INSTITUTE OF DEVELOPING ECONOMIES

IDE Discussion Papers are preliminary materials circulated
to stimulate discussions and critical comments

IDE DISCUSSION PAPER No. 403

Agglomeration and firm-level productivity: A Bayesian Spatial approach

Yoshihiro HASHIGUCHI* and Kiyoyasu TANAKA

Abstract

This paper estimates the impact of industrial agglomeration on firm-level productivity in Chinese manufacturing sectors. To account for spatial autocorrelation across regions, we formulate a hierarchical spatial model at the firm level and develop a Bayesian estimation algorithm. A Bayesian instrumental-variables approach is used to address endogeneity bias of agglomeration. Robust to these potential biases, we find that agglomeration of the same industry (i.e. localization) has a productivity-boosting effect, but agglomeration of urban population (i.e. urbanization) has no such effects. Additionally, the localization effects increase with educational levels of employees and the share of intermediate inputs in gross output. These results may suggest that agglomeration externalities occur through knowledge spillovers and input sharing among firms producing similar manufactures.

Keywords: agglomeration economies, spatial autocorrelation, Bayes, Chinese firm-level data, GIS

JEL classification: C21, C51, R10, R15

---

* Corresponding author. Development Studies Center, IDE-JETRO

The Institute of Developing Economies (IDE) is a semigovernmental, nonpartisan, nonprofit research institute, founded in 1958. The Institute merged with the Japan External Trade Organization (JETRO) on July 1, 1998. The Institute conducts basic and comprehensive studies on economic and related affairs in all developing countries and regions, including Asia, the Middle East, Africa, Latin America, Oceania, and Eastern Europe.

INSTITUTE OF DEVELOPING ECONOMIES (IDE), JETRO
3-2-2, WAKABA, MIHAMA-KU, CHIBA-SHI
CHIBA 261-8545, JAPAN

Agglomeration and Firm-level Productivity: A Bayesian Spatial Approach*

Yoshihiro HASHIGUCHI†
Institute of Developing Economies, JETRO

Kiyoyasu TANAKA
Institute of Developing Economies, JETRO

March 2013

Abstract

This paper estimates the impact of industrial agglomeration on firm-level productivity in Chinese manufacturing sectors. To account for spatial autocorrelation across regions, we formulate a hierarchical spatial model at the firm level and develop a Bayesian estimation algorithm. A Bayesian instrumental-variables approach is used to address endogeneity bias of agglomeration. Robust to these potential biases, we find that agglomeration of the same industry (i.e. localization) has a productivity-boosting effect, but agglomeration of urban population (i.e. urbanization) has no such effects. Additionally, the localization effects increase with educational levels of employees and the share of intermediate inputs in gross output. These results may suggest that agglomeration externalities occur through knowledge spillovers and input sharing among firms producing similar manufactures.

Keywords: agglomeration economies, spatial autocorrelation, Bayes, Chinese firm-level data, GIS

JEL classification: C21, C51, R10, R15

1 Introduction

A spatial concentration of economic activity in a number of industries has attracted a considerable attention and posed the question of why firms agglomerate over space since the seminal work of Marshall (1920). It has long been argued that firms and workers locate in the agglomerated area to benefit from productivity advantages generated by agglomeration economies through more efficient sharing of local suppliers, better matching between employers and workers, and technology or knowledge spillovers among firms and workers (Duranton and Puga 2004). However, agglomeration also creates heavy congestion and raise factor prices such as

*We would like to thank Toshiyuki Matsuura for his data assistance and constructive comments. All remaining errors are our own. This work is partially supported by the Grants-in-Aid for Scientific Research (C): No. 25350319.

†Corresponding author.
Address: 3-2-2 Wakaba, Mihama-ku, Chiba-shi, Chiba 261-8545, Japan.
E-mail: Yoshihiro_Hashiguchi@ide.go.jp

wages and land prices. Because congestion reduces the productivity advantage from agglomeration economies, these opposing effects render agglomeration economies an empirical question. Moreover, it is emphasized that “for policy makers the challenge is to best relax the constraints generated by the congestion and overcrowding of land and resources so that the benefits of agglomeration can be maximized (World Bank, 2008, p. 144).” To evaluate regional policy for maximizing the benefit of agglomeration effects, it is crucial to quantitatively assess the net agglomeration effects.[1]

In this paper, we seek to estimate the impact of agglomeration on firm-level productivity in manufacturing using the Chinese industrial statistics in 2004. As the China’s coastal area has achieved remarkable economic growth since the economic reform in 1978, the Yangzi River Delta has become one of the largest industrial agglomeration areas in the world. As a prominent example of industrial agglomeration, we focus on the region and make the methodological contributions to the literature. Specifically, we introduce spatial autocorrelation into the empirical model based on firm-level data and employ the hierarchical Bayesian method with an endogenous regressor.

There are large number of firm-level studies on agglomeration economies such as Mitra (1999), Henderson (2003), Graham (2008), and Greenstone et al. (2010). Because more productive firms may self-select to locate themselves in agglomerated areas (Baldwin and Okubo, 2006), these studies carefully address endogeneity bias in estimation. However, these studies have not paid attention to spatial autocorrelation resulting, for example, from correlation between local natural advantages in nearby regions (Anselin, 1988 and 2001; LeSage and Pace, 2009). Thus, agglomeration effects may be estimated with bias when spatial autocorrelation is not considered. By contrast, the previous studies such Ke (2010), Artis et al. (2012) and Hashiguchi and Chen (2012) use the aggregate data at the region level and take into account spatial autocorrelation when estimating agglomeration effects by maximum likelihood estimation (MLE), generalized method of moments (GMM), and Bayesian estimation. Although these estimation methods resolve spatial autocorrelation, they have been applied only for the analysis on regional-level productivity.

To identify agglomeration effects on individual firms, one must account for endogeneity bias and spatial autocorrelation in a firm-level empirical model. However, the prior work has not fully addressed these multiple issues together for the methodological limitation. For instance, it is computationally difficult to take into account spatial autocorrelation among individual firms. Alternatively, spatial autocorrelation may be controlled for at the region-level in the firm-level specification with an instrumental variable (IV). However, such a hierarchical model contains a complex structure with large number of unknown parameters, causing the difficulty in implementing MLE and GMM. Therefore, we develop a Bayesian method to estimate a hierarchical spatial model at the firm level. The Bayesian approach allows us to assume various prior distributions of unknown parameters in a flexible manner and estimate the spatial econometric model with a hierarchical structure. Based on the posterior distribution of unknown parameters, we conduct Bayesian statistical inference on the estimate of agglomeration economies. Additionally, we develop a Bayesian instrumental-variable approach to address an endogeneity of agglomeration (Rossi et al., 2005). To the best of our knowledge, our work is the first to apply Bayesian estimation to estimate agglomeration effects on firm-level productivity in a

[1]The magnitude of agglomeration economies has been estimated in various countries. For the details, see for example Eberts and McMillen (1999, Section 3), Rosenthal and Strange (2004, Section 2), Graham (2008, pp. 65–67), Cohen and Paul (2009), Broersma and Oosterhaven (2009, pp. 487–489), and Puga (2010).

hierarchical spatial model.

Our empirical results can be summarized as follows. First, same industry agglomeration (i.e., localization) has a productivity-boosting effect in this region, but urban population agglomeration (i.e., urbanization) has no such effect. The estimates of spatial autocorrelation are positive and significant, indicating that our hierarchical spatial modeling is effective and meaningful. Second, we investigate the difference between IV- and non-IV based posterior distributions, and find that IV-based posterior distributions tend to be located more on the left-hand side than the non-IV distribution, implying that the endogenous bias appears to be upward in direction, consistent with the implication in Baldwin and Okubo (2006). Moreover, the IV estimates of localization economies are significantly positive, and the above conclusion is robust to the IV-based methods.

Finally, we shed light on the nature of agglomeration externalities by analyzing whether localization effects depend on firm characteristics, such as educational levels of employees, the degree of input-sharing, and the experience of export business. Such analysis enhances our understanding of the relationship between agglomeration effects and firm heterogeneity.[2] We find that the localization effects have a positive relationship with educational levels of employees and the degree of input sharing, indicating that the localization effects generate larger productivity gains for the firms that employ more well-educated workers and depend more on intermediate inputs. These results may suggest that absorptive capacity for technology spillovers and reliance on local procurement are crucial to benefit from industrial agglomeration. Thus, we highlight the importance of input sharing and knowledge spillovers in accounting for the nature of agglomeration economies in the Yangzi River Delta.

Before proceeding to the details of our analysis, we must emphasize that our analysis contributes to the literature on agglomeration effects in China. Several empirical studies have investigated agglomeration economies in China and found the inconsistent results. For example, Fan (2007) analyzed nonagricultural industrial sectors in 261 prefecture-level regions for 2004 whereas Li et al. (2011) examined firm-level panel data in textile industry for 2000-2005. They found that the net agglomeration effects are significantly positive. By contrast, Ke (2010) investigated nonagricultural industrial sectors in 617 cities for 2005 while Hashiguchi and Chen (2012) examined the county-level data in Shanghai, Jiangsu, and Zhejiang for 2009. They found that the agglomeration effects are almost zero or even negative. Although it is challenging to identify the source of varying impacts of agglomeration economies in these previous works, our paper highlights the role of firm heterogeneity in determining agglomeration externalities.

The rest of this paper is structured as follows. Section 2 describes our empirical model and estimation method. Section 3 presents data sources and shows the geographic distribution of manufacturing firms and urban population in the Yangzi River Delta, using the geographical information system (GIS). Section 4 reports the empirical results, and Section 5 concludes.

[2] Rosenthal and Strange (2001) and Ellison et al. (2010) empirically analyzed the causes of industrial agglomeration using the Ellison and Glaeser (1997, hereafter EG) agglomeration index. However, their analysis focused on the relationship between the EG index and proxy variables of input sharing, labor market pooling, and knowledge spillovers at the industry level.

2 Model and estimation methods

2.1 Model

Let us consider that firm f produces output Y_f using the following production technology:

$$Y_f = T_f K_f^{\beta_K} L_f^{\beta_L}, \quad f = 1, 2, \ldots, N_F \,, \tag{1}$$

where K_f and L_f denote capital and labor, respectively. β_K and β_L are parameters. N_F and T_f denote the total number of firms and total factor productivity (TFP). We assume that T_f depends on the following factors:

$$\log T_f = \beta_A \log A_f + \beta_h h_f + \beta_{EX} D_f^{EX} + \sum\nolimits_{c=1}^{N_C} \alpha_c D_{cf}^C + \varepsilon_f \,, \tag{2}$$

where A_f denotes the degree of localized industrial agglomeration faced by a firm f. Its elasticity β_A indicates the magnitude of agglomeration effects (localization economies). h_f denotes the average educational level of firm's employees. D_f^{EX} represents the export dummy variable such that $D_f^{EX} = 1$ if firm f engages in exporting. The export dummy is introduced because several previous studies have argued that the more productive firms are likely to conduct export business (Bernard et al., 2007).

D_{cf}^C is a dummy variable such that $D_{cf}^C = 1$ if firm f is located in county c. Its parameter α_c indicates the degree of unobserved local advantage in county c. We add this term to capture the differences among counties' productivity, which arise in part from differences in regional comparative advantages. Further, α_c can be spatially autocorrelated if the productivity level in county c is more similar to that in its neighboring regions than to that in regions far removed from county c. For example, local advantages in one county, arising from local climate, infrastructure, natural resources and so on, are likely to have an impact on the productivity not only within that county, but also in the other nearby counties. This implies that local advantages increase productivity beyond a county's border, causing spatial autocorrelation. In addition, if there exist positive external effects from urbanization, called urbanization economies, the degree of urbanization in county c positively affects the level of α_c.

To represent these effects of spatial autocorrelation and urbanization economies, we specify the following spatially autoregressive model:

$$\alpha_c = \mu_0 + \mu_U \log U_c + \rho \sum\nolimits_{j=1}^{N_C} w_{cj} \alpha_j + u_c, \quad c = 1, 2, \ldots, N_C \,, \tag{3}$$

where N_C and $\log U_c$ denote the number of counties and degree of urbanization in county c, respectively. u_c is an error term. μ_0, μ_U, and ρ are parameters. μ_U and ρ represent the magnitudes of urbanization economies and spatial autocorrelation, respectively. w_{cj} is specified as $w_{cj} = d_{cj} / \sum_j^{N_C} d_{cj}$; d_{cj} is a binary variable such that $d_{cj} = 1$ if counties c and j are neighbors, and is zero otherwise. We use the concept of queen type binary contiguity (Anselin 1988, p. 18) as the definition of neighbors, such that counties c and j are regarded as neighbors ($d_{cj} = 1$) if they have a common land or river border.[3]

[3]Chongming is an island county in Shanghai. As it dose not have a land border, we assume that Chongming neighbors on Pudongxin (Shanghai) and Qidong (Jiangsu). Zhejiang also contains four island counties: Zhoushan district, Daishan, Shengsi and Dongtou. We regard the closest continental or island county as these counties' neighbors:

Finally, we introduce fixed effects of industry and ownership as follows:

$$\varepsilon_f = \sum\nolimits_{i=2}^{N_I} \phi_{Ii} D_{if}^I + \phi_{FO} D_f^{FO} + \eta_f \,, \tag{4}$$

where ϕ_{Ii} and ϕ_{FO} are parameters. η_f is an error term. D_{if}^I and $D_f^{FO}$are dummy variables such that

$D_{if}^I = 1$ if f belongs to industry i, and zero otherwise;
$D_f^{FO} = 1$ if f receives foreign capital or capital from Hong Kong, Macao, or Taiwan (HMT), and zero otherwise.

N_I denotes the number of industries. η_f is an unobservable random effect on productivity.

Taking logarithms on both sides of Equation (1), and substituting Equations (2) and (4) for (1), we obtain our empirical model as follows:

$$\begin{aligned} y_f &= \beta_A x_f + \mathbf{z}_f \boldsymbol{\beta} + \mathbf{D}_f^C \boldsymbol{\alpha} + \eta_f \\ \boldsymbol{\alpha} &= \mathbf{H}\boldsymbol{\mu} + \rho \mathbf{W} \boldsymbol{\alpha} + \mathbf{u} \end{aligned} \tag{5}$$

where $y_f = \log Y_f$, $x_f = \log A_f$, and

$$\begin{aligned}
\mathbf{z}_f &= \begin{bmatrix} h_f & D_f^{EX} & \log K_f & \log L_f & \mathbf{D}_f \end{bmatrix}, \\
\mathbf{D}_f^C &= \begin{bmatrix} D_{1f}^C & D_{2f}^C & \dots & D_{N_C f}^C \end{bmatrix}, \\
\mathbf{D}_f &= \begin{bmatrix} D_{2f}^I & D_{3f}^I & \dots & D_{N_I f}^I & D_f^{FO} \end{bmatrix}, \\
\boldsymbol{\beta} &= \begin{bmatrix} \beta_h & \beta_{EX} & \beta_K & \beta_L & \boldsymbol{\phi}' \end{bmatrix}', \\
\boldsymbol{\alpha} &= \begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_{N_C} \end{bmatrix}', \\
\mathbf{H} &= \begin{bmatrix} \mathbf{1} & \log \mathbf{U} \end{bmatrix}, \\
\boldsymbol{\mu} &= \begin{bmatrix} \mu_0 & \mu_U \end{bmatrix}', \\
\mathbf{u} &= \begin{bmatrix} u_1 & u_2 & \dots & u_{N_C} \end{bmatrix}', \\
\mathbf{W} &= \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1N_C} \\ w_{21} & w_{22} & \dots & w_{2N_C} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N_C 1} & w_{N_C 2} & \dots & w_{N_C N_C} \end{bmatrix}.
\end{aligned}$$

Equation (5) is a spatial econometric model with a hierarchical structure of the parameter $\boldsymbol{\alpha}$.

- Shengsi neighbors on Daishan,
- Daishan neighbors on Zhoushan district,
- Zhoushan district neighbors on Ningbo district (a continental county), and
- Dongtou neighbors on Yuhuan (a continental county).

2.2 Bayesian estimation

Spatial econometric models based on regional aggregate data can be generally estimated by MLE, GMM, or Bayesian methods (Anselin, 1988, 2001; LeSage and Pace, 2009). Because we estimate the complicated spatial model for firm-level data, it is difficult to apply MLE or GMM. Thus, we develop a Bayesian method to estimate the hierarchical spatial model.

To identify the impact of agglomeration economies, we need to carefully take into account endogeneity issues. Previous studies on agglomeration economies have addressed potential correlation between the proxy of industrial agglomeration ($x_f = \log A_f$) and the error term (η_f) because regression analysis may suffer from endogeneity resulting from omitted variables and reverse causality between productivity and agglomeration.[4] For example, unobserved local endowments (e.g., local climate, social infrastructure, and natural resources) may increase firm's productivity as well as the degree of agglomeration, leading to omitted-variables bias. Also, high productive firms may self-select to locate their production base in agglomerated area, causing reverse causality. To minimize the endogeneity bias, we introduce a number of fixed effects to control for unobserved local endowments in Equation (3), but a concern regarding reverse causality remains in our model. Because IV methods offer a solution for such endogeneity, we adopt the Bayesian IV approach proposed by Rossi et al. (2005, pp. 185–206).

Consider that x_f is a potentially endogenous variable and has a linear relationship with a set of instruments ($q_f,\ \mathbf{z}_f,\ \mathbf{D}_f^C$) and an idiosyncratic shock ξ_f, where q_f denotes a variable related to x_f but independent of the error terms η_f and ξ_f. We discuss the data used for q_f in Section 3.1. Following Rossi et al. (2005), we specify the system of equations as follows.

$$x_f = q_f\gamma_0 + \mathbf{z}_f\boldsymbol{\gamma}_1 + \mathbf{D}_f^C\boldsymbol{\gamma}_2 + \xi_f \tag{6}$$

$$\begin{aligned} y_f &= \beta_A x_f + \mathbf{z}_f\boldsymbol{\beta} + \mathbf{D}_f^C\boldsymbol{\alpha} + \eta_f \\ \boldsymbol{\alpha} &= \mathbf{H}\boldsymbol{\mu} + \rho\mathbf{W}\boldsymbol{\alpha} + \mathbf{u} \end{aligned} \tag{7}$$

This system shows the case of structural Equations (7) with one endogenous variable and multiple instruments. In this paper, Equations (6) and (7) are referred to as the IV model and Equation (5) as the Non-IV model.

We assume that ξ_f, η_f and $\mathbf{u}$ have the following multivariate normal distributions:

$$\begin{pmatrix}\xi_f\\ \eta_f\end{pmatrix} \sim MVN(\mathbf{0},\ \boldsymbol{\Sigma}), \quad \mathbf{u} \sim MVN(\mathbf{0},\ \tau^2\mathbf{I}), \tag{8}$$

where $\boldsymbol{\Sigma}$ is a (2×2) covariance matrix and $\mathbf{I}$ is an identity matrix. Independent priors for the unknown parameters are specified as

$$\begin{aligned} &\begin{pmatrix}\gamma_0\\ \boldsymbol{\gamma}_1\end{pmatrix} \sim MVN(\bar{\mathbf{b}}_\gamma, \bar{\mathbf{B}}_\gamma), \quad \boldsymbol{\gamma}_2 \mid \mu_{\gamma_2}, \sigma^2_{\gamma_2} \sim N(\mu_{\gamma_2}\mathbf{1}, \sigma^2_{\gamma_2}\mathbf{I}), \\ &\begin{pmatrix}\beta_A\\ \boldsymbol{\beta}\end{pmatrix} \sim MVN(\bar{\mathbf{b}}_\beta, \bar{\mathbf{B}}_\beta), \quad \boldsymbol{\Sigma} \sim IW(\bar{b}_\Sigma,\ \bar{\mathbf{B}}_\Sigma), \\ &\boldsymbol{\mu} \sim MVN(\bar{\mathbf{b}}_\mu, \bar{\mathbf{B}}_\mu), \quad \rho \sim U(\lambda^{-1}_{\min}, \lambda^{-1}_{\max}), \quad \tau^2 \sim IG(\bar{\nu}_\tau/2,\ \bar{\omega}_\tau/2), \\ &\mu_{\gamma_2} \sim N(\bar{b}_{\mu\gamma},\ \bar{\mathbf{B}}_{\mu\gamma}), \quad \sigma^2_{\gamma_2} \sim IG(\bar{\nu}_\gamma/2,\ \bar{\omega}_\gamma/2), \end{aligned} \tag{9}$$

[4] For detailed discussion about the endogeneity problems in this literature, refer to Eberts and McMillen (1999), Rosenthal and Strange (2004), Cohen and Paul (2009), and Puga (2010).

where $IW()$, $IG()$ and $U()$ denote the inverted wishart distribution, the inverted gamma distribution, and the uniform distribution, respectively. The prior parameters are $\bar{\mathbf{b}}_\gamma$, $\bar{\mathbf{B}}_\gamma$, $\bar{\mathbf{b}}_\beta$, $\bar{\mathbf{B}}_\beta$, $\bar{\mathbf{b}}_\mu$, $\bar{\mathbf{B}}_\mu$, $\bar{b}_{\mu\gamma}$, $\bar{\mathbf{B}}_{\mu\gamma}$, $\bar{b}_\Sigma$, $\bar{B}_\Sigma$, $\bar{\nu}_\gamma$, $\bar{\omega}_\gamma$, $\bar{\nu}_\tau$, $\bar{\omega}_\tau$, $\lambda_{\min}$ and $\lambda_{\max}$. The $\lambda_{\min}$ and $\lambda_{\max}$ denote the smallest and the largest eigenvalue of $\mathbf{W}$, respectively. We limit the parameter space of ρ such as $\lambda_{\min}^{-1} < \rho < \lambda_{\max}^{-1}$.[5] The values of the other prior parameters are assumed as follows:

$$\bar{\mathbf{b}}_\gamma = \mathbf{0},\ \bar{\mathbf{b}}_\beta = \mathbf{0},\ \bar{\mathbf{b}}_\mu = \mathbf{0},\ \bar{b}_{\mu\gamma} = 0,\ \bar{b}_\Sigma = 2,$$
$$\bar{\mathbf{B}}_\gamma = 100\mathbf{I},\ \bar{\mathbf{B}}_\beta = 100\mathbf{I},\ \bar{\mathbf{B}}_\mu = 100\mathbf{I},\ \bar{B}_{\mu\gamma} = 100,\ \bar{\mathbf{B}}_\Sigma = 2\mathbf{I},$$
$$\bar{\nu}_\gamma = \bar{\omega}_\gamma = \bar{\nu}_\tau = \bar{\omega}_\tau = 0.001.$$

We set these priors to have a very large variance in order to ensure that our prior beliefs for unknown parameters are non-informative.

Based on the technique proposed by Rossi et al. (2005), we develop a Gibbs-within-Metropolis sampler, which is a Markov Chain Monte Carlo (MCMC) method (Gamerman and Lopes, 2006, p. 213) for the model in Equations (6)–(9). The sampler allows us to generate samples from the joint posterior distribution and the marginal posterior distributions for each parameter. Using the generated samples, we can statistically infer the unknown parameters.

The Gibbs-within-Metropolis sampling for our model is based on the following *full* conditional distributions:

$$\begin{gathered}
\beta_A, \boldsymbol{\beta} \mid \Theta_{-(\beta_A,\, \boldsymbol{\beta})}, \mathbf{Data} \\
\boldsymbol{\alpha} \mid \Theta_{-\boldsymbol{\alpha}}, \mathbf{Data} \\
\gamma_0, \boldsymbol{\gamma}_1 \mid \Theta_{-(\gamma_0,\, \boldsymbol{\gamma}_1)}, \mathbf{Data} \\
\boldsymbol{\gamma}_2 \mid \Theta_{-\boldsymbol{\gamma}_2}, \mathbf{Data} \\
\boldsymbol{\Sigma} \mid \Theta_{-\boldsymbol{\Sigma}}, \mathbf{Data} \\
\boldsymbol{\mu} \mid \Theta_{-\boldsymbol{\mu}}, \mathbf{Data} \\
\rho \mid \Theta_{-\rho}, \mathbf{Data} \\
\tau^2 \mid \Theta_{-\tau^2}, \mathbf{Data} \\
\mu_{\gamma_2} \mid \Theta_{-\mu_{\gamma_2}}, \mathbf{Data} \\
\sigma^2_{\gamma_2} \mid \Theta_{-\sigma^2_{\gamma_2}}, \mathbf{Data}
\end{gathered} \tag{10}$$

where $\Theta_{-\beta}$ denotes the set of the unknown parameters except for β. All the above distributions except for ρ can be straightforwardly derived and belong to well-known distribution families, enabling us to apply the Gibbs sampling technique. However, ρ's full conditional density function is not standard; it can be written as

$$\begin{aligned}
&P(\rho \mid \Theta_{-\rho}, \mathbf{Data}) \\
&\quad \propto |\mathbf{I} - \rho\mathbf{W}| \exp\left\{-\frac{1}{2\hat{\sigma}_\rho^2}(\rho - \hat{\rho})^2\right\} I\left[\rho \in (\lambda_{\min}^{-1}, \lambda_{\max}^{-1})\right],
\end{aligned} \tag{11}$$

where

$$\hat{\sigma}_\rho^2 = \left[\tau^{-2}\boldsymbol{\alpha}'\mathbf{W}'\mathbf{W}\boldsymbol{\alpha}\right]^{-1}, \tag{12}$$
$$\hat{\rho} = \hat{\sigma}_\rho^2\tau^{-2}\boldsymbol{\alpha}'\mathbf{W}'(\boldsymbol{\alpha} - \mathbf{H}\boldsymbol{\mu}). \tag{13}$$

[5]This restriction is necessary to ensure that the determinant of $\mathbf{I} - \rho\mathbf{W}$ is positive.

$I[\rho \in (\lambda_{\min}^{-1}, \lambda_{\max}^{-1})]$ is an indicator function that takes on unity if $\rho \in (\lambda_{\min}^{-1}, \lambda_{\max}^{-1})$. To generate a sample from this density, we use the Metropolis–Hastings (MH) technique.[6] The candidate generating function used in the MH algorithm is a normal distribution truncated on the interval $(\lambda_{\min}^{-1}, \lambda_{\max}^{-1})$, with the mean $\hat{\rho}$ and the variance $\hat{\sigma}_{\rho}^2$. Sampling from these conditionals in Equation (10) is repeated 400,000 times; the initial 50,000 replications were discarded and a statistical inference was made from the remaining 350,000 replications.

3 Data and choropleth maps

3.1 Data description

Our data source is the Chinese industry statistical database by China's National Bureau of Statistics. This database is based on the annual survey of industrial enterprises in mainland China with sales of more than five million yuan, including state-owned enterprises, privately-owned firms, and foreign-invested enterprises. As the survey is mandatory for firms to respond, the sample coverage is comprehensive.[7] This dataset has been used in prior research such as Brandt et al. (2011).

The 2004 survey is more comprehensive than other years' in that the survey information includes the level of employee education. To measure employees' educational level h_f, we exclusively use 2004 data. Our analysis focuses on manufacturing industries in four province level regions in east-central China: Shanghai city and Jiangsu, Zhejiang, and Anhui provinces. These regions constitute China's three major areas of industrial agglomeration; other areas include the Bohai Economic Rim around Beijing and Tiangjin and the Pearl River Delta area in Guangdong. As the remarkable growth of industrial activities and international trade in these areas has driven the Chinese economy, our sample is ideal to analyze the impact of industrial clusters on firm-level performance.

Using the 2004 database, we construct the variables Y_f, K_f, L_f, h_f, U_f, A_f and q_f as follows. Y_f is constructed as firm's value added. We exclude sample firms with negative values of value added and out-of-operation status. K_f is the sum of fixed and intangible assets. L_f is the number of employees. To measure h_f, we use the average years of schooling as follows:

$$h_f = 18\left(\frac{L_f^G}{L_f}\right) + 16\left(\frac{L_f^U}{L_f}\right) + 14\left(\frac{L_f^C}{L_f}\right) + 12\left(\frac{L_f^H}{L_f}\right) + 9\left(\frac{L_f^J}{L_f}\right), \tag{14}$$

where L_f^G, L_f^U, L_f^C, L_f^H, and L_f^J denote the number of graduate school-educated, university-educated, community college-educated, high school-educated, and junior high school- or under-educated employees, respectively. The county-level urbanization U_c is measured by county urban population divided by land area (urban population density). The localization index A_f is alternatively proxied by the variables FD_f, LD_f or FN_f. Specifically, FD_f and LD_f denote the density of the number of firms and labor in the same county and industry as firm f. FN_f denotes the number of firms in the same county and industry as firm f. We construct the industry-dummy variables D_{if}^I using China's two-digit industrial classification, which contains

[6]For more details on the Metropolis–Hastings and Gibbs sampling techniques, refer to Gamerman and Lopes (2006, Chapters 5 and 6).

[7]For example, the survey includes 330,000 enterprises in China for 2007, which accounted for nearly 90% of total industrial output as reported in the China statistical yearbook.

$N_I = 29$ industries excluding the tobacco industry. The county-dummy variables D^C_{cf} are based on $N_C = 223$ counties, and Tables 4–6 contain the list of county names. The data on total land area at the county-level region are obtained from China's 2005 county statistical yearbook and China's individual statistical yearbooks of Shanghai, Jiangsu, Zhejiang, and Anhui.

Finally, we explain q_f which is an instrument of x_f. The data on q_f is the logarithm of the 1990 density of labor in the same county and industry as firm f. This data should be valid as an instrument because it is natural to consider that the labor density 14 years ago does not affect 2004 firm-level productivity, but the formulation process of industrial agglomeration from 1990 to 2004. The 1990 labor density may be correlated with 1990 local advantages which probably affect the 2004 firm-level productivity, and thus q_f may indirectly affect the 2004 firm-level productivity. However, we control for this indirect effect by the county-level fixed effects in Equation (3), making our data on q_f a valid instrument of x_f.

[Table 1 around here]

Table 1 reports summary statistics of the sample used for analysis, which includes 97,947 manufacturing firms after excluding 23 firms in the tobacco industry. There are nearly 40,000 firms in Jiangsu and Zhejiang, 14,554 firms in Shanghai, and 4,125 firms in Anhui. The average production measured by $\log Y$ is relatively lower in Jiangsu and Zhejiang, indicating that smaller firms tend to locate in these areas. This observation appears to be supported by the average levels of measured capital. The variable IS is the value of a firm's intermediate inputs divided by its gross output, which we use in Section 4.2 in the analysis of heterogeneity in localization effects. The average of $\log IS$ reveals that Zhejiang has the highest value, followed by Jiangsu, Anhui, and Shanghai. This outcome indicates that firms in Jiangsu and Zhejiang are relatively small and have greater input sharing. Regional differences appear to be relatively small in the number of employees ($\log L$) and their educational attainment (h). Our measure of localization indicates a large difference between coastal and inland provinces. Shanghai has the highest degree of localization for all three variables whereas Anhui has the lowest. There are relatively minor differences between Jiangsu and Zhejiang. Finally, Shanghai has the highest degree of urbanization, followed by Jiangsu, Anhui, and Zhejiang.

3.2 Geographic distribution of firms and urban population

[Figure 1 around here]

Figure 1 illustrates the number of firms per county land area (firm density) and urban population density.[8] There appears to be substantial agglomeration of manufacturing firms along the coastal provinces such as Jiangsu, Shanghai, and Zhejiang. Huangpu[9] (id = 1) exhibits the highest firm density and a highly agglomerated area spans from Shanghai to southern Jiangsu province (around the Wuxi district (15)) and from Shanghai to northeast Zhejiang province, including Hangzhou (77), Shaoxing (108), and Ningbo (83) districts. The Taizhou (130, 131, 135) and Wenzhou (89, 96, 97) areas and Yiwu (119), Yonkang (121), Hefei district (146), and Wuhu

[8] China's county-level GIS map (*shapefile*) is obtained from All China Marketing Research Co., Ltd. (2005), and the polygon of the Yangzi River is downloaded from the DIVA-GIS web-site: `http://www.diva-gis.org/` (accessed July 18, 2012). These maps are drawn with R version 2.9.2 (R Development Core Team, 2009).

[9] Huangpu includes the following nine county-level regions: Huangpu, Luwan, Xuhui, Changning, Jingan, Putuo, Zhabei, Hongkou, and Yangpu. Because these regions' land areas are quite small, we merge them into one composite called Huangpu et al.

district (150) exhibit relatively high firm agglomeration. In contrast, firms in Anhui province face a relatively smaller degree of manufacturing agglomeration, but Anhui's capital (Hefei district, 146) and counties along the Yangzi river (150, 152, 160, 161, 164, 166) exhibit a relatively high degree of firm density.

As with firm density, the figure of urban population density indicates that Shanghai, southern Jiangsu, and northeast and southeast Zhejiang exhibit high urban population density.[10] However, counties in northwest Anhui have relatively greater urban population density, which appears to be more dispersed than that of manufacturing firms.

[Figure 2 around here]

To assess whether the degree of spatial concentration differs between manufacturing firms and urban population, we draw the Moran scatterplots and calculate Moran's I statistics (Figure 2). The dashed lines represent the averages of each axis, and the solid line is the regression line. The horizontal axis (x_c) is the deviation from the mean of firm density or urban population density, and the vertical axis (y_c) is the neighboring averages of the deviation values, calculated by $y_c = \sum_j^{N_C} w_{cj} x_j$. When using a low-standardized spatial weight $\sum_j^{N_C} w_{cj} = 1$, the coefficient of the regression line is equivalent to Moran's I statistic which indicates the degree of geographic concentration.[11] The Moran scatterplots quadrants separated by the dashed lines correspond to four relationships between a region and its neighbors. More specifically,

1. HH: a region plotted in the first quadrant (HH area) has a high value (above the mean) and is surrounded by regions with high values;

2. LH: a region in the second quadrant (LH area) has a low value but is surrounded by regions with high value;

3. LL: a region in the third quadrant (LL area) has a low value and is surrounded by regions with low value; and

4. HL: a region in the fourth quadrant (HL area) has a high value but is surrounded by regions with low values.

As Figure 2 shows, Moran's I coefficient is 0.7857 for firm density and 0.6403 for urban population density, indicating that the degree of spatial clustering in the firms' density is greater than that of urban population density. Intuitively speaking, manufacturing firms concentrate over space more than urban inhabitants.

Because Moran's I does not provide the information about the location of clusters, we calculated the local Moran statistics (Anselin, 1995). The local Moran test is a location-specific statistical test, and enables us to investigate whether the local relationship (HH, LH, LL, or HL) for each location c differs significantly from the random location relationship.[12]

[Figure 3 around here]

[10] The correlation coefficient between firm density and urban population density is 0.683.

[11] The Moran's I coefficients are calculated using the spatial weighting matrix **W** defined in Equation (5).

[12] Because the exact distribution of the local Moran statistic under the randomization (or random location) hypothesis is unknown, we used the random permutation approach proposed by Anselin (1995). The randomization test was performed using 9,999 permutations and we obtained pseudo significance levels from the 2.5% and 97.5% quantile points of this simulated distribution.

Figure 3 depicts the results of the local Moran test. The 5% significance means that the random location hypothesis is rejected at the 5% significance level. The regions around Shanghai have a significant HH relationship for both maps, confirming a local, significant high-value cluster around Shanghai. Compared to urban population density, the firm density covers a broader range. However, the location of the local significant low-value cluster (LL) differs between the maps. The firm density map depicts the main low-value cluster west of Anhui, whereas the urban population density map shows the low-value cluster ranging from southern Anhui to western Zhejiang. In sum, the geographic distributions of manufacturing activities and urban population differs in the degree of spatial concentration and the range and location of clusters, implying that these two may have a different impact on firm-level productivity.

4 Results

4.1 Estimation results

[Table 2 around here]

Table 2 reports the estimation results of the model in Equations (6)–(9).[13] We perform the estimation separately for three localization variables: the firm density ($\log FD_f$), the density of labor ($\log LD_f$), and the number of firms ($\log FN_f$). The posterior means of β_K and β_L are roughly 0.21 and 0.60, respectively. The coefficients of the export dummy and average years of schooling (β_{EX} and β_h) are also positive and significant, consistent with our prediction. Furthermore, the posterior means of ρ are positive (roughly 0.53) and significant, demonstrating that a positive spatial autocorrelation exists in the unobserved local advantage ($\boldsymbol{\alpha}$) and that our hierarchical spatial modeling is effective and meaningful.

The posterior means of β_A are 0.022, 0.009, and 0.033 across alternative specifications. The 95% credible intervals for $\log FD$ and $\log FN$ do not contain zero. On the other hand, $\log LD$ intervals contain zero, but the posterior probability of $\beta_A > 0$ is 93.81%. These results indicate that agglomeration of the same industry has a significant positive effect on firm-level productivity. The magnitude of these estimates are similar to that of Au and Henderson (2006), Henderson (2003) and other previous studies. Their estimates of agglomeration economies were 0.013 in Au and Henderson (2003) and 0.024 in Henderson (2003) on average. According to Melo et al. (2009)'s meta-analysis, an average of estimates in manufacturing taken from early studies is 0.0175 (Melo et al., 2009, Table 1). Our estimates (0.009–0.033) imply that a doubling of firm's agglomeration increases firm's productivity by around 0.9–3.3%.

However, the posterior means of μ_U are -0.0007, 0.0274, and 0.015. Their 95% credible intervals contain zero for all three cases. In addition, the posterior probabilities of $\mu_U > 0$ are 48.14%, 92.79%, and 80.83%, respectively. These results do not provide conclusive evidence of the positive relationship between agglomeration of urban population and a county-level productivity. Agglomeration economies for manufacturing firms are relatively localized within the same industry, but urbanization does not necessarily increase manufacturing firm productivity. These results are consistent with the study by Henderson (2003). The reason is that urbanization can greatly increase production costs and reduce the competitiveness in manufacturing sectors.

[13] The estimation is implemented with *Ox* version 6.20 (Doornik, 2006).

We proceed to compare the estimation results of IV- and Non-IV-based models. Because the posteriors of all parameters except β_A are similar among these models, we report only the results of β_A.

[Figure 4 around here]

Figure 4 depicts IV- and Non-IV-based posterior distributions of β_A. The IV-based posterior is located more to the left and exhibits higher dispersion. More specifically, the posterior means of the Non-IV model are 0.036 for $\log FD$, 0.022 for $\log LD$, and 0.036 for $\log FN$. Their standard deviations are 0.0034, 0.0030, and 0.0034, respectively. These results indicate that the endogenous bias in the posterior means is upward, and the bias in the posterior dispersion is downward. Thus, our Non-IV estimates may be over-estimated. The upward bias suggests that more productive firms are attracted to the agglomeration area as is predicted in the economic geography model of firm heterogeneity in Baldwin and Okubo (2006). Nevertheless, the IV-based posteriors remain largely distributed in the positive area, indicating that agglomeration of the same industry has a causal positive effect on firm-level productivity.

4.2 Firm heterogeneity in localization effects

In this subsection, we investigate whether the magnitude of localization economies depends on firm specific characteristics such as average years of employee education h_f, engagement in export business D_f^{EX}, degree of input sharing IS_f, period (months) of operation MO_f, and ownership D_f^{FO}. Based on the Non-IV model, we introduce the following cross product:

$$\log T_f = \beta_h h_f + \beta_{EX} D_f^{EX} + \beta_{IS} \log IS_f + \beta_{MO} \log MO_f + \sum_{c=1}^{N_C} \alpha_c D_{cf}^C + \left(\beta_A + \beta_{A.h} h_f + \beta_{A.EX} D_f^{EX} + \beta_{A.IS} \log IS_f + \beta_{A.MO} \log MO_f + \beta_{A.FO} D_f^{FO}\right) \log A_f + \varepsilon_f \,, \tag{2'}$$

where $\beta_{A.h}$, $\beta_{A.EX}$, $\beta_{A.IS}$, $\beta_{A.FO}$, and $\beta_{A.MO}$ are parameters. D_f^{EX} and D_f^{FO} denote export and foreign-ownership dummies, respectively.[14] Following previous studies such as Holmes (1999), Rosenthal and Strange (2001), and Ellison et al. (2010), we construct the index of input sharing IS_f from the value of a firm's intermediate inputs divided by its gross output. MO_f represents the number of months in operation.

$\beta_{A.h}$ should be positive if a firm's ability to absorb positive localization effects is significantly important. We predict that $\beta_{A.IS}$ should be positive because industrial agglomeration can reduce the costs of obtaining inputs, called input sharing effects (Duranton and Puga 2004). That is, firms highly dependent on intermediate inputs are more likely to gain such positive agglomeration effects. For the number of months in operation MO_f, we hypothesize that older firms are more likely than newer firms to possess business know-how and an efficient (cost-saving) network of inter-firm relationships to survive in the market. If these advantages accumulate through experience and are significantly important to access localization economies, the coefficient $\beta_{A.MO}$ would be positive.

[Table 3 around here]

[14]The analysis of firm heterogeneity in localization effects does not use instrumental variable techniques because valid instruments for the cross products are not available.

To save space, Table 3 reports the estimation results only for the coefficients of the cross products in Equation (2′). The posterior probabilities of $\beta_{A.h} > 0$ and of $\beta_{A.IS} > 0$ exceed 90% for all three cases, indicating that agglomeration effects from localization positively correlate with employee educational levels and the ratio of intermediate inputs to gross output. The posterior probabilities of $\beta_{A.EX} > 0$ are 37.95% for $\log FD$, 97.42% for $\log LD$, and 29.82% for $\log FN$. There is no clear relationship between localization effects and export. Moreover, the posterior of $\beta_{A.FO}$ is inconsistent among the three cases, such that the probability of $\beta_{A.FO} > 0$ is significantly high for $\log LD$ (99.69%), but 91.45% and 76.28% for the other two cases, and their 95% credible intervals contain zero. We did not observe strong evidence for the relationship between localization effects and firm ownership.

Furthermore, it is noteworthy that the posterior probabilities of $\beta_{A.MO} > 0$ are only approximately 0.01%, 28.8%, and 1.36%, respectively, and the 95% credible intervals do not contain zero for $\log FD$ and $\log FN$. These posterior distributions are largely distributed in the negative area, indicating a negative relationship between months of operation and degree of localization effects. This result suggests that new firms are more likely to obtain higher positive localization effects. However, the posterior means of β_{MO}, which is the coefficient of $\log MO_f$ (not shown in Table 3), are 0.0165 for $\log FD$, 0.0368 for $\log LD$, and 0.0524 for $\log FN$. These estimates are significant for all three cases. These results indicate that old firms are more productive than new firms.[15] Although new firms have relatively lower productivity than old firms, they are more likely to obtain the positive externalities from localization.

In sum, the analysis highlights that localization effects differ by individual firms. Industrial agglomeration yields larger productivity-boosting effects for the firms that use more intermediate inputs, employ more educated workers, and are young in business operations. By contrast, there is no evidence that exporting and foreign ownership enable firms to benefit more from industrial clusters.

5 Conclusions

This paper estimates the magnitude of agglomeration effects in the Yangzi River Delta, China, using 2004 manufacturing firm-level data. We employ a hierarchical spatial model to deal with spatial autocorrelation of unobserved local endowments. To resolve potential endogeneity bias in agglomeration, a Bayesian instrumental-variables technique is applied for the hierarchical spatial model. These approaches enable us to investigate the difference between IV- and non-IV based posterior distributions. Our main findings show that the estimates of spatial autocorrelation are significantly positive, indicating that our hierarchical spatial modeling is effective and meaningful. Robust to the Bayesian IV estimation, agglomeration economies for manufacturing firms are largely localized within the own industry whereas urbanization does not necessarily have a productivity-boosting effect on manufacturing firms. Additionally, the localization effects increase with educational levels of employees and the degree of input sharing, highlighting that technology spillovers and input sharing can serve as crucial channels through which localization effects boost firm-level productivity.

Although our paper sheds light on agglomeration economies in China, there are remaining

[15] The estimation results of β_{MO} are inconsistent with Lin et al. (2011). Their estimations are based on firm-level panel data of China's textile industry from 2000 to 2005; they found a negative relationship between years of operation and firm's productivity.

issues for future research. We can make an extension to analyze firm-level panel data, which would allow us to examine a short-run effect of agglomeration across years. Also, long-period panel data would enable us to compare net agglomeration economies across the stages of economic development in China. As Brülhart and Sbergami (2009) find evidence for the inverse U-shaped relationship between agglomeration and growth, the net magnitude of agglomeration effects may recently change in the Yangzi River Delta. Furthermore, the use of panel data would allow us to control for unobserved shocks to production and self-selection bias due to decisions on plant location. It is a challenging task to take into account these dynamic issues in a spatial econometric model, but crucial for identifying dynamic impacts of agglomeration processes on firm's productivity.

References

All China Marketing Research Co. 2005. China historical county population census data with GIS maps (1953, 1964, 1982, 1990, 2000). CD-ROM distributed by China Data Center, University of Michigan.

Anselin, L., 1995. Local indicators of spatial association: LISA. Geographic Analysis 27(2), 93–115.

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kluwer Academic Publishers, Dordrecht.

Anselin, L., 2001. Spatial econometrics. In: Baltagi, B. (Eds.), A Companion to Theoretical Econometrics. Blackwell Publishers, Malden, Mass, pp. 310–330.

Artis, M.J., Miguélez, E., Moreno, R., 2012. Agglomeration economies and regional intangible assets: an empirical investigation. Journal of Economic Geography 12(6), 1167–1189.

Au, C.C., Henderson, J.V., 2006. Are Chiese cities too small? Review of Economic Studies 73, 549–576.

Baldwin, R., Okubo, T., 2006. Heterogeneous firms, agglomeration and economic geography: spatial selection and sorting. Journal of Economic Geography 6(3), 323–346.

Ciccone, A., 2002. Agglomeration effects in Europe. European Economic Review 46, 213–227.

Ciccone, A., Hall, R.E., 1996. Productivity and the density of economic activity. American Economic Review 86, 54–70.

Cohen, J.P., Paul, C.J.M., 2009. Agglomeration, productivity and regional growth: production theory approaches. In: Capello, R., Nijkamp, P., (Eds.), Handbook of Regional Growth and Development Theories. Edward Elgar, Cheltenham, UK, pp. 101–117.

Bernard, A.B., Jensen, J.B., Redding, S.J., Schott, P.K., Firms in international trade. Journal of Economic Perspective 21(3), 105–130.

Brandt, L., Biesebroeck, J.V., Zhang, Y., 2011. Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. Journal of Development Economics 97, 339–351.

Broersma, L., Oosterhaven, J., 2009. Regional labor productivity in the Netherlands: evidence of agglomeration and congestion effects. Journal of Regional Science 49, 483–511.

Brülhart, M., Federica, S., 2009. Agglomeration and growth: cross-country evidence. Journal of Urban Economics 65(1), 48–63.

Brülhart, M., Mathys, N.A., 2008. Sectoral agglomeration economies in a panel of European

regions. Regional Science and Urban Economics 38, 348–362.
Doornik, J.A., 2006. Ox: An Objected-Oriented Matrix Programming Language. Timberlake Consultants, London.
Duranton, G., Puga, D., 2004. Micro-foundations of urban agglomeration economies. In: Henderson, J.V., Thisse, J.F., (Eds.), Handbook of Regional and Urban Economics, vol. 4. North-Holland, New York, pp. 2063–2117.
Eberts, R.W., McMillen, D.P., 1999. Agglomeration economies and urban public infrastructure. In: Cheshire P., Mills E.S., (Eds.), Handbook of Regional and Urban economics, vol. 3. North-Holland, New York, pp. 1455–1495.
Ellison, G., Glaeser, E.L., Kerr, W.R., 2010. What causes industry agglomeration? Evidence from coagglomeration patterns. American Economic Review 100, 1195–1213.
Gamerman, D., Lopes, H.F., 2006. Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference, second edition. Chapman and Hall/CRC, Boca Raton.
Graham, D.J., 2008. Identifying urbanisation and localisation externalities in manufacturing and service industries. Papers in Regional Science, 88, 63-84.
Hashiguchi, Y., Chen, K.-h., 2012. Assessing agglomeration economies in the Yangzi river delta, China: a Bayesian spatial econometric approach. IDE Discussion Paper, No. 339.
Henderson, J.V., 2003. Marshall's scale economies. Journal of Urban Economics 53, 1–23.
Holmes, T.J., 1999. Localization of industry and vertical disintegration. Review of Economic and Statistics 81(2), 314–325.
Fan, J. 2007. Industrial agglomeration and difference of regional productivity. Frontiers of Economics in China 2, 346–361.
Graham, D.J., 2008. Identify urbanisation and localisation externalities in manufacturing and service industries. Papers in Regional Science 88, 63–84.
Greenstone, M., Hornbeck, R., Moretti, E., 2010. Identifying agglomeration spillovers: evidence from winners and losers of large plant openings. Journal of Political Economy 118, 536–598.
Ke, S., 2010. Agglomeration, productivity, and spatial spillovers across Chinese cities. Annals of Regional Science 45, 157–179.
LeSage, J., Pace, P.K., 2009. Introduction to Spatial Econometrics. CRC press, Boca Raton.
Lin, H.-L., Li, H.-Y., Yang, C.-H., 2011. Agglomeration and productivity: firm-level evidence from China's textile industry. China Economic Review 22, 313–329.
Martin, P., Mayer, T., Mayneris, F., 2011. Spatial concentration and plant-level productivity in France. Journal of Urban Economics 69, 182–195.
Marshall, A., 1920. Principles of Economics, 8th edition (Reprinted in 1959). Macmillan, London.
Melo, P.C., Graham, D.J., Noland, R.B., 2009. A meta-analysis of estimates of urban agglomeration economies. Regional Science and Urban Economics 39, 332–342.
Mitra, A., 1999. Agglomeration economies as manifested in technical efficiency at the firm level. Journal of Urban Economics 45, 490–500.
Morikawa, M., 2011. Economies of density and productivity in service industries: an analysis of personal service industries based on establishment-level data. Review of Economics and Statistics 993, 179–192.
Puga, D., 2010. The magnitude and causes of agglomeration economies. Journal of Regional Science 50, 203–219.
R Development Core Team, 2009. R: A Language and Environment for Statistical Computing.

Vienna: R Foundation for Statistical Computing. http://www.R-project.org.

Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. Journal of Urban Economics 50, 191–229.

Rosenthal, S.S., Strange, W.C., 2004. Evidence on the nature and sources of agglomeration economies. In: Henderson, J.V., Thisse, J.F., (Eds.), Handbook of Regional and Urban Economics, vol. 4. Elsevier, Amsterdam, pp. 2119–2171.

Rossi, P.E., Allenby, G.M., McCulloch, R., 2005. Bayesian Statistics and Marketing. John Wiley and Sons Lts, West Sussex.

World Bank, 2008. World Development Report 2009: Reshaping Economic Geography. World Bank, Washington DC.

Table 1: Summary statistics

	Mean	Stdev	Min	Max
Value-added (log Y)	8.352	1.216	0.000	17.155
Number of Labor (log L)	4.475	1.032	0.000	10.698
Capital stock value (log K)	8.081	1.690	0.000	17.635
Average years of schooling (h)	10.413	1.138	9.000	17.052
Share of intermediate inputs (log IS)	-0.297	0.205	-8.229	-0.00003
Months in operation (log MO)	4.139	0.893	0.000	7.824
Localization (log A)				
log FD	-2.831	1.513	-8.401	-0.217
log LD	2.145	1.651	-5.950	5.157
log FN	4.289	1.429	0.000	6.930
Urbanization				
log U	395.813	852.654	39.148	12295.540
Sample size: 97947				

By region:	Shanghai	Jiangsu	Zhejiang	Anhui	All
Number of firms	14554	39591	39677	4125	97947
Number of county-level regions	11	65	69	78	223
Total land area (sq. km)	6340.5	100874	103641	138790	349645.5
Mean of log Y	8.480	8.444	8.200	8.470	8.352
Mean of log K	8.143	8.018	8.096	8.336	8.081
Mean of log L	4.430	4.460	4.484	4.677	4.475
Mean of h	10.687	10.534	10.155	10.778	10.413
Mean of log IS	-0.355	-0.307	-0.259	-0.350	-0.297
Mean of log MO	4.325	4.111	4.106	4.071	4.139
Localization (log A)					
Mean of log FD	-1.777	-2.933	-2.875	-5.156	-2.831
Mean of log LD	3.211	2.071	2.058	-0.061	2.145
Mean of log FN	4.413	4.374	4.405	1.914	4.289
Urbanization					
Mean of log U	1967.737	394.010	277.143	280.539	395.813

Notes: Stdev denotes standard deviation. Value-added (Y) and capital (K) is in thousand yuan; labor (L) is in persons; h is average years of employee schooling. FD and LD are the density (per unit county's land area) of total number of firms and labor, respectively, in the same industry and county as firm f. FN is the number of firms in the same industry and county as firm f. U is county-level urban population density (per county land area unit).

Table 2: Estimation results

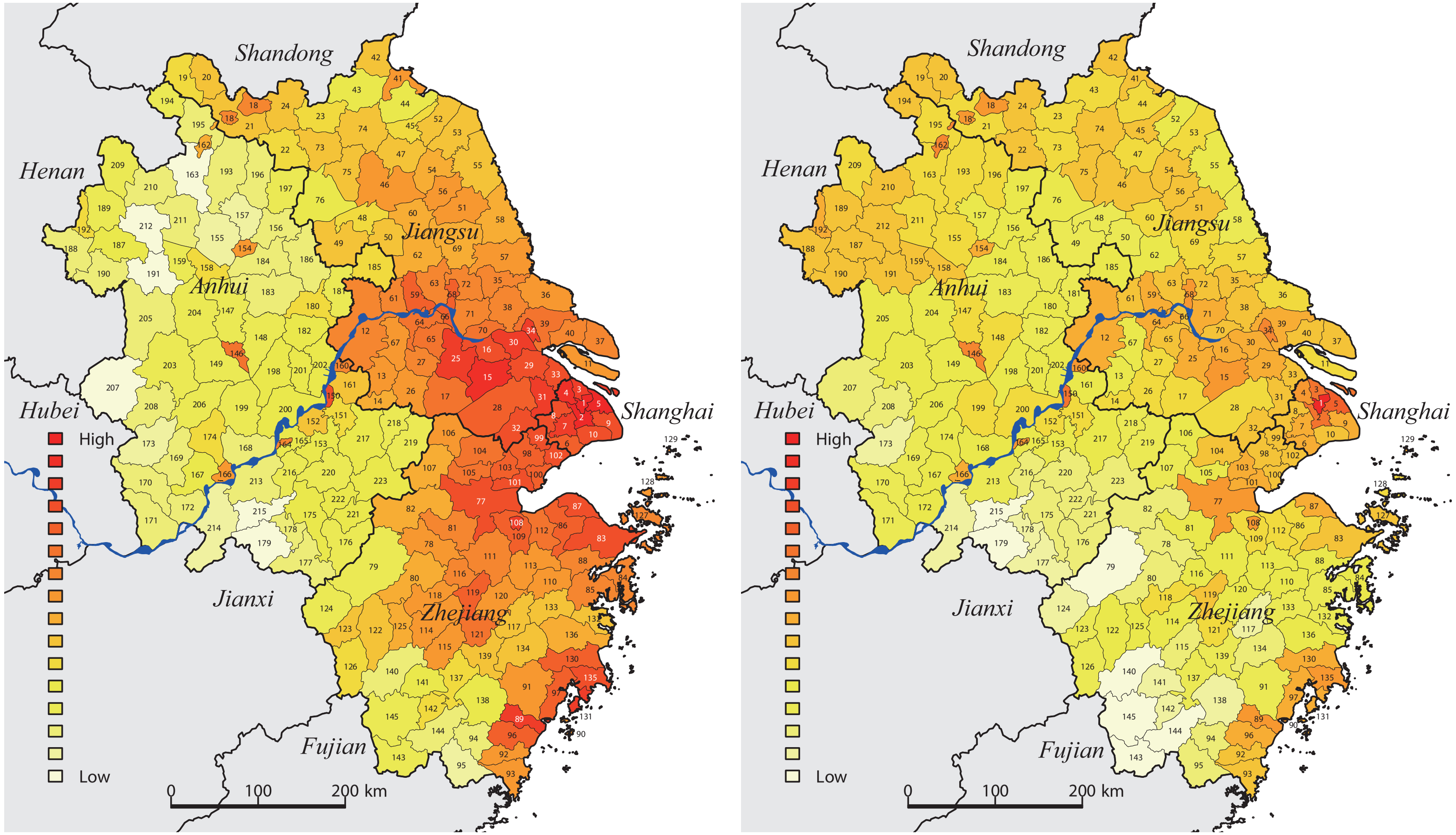
	Mean	Stdev	95%CI
$\log A_f = \log FD_f$			
β_A	0.0220	(0.0052)	[0.012, 0.032]
β_h	0.1487	(0.0032)	[0.142, 0.155]
β_{EX}	0.0260	(0.0064)	[0.013, 0.039]
β_K	0.2102	(0.0021)	[0.206, 0.214]
β_L	0.5941	(0.0037)	[0.587, 0.601]
μ_0	1.2737	(0.2410)	[0.813, 1.765]
μ_U	-0.0007	(0.0165)	[-0.033, 0.032]
ρ	0.5265	(0.0847)	[0.351, 0.688]
$\log A_f = \log LD_f$			
β_A	0.0099	(0.0065)	[-0.003, 0.023]
β_h	0.1527	(0.0028)	[0.147, 0.158]
β_{EX}	0.0259	(0.0065)	[0.013, 0.039]
β_K	0.2104	(0.0021)	[0.206, 0.215]
β_L	0.5954	(0.0035)	[0.588, 0.602]
μ_0	0.9780	(0.2249)	[0.559, 1.445]
μ_U	0.0274	(0.0190)	[-0.009, 0.065]
ρ	0.5559	(0.0813)	[0.386, 0.709]
$\log A_f = \log FN_f$			
β_A	0.0329	(0.0057)	[0.021, 0.044]
β_h	0.1556	(0.0027)	[0.150, 0.161]
β_{EX}	0.0230	(0.0064)	[0.010, 0.036]
β_K	0.2099	(0.0021)	[0.206, 0.214]
β_L	0.5976	(0.0035)	[0.591, 0.604]
μ_0	1.0414	(0.2230)	[0.620, 1.498]
μ_U	0.0159	(0.0182)	[-0.020, 0.052]
ρ	0.5340	(0.0841)	[0.360, 0.692]
Sample size: 97947			

Notes: Stdev and 95%CI denote standard deviation and 95% credible interval, respectively, of the posterior distribution for each parameter. The inverse of minimum and maximum eigen values of the spatial weight matrix **W** are $\lambda_{min}^{-1} = -1.110$ and $\lambda_{max}^{-1} = 1.000$, respectively.

Table 3: Estimates of cross products

	Mean	95%CI	Prob ($x > 0$) (%)
$\log A_f = \log FD_f$			
$\beta_{A.h}$	0.0028	[-0.0001, 0.006]	97.05
$\beta_{A.IS}$	0.0451	[0.029, 0.061]	100.00
$\beta_{A.EX}$	-0.0011	[-0.008, 0.006]	37.95
$\beta_{A.MO}$	-0.0063	[-0.010, -0.003]	0.01
$\beta_{A.FO}$	0.0063	[-0.003, 0.015]	91.45
$\log A_f = \log LD_f$			
$\beta_{A.h}$	0.0080	[0.005, 0.011]	100.00
$\beta_{A.IS}$	0.0767	[0.062, 0.091]	100.00
$\beta_{A.EX}$	0.0068	[-0.00005, 0.014]	97.42
$\beta_{A.MO}$	-0.0009	[-0.004, 0.002]	28.80
$\beta_{A.FO}$	0.0117	[0.003, 0.020]	99.69
$\log A_f = \log FN_f$			
$\beta_{A.h}$	0.0024	[-0.001, 0.006]	92.47
$\beta_{A.IS}$	0.0353	[0.017, 0.053]	100.00
$\beta_{A.EX}$	-0.0021	[-0.010, 0.006]	29.82
$\beta_{A.MO}$	-0.0040	[-0.008, -0.0004]	1.36
$\beta_{A.FO}$	0.0035	[-0.006, 0.013]	76.28
Sample size: 97947			

Notes: 95%CI denotes 95% credible interval of the posterior distribution for each parameter. Prob ($x > 0$) denotes the positive area of each posterior distribution.



Number of manufacturing firms per unit area (logarithm) | Urban population per unit area (logarithm)

Figure 1: Geographic distribution of firms and urban population

Notes: The names of counties in the map are listed in Tables 4-6. This colored map was created by equally dividing the range of data into 16 equal parts and categorizing counties into 16 classes.

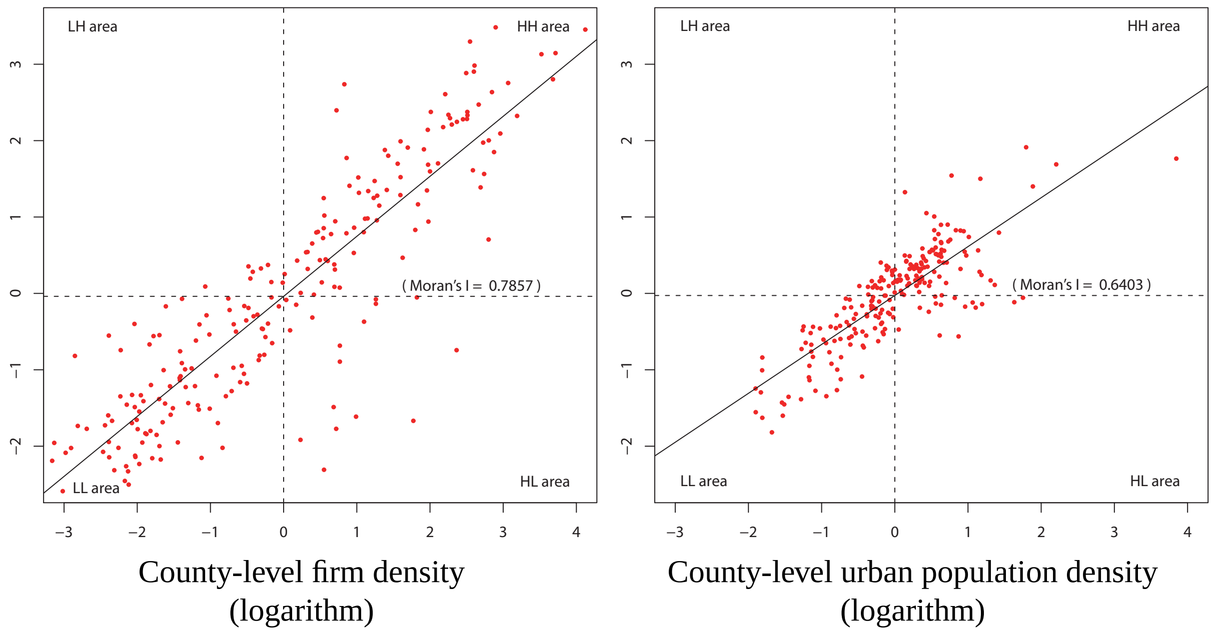
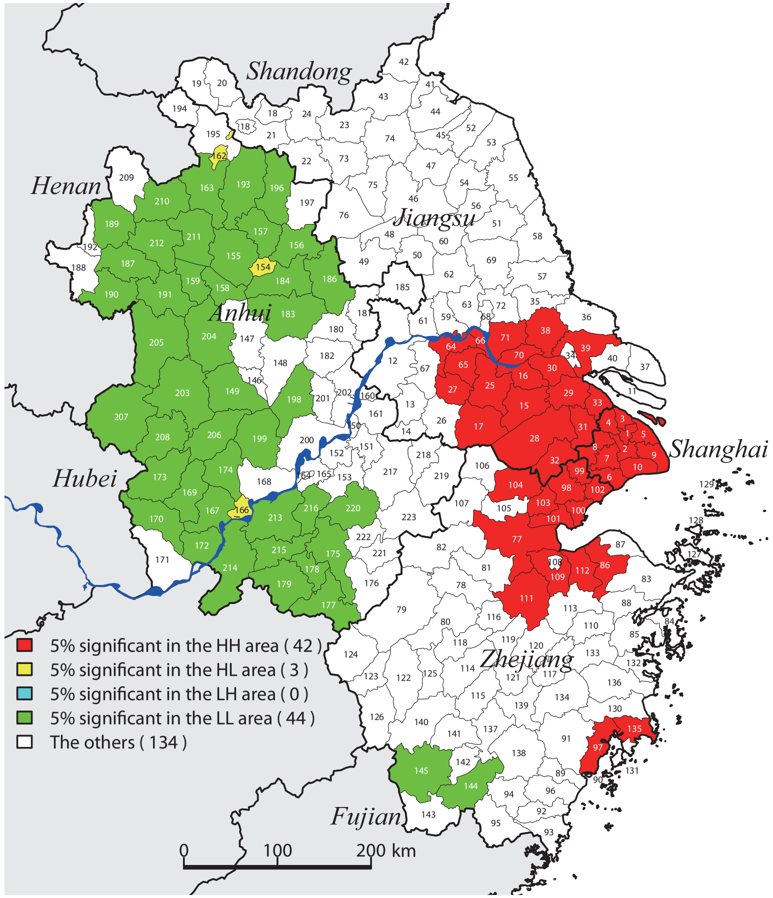
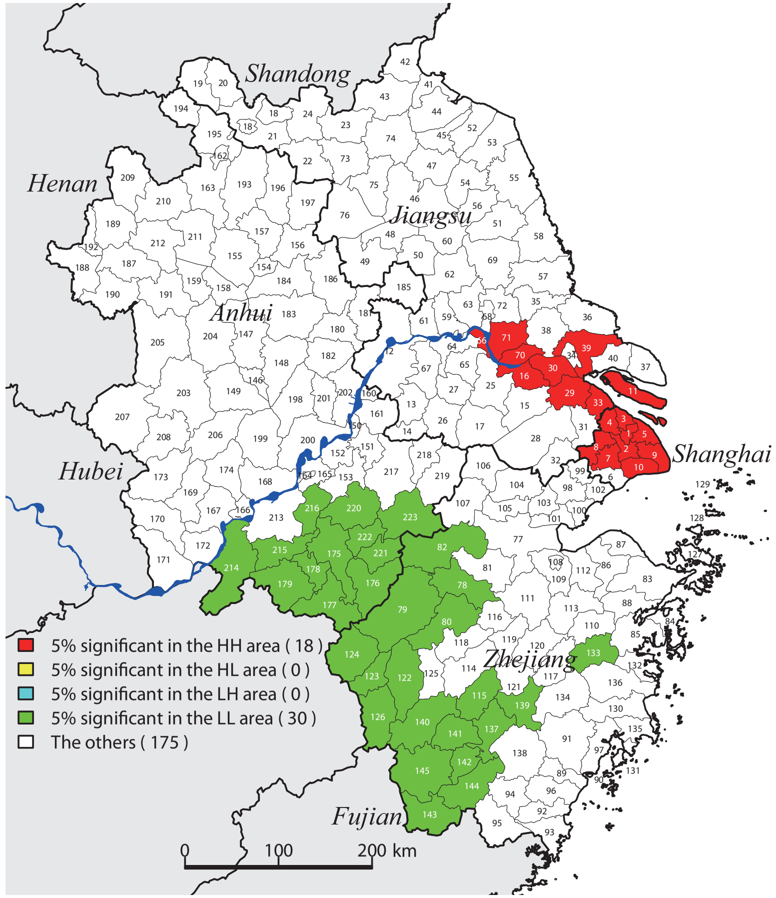


Figure 2: Moran scatterplots

Notes: The firm density is defined as the number of manufacturing firms in a county-level region divided by the region's land area. The urban population density is defined as urban population in a county-level region divided by the region's land area.



Number of manufacturing firms per unit area (logarithm)



Urban population per unit area (logarithm)

Figure 3: Significant local cluster (local Moran statistics)

Notes: The names of counties in the map are listed in Tables 4-6. The HH, LH, LL, and HL areas indicate the first, second, third, and fourth quadrants depicted in Figure 2.

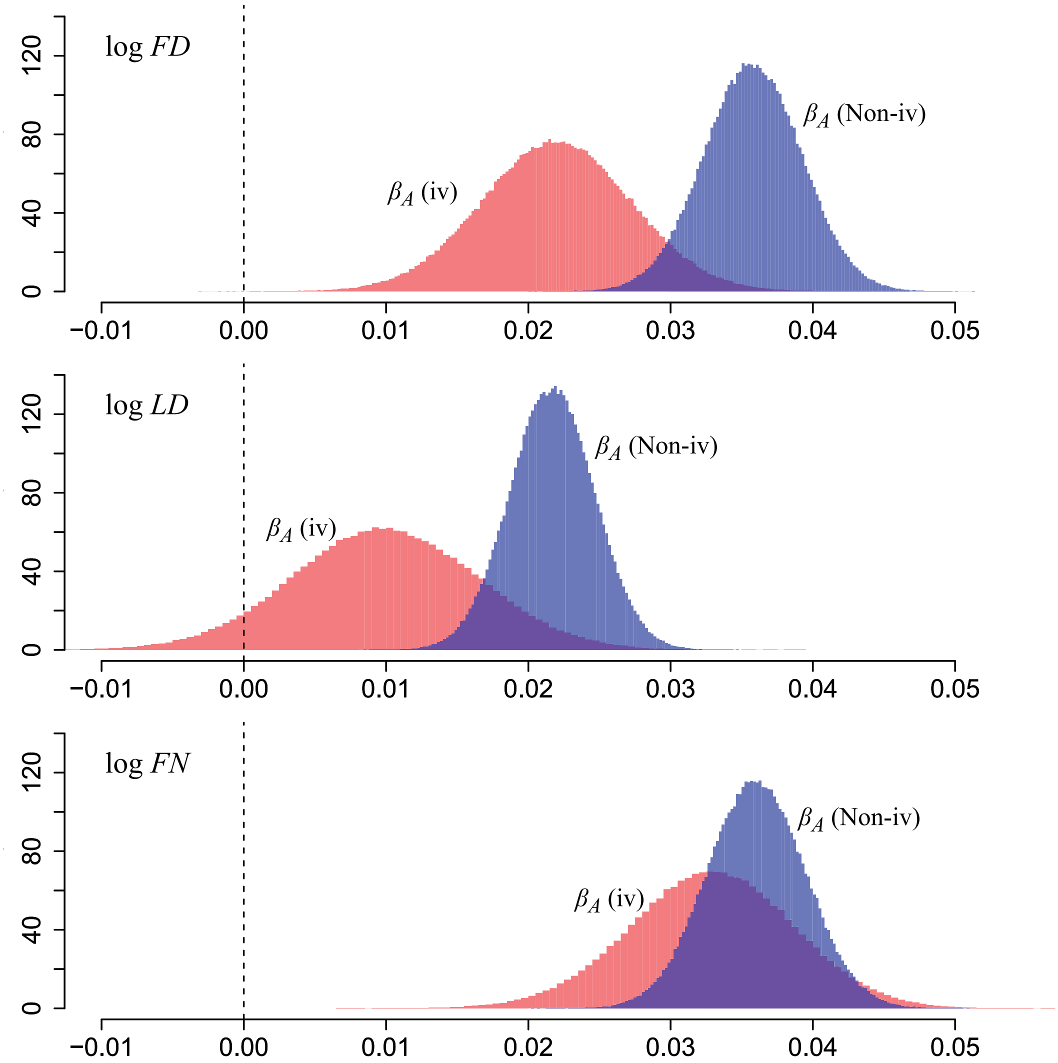


Figure 4: IV- and Non-IV-based posterior distributions of β_A

Table 4: List of county names

id	Prefecture-level	County-level	id	Prefecture-level	County-level
1	Shanghai	Huangpu*	41	Lianyungang	Lianyungang Dist.
2	Shanghai	Minxing	42	Lianyungang	Ganyu
3	Shanghai	Baoshan	43	Lianyungang	Donghai
4	Shanghai	Jiading	44	Lianyungang	Guanyun
5	Shanghai	Pudongxin	45	Lianyungang	Guannan
6	Shanghai	Jinshan	46	Huaiyin	Huaiyin Dist.
7	Shanghai	Songjiang	47	Huaiyin	Lianshui
8	Shanghai	Qingpu	48	Huaiyin	Hongze
9	Shanghai	Nanhui	49	Huaiyin	Xuyi
10	Shanghai	Fengxian	50	Huaiyin	Jinhu
11	Shanghai	Chongming	51	Yancheng	Yancheng Dist.
12	Nanjing	Nanjing Dist.	52	Yancheng	Xiangshui
13	Nanjing	Lishui	53	Yancheng	Binhai
14	Nanjing	Gaochun	54	Yancheng	Funing
15	Wuxi	Wuxi Dist.	55	Yancheng	Sheyang
16	Wuxi	Jiangyin	56	Yancheng	Jianhu
17	Wuxi	Yixing	57	Yancheng	Dongtai
18	Xuzhou	Xuzhou Dist.	58	Yancheng	Dafeng
19	Xuzhou	Fengxian	59	Yangzhou	Yangzhou Dist.
20	Xuzhou	Peixian	60	Yangzhou	Baoying
21	Xuzhou	Tongshan	61	Yangzhou	Yizheng
22	Xuzhou	Suining	62	Yangzhou	Gaoyou
23	Xuzhou	Xinyi	63	Yangzhou	Jiangdu
24	Xuzhou	Pizhou	64	Zhenjiang	Zhenjiang Dist.
25	Changzhou	Changzhou Dist.	65	Zhenjiang	Danyang
26	Changzhou	Liyang	66	Zhenjiang	Yangzhong
27	Changzhou	Jintan	67	Zhenjiang	Jurong
28	Suzhou	Suzhou Dist.	68	Taizhou	Taizhou Dist.
29	Suzhou	Changshu	69	Taizhou	Xinghua
30	Suzhou	Zhangjiagang	70	Taizhou	Jingjiang
31	Suzhou	Kunshan	71	Taizhou	Taixing
32	Suzhou	Wujiang	72	Taizhou	Jiangyan
33	Suzhou	Taicang	73	Suqian	Suqian Dist.
34	Nantong	Nantong Dist.	74	Suqian	Shuyang
35	Nantong	Haian	75	Suqian	Siyang
36	Nantong	Rudong	76	Suqian	Sihong
37	Nantong	Qidong	77	Hangzhou	Hangzhou Dist.
38	Nantong	Rugao	78	Hangzhou	Tonglu
39	Nantong	Tongzhou	79	Hangzhou	Chunan
40	Nantong	Haimen	80	Hangzhou	Jiande

Notes: * Huangpu (id = 1) includes the following nine county-level regions: Huangpu, Luwan, Xuhui, Changning, Jingan, Putuo, Zhabei, Hongkou, and Yangpu. Because the areas of these regions are quite small, we aggregate them into one composite. "Dist." is the abbreviation for "District."

Table 5: List of county names (continued)

id	Prefecture-level	County-level	id	Prefecture-level	County-level
81	Hangzhou	Fuyang	121	Jinhua	Yongkang
82	Hangzhou	Linan	122	Quzhou	Quzhou Dist.
83	Ningbo	Ningbo Dist.	123	Quzhou	Changshan
84	Ningbo	Xiangshan	124	Quzhou	Kaihua
85	Ningbo	Ninghai	125	Quzhou	Longyou
86	Ningbo	Yuyao	126	Quzhou	Jiangshan
87	Ningbo	Cixi	127	Zhoushan	Zhoushan Dist.
88	Ningbo	Fenghua	128	Zhoushan	Daishan
89	Wenzhou	Wenzhou Dist.	129	Zhoushan	Shengsi
90	Wenzhou	Dongtou	130	Taizhou	Taizhou Dist.
91	Wenzhou	Yongjia	131	Taizhou	Yuhuan
92	Wenzhou	Pingyang	132	Taizhou	Sanmen
93	Wenzhou	Cangnan	133	Taizhou	Tiantai
94	Wenzhou	Wencheng	134	Taizhou	Xianju
95	Wenzhou	Taishun	135	Taizhou	Wenling
96	Wenzhou	Ruian	136	Taizhou	Linhai
97	Wenzhou	Leqing	137	Lishui	Lishui Dist.
98	Jiaxing	Jiaxing Dist.	138	Lishui	Qingtian
99	Jiaxing	Jiashan	139	Lishui	Jinyun
100	Jiaxing	Haiyan	140	Lishui	Suichang
101	Jiaxing	Haining	141	Lishui	Songyang
102	Jiaxing	Pinghu	142	Lishui	Yunhe
103	Jiaxing	Tongxiang	143	Lishui	Qingyuan
104	Huzhou	Huzhou Dist.	144	Lishui	Jingning
105	Huzhou	Deqing	145	Lishui	Longquan
106	Huzhou	Changxing	146	Hefei	Hefei Dist.
107	Huzhou	Anji	147	Hefei	Changfeng
108	Shaoxing	Shaoxing Dist.	148	Hefei	Feidong
109	Shaoxing	Shaoxing	149	Hefei	Feixi
110	Shaoxing	Xinchang	150	Wuhu	Wuhu Dist.
111	Shaoxing	Zhuji	151	Wuhu	Wuhu
112	Shaoxing	Shangyu	152	Wuhu	Fanchang
113	Shaoxing	Shengzhou	153	Wuhu	Nanling
114	Jinhua	Jinhua Dist.	154	Bengbu	Bengbu Dist.
115	Jinhua	Wuyi	155	Bengbu	Huaiyuan
116	Jinhua	Pujiang	156	Bengbu	Wuhe
117	Jinhua	Panan	157	Bengbu	Guzhen
118	Jinhua	Lanxi	158	Huainan	Huainan Dist.
119	Jinhua	Yiwu	159	Huainan	Fengtai
120	Jinhua	Dongyang	160	Maanshan	Maanshan Dist.

Note: "Dist." is the abbreviation for "District."

Table 6: List of county names (continued)

id	Prefecture-level	County-level	id	Prefecture-level	County-level
161	Maanshan	Dangtu	201	Chaohu	Hanshan
162	Huaibei	Huaibei Dist.	202	Chaohu	Hexian
163	Huaibei	Suixi	203	Liuan	Liuan Dist.
164	Tongling	Tongling Dist.	204	Liuan	Shouxian
165	Tongling	Tongling	205	Liuan	Huoqiu
166	Anqing	Anqing Dist.	206	Liuan	Shucheng
167	Anqing	Huaining	207	Liuan	Jinzhai
168	Anqing	Zongyang	208	Liuan	Huoshan
169	Anqing	Qianshan	209	Bozhou	Bozhou Dist.
170	Anqing	Taihu	210	Bozhou	Guoyang
171	Anqing	Susong	211	Bozhou	Mengcheng
172	Anqing	Wangjiang	212	Bozhou	Lixin
173	Anqing	Yuexi	213	Chizhou	Chizhou Dist.
174	Anqing	Tongcheng	214	Chizhou	Dongzhi
175	Huangshan	Huangshan Dist.	215	Chizhou	Shitai
176	Huangshan	Shexian	216	Chizhou	Qingyang
177	Huangshan	Xiuning	217	Xuancheng	Xuancheng Dist.
178	Huangshan	Yixian	218	Xuancheng	Langxi
179	Huangshan	Qimen	219	Xuancheng	Guangde
180	Chuzhou	Chuzhou Dist.	220	Xuancheng	Jingxian
181	Chuzhou	Laian	221	Xuancheng	Jixi
182	Chuzhou	Quanjiao	222	Xuancheng	Jingde
183	Chuzhou	Dingyuan	223	Xuancheng	Ningguo
184	Chuzhou	Fengyang			
185	Chuzhou	Tianchang			
186	Chuzhou	Mingguang			
187	Fuyang	Fuyang Dist.			
188	Fuyang	Linquan			
189	Fuyang	Taihe			
190	Fuyang	Funan			
191	Fuyang	Yingshang			
192	Fuyang	Jieshou			
193	Suzhou	Suzhou Dist.			
194	Suzhou	Dangshan			
195	Suzhou	Xiaoxian			
196	Suzhou	Lingbi			
197	Suzhou	Sixian			
198	Chaohu	Chaohu Dist.			
199	Chaohu	Lujiang			
200	Chaohu	Wuwei			

Note: “Dist.” is the abbreviation for “District.”