ADBI Working Paper Series

THE DIRECT AND INDIRECT EFFECTS OF INFRASTRUCTURE ON FIRM PRODUCTIVITY: EVIDENCE FROM MANUFACTURING IN THE PEOPLE'S REPUBLIC OF CHINA

Guanghua Wan and Yan Zhang

No. 714
April 2017

Asian Development Bank Institute

Guanghua Wan is director of research and Yan Zhang is research project consultant at the Asian Development Bank Institute.

Suggested citation:

Wan, G. and Y. Zhang. 2017. The Direct and Indirect Effects of Infrastructure on Firm Productivity: Evidence from Manufacturing in the People's Republic of China. ADBI Working Paper 714. Tokyo: Asian Development Bank Institute. Available: https://www.adb.org/publications/direct-and-indirect-effects-infrastructure-firm-productivity-prc

Please contact the authors for information about this paper.

Email: yanzhang@adbi.org, gwan@adbi.org

Asian Development Bank Institute
Kasumigaseki Building, 8th Floor
3-2-5 Kasumigaseki, Chiyoda-ku
Tokyo 100-6008, Japan

Tel: +81-3-3593-5500
Fax: +81-3-3593-5571
URL: www.adbi.org
E-mail: info@adbi.org

Abstract

This paper attempts to distinguish and estimate the direct and indirect effects of infrastructure on firm productivity. The latter arises from the infrastructure–agglomeration link and has been largely overlooked in the literature on infrastructure. An analytical framework is then developed to estimate both effects. Finally, empirical results are obtained using large-scale firm-level survey data from the People’s Republic of China (PRC). Major findings include: (1) all the three kinds of infrastructure—road, telecommunication servers, and cable—are found to directly promote firm productivity; (2) they also exert a positive indirect effect on firm productivity through the agglomeration channel; and (3) the empirical results are robust to different agglomeration indicators and different subsamples.

Keywords: infrastructure, indirect effect, productivity, agglomeration, PRC

JEL Classification: D21, D23, L1, R12, R3

Contents

1. INTRODUCTION

Existing studies on the productivity impacts of infrastructure mostly focus on its role in providing physical connectivity and information (Démurger 2001; Mikelbank and Jackson 2000). For example, infrastructure helps firms expand catchment areas (Helpman and Krugman 1985; Fujita, Krugman, and Mori 1999), access larger labor markets (Duranton and Turner 2012), and reduce logistic costs (Holl 2006; Röller and Waverman 2001; McCann and Shefer 2003). These benefits accrue from the direct use of infrastructure by firms and households and can be called the direct effect.

At the same time, infrastructure is known to help promote agglomeration of economic activities (Lewis and Bloch 1998; McCann and Shefer 2003). The proximity of industries and businesses favors cost reduction and productivity improvement by allowing firms in the region to share a similar labor pool and enjoy knowledge spillovers (Fujita, Krugman, and Venables 2001; Fujita and Thisse 2013). This effect arising from agglomeration can be defined as the indirect infrastructure effect. Whereas the literature on the direct impact is sizable, few studies pay attention to the indirect effects. McCann and Shefer (2003) were the first to raise the concept of the indirect effect. They argue that infrastructure users will plow profits back into the regional economy by providing better services and cheaper products. However, McCann and Shefer (2003) do not conduct any theoretical and empirical analyses of the indirect effect. According to the survey of infrastructure literature by Eberts and McMillen (1999), most studies tend to mix the indirect and direct effects as the total effect when estimating the impact of infrastructure.

This is regrettable because distinguishing the two effects has significant meanings. First, the total effect could mask the discrepancies between its two components, which often have different policy implications. Let us compare two scenarios—in both infrastructure exerts a limited influence on productivity, such as 1%. In scenario 1, the direct effect is 5% and the indirect effect is –4%. The total effect turns out to be low owing mainly to the indirect effect, which suggests a scenario of agglomeration diseconomy, perhaps as a result of the adverse externalities of clustering, such as air and water pollution or transportation congestion. In this case, policies to mitigate negative impacts should be implemented along with infrastructure investments. For example, pollution emission standards should be established and strictly carried out; health and education services should be provided to accommodate agglomeration. In contrast, in scenario 2, the direct and indirect effects are both 0.5%. Such small effects from both channels suggest the inefficiency of infrastructure investments in this region. Hence, the government should stop similar projects to avoid overinvestment. Without separating the two effects, the government cannot get accurate information on the sources of impacts and cannot take proper actions accordingly.

Second, analyzing these two effects is meaningful for financing new projects. Investors will mostly be refunded from fees paid by users. If the direct effect of an existing project is large, investors can easily get profits as a large direct effect implies more usage. The government could then consider co-financing similar projects through public–private partnerships because private investors would be attracted to participate. However, not all infrastructure projects are suitable for such partnerships. If the indirect effect is dominant, the major impact would be the social consequences stemming from agglomeration. These projects would not be as profitable as those with a large direct effect but would benefit the whole region. In this case, as the representative of social interests, the government should take more responsibility by either solely financing or subsidizing private investors. To work out feasible financing plans for future projects, the government must evaluate the direct and indirect effects of those currently existing.

Third, information on the two effects helps government invite suitable industries to invest. If the infrastructure in this area has a relatively large direct effect, the government should approach facility users, such as the food industry, which pays careful attention to transportation systems and seeks to reduce inventories and logistics costs. If the infrastructure has a relatively significant indirect effect, it is better to invite industries that more easily enjoy spillovers from other firms, such as the high-tech industries clustering in the Silicon Valley in California. Based on the different impacts of infrastructure, governments should find industries suited to the respective local areas. Otherwise, even if businesses come to invest, they cannot survive in the long run.

Motivated by the issues mentioned above, this paper distinguishes the direct and indirect effects of infrastructure on firm productivity. The latter arises from the infrastructure–agglomeration link that has been widely overlooked in the literature on infrastructure. An analytical framework is then developed to estimate both effects. Finally, empirical results are obtained using large-scale firm-level survey data from the People’s Republic of China (PRC). Key findings include: (1) all the three kinds of infrastructure—road, telecommunication servers and cable—are found to directly promote firm productivity; (2) they also exert a positive indirect effect on firm productivity through the agglomeration channel; and (3) the empirical results are robust to different estimates of firm productivity, agglomeration indicators and different subsamples.

Our contribution is two-fold: 1) this is one of few papers distinguishing infrastructure effects arising from direct use and agglomeration economies; and 2) We employ large-scale panel data from PRC manufacturing firms. The PRC has made massive infrastructure investments since the 1990s. In our sample period 2002–2007, the total length of roads in the PRC more than doubled from 1.77 million kilometers to 3.58 million kilometers. Therefore, the PRC offers a good dataset to estimate the impacts of infrastructure investments. Also, we use the micro-level survey data from PRC manufacturing. Firm-level data has an advantage over the aggregate level, such as province or city, in that they help circumvent the potential reserved causality from productivity to infrastructure, i.e., the aggregate productivity growth may raise the demand for infrastructure (Fernald 1999; Li and Li 2013).

The paper is organized as follows. Section 2 reviews related literature. Section 3 describes the analytical framework. Section 4 discusses data and the empirical results. Section 5 reports the robustness checks. Section 6 provides policy implications and concludes.

2. LITERATURE REVIEW

The literature on the positive impacts of public infrastructure stock on productivity growth is extensive. Two pioneering works of Aschauer (1989a; 1989b) found that non-military public capital stock, e.g., transportation and water systems, had a strong positive effect on productivity in the United States (US) and other G-7 developed countries during the 1950s and 1960s. Following Aschauer, many studies have confirmed this effect from different perspectives, including theoretical works on growth (Barro 1990; Lynde and Richmond 1992; Gramlich 1994), empirical evidences on long-term productivity (Easterly and Rebelo 1993; Canning and Pederoni 2008), cross-country analyses (Adam and Bevan 2005), and evidence from various kinds of infrastructures (Cronin et al. 1991; Morrison and Schwartz 1996; Demetriades and Mamuneas 2000).

However, when explaining the infrastructure impact, the literature has mainly focused on the direct effect, as indicated by the following points.

1. Infrastructure helps expand market catchment. Hurd (1975) studied infrastructure investment in India from 1861 to 1921 and found that prices across India began to converge and the India-wide market in grains developed as a result of investments. Baum–Snow et al. (2015) analyzed investments in the road and railroad networks in PRC cities and regions. They reached the conclusion that expansions of road infrastructure improved access to nearby markets and significantly promoted local growth after 1990.
2. Infrastructure helps to reduce transaction costs. Using data from the European Union, Bougheas, Demetriades, and Morgenroth (1999) discovered that trade costs substantially declined in association with transportation infrastructure construction. Similar evidence was provided by Röller and Waverman (2001) when estimating the effect of telecommunications infrastructure in 21 Organisation for Economic Co-operation and Development (OECD) countries. Li and Li (2013) analyzed a large panel data set of PRC manufacturing firms and concluded that a dollar in road investment saved two cents in inventory costs.
3. Infrastructure helps access broader labor markets. Duranton and Turner (2012) found that a 10% increase in a city's initial stock of highways led to an employment increase of about 1.5% in the US between 1983 and 2003.

Another strand of literature provides substantial evidence on agglomeration economies. Ciccone and Hall (1996) were the first to examine agglomeration impacts on productivity and found that doubling the employment density in a US county increased average labor productivity by 6%. More recent literature looked at productivity effects from different types of agglomeration externalities. For example, agglomeration in the same industry (termed as localization, Marshall 1890; Romer 1986) and with more diversified industries (termed as urbanization, Jacobs 1969). Henderson (2003) used firm-level panel data from the machinery and high-tech industries and found that a 10-fold increase in the number of local firms in a high-tech industry improved labor productivity by over 20%. Maré and Timmins (2006) concluded that both localization and urbanization lead to higher labor productivities in New Zealand, after controlling for heterogeneity among industries, locations, and companies. Similarly, Lee, Jang, and Hong (2010) revisited agglomeration economies by estimating the effects of localization, urbanization, and local competition on labor productivity in the Republic of Korea and found that firms were more productive if located in more localized, more urbanized, and more competitive areas. In contrast, some studies also came up with evidence of agglomeration diseconomies. Lin, Li, and Yang (2011) reported an inverted U-shape relationship between agglomeration and productivity in the PRC's textile industry, which suggests agglomeration diseconomies appeared if there is too much concentration.

Although the literature on infrastructure and agglomeration is considerable, only in a few cases are these studied together and even fewer studies worked on the indirect infrastructure effect through the agglomeration channel (Eberts and McMillen 1999). A few exceptions include Mera (1973), who discussed the determinants of agglomeration economies and highlighted the role of transportation and communication facilities. Using employment density as the agglomeration measure, he concluded that the high per capita income in high-density areas could be explained by savings in social overhead capital costs, such as public infrastructure. Moomaw (1983) found that after controlling for population agglomeration, transport infrastructure still had a positive effect on productivity in five of six selected industries. Calem and Carlino (1991) recognized that infrastructure and resources endowment were likely to influence urban productivity when analyzing agglomeration economies in US cities. They highlighted the potential bias of omitting the infrastructure effects but also acknowledged the difficulties in measuring them.

3. METHODOLOGY

3.1 Baseline Model

To evaluate the effects of infrastructure on firm productivity, we start with the following conventional model used in previous studies:

$$prod_{i,j,k,t} = \beta_0 + \beta_1 inf_{k,t} + \delta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + u_{i,j,k,t} \quad (1)$$

where the subscripts *i, j*, *k* and *t* index the firm, industry, province, and years, respectively; $prod$ is the total factor productivity (TFP) calculated from a Cobb–Douglas production function and Solow's growth accounting framework;[1] inf stands for the infrastructure in the given province—the three types of infrastructure examined are the length of road in kilometers, the number of telecommunication servers, and the length of cable in kilometers (all in logarithms); X is a vector of firm-specific control variables from existing studies, including firm size (*asset*) measured by the logarithm of total assets; capital intensity (*cap*) measured by capital per employee; export status (*export*) measured by the share of exports to total sales; ownership dummy variable about foreign owned or privately owned (*poe* and *foe*); and firm age (*age*) measured by years of establishment. Further, ρ, π and θ are employed to indicate the fixed effects of the k-th province, the j-th industry and the t-th year, respectively. Finally, u is the random error term.

Conventional studies tend to ignore agglomeration when examining the infrastructure impacts. However, agglomeration economies may affect the production of all firms within a region. If the agglomeration is led by infrastructure investment $(Cov(inf, agg) \neq 0)$, omitting it contaminates the estimation of the infrastructure effect. Therefore, we incorporate a province–industry concentration indicator *agg* into Model (1), as follows:

$$prod_{i,j,k,t} = \alpha_0 + \alpha_1 inf_{k,t} + \alpha_2 agg_{j,k,t} + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} \quad (2)$$

where agg is defined as the ratio of the sales of industry j in province k to the sales of the industry j in the country. In the robustness check, we use the assets ratio as an alternative indicator. The rationale here is to measure the economic density using the

[1] TFP is estimated by a panel data from 2002–2007 and two-digit level industry, and defined as $TFP_{it}^{LP} = exp(VA_{it} - \hat{\beta}_K^{LP} K_{it} - \hat{\beta}_L^{LP} L_{it})$, where subscripts *i* and *t* denote firm and year, respectively; $\hat{\beta}_K^{LP}, \hat{\beta}_L^{LP}$ denote the estimators of K_{it} and L_{it} by the Levinsohn–Petrin semiparametric estimation approach (Levinsohn and Petrin 2003); VA_{it} and K_{it} are calculated and deflated as defined in Appendix A. TFP is used in logs.

business activities of an industry in one province to the country total (Calem and Carlino 1991).[2]

3.2 The Endogeneity Issue

Previous literature finds that firms with higher productivity might self-select to locate in business-intense places (Melitz 2003; Combes et al. 2012), whereas less productive firms tend to exit these areas to avoid tough competition (Melitz and Ottaviano 2008). Therefore, causality might flow from firm productivity to agglomeration rather than the other way. To alleviate this potential endogeneity, we implement the two-stage least squares (2SLS) method with a first-stage regression explaining industrial agglomeration, which is specified as:

$$agg_{j,k,t} = \gamma_0 + \gamma_1 inf_{k,t} + \beta' Z_{j,k,t} + \varepsilon_{j,k,t}. \quad (3)$$

In Model (3), we incorporate the province-level infrastructure and a group of variables from a previous study on PRC industrial agglomeration (Ge 2009), including export intensity (*ind_export*), industry scale (*size*), tax-plus-profit margin (*taxprofit*), share of state-owned assets (*ind_soe*), and share of foreign-owned assets (*ind_fdi*). ε is the error term. Based on the first stage result, we predict the agglomeration indicator and substitute it into the second-stage model.

3.3 The Direct and Indirect Effects of Infrastructure

From the 2SLS model, we can distinguish the direct and indirect infrastructure effects. By substituting Eq. (3) into Eq. (2), we get Eq. (4):

$$\begin{aligned} prod_{i,j,k,t} &= \alpha_0 + \alpha_1 inf_{k,t} + \alpha_2(\gamma_0 + \gamma_1 inf_{k,t} + \beta' Z_{j,k,t} + \varepsilon_{j,k,t}) + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} \\ &= \alpha_0 + \alpha_1 inf_{k,t} + \alpha_2\gamma_0 + \alpha_2\gamma_1 inf_{k,t} + \alpha_2\beta' Z_{j,k,t} + \alpha_2\varepsilon_{j,k,t} + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} \\ &= \alpha_0 + \alpha_2\gamma_0 + (\alpha_1 + \alpha_2\gamma_1) inf_{k,t} + \alpha_2\beta' Z_{j,k,t} + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} + \alpha_2\varepsilon_{j,k,t} \end{aligned} \quad (4)$$

We note several points here: first, the last line of Eq. (4) clearly shows the two different sources of the total effect, the direct effect (α_1 in Eq. 2) and the indirect effect ($\alpha_2\gamma_1$) arising from the infrastructure–agglomeration link (α_2 in Eq. 2 and γ_1 in Eq. 3). Second, the sign and size of the indirect effect depends on two components, the impact of agglomeration (α_2) and the scale of agglomeration caused by infrastructure (γ_1). The indirect effect could be negative in the case of either agglomeration diseconomies ($\alpha_2 < 0$) or if infrastructure fails to promote agglomeration ($\gamma_1 < 0$). If both of them happen, the indirect effect could become positive. This is because the infrastructure alleviates over-concentration of economic activities in local areas, perhaps by providing transportation services and information for businesses and households to migrate to other places. As a result, the diseconomies would be mitigated. Third, the two different effects cannot be disentangled by the conventional model, where agglomeration channel is ignored. Conventional studies might obtain the same result with our analysis on the total effect ($\beta_1 = \alpha_1 + \alpha_2\gamma_1$), however, without considering the agglomeration–infrastructure link the two different sources are concealed and the overall effect is likely to be explained as the direct effect.

---

[2] There are several widely used agglomeration indicators, such as spatial Gini coefficients (Krugman 1991) and Ellison–Glaeser index (Ellison and Glaeser 1997). However, these indexes are not applicable to our situation. They are employed as country-level agglomeration indicators to measure industrial agglomeration within a country. Here, we need a province-level industry agglomeration indicator, which must be comparable across different provinces.

4. DATA AND EMPIRICAL RESULTS

4.1 Data

We use firm-level data from the China Industrial Enterprise Database (CIED), which contains data from the annual enterprise census conducted by the National Bureau of Statistics of China. CIED includes all state-owned enterprises and private-owned enterprises with a criterion of sales above five million yuan. Our dataset involves 440,490 firms over the period 2002–2007. Infrastructure data are from the Statistical Yearbooks of 31 provinces in the PRC. The industrial variables used in Eq. (3) are aggregated from our dataset.

Table 1 Panel A reports descriptive statistics for variables used in empirical analyses. The numbers of observations vary due to the different aggregate levels for each variable. Particularly, infrastructure is at province level; agglomeration indicators are at industry level; and productivity measures are at firm level. Firms are classified into 30 manufacturing industries by two-digit industrial classification, which incorporates most forward and backward linkages other than three-digit and four-digit classifications and thereby enables us to estimate the agglomeration effects arising from the widest spillovers channels. The last line of Panel A shows a regional dummy variable, *westmid*, which is defined to have value 1 for firms located in Western and Central provinces and is equal to 0 if in the Eastern provinces.[3] Because of the asset criteria for the CIED database, it includes more firms from Eastern provinces, where the economy is relatively developed.

Table 1: Descriptive Statistics

Panel A: Descriptive Statistics of Variables Used in Regression Models					
Variable	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
road	186	77,510.58	51,214.44	6,286	238,676
tele	186	1,588.996	1,555.583	31	11,365.8
cable	186	20,241.34	10,900.45	618	55,910
agg_sal	5,375	0.03	0.05	0.00	0.38
agg_ast	5,375	0.03	0.05	0.00	0.40
tfp	1,314,378	6.48	1.13	–3.85	12.93
asset	1,335,926	76,571.59	681,988.60	3.00	155,000,000.00
cap	1,335,926	74.37	109.08	0.98	691.60
export	1,335,589	0.18	0.35	0.00	1.00
age	1,335,311	9.69	9.48	1.00	51.00
poe	1,335,926	0.46	0.50	0.00	1.00
foe	1,335,926	0.22	0.42	0.00	1.00
westmid	1,335,926	0.23	0.42	0.00	1.00

continued on next page

[3] Eastern PRC covers 11 provinces and regions: Beijing, Tianjin, Shanghai, Hebei, Liaoning, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, and Hainan. Central and Western PRC covers 21 provinces and regions: Shanxi, Inner Mongolia, Jilin, Heilongjiang Anhui, Jiangxi, Henan, Hubei, Hunan, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Tibet, Ningxia, and Xinjiang.

Table 1 *continued*

Panel B: Infrastructure Variables						
	2002	**2003**	**2004**	**2005**	**2006**	**2007**
All country						
road	0.38	0.39	0.41	0.42	0.72	0.74
tele	0.02	0.02	0.02	0.02	0.03	0.04
cable	0.10	0.12	0.13	0.14	0.14	0.15
Eastern						
road	0.57	0.59	0.62	0.64	0.96	0.99
tele	0.04	0.04	0.05	0.06	0.07	0.08
cable	0.15	0.19	0.20	0.22	0.20	0.21
Western and Central						
road	0.26	0.26	0.27	0.28	0.56	0.58
tele	0.00	0.00	0.00	0.01	0.01	0.01
cable	0.07	0.08	0.09	0.09	0.10	0.11

Panel C: Agglomeration Variable						
	2002	**2003**	**2004**	**2005**	**2006**	**2007**
All country	0.03	0.03	0.03	0.03	0.03	0.03
Eastern	0.06	0.07	0.07	0.06	0.06	0.06
Western&Central	0.014	0.013	0.012	0.013	0.013	0.014

Infrastructure variables exhibit great variability in our dataset, indicating that the infrastructure investments are quite different among PRC provinces. We further summarize the infrastructure density[4] by regions in Panel B of Table 1. On country average, the three kinds of infrastructure increased steadily from 2002 to 2007. In particular, the length of roads and the number of telecommunication servers almost doubled, showing the huge investment in the PRC in transportation. This trend can be found in both the sub-regions as well. Moreover, the Eastern provinces had a higher density in all three measurements of infrastructure, almost double those for the Western and Central provinces. In addition, businesses and households are more concentrated in the Eastern provinces. Panel C of Table 1 presents the averaged agglomeration index for all industries in the time period considered. Compared with the West and Central provinces, the Eastern provinces had a much higher degree of agglomeration. Based on these spatial variances, we separate the dataset by regions and test the different infrastructure impacts as the robustness check.

4.2 Baseline Results

The two baseline models (1) and (2) are estimated using the fixed effects method. Table 2 presents the empirical results. Columns 1–3 are based on Model (1) and columns 4–6 are based on Model (2). The coefficients of the three infrastructure indicators are estimated to be positive values that are statistically significant at the 1% level in all models. After incorporating agglomeration in columns 4–6, the coefficients of road and telecommunication server decrease by 6.5% ([0.2728–0.2552]/0.2728) and 11.5% ([0.3559–0.3149]/0.3559), respectively, and the coefficient of cable increases only slightly by 3.0%. All agglomeration indicators in columns 4–6 are positive,

[4] We divide the infrastructure variables by territory size of each province to obtain infrastructure density.

which suggests significant agglomeration economies on firm productivity. Ignoring this agglomeration impact leads to an overestimation of direct infrastructure effect.

The coefficients of the control variables are entirely consistent with previous studies (Huang and Zhang 2016). Larger companies and longer-established companies have relatively stronger abilities in innovation, and, therefore, exhibit higher productivities. The foreign and private-owned firms usually possess advanced technologies and are more willing to adopt new technologies, so they have relatively higher productivities than local and state-owned companies.

Table 2: Baseline Results

	Dependent Variable: Firm-level TFP					
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**
lnroad	0.2728***			0.2552***		
	(0.0069)			(0.0069)		
lntele		0.3559***			0.3149***	
		(0.0069)			(0.0070)	
lncable			0.1670***			0.1720***
			(0.0068)			(0.0068)
agg				0.499***	0.363***	0.564***
				(0.014)	(0.014)	(0.014)
asset	0.3862***	0.3881***	0.3881***	0.3866***	0.3884***	0.3883***
	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0025)	(0.0025)
cap	–0.1890***	–0.1885***	–0.1881***	–0.1887***	–0.1882***	–0.1878***
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
export	–0.0042	–0.002	–0.0033	–0.0023	–0.0008	–0.0011
	(0.0050)	(0.0050)	(0.0050)	(0.0050)	(0.0050)	(0.0050)
age	0.1265***	0.1259***	0.1272***	0.1265***	0.1261***	0.1271***
	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0027)
poe	0.0216***	0.0232***	0.0225***	0.0217***	0.0233***	0.0224***
	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)	(0.0039)
foe	0.0399***	0.0385***	0.0373***	0.0417***	0.0400***	0.0393***
	(0.0098)	(0.0098)	(0.0098)	(0.0098)	(0.0098)	(0.0098)
cons	0.04	0.14	1.29***	0.18	0.41	1.23**
	(0.48)	(0.49)	(0.47)	(0.49)	(0.49)	(0.48)
province_dummy	yes	yes	yes	yes	yes	yes
year_dummy	yes	yes	yes	yes	yes	yes
industry_dummy	yes	yes	yes	yes	yes	yes
N	1,313,465	1,313,465	1,313,465	1,313,465	1,313,465	1,313,465
adj. R-sq	0.211	0.212	0.209	0.212	0.212	0.211

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5% and 1% levels, respectively. Robust standard errors are employed for statistical inference.

4.3 2SLS Results

Table 3 presents results for the first stage of the 2SLS regressions. The infrastructure indicators are estimated to have positive coefficients in each of the cases, but only road and telecommunication servers are significant at the 10% and 5% level, respectively, implying that the construction of transport and information infrastructure in a province promotes the business concentration for all its industries. This might be because new firms are established in the areas or the operation conditions of existing firms are improved along with infrastructure construction. Adjusted R^2s in Table 3 are as high as 0.74, which suggests a good fitness of the first-stage model.

Table 3: First-stage Results of 2SLS: Agglomeration Impact on Infrastructure

	Dependent Variable: Agglomeration		
	(1)	**(2)**	**(3)**
lnroad	0.0053*		
	(0.0032)		
lntele		0.0100**	
		(0.0044)	
lncable			0.0044
			(0.0031)
ind_export	–0.0187***	–0.0188***	–0.0188***
	(0.0043)	(0.0043)	(0.0043)
taxprofit	0.074***	0.074***	0.074***
	(0.020)	(0.020)	(0.020)
ind_soe	–0.003	–0.003	–0.0032
	(0.0027)	(0.0027)	(0.0027)
size	0.0240***	0.0240***	0.0240***
	(0.0006)	(0.0006)	(0.0006)
ind_fdi	0.0070**	0.0068**	0.0068**
	(0.0029)	(0.0029)	(0.0029)
cons	–0.280***	–0.298***	–0.265***
	(0.031)	(0.031)	(0.026)
province_dummy	yes	yes	yes
year_dummy	yes	yes	yes
industry_dummy	yes	yes	yes
N	4495	4495	4495
adj. R-sq	0.739	0.739	0.739

2SLS = two-stage least squares.

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5% and 1% levels, respectively. Robust standard errors are employed for statistical inference.

Table 4 summarizes the second-stage results of 2SLS estimation. The predicted agglomeration indicators are positive in all three models at the 1% significance level and the coefficients are uniformly larger than those in Table 2. The coefficients for three infrastructure measures are still smaller than those in columns 1–3 of Table 2, which confirms that the direct infrastructure impacts are lower with the agglomeration effect accounted for.

Table 4: Second-stage Results of 2SLS

	Dependent Variable: Firm-level TFP		
	(1)	**(2)**	**(3)**
agg_sal_road	1.18***		
	(0.19)		
lnroad	0.2668***		
	(0.0070)		
agg_sal_tele		1.07***	
		(0.19)	
lntele		0.3451***	
		(0.0072)	
agg_sal_cable			1.15***
			(0.19)
lncable			0.1622***
			(0.0069)
asset	0.3857***	0.3878***	0.3877***
	(0.0025)	(0.0025)	(0.0025)
cap	–0.1887***	–0.1883***	–0.1878***
	(0.0015)	(0.0015)	(0.0015)
export	–0.0044	–0.0022	–0.0035
	(0.0050)	(0.0050)	(0.0050)
age	0.1271***	0.1266***	0.1279***
	(0.0027)	(0.0027)	(0.0027)
poe	0.0218***	0.0233***	0.0227***
	(0.0039)	(0.0039)	(0.0039)
foe	0.0400***	0.0388***	0.0376***
	(0.0098)	(0.0098)	(0.0098)
cons	0.22	0.34	1.45***
	(0.51)	(0.51)	(0.50)
province_dummy	yes	yes	yes
year_dummy	yes	yes	yes
industry_dummy	yes	yes	yes
N	1,304,472	1,304,472	1,304,472
adj. R–sq	0.211	0.212	0.21

2SLS = two-stage least squares; TFP = total factor productivity.

Notes*:* Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5% and 1% levels, respectively. Robust standard errors are employed for statistical inference.

To make a clear distinction between the direct and indirect effects, we summarize the above results in Table 5. Columns 1 and 2 show the different results of direct effect estimated in the conventional model and our 2SLS analysis, whereas column 3 reports our results for indirect effect. Two points are worth noting: *i*) the estimates for the direct effect in Model 1 (β_1) are equal to our total effect in column 6. However, existing studies could not disentangle different sources. In fact, β_1 consists of a "real" direct effect and an indirect effect. Ignoring the agglomeration–infrastructure link, the existing studies overestimate the direct effect. For example, the direct effect of road investments is 0.273 in column 1 but it is 0.267 in our analysis. The differential of 0.006 is the indirect effect from the agglomeration channel. An exaggerated direct effect would lead to over-expectations on benefits to users. *ii)* Columns 4 and 5 show the two components of the indirect effect, i.e., the agglomeration effect on productivity, α_2, and the infrastructure effect on agglomeration, γ_1. The large magnitude of α_2 suggests strong agglomeration economies. The estimate for γ_1 is small for all three infrastructure measurements, implying the infrastructure-led agglomeration is relatively small in the PRC.

Table 5: Direct and Indirect Effects of Infrastructure

	(1) Direct Effect in Eq. (1)	(2) Direct Effect in Eq. (4)	(3) Indirect in Eq. (4)	(4)	(5)	(6) (2)+(3)
	β_1	α_1	$\alpha_2\gamma_1$	α_2	γ_1	$\alpha_1 + \alpha_2\gamma_1$
Road	0.273	0.267	0.006	1.181	0.005	0.272905
Tel	0.356	0.345	0.011	1.069	0.01	0.35569
Cable	0.167	0.162	0.005	1.148	0.004	0.166592

Notes*:* β_1 is coefficient from equation (1) and its estimate is from Table 2; α_1, α_2, γ_1 are coefficients from equation (4); estimates of α_1 and α_2 are from Table 4 and estimate of γ_1 is from Table 3.

5. ROBUSTNESS CHECK

5.1 Different Agglomeration Measures

The agglomeration indicator in our main text is the sales concentration. To check for robustness of our results, in this part we use industrial-level total assets to measure business activities. The results are reported in Table 6.

Table 6 reports similar results to Table 4, in that: 1) all agglomeration indicators are positive and significant. Although the magnitudes are smaller than those in Table 4, they are still higher than the direct infrastructure effect, implying significant agglomeration economies on firm productivity. 2) The coefficients for three kinds of infrastructure are smaller than in the first three columns of Table 2, confirming the overestimation of direct infrastructure effect in conventional models.

5.2 Different Infrastructure Effects between Eastern and Western and Central Provinces in the PRC

Our dataset covers firms from different regions in the PRC, resulting in a quite heterogeneous dataset regarding infrastructure investment and economic agglomerations. To estimate the different impacts with spatial variability, we add a regional dummy (*westmid)* and its interactive term with different infrastructure effect in Eq. (4).

Table 6: Using different Agglomeration Measures

	Dependent Variable: Firm-level TFP		
	(1)	**(2)**	**(3)**
agg_ast_road	0.73***		
	(0.18)		
lnroad	0.2720***		
	(0.0069)		
agg_ast_tele		0.65***	
		(0.18)	
lntele		0.3538***	
		(0.0070)	
agg_ast_cable			0.64***
			(0.18)
lncable			0.1645***
			(0.0069)
asset	0.3859***	0.3879***	0.3879***
	(0.0025)	(0.0025)	(0.0025)
cap	–0.1888***	–0.1883***	–0.1878***
	(0.0015)	(0.0015)	(0.0015)
export	–0.0045	–0.0023	–0.0036
	(0.0050)	(0.0050)	(0.0050)
age	0.1272***	0.1268***	0.1281***
	(0.0027)	(0.0027)	(0.0027)
poe	0.0217***	0.0233***	0.0226***
	(0.0039)	(0.0039)	(0.0039)
foe	0.0401***	0.0388***	0.0376***
	(0.0098)	(0.0098)	(0.0098)
cons	0.1745	0.2878	1.4388***
	–0.5118	–0.514	–0.5003
dummy_provice	yes	yes	yes
dummy_year	yes	yes	yes
dummy_industry	yes	yes	yes
N	1,304,472	1,304,472	1,304,472
adj. R-sq	0.211	0.212	0.21

TFP = total factor producticvity.

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5%, and 1% levels, respectively. Robust standard errors are employed for statistical inference.

5.2.1 Direct Effect of Infrastructure Investment

First, we estimate the differences of direct effects between the two sub-regions by incorporating the interaction between infrastructure investment and a regional dummy.

$$prod_{i,j,k,t} = \alpha_0 + \alpha_1 inf_{k,t} + \alpha_2 agg_{i,k,t} + \alpha_3 inf_{k,t} * westmid_k + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} \quad (5)$$

Table 7: Different direct Infrastructure Effects on Productivity between Eastern and Central and Western Provinces (Central and Western = 1, Eastern = 0)

	Dependent Variable: Firm-level TFP		
	(1)	**(2)**	**(3)**
agg_sal_road	0.48***		
	(0.15)		
lnroad	0.2034***		
	(0.0079)		
westmid*road	0.1122***		
	(0.0061)		
agg_sal_tele		0.67***	
		(0.15)	
lntele		0.2836***	
		(0.0074)	
westmid*tele		0.1820***	
		(0.0057)	
agg_sal_cable			0.52***
			(0.15)
lncable			0.1107***
			(0.0079)
westmid*cable			0.249***
			(0.021)
asset	0.3860***	0.3876***	0.3883***
	(0.0025)	(0.0025)	(0.0025)
cap	–0.1887***	–0.1889***	–0.1876***
	(0.0015)	(0.0015)	(0.0015)
export	–0.004	–0.0024	–0.0026
	(0.0051)	(0.0051)	(0.0051)
age	0.1271***	0.1287***	0.1281***
	(0.0027)	(0.0027)	(0.0027)
poe	0.0215***	0.0222***	0.0228***
	(0.0039)	(0.0039)	(0.0039)
foe	0.0385***	0.0380***	0.0380***
	(0.0099)	(0.0099)	(0.0100)
cons	0.369***	0.490***	1.221***
	(0.082)	(0.056)	(0.072)
province_dummy	yes	yes	yes
year_dummy	yes	yes	yes
industry_dummy	yes	yes	yes
N	1,304,472	1,304,472	1,304,472
adj. R-sq	0.197	0.199	0.196

TFP = total factor productivity.

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5%, and 1% levels, respectively. Robust standard errors are employed for statistical inference.

Table 7 presents the results for Model (5). Consistent with our main results, infrastructure and agglomeration positively affect firm productivity. Moreover, the interaction term of infrastructure and the regional dummy is significantly positive, implying a greater direct effect in Western and Central provinces than for the Eastern provinces. One unit of investment in road, telecommunication, and cable enhances 0.112, 0.182, and 0.249 units of firm productivity more, respectively, in Western and Central regions than in the Eastern provinces. These sharp differences might be because Western and Central provinces are short of infrastructure facilities and have more remote areas unconnected to the country's main transportation and information system. Infrastructure constructions would bring more direct benefit by providing more services to the users in these regions. For example, roads help firms access large markets; communication infrastructure aids the acquisition of information and technologies. For the remote areas, construction of an infrastructure project might be a "zero to one" change and thereby could bring huge benefits to users. In contrast, infrastructure in the Eastern provinces is relatively abundant, where the transportation and communication facilities are much more developed. An extra construction just provides one alternative way to users. The marginal benefit could be much lower in terms of usage.

5.2.2 Indirect Effect of Infrastructure Investment

To estimate the difference effect through the agglomeration channel in different regions, we incorporate the interactive term of agglomeration and regional dummy in the second-stage regression of the 2SLS model.

$$prod_{i,j,k,t} = \alpha_0 + \alpha_1 inf_{k,t} + \alpha_2 agg_{i,k,t} + \alpha_3 agg_{i,k,t} * westmid_k + \zeta' X_{i,j,k,t} + \rho_k + \pi_j + \theta_t + \nu_{i,j,k,t} \quad (6)$$

Table 8 shows the empirical results from estimating the model of Eq. (6). The coefficient of the interactive term is significantly positive, implying a larger agglomeration effect for the Western and Central provinces. This is in line with the current situation in the PRC that economic density is much higher in Eastern provinces, regarding more businesses and households clustering. Agglomeration economies in these provinces might not be as high as in Western and Central regions because the benefits might be offset by the negative consequence of agglomeration. For example, transportation congestion and pollution might lead to diseconomies (Lin, Li, and Yang 2011). In contrast, the businesses and households are relatively scattered in Western and Central PRC and the adverse effects of agglomeration might not be high at this stage.

In addition, to test the effect of infrastructure on promoting agglomeration in different regions, we incorporate the interaction term of infrastructure and the regional dummy in the first-stage estimation of the 2SLS model.

$$agg_{j,k,t} = \gamma_0 + \gamma_1 inf_{k,t} + \gamma_2 inf_{k,t} * westmid_k + \beta' Z_{j,k,t} + \varepsilon_{j,k,t} \quad (7)$$

Table 8: Different Agglomeration Effects on Productivity between Eastern and Central and Western Provinces (Central and Western = 1, Eastern = 0)

	Dependent Variable: Firm-level TFP		
	(1)	**(2)**	**(3)**
agg_sal_road	–0.09		
	(0.10)		
lnroad	0.2672***		
	(0.0070)		
westmid*road	8.69***		
	(0.77)		
agg_sal_tele		–0.16	
		(0.10)	
lntele		0.3540***	
		(0.0070)	
westmid*tele		9.33***	
		(0.78)	
agg_sal_cable			–0.024
			(0.099)
lncable			0.1588***
			(0.0070)
westmid*cable			9.26***
			(0.77)
asset	0.3853***	0.3872***	0.3872***
	(0.0025)	(0.0025)	(0.0025)
cap	–0.1884***	–0.1879***	–0.1873***
	(0.0015)	(0.0015)	(0.0015)
export	–0.0033	–0.0011	–0.0023
	(0.0051)	(0.0051)	(0.0051)
age	0.1256***	0.1251***	0.1263***
	(0.0027)	(0.0027)	(0.0027)
poe	0.0212***	0.0227***	0.0222***
	(0.0039)	(0.0039)	(0.0039)
foe	0.0382***	0.0370***	0.0360***
	(0.0099)	(0.0099)	(0.0099)
cons	–0.026	0.286***	1.331***
	(0.078)	(0.056)	(0.070)
province_dummy	yes	yes	yes
year_dummy	yes	yes	yes
industry_dummy	yes	yes	yes
N	1,304,472	1,304,472	1,304,472
adj. R-sq	0.196	0.198	0.195

TFP = total factor productivity.

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and ***, indicate that the coefficients are significantly different from zero at the 10%, 5%, and 1% levels, respectively. Robust standard errors are employed for statistical inference.

Table 9 tabulates the empirical results obtained by estimating the model of Eq. (7). The coefficient of the interaction term is significantly negative, which implies that the infrastructure more easily promotes agglomeration in the Eastern provinces than in the Western and Central provinces. It is interesting that one unit investment in the Eastern region will attract more clustering than in the Western and Central region. It might be because businesses and households are more willing to go to the Eastern provinces than to the Western and Central provinces. The construction of infrastructure might aid migrations outflow in the less developed areas.

Table 9: Different Infrastructure Effects on Agglomeration between Eastern and Central and Western Provinces (Central and Western = 1, Eastern = 0)

	Dependent Variable: Agglomeration		
	(1)	**(2)**	**(3)**
lnroad	0.0124***		
	(0.0008)		
westmid*lnroad	–0.0033***		
	(0.0001)		
lntele		0.0189***	
		(0.0010)	
westmid*lntele		–0.0037***	
		(0.0002)	
lncable			0.0105***
			(0.0008)
westmid*lncable			–0.0037***
			(0.0002)
ind_export	0.0023	0.0013	0.0025
	(0.0029)	(0.0029)	(0.0030)
taxprofit	0.083***	0.065***	0.078***
	(0.014)	(0.014)	(0.015)
ind_soe	–0.0169***	–0.0143***	–0.0175***
	(0.0019)	(0.0019)	(0.0019)
size	0.0193***	0.0165***	0.0191***
	(0.0005)	(0.0005)	(0.0005)
ind_fdi	0.0015	–0.0034	–0.0005
	(0.0022)	(0.0022)	(0.0022)
cons	–0.2809***	–0.2494***	–0.2456***
	(0.0091)	(0.0068)	(0.0076)
province_dummy	yes	yes	yes
year_dummy	yes	yes	yes
industry_dummy	yes	yes	yes
N	4,495	4,495	4,495
adj. R-sq	0.6	0.615	0.591

Notes: Standard errors are reported in parentheses. The asterisks, *, **, and *** indicate that the coefficients are significantly different from zero at the 10%, 5%, and 1% levels, respectively. Robust standard errors are employed for statistical inference.

6. CONCLUSIONS

This paper distinguishes and estimates the direct and indirect effects of infrastructure on firm productivity. Using data from PRC manufacturing and the infrastructure construction from PRC provinces, we find both effects are significantly positive in the PRC. Although existing studies were able to reach the same results regarding the total effect, ignoring of infrastructure–agglomeration link covers the channels of infrastructure effects and leads to overestimation of the direct effect.

Our results have profound policy implications for governments: 1) along with infrastructure investment, the government should pay attention to increasing agglomeration, especially in the relatively developed areas, like the Eastern provinces of the PRC. Otherwise the adverse impacts of clustering might offset the direct infrastructure benefits. 2) For regions relatively short of infrastructures, the direct effect would be high. When financing projects in these regions, the government could consider public–private partnerships, because the investors could get higher returns from the users. 3) Infrastructure in remote areas would lead to less agglomeration than in the developed areas. When inviting firms to invest, the government should focus more on the infrastructure users than on the direct beneficiaries from agglomerations.

REFERENCES

Adam, C. S., and D. L. Bevan. 2005. Fiscal Deficits and Growth in Developing Countries. *Journal of Public Economics* 89: 571–597.

Aschauer, D. A. 1989a. Is Public Expenditure Productive? *Journal of Monetary Economics* 23(2): 177–200.

———. 1989b. Does Public Capital Crowd out Private Capital? *Journal of Monetary Economics* 24(2): 171–188.

Barro, R. J. 1990. Government Spending in a Simple Model of Endogeneous Growth. *Journal of Political Economy* 98(5): 103–125.

Baum–Snow, N., L. Brandt, V. Henderson, M. Turner, and Q. Zhang. 2015. Transport Infrastructure, Urban Growth and Market Access in China. *ERSA Conference Papers*. No. ersa15p1177. European Regional Science Association.

Bougheas, S., P. O. Demetriades, and E. L. W. Morgenroth. 1999. Infrastructure, Transport Costs and Trade. *Journal of International Economics* 47: 169–189.

Calem, P. S., and G. A. Carlino. 1991. Urban Agglomeration Economies in the Presence of Technical Change. *Journal of Urban Economics* 29(1): 82–95.

Canning, D., and P. Pederoni. 2008. Infrastructure, Long-run Economic Growth and Causality Tests for Cointergrated Panels. *Manchester School* 76(5): 504–527.

Ciccone, A., and R. E. Hall. 1996. Productivity and the Density of Economic Activity Productivity and the Density of Economic Activity. *The American Economic Review* 86(1): 54–70.

Combes, P.-P., G.Duranton, L.Gobillon, D. Puga, B. Pierre–Philippe Combes, and S. Roux. 2012. The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection. *Econometrica* 80(6): 2543–2594.

Cronin, F. J., E. B. Parker , E. K. Colleran, and M. A. Gold. 1991. Telecommunications Infrastructure and Economic Growth: An Analysis of Causality. *Telecommunications Policy* 15(6): 529–535.

Demetriades, P. O., and T. P. Mamuneas. 2000. Intertemporal Output and Employment Effects of Public Infrastructure Capital: Evidence from 12 OECD Economies. *The Economic Journal* 110(465): 687–712.

Démurger, S. 2001. Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China? *Journal of Comparative Economics* 29(1): 95–117.

Duranton, G., and M. A. Turner. 2012. Urban Growth and Transportation. *Review of Economic Studies* 79: 1407–1440.

Easterly, W., and S. Rebelo. 1993. Fiscal Policy and Economic Growth An Empirical Investigation. *Journal of Monetary Economics* 32: 417–458.

Eberts, R. W., and D. P. McMillen. 1999. Agglomeration Economies and Urban Public Infrastructure. *Handbook of Regional and Urban Economics* 3: 1455–1495.

Ellison, G., and E. L. Glaeser. 1997. Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy* 105(5): 889–927.

Fernald, J. G. 1999. Roads to Prosperity? Assessing the Link between Public Capital and Productivity. *The American Economic Review 89*(3): 619–638.

Fujita, M., P. Krugman, and T. Mori. 1999. On the Evolution of Hierarchical Urban Systems. *European Economic Review* 43(2): 209–251.

Fujita, M., P. Krugman, and A. J. Venables. 2001. *The Spatial Economy: Cities, Regions, and International Trade* (Edition 1). Cambridge, MA: MIT Press.

Fujita, M., and J.-F. Thisse. 2013. *Economics of Agglomeration*. Cambridge University Press.

Ge, Y. 2009. Globalization and Industry Agglomeration in China. *World Development* 7(3): 550–559.

Gramlich, E. M. 1994. Infrastructure Investment: A Review Essay. *Journal of Economic Literature* 32(3): 1176–1196.

Helpman, E., and P. Krugman. 1985. *Market Structure and Foreign Trade Increasing Returns, Imperfect Competition, and the International Economy*. Cambridge, MA: MIT Press.

Henderson, V. 2003. The So-What Question. *Journal of Economic Growth*, *8*(8): 47–71.

Holl, A. 2006. A Review of the Firm-Level Role of Transport Infrastructure with Implications for Transport Project Evaluation. *Journal of Planning Literature* 21 (August): 3–14.

Huang, Y., and Y. Zhang. 2016. Wage, Foreign-owned Firms, and Productivity Spillovers via Labour Turnover: a Non-linear Analysis Based on Chinese Firm-level Data. *Applied Economics* (forthcoming).

Hurd, J. 1975. Railways and the Expansion of Markets in India, 1861–1921. *Explorations in Economic History* 12(3): 263–288.

Jacobs, J. 1969. *The Economy of Cities.* New York: Random House.

Krugman, P. 1991. *Geography and Trade*. Cambridge, MA: MIT Press.

Lee, S. B., S. Jang, and H. S. Hong. 2010. Marshall's Scale Economies and Jacobs' Externality in Korea: the Role of Age, Size and the Legal Form of Organisation of Establishments.

Levinsohn, J., and A. Petrin. 2003. Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies* 70(2): 317–341.

Lewis, D., and R. Bloch. 1998. SDIs: Infrastructure, Agglomeration and the Region in Industrial Policy. *Development Southern Africa* 15(5): 727–755.

Li, H., and Z. Li. 2013. Road Investments and Inventory Reduction: Firm Level Evidence from China. *Journal of Urban Economics* 76: 43–52.

Lin, H. L., H. Y. Li, and C. H. Yang. 2011. Agglomeration and Productivity: Firm-level Evidence from China's Textile Industry. *China Economic Review* 22: 313–329.

Lynde, C., and J. Richmond. 1992. The Role of Public Capital in Production. *The Review of Economics and Statistics* 74(1): 37–44.

Maré, D. C., and J. Timmins. 2006. *Geographic Concentration and Firm Productivity*. Motu Working Paper 06–08.

Marshall, A. 1890. *Principles of Economics*. London: Macmillan.

McCann, P., and D. Shefer. 2003. Location, Agglomeration and Infrastructure. *Papers in Regional Science* 83: 177–196.

Melitz, M. J. 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71(6): 1695–1725.

Melitz, M. J., and G. I. P. Ottaviano. 2008. Market Size, Trade, and Productivity. *Review of Economic Studies* 75: 295–316.

Mera, K. 1973. Regional Production Functions and Social Overhead Capital: An Analysis of the Japanese Case. *Regional and Urban Economics* 3(2): 157–186.

Mikelbank, B. A., and R. W. Jackson. 2000. The Role of Space in Public Capital Research. *International Regional Science Review* 23(3): 235–258.

Moomaw, R. L. 1983. Is Population Scale a Worthless Surrogate for Business Agglomeration Eeconomies? *Regional Science and Urban Economics* 13: 525–545.

Morrison, C. J., and A. E. Schwartz. 1996. State Infrastructure and Productive Performance. *American Economic Review* 6(5): 1095–1111.

Röller, L.-H., and L. Waverman. 2001. Telecommunications Infrastructure and Economic Development: A Simultaneous Approach. *The American Economic Review* 91: 909–923.

Romer, P. M. 1986. Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94(5): 1002–1037.

APPENDIX A: DEFINITIONS OF VARIABLES

Variable	Definition
lntfp	Log value of Total Factor Productivity using Levinsohn–Petrin semi-parametric estimation approach
agg_sal	Share of sales in a province to the country total of one industry.
agg_ast	Share of asset in a province to the country total of one industry.
agg_sal_road	Predicted sales agglomeration from first stage of 2SLS using road
agg_sal_tele	Predicted sales agglomeration from first stage of 2SLS using telecommunication server
agg_sal_cable	Predicted sales agglomeration from first stage of 2SLS using cable
agg_ast_road	Predicted assets agglomeration from first stage of 2SLS using road
agg_ast_tele	Predicted assets agglomeration from first stage of 2SLS using telecommunication server
agg_ast_cable	Predicted assets agglomeration from first stage of 2SLS using cable
lnroad	Log value of the length of road in one province
lntele	Log value of the number of telecommunication servers in one province
lncable	Log value of the length of cable in one province
asset	Firm size, measured by log value of firm's total assets
cap	Capital intensity, measured by log value of firm's capital per employee
export	Export status, measured by firm's share of exports to total sales
age	Firm age, measured by years of establishment
poe	Dummy=1, if private owned company
foe	Dummy=1, if foreign owned company
ind_exp	Industry export intensity
taxprofit	Tax-plus-profit margin
ind_soe	Share of state owned assets
size	Industry scale
ind_fdi	Share of foreign owned assets
prov_east	Dummy=1, if firm is located in East provinces of the PRC
westmid	Dummy=1, if firm is not located in East provinces of the PRC