obtained using the RE or FE models are similar.[15] The coefficient of size is significantly negative, indicating that larger parks, in terms of the number of employees, have lower R&D misallocation. As discussed previously, larger parks have potential R&D externality and a spillover effect; a cluster of more high-tech firms results in parks experiencing a lower R&D misallocation than their smaller counterparts. Albahari et al. (2018) argue that firms in larger parks have better innovation outputs; this implies that their R&D misallocation could be lower.

We find a significantly positive relation between *New_Park* and R&D misallocation in Columns (1) and (2), echoing the finding demonstrated in the previous section that there is a lower R&D efficiency for newly established parks. This result is also consistent with findings in Albahari et al. (2018) that older parks better perform on innovation outputs, because managerial experience may help tenants promote R&D efficiency through administrative efficiency.

Exports, in terms of export value or export intensity, exhibit an insignificant relation with R&D misallocation. In other words, export behaviors have no significant influence on raising or lowering R&D efficiency. Yang (2018) claims that serving the international market can facilitate innovation through learning, competition, and customer feedback; while ordinary and process has a converse influence on R&D in China, which are positive and negative, respectively. China's exports concentrate on processing trade that exploits the advantages of low-cost labor and tariff-free imported intermediate goods. Although the ratio of processing exports to total exports

[15] Through Monte Carlo simulations, Bell and Jones (2015) show that the RE model is more adequate than FE model when the time-invariants (rarely changed variables) are incorporated, such as the province dummy in Columns (1) and (3).

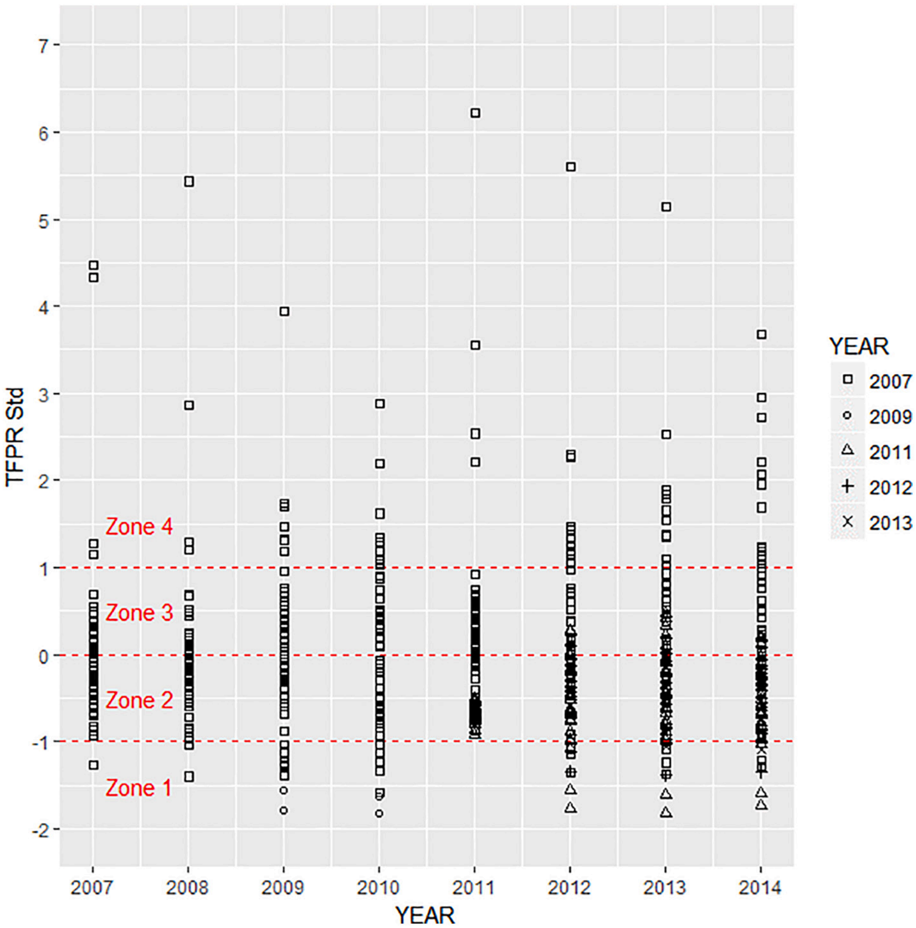

Fig. 3H. Distribution of science parks' demeaned TFPRs: Counter-Factual Analysis.

Table 5H
Distribution of demeaned TFPRs: Counter-factual Analysis.

Distribution of demeaned TFPRs across different science parks (number)								
Zone	2007	2008	2009	2010	2011	2012	2013	2014
Zone1	1	2	9	8	0	7	7	7
Zone2	31	31	23	20	48	45	47	49
Zone3	18	17	17	18	36	23	23	19
Zone4	4	4	7	10	4	12	11	13
Total	54	54	56	56	88	87	88	88
Distribution of demeaned TFPRs across different science parks (ratio %)								
Zone	2007	2008	2009	2010	2011	2012	2013	2014
Zone1	2%	4%	16%	14%	0%	8%	8%	8%
Zone2	57%	57%	41%	36%	55%	52%	53%	56%
Zone3	33%	31%	30%	32%	41%	26%	26%	22%
Zone4	7%	7%	13%	18%	5%	14%	13%	15%
Total	100%	100%	100%	100%	100%	100%	100%	100%

demonstrated a decreasing trend in 2000s, it was still as high as 48.1% in 2008 (Koopman, Wang, & Wei, 2008), and this proportion might be higher in high-tech industries. Therefore, export activity has no specific relation with R&D misallocation.

Human capital is relevant to R&D efficiency. Human capital not only has a positive impact on technological innovation (Gallie & Legros, 2012), but also be complementary to firm R&D through absorption capability embedded in human capital. Moreover, higher educated employees also help coordinate with R&D personnel. Parks with a higher ratio of higher educated employees use R&D resource more efficiently, resulting in a lower R&D misallocation.

Regarding the linkage effect among triple helix constituent parties, Columns (3)–(6) illustrate results of various specifications. The

Table 6
Determinants of R&D Misallocation across Science Parks.

	(1)	(2)	(3)	(4)	(5)	(6)
	RE	FE	RE	FE	FE	FE
lnSIZE (EMP)	−0.231* (0.120)	−0.380* (0.197)	−0.233* (0.120)	−0.382* (0.197)	−0.575*** (0.133)	−0.581*** (0.134)
New_Park	1.108*** (0.215)	3.235*** (0.402)	1.109*** (0.216)	3.235*** (0.402)	2.952*** (0.332)	2.933*** (0.332)
lnEXP	0.054 (0.050)		0.054 (0.050)			
EXP_ratio		0.241 (0.507)		0.241 (0.508)	0.161 (0.368)	0.171 (0.368)
UNIV_ratio	−0.764*** (0.276)	−1.684*** (0.577)	−0.840*** (0.282)	−1.665*** (0.579)	−1.051*** (0.288)	−1.138*** (0.294)
RES_IND			0.003 (0.007)	−0.005 (0.200)		0.001 (0.007)
UNI_IND			−0.009* (0.005)	0.013 (0.047)		−0.009* (0.005)
Prov_distortion					−0.292* (0.163)	−0.298* (0.164)
Constant	1.364 (1.830)	4.298* (2.409)	−0.130 (1.077)	3.972*** (2.533)	6.268*** (1.530)	6.719*** (1.561)
Year FE	Yes	No	Yes	No	Yes	Yes
Province FE	Yes	No	Yes	No	No	No
Province-year FE	No	Yes	No	Yes	No	No
R-square	0.298	0.539	0.299	0.539	0.227	0.231
Observations	571	571	571	571	571	571

Note: Figures in parentheses are heteroscedastic-consistent standard deviations. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

estimated coefficient of *RES_IND* is insignificant, whereas *UNI_IND* is, as expected, associated with a significantly negative coefficient in most estimations. These results are similar to findings in Jongwanich et al. (2014) that there is a positive innovation-enhancing effect brought about by R&D cooperation between industries and universities, rather than research institutes. Frequent interaction through R&D collaborations enables high-tech firms in science parks to exchange ideas and research experiences with their R&D partners. However, this positive network effect on R&D efficiency is not seen in R&D collaboration with research institutes. Contrastingly, the university–industry collaborations (UIC) seem to facilitate the use of R&D resources more efficiently, lowering R&D misallocation.

Firms may prefer R&D collaboration with universities in China for two principal reasons. First, new knowledge generated by universities can be treated as a public goods, and universities are recognized as a source of knowledge for firms. In particular, the information generally complements a firms' R&D (Baba, Shichijo, & Sedita, 2009). This cross-function non-competitive alliance helps firms to promote R&D efficiency. Second, UIC gives firms inexpensive and low-risk access to universities' specialist knowledge (Azagra-Caro, Pardo, & Rama, 2014). The lower transaction cost in R&D helps firms raise their R&D efficiency.

As outlined above, the cooperation between a company's R&D department and a research institute is advantageous. However, it is not clear why there is no significant effect on lowering R&D misallocation. One major issue of the UIC linkage or research institute–industry collaboration is that it is fraught with risks due to the uncertainty of innovation commercialization, which is the main aim of profit-maximizing firms. In China, universities may conduct profit-oriented R&D projects, because university-operated or spin-off companies are allowed to create profit for universities. Consequently, universities are profit motivated. Alternatively, research institutes may focus on more generic and basic knowledge and lack strong profit incentives; preventing high-tech firms from commercializing R&D output in the short-term with no significant stimulating effect on R&D efficiency.

5.2. Robustness checks

We have thus far illustrated some influential factors of R&D misallocation in science parks. In this section, robustness checks are implemented to deal with several potential problems. First, a park's characteristics and its R&D activity could have a causal relationship, or the R&D misallocation might be persistent. To address this potential endogeneity issue in conducting the robustness check, we adopt the technique of dynamic generalized method of moment (GMM). Second, we use province-specific wages to capture regional price distortion, which could act as a robustness check. Third, instead of using the standardized TFPR (RD_mis) as the dependent variable, we run direct regressions on distortion parameters ($TFPR_i$). Although it is not apparent in the literature, this robustness check helps facilitate understanding of how those distortions emerge. Table 7 summarizes the results of the robustness checks.

Most estimates in Table 7 are similar to those in Table 6, although interesting results emerge and are worth noting, particularly the triple helix linkages. To begin with, results obtained using the dynamic GMM in Columns (1) and (2) show that park R&D misallocation seems to be persistent. This implies that the correction of R&D misallocation might be difficult to finalize in the short run. Crucially, we find that the estimated coefficient of RES_IND becomes significantly negative in Column (1), suggesting that R&D cooperation with research institutes may also help lower science park R&D misallocation. Next, using the distortions measure ($TFPR_i$) as the dependent variable, the influences of park characteristics remain the same, while RES_IND and UNI_IND turn out to have significantly negative and insignificant relationships with R&D misallocation, respectively.

Drawn from estimates of RES_IND and UNI_IND in Tables 6 and 7, both R&D collaboration with research institutes or universities can help raise the R&D efficiency of science parks, thereby lowering R&D misallocation. It is likely due to weak significance, resulting in the varied significance of RES_IND and UNI_IND in different empirical specifications. Overall, cross-function, non-competitive alliance with research institutes and/or universities help firms to promote R&D efficiency, highlighting the importance of national and/or regional innovation systems.

Table 7
Robustness Checks.

	(1)	(2)	(3)	(4)	(5)
	Dynamic GMM	Dynamic GMM	Distortion: wage	Y = $TFPR_i$	Y = $TFPR_i$
lag. RD_miss	0.578*** (0.019)	0.585*** (0.019)			
lnSIZE	−0.067*** (0.025)	−0.065*** (0.026)	−0.554*** (0.134)	−0.320*** (0.063)	−0.325*** (0.061)
New_Park	0.332*** (0.081)	0.195*** (0.072)	3.026*** (0.329)	0.650*** (0.154)	0.621*** (0.151)
lnEXP	−0.046* (0.024)			−0.008 (0.023)	
EXP_ratio		0.194 (0.198)	0.140 (0.370)		0.133 (0.168)
UNIV_ratio	−0.778*** (0.129)	−0.708*** (0.135)	−1.111*** (0.296)	−0.338*** (0.130)	−0.314** (0.134)
RES_IND	−0.007* (0.0036)	−0.006 (0.0038)	0.002 (0.007)	−0.006* (0.003)	−0.006* (0.003)
UNI_IND	−0.015*** (0.003)	−0.015*** (0.002)	−0.009* (0.005)	−0.002 (0.002)	−0.002 (0.002)
Prov_distortion	−0.459*** (0.171)	−0.508*** (0.168)		−0.045 (0.074)	−0.052 (0.075)
Province wage			−0.421 (0.663)		
Constant	2.322*** (0.553)	1.661*** (0.309)	10.358 (6.809)	7.315*** (0.723)	7.248*** (0.712)
Year FE	Yes	Yes	Yes	Yes	Yes
Province FE	No	No	No	No	No
R-square			0.266	0.270	0.271
Arellano-Bond test	0.001	0.001			
Hansen Sargan	0.014	0.025			
Observations	482	482	571	571	571

Note: Figures in parentheses are heteroscedastic-consistent standard deviations. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 8
R&D Growth and R&D Misallocation.

Dep. var.	(1)	(2)	(3)	(4)
	RD_mis	*RD_mis*	$TFPR_i$	$TFPR_i$
Δ RD expenditure	−0.091*** (0.030)		−0.047*** (0.013)	
Δ RD personnel		−0.119*** (0.041)		−0.065*** (0.017)
Constant	−0.068 (0.055)	−0.070 (0.055)	3.754*** (0.023)	3.754*** (0.023)
Province-Year FE	Yes	Yes	Yes	Yes
R-square	0.430	0.429	0.541	0.541
Observations	568	570	568	570

Note: Figures in parentheses are heteroscedastic-consistent standard deviations. ***$p < 0.01$.

5.3. *R&D scale and R&D misallocation*

The above analysis identifies how park characteristics relate to R&D misallocation. It is clear that R&D misallocation is affected by R&D scale. Jones and Williams (1998) indicate that R&D can be increasing or decreasing return to scale, depending on the way past ideas have affected current R&D productivity. R&D investment risk also declines with scale (Ciftci & Cready, 2011), implying that R&D productivity increases with scale. In other words, R&D misallocation could fall as R&D inputs grow. Using return-to-scale as a conceptual foundation, we explore how R&D growth affects parks' R&D misallocation. Table 8 shows the estimation results.

Results using R&D misallocation and TFPR as the dependent variable are shown in Columns (1)–(2) and Columns (3)–(4), respectively.[16] We find a significantly negative relationship between the growth of R&D expenditure (growth of R&D personnel) and R&D misallocation. From the view of innovation production function, the increases in R&D inputs seem to generate a non-proportional, larger increase in R&D output, suggesting an increasing return-to-scale with NHIZs. The agglomeration and network effects, as well as institute-university-park linkages, can also facilitate R&D efficiency. Thus, increases in R&D help lower R&D misallocation, particularly in the old group of NHIZs.

6. Concluding remarks and policy implication

Establishing science parks is widely adopted as a key strategy for developing high-tech industries and creating jobs to facilitate regional development across countries. The success of science parks, in terms of R&D and innovation performance, relies heavily on the park characteristics and the agglomeration effect of various R&D actors: mainly research institutes, universities, and industries. Adequate locations and the number of science parks are particularly relevant to developing countries where R&D resources are relatively limited.

Along with economic development, China has established more than 100 national-class science parks in all provinces except for

[16] Inasmuch as growth of R&D expenditure and R&D personnel are highly correlated, they are included separately here as covariates. We have also separated growth of outputs and exports and found no significant influence on R&D misallocation.

Tibet, over the past two decades. There are considerable differences in the degree of economic development across provinces, which lead to the emergence of crucial issues. Does allocating R&D resources to numerous science parks result in serious R&D misallocations? How do park characteristics and the triple helix of university–industry–government linkages affect the efficiency of R&D in the science parks? This study aimed to evaluate R&D misallocation and its determinants across science parks in China. Based on an unbalanced panel data of 145 science parks from 2007 to 2014, we applied Hsiech and Klenow's (2009) methodology to calculate R&D efficiency (misallocation) of science parks and then examined the determinants of R&D efficiency across science parks.

Crucial findings are summarized as follow. First, we found a decreasing trend for R&D efficiency across science parks in China. Particularly, the average R&D efficiency had decreased sharply since 2011 when China began to start up many new science parks, suggesting a considerable R&D misallocation caused by establishing science parks in all areas. Second, as argued in Albahari et al. (2018), the heterogeneity of science parks is relevant to their R&D performance. We found science parks which are larger, older, or with higher quality of human capital experience lower R&D misallocation. Export activity seems to less relevant to the degree of R&D misallocation efficiency across parks. Third, the triple helix of university–industry–government linkages matter to science parks' R&D efficiency, and the UIC seems to more relevant than the collaboration between the research institutes and industry.

Two crucial policy implications arise from our analyses. First, most of the less R&D efficient science parks are newly established or are upgraded from the provincial level industrial parks. They are generally located in cities without nationally renowned universities and research institutes, whereas the first science parks are located in big cities, such as Beijing, Shanghai, and Nanjing, among others. Although science parks at a national level are nominal, they are indeed managed by the provincial government. Therefore, for some R&D efficiency parks, the authority may have to coordinate with the central government to improve their R&D efficiency, lowering the problem of R&D misallocation.

Second, although the importance of the triple helix linkages in the regional innovation system is widely recognized, the research institute–industry R&D collaboration seems to have less significant R&D efficiency-enhancing effect for science parks. The resources for R&D that were allocated to public research institutes is higher than that of universities in China. For example, research institutes and universities accounted for 15.04% and 10.47% of total R&D expenditure in 2011, respectively. By contrast, the collaborative R&D accounts for a much smaller ratio of total R&D for research institutes than that of universities, as shown in Table 4. The authorities may consider how to facilitate the technology transfer from public research institutes and how to enhance R&D collaboration with industries by aiming to complement the industries' R&D. It would help promote the science parks' R&D efficiency.

Acknowledgement

Chih-Hai Yang gratefully acknowledges financial support from the Ministry of Science and Technology of Taiwan (NSC107-2410-H-008-004). Helpful comments and suggestions from participants at the 2019 Asia Pacific Innovation Conference (APIC) are highly appreciated. Miss. Yi-Fei Li's research assistant on collecting the science parks information is also appreciated.

References

Albahari, A., Barge-Gil, A., Pérez-Canto, S., & Modrego-Rico, A. (2018). The influence of science and technology park characteristics on firms' innovation results. *Papers in Regional Science, 97*(2), 253–279.
Atkeson, A., & Kehoe, P. (2005). Modeling and measuring organization capital. *Journal of Political Economy, 113*, 1026–1053.
Azagra-Caro, J., Pardo, R., & Rama, R. (2014). Not searching, but finding: How innovation shapes perceptions about universities and public research organisations. *Journal of Technology Transfer, 39*(3), 454–471.
Baba, Y., Shichijo, N., & Sedita, S. (2009). How do collaborations with universities affect firm's innovative performance? The role of "Pasteur scientists" in the advanced materials field. *Research Policy, 38*, 756–764.
Bai, J. (2013). On regional innovation efficiency: Evidence from panel data of China's different provinces. *Regional Studies, 47*, 773–788.
Bell, A., & Jones, K. (2015). Explaining fixed effects: Random effects Modeling of time-series cross-sectional and panel data. *Political Science Research and Methods, 3*(1), 133–153.
Bigliardi, B., Dormio, A. I., Nosella, A., & Petroni, G. (2006). Assessing science Parks' performances: Directions from selected Italian case studies. *Technovation, 26*(4), 489–505.
Ciftci, M., & Cready, W. M. (2011). Scale effects of R&D as reflected in earnings and returns. *Journal of Accounting and Economics, 52*(1), 6–80.
Colombo, M. G., & Delmastro, M. (2002). How effective are technology incubators? Evidence from Italy. *Research Policy, 31*, 1103–1122.
Crépon, B., & Duguet, E. (1997). Estimating the innovation function from patent number: GMM on count panel data. *Journal of Applied Econometrics, 12*, 243–263.
Diez-Vial, I., & Fernández-Olmos, M. (2017). The effect of science and technology parks on a Firm's performance: A dynamic approach over time. *Journal of Evolutionary Economics, 27*(3), 413–434.
Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From National Systems and "mode 2" to a triple Helix of university–industry–government relations. *Research Policy, 29*(2), 109–123.
Fan, C. C., & Scott, A. J. (2003). Industrial agglomeration and development: A survey of spatial economic issues in East Asia and a statistical analysis of Chinese regions. *Economic Geography, 79*(3), 295–319.
Foster, L., Haltiwanger, J., & Syverson, C. (2008). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability. *American Economic Review, 98*, 394–425.
Fritsch, M., & Franke, G. (2004). Innovation, regional knowledge Spillovers and R&D cooperation. *Research Policy, 33*(2), 245–255.
Fukugawa, N. (2006). Science parks in Japan and their value-added contributions to new technology-based firms. *International Journal of Industrial Organization, 24*, 381–400.
Gallie, E. P., & Legros, D. (2012). Firms' human capital, R&D and innovation: A study on French firms. *Empirical Economics, 43*, 581–596.
Griliches, Z. (1990). Patent statistics as economic indicators. *Journal of Economic Literature, 28*, 1661–1707.
Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *Quarterly Journal of Economics, 124*(4), 1403–1448.
Hu, A. G. (2007). Technology parks and regional economic growth in China. *Research Policy, 36*(1), 76–87.
Hu, J. L., Han, T. F., Yeh, Y. F., & Lu, C. L. (2010). Efficiency of science and technology industrial parks in China. *Journal of Management Research, 10*(3), 151–166.
Hu, A. G., & Jefferson, G. H. (2009). A Great Wall of patents: What is behind China's recent patent explosion? *Journal of Development Economics, 90*(1), 57–68.
Izadikhah, M., & Saen, R. F. (2015). A new data envelopment analysis method for ranking decision making units: An application in industrial parks. *Expert Systems, 32*(5), 596–608.

Jefferson, G. H., Bai, H., Guan, X., & Yu, X. (2006). R&D performance in Chinese industry. *Economics of Innovation and New Technology, 15*(4/5), 345–366.
Jin, H., Qian, Y., & Weingast, B. R. (2005). Regional decentralization and fiscal incentives: Federalism, Chinese Style. *Journal of Public Economics, 89*, 1719–1742.
Jones, C. I., & Williams, J. C. (1998). Measuring the social return to R&D. *Quarterly Journal of Economics, 113*(4), 1119–1135.
Jones, C. I., & Williams, J. C. (2000). Too much of a good thing? The economics of Investment in R&D. *Journal of Economic Growth, 5*(1), 65–85.
Jongwanich, J., Kohpaiboon, A., & Yang, C. H. (2014). Science Park, triple Helix and regional innovative capacity: Province-level evidence from China. *Journal of Asian-Pacific Economy, 19*(2), 333–352.
Koopman, R., Wang, Z., & Wei, S. J. (2008). How much of Chinese exports is really made in China?. In *Assessing domestic value-added when processing trade is pervasive, NBER Working Paper No. 14109*.
Li, X. (2009). China's regional innovation capacity in transition: An empirical approach. *Research Policy, 38*, 338–357.
Li, H. C., Lee, W. C., & Ko, B. T. (2017). What determines misallocation in innovation? A study of regional innovation in China. *Journal of Macroeconomics, 52*, 221–237.
Liberati, D., Marinucci, M., & Tanzi, G. M. (2015). Science and technology parks in Italy: Main features and analysis of their effects on the firms hosted. *Journal of Technology Transfer, 41*(4), 694–729.
Link, A., & Scott, J. (2003). U.S. science parks: The diffusion of an innovation and its effects on the academic missions of universities. *International Journal of Industrial Organization, 21*, 1323–1356.
Löfsten, H., & Lindelöf, P. (2002). Science parks and the growth of new technology-based firms: Academic-industry links, Innovation and Markets. *Research Policy, 31*, 859–876.
Luger, M. I., & Goldstein, H. A. (1991). *Technology in the Garden: Research Parks and Regional Economic Development*. Chapel Hill, NC, U.S.A.: The University of North Carolina Press.
Macdonald, S. (1987). British science parks: Reflections on the politics of high technology. *R&D Management, 17*(1), 25–37.
Massey, D., Quintas, P., & Wield, D. (1992). *High tech fantasies: Science parks in society, science and space*. London: Routhledge.
Mccann, B. T., & Folta, T. B. (2011). Performance differentials within geographic clusters. *Journal of Business Venturing, 26*, 104–123.
Nadiri, M. I., & Prucha, I. R. (1996). Estimation of the depreciation rate of physical and R&D Capital in the US. Total Manufacturing Sector. *Economic Inquiry, 34*(1), 43–56.
Nagaoka, S., Motohashi, K., & Goto, A. (2010). Patent statistics as an innovation indicator. In B. H. Hall, & N. Rosenberg (Eds.)*, Vol. 2. Handbook of the Economics of Innovation* (pp. 1083–1127). Amsterdam, Netherland: Elsevier.
Quintas, P., Wield, D., & Massey, D. (1992). Academic-industry links and innovation - questioning the Science Park model. *Technovation, 12*, 161–175.
Ramani, S. V., EI-Aroui, M. A., & Carrere, M. (2008). On estimating a knowledge production function at the firm and sector level using patent statistics. *Research Policy, 37*(9), 1568–1758.
Restuccia, D., & Rogerson, R. (2008). Policy distortions and aggregate productivity with heterogeneous plants. *Review of Economic Dynamics, 11*, 707–720.
Syverson, C. (2011). What determines productivity? *Journal of Economic Literature, 49*(2), 326–365.
Vásquez-Urriago, A. R., Barge-Gil, A., Rico, A. M., & Paraskevopoulou, E. (2014). The impact of science and technology parks on Firms' product innovation: Empirical evidence from Spain. *Journal of Evolutionary Economics, 24*(4), 825–873.
Wainova. (2009). *Wainova atlas of innovation: Science/technology/research parks and business incubators in the world*. Cheshire: Ten Alps Publishing.
Weil, D. N. (2013). *Economic growth* (3rd ed.). Boston: Pearson.
Westhead, P., & Batstone, S. (1998). Independent technology-based firms: The perceived benefits of a Science Park location. *Urban Studies, 35*(12), 2197–2219.
Yang, C. H. (2018). Exports and innovation: The role of heterogeneity in exports. *Empirical Economics, 55*(3), 1065–1087.
Yang, C. H., Motohashi, K., & Chen, J. R. (2009). Are new technology-based firms located on science parks really more innovative? Evidence from Taiwan. *Research Policy, 38*(1), 77–85.