Foreign Direct Investment and Full Factor Productivity in China

Eunsuk Hong[a] and Laixiang Sun[b, c, d, e]

[a] Management School, Queen's University Belfast, M112, 25 University Square, Belfast BT7 1NN, Northern Ireland, UK, Phone: +44 28 9097 1414. Fax: +44 28 9097 5156. Email: e.hong@qub.ac.uk.

[b] Department of Financial & Management Studies (DeFiMS), SOAS, University of London Thornhaugh Street, Russell Square, London WC1H 0XG, United Kingdom. Phone: +44 20 7898 4821. Fax: +44 20 7898 4089. Email: LS28@soas.ac.uk.

[c] Institute of Geographic Sciences & Natural Resources, Chinese Academy of Sciences, Beijing, China.

[d] International Institute for Applied Systems Analysis, Laxenburg, Austria.

[e] Guanghua School of Management, Peking University, Beijing, China.

(This version: May 2010)

Corresponding Author: Prof. Laixiang Sun, DeFiMS, SOAS, University of London, Thornhaugh Street, Russell Square, London WC1H 0XG, United Kingdom. Phone: +44 20 7898 4821. Fax: +44 20 7898 4089. Email: LS28@soas.ac.uk.

---

[*] We gratefully acknowledge the support of the Sixth Framework Programme of the European Commission (contract number: 044255, the CATSEI project). We thank Tao Li, Pilsoo Kim, Xiaming Liu, Damian Tobin, Yang Yao, Jihai Yu and the seminar and workshop participants at SOAS of University of London, University of Oxford, and Management School of Queen's University Belfast.

Foreign Direct Investment and Full Factor Productivity in China

Abstract

This paper develops a spatial dynamic model to assess the total-factor-productivity (TFP) effects of externalities generated by foreign direct investment (FDI). The model is capable of disentangling TFP effects from capital accumulation effects and introducing spatial interdependence based on theoretical derivation rather than spatial statistical tests. An application of this model to a panel dataset of China at the provincial level over 1980-2005 shows a strong positive impact of FDI externalities on TFP. Moreover, the incorporation of FDI externalities, structural change, and spatial interdependence into neoclassical growth empirics provides a much improved account of interprovincial variations in income levels and economic growth.

JEL classification code: D62; F21; F23; O33; O47 ; P33.
Key Words: Foreign direct investment; Total factor productivity; Spatial econometrics; China.

1. Introduction

Since 1978, China has utilized over US$940 billion in FDI.[1] The rapid rise of China as a global trade power since the early 1990s has been closely associated with the FDI inflow on such a large scale. The share of exports by foreign-funded enterprises (FFEs) in the total has increased speedily from 26% in 1992 to 50% in 2000 and further to the peak of 60% in 2006 (*China Statistical Yearbook*, 1993-2009). The dominance of foreign firms in China's exports is even more apparent in advanced industrial sectors. For example, while exports of industrial machinery grew twenty-fold in real terms over 1993-2003, the share produced by FFEs in these exports grew from 35% to 79%. The exports of electronics and telecom grew seven-fold in the same period, with the FIEs' share rising from 45% to 74% (Gilboy, 2004). As a consequence of this dominance, China has become known as the "world's factory".

Despite the world's factory at the national level, the regional distribution of FDI 'workshop' has been highly uneven. FDI has been mainly located in relatively advanced coastal provinces where a number of locational advantages can be reaped by foreign invested firms. The concentration of FDI would in turn contribute to the spatial density of economic activity which may explain much of the variation of productivity (Ciccone and Hall, 1996) across China. In contrast to the large body of literature on the location choice of foreign investors in the context of China, there is a shortage of research on the productivity and income effect of FDI and on the role of spatial factors in shaping these effects.

An important exception is Madariaga and Poncet (2007). They estimate the autoregressive form of the augmented Solow growth model as proposed by Mankiw et al. (1992), directly add the FDI/GDP rate into the growth model to determine its influence on growth, and include spatially lagged levels of FDI/GDP and per capita GDP in the model to catch the effect of spatial dependence in these two variables which is detected by spatial statistical tests. Their analysis covers 180 Chinese cities over the period 1990-2002 and finds that FDI received by the host city makes significant and positive contribution to the growth of per capita GDP, so does that received by its neighbouring cities. Nevertheless, there are several weaknesses in their research. First, the saving, proxied by fixed-capital investment in the research, also includes FDI. In many coastal cities (i.e., Shenzhen,

[1] China in this paper is referred as the mainland China. Data sources: *China Statistical Yearbook* 2009.

Dongguan and others) FDI accounts for majority of fixed investment.[2] The inclusion of both saving rate and FDI/GDP rate in the model would lead to double-counting of FDI and thus lead to an upward "aggregation bias" in the estimation as indicated in Hale and Long (2006) and Hanousek et al. (2010). Second, the introduction of the FDI/GDP rate is based on descriptive argument and the incorporation of spatial dependence is based on statistical diagnostic tests carried out on the data, rather than derivations from a theory.

The current paper aims to disentangle the capital accumulation effects of FDI from the total-factor-productivity (TFP) effects of externalities generated by FDI and to introduce spatial interdependence based on theoretical derivation rather than spatial statistical tests. From the perspective of national income accounting, FDI is a direct source of capital accumulation and accounted as an integral part of fixed capital investment/formation in the expenditure-based GDP statistics. We follow this convenient approach and do not treat the capital accumulation effect of FDI differently from that of domestic fixed investment. On the other hand, we introduce the FDI intensity variable, which is measured as the ratio of FDI to the total fixed-capital formation, into the specification of TFP to catch the effects of FDI externality on total factor productivity. We note the widely acknowledged perception that the FDI-related spillovers of technological and managerial knowledge to local firms would provide the major transmission mechanism through which FDI promote productivity and income growth. We are also aware of the possible "no spillovers" or even negative effect of FDI externalities on the productivity of local firms due to the lack of absorptive capacity, the large gap problem between domestic and foreign firms in terms of technology and human capital (Fagerberg, 1994; Hanousek et al., 2010), and the "market-stealing" effect and unfair competition (Aitken and Harrison, 1999). We also control for the endogeneity of the FDI intensity variable.

Our theoretical model, which is an extension of Yu (2007) and Sun at al. (2009), starts from a labour-augmented Cobb-Douglas production function, explicitly incorporates the FDI intensity variable and inter-provincial technological spillover and factor mobility into the specification of the TFP function. It further allows the traditionally "same" exogenous rate of technological progress to vary across provinces so that significant structural changes which have happened in China's rapid industrialization can be

[2] For example, the maximum of the FDI/GDP rate in Madariaga and Poncet (2007) is 0.62, which implies that virtually all fixed investment in the city-year is FDI.

introduced in a manner similar to Temple and Wößmann (2006). Mathematical derivation of the model directly leads to an estimation equation with both spatial and serial lag of the dependent variable (i.e., per capita income level). The equation also provides comprehensive and consistent control for influential variables and specification errors which may alter the relationship between FDI intensity and TFP. We apply this augmented spatial dynamic model to the panel data on Chinese provinces over the period 1980-2005. We estimate the model using the most representative estimators adopted in the literature, including the pooling regression with OLS, fixed effects estimator, spatial panel regression with maximum likelihood (ML) estimator, and the combined spatial and dynamic panel regression with system generalized method of moments (GMM) estimator. The system GMM is capable of dealing with the joint endogeneity problem of the serially and spatially lagged dependent variable, corrects for the potential endogeneity of other explanatory variables as being in our case the endogeneity of FDI intensity, and allows for unobserved region-specific effects and measurement errors.

The estimation results indicate a significantly positive impact of FDI externalities on TFP. Moreover, in comparison with the results in Sun et al. (2009), the incorporation of FDI externalities, together with structural change and spatial interdependence, into the specification of TFP in the neoclassical growth model provides a further improved account of interprovincial variations in income levels and economic growth. The incomes remains spatially correlated mainly due to inter-regional technological spillover and factor mobility. The speed of convergence increases from about 2.1% to about 3%. The findings suggest that capital investments, FDI intensity in capital investment, changes in the structure of employment, force of conditional convergence, and population growth are the main sources of the income and growth difference across Chinese provinces. These findings shed new light on the growth empirics in the world's biggest developing and transitional economy, China, during the vital stage of its economic take-off.

The rest of the paper is organized as follows. Section 2 derives the model from the neoclassical growth theory and reviews the literature on spillovers from FDI. Sections 3 and 4 apply the augmented spatial dynamic model to the case of China. Section 3 introduces data and variables and discusses estimation issues. Section 4 reports the empirical results and discusses how China's story is either different or similar to that of others. Section 6 concludes.

2. Theory

2.1. The Underlying Model

In this section we extend the spatial dynamic growth model of Sun et al. (2009) with the aim of incorporating the spillover effects of FDI in the TFP specification. For the sake of self containing, we also draw on the relevant text. The production function takes the labour-augmenting Cobb-Douglas form:

$$Y_{it} = K_{it}^{\alpha}\left(A_{it}L_{it}\right)^{1-\alpha}, \qquad 0 < \alpha < 1, \tag{1}$$

where Y is output, K is capital, L is labour, and A is labour augmenting technological progress. The subscript i stands for provinces and t for years. For notational convenience, define output and capital per unit of effective labour as $\hat{y} = Y/AL$ and $\hat{k} = K/AL$ respectively and let bold letter denote a vector such as $\mathbf{A}_t = (A_{1t}, A_{2t}, \cdots, A_{Nt})'$ and $\mathbf{1}_N = (1, 1, \cdots, 1)'$ with N indicating the number of elements in the vector.

TFP is affected by technological externalities of FDI in the province, technology spillover across provinces, and the additional exogenous technological progress rate. More specifically, we model TFP as

$$A_{it} = A_{i0}e^{\xi d_{it} + g_i t}\prod_{j \neq i}^{N} A_{jt}^{\rho w_{ij}}, \qquad 0 \leq \rho < 1. \tag{2}$$

The expression (2) says that the technology level in province i, A_{it}, depends on not only its own initial level A_{i0} and exogenous progress rate g_i, but also on FDI intensity, d_{it}, and a Cobb-Douglas combination of the levels of its neighbours' A_{jt}. The coefficient on d_{it} can be positive, zero, or negative because the intraregional spillover effect of FDI on TFP can be positive, zero, or negative as we will discussed in Section 2.2. Technology levels in other provinces, which include the effect of d_{jt} in province j, spill over to the province i with an elasticity of $\rho \cdot w_{ij}$, where w_{ij} depends on the distance between province i and j. Such a specification allows the spillover effect of A_{jt} on A_{it} to be different across j, depending on the distance between j and i. It means that interregional technology spillovers decay rapidly across geographic space (Adams and Jaffe, 1996; Bottazi and Peri, 2003).

The expression (2) can be transformed into $\ln \mathbf{A}_t = \ln \mathbf{A}_0 + \mathbf{g}t + \xi \mathbf{d}_t + \rho \mathbf{W} \ln \mathbf{A}_t$, where $\mathbf{g} = (g_1, g_2, \cdots, g_N)'$ and $\mathbf{W} = (w_{ij})$ is a row-normalized $N \times N$ matrix of inverse distance between province i and j. This $\mathbf{W}$ is called the spatial lag weighting matrix in spatial econometrics and we will specify it in details in the next section. Let $\bar{g}$ denote the mean of **g** vector. Noting that $(\mathbf{I}_N - \rho\mathbf{W})^{-1}\bar{g}\mathbf{1}_N = \frac{1}{1-\rho}\bar{g}\mathbf{1}_N$ as **W** is row-normalized and assuming that $(\mathbf{I}_N - \rho\mathbf{W})'(\mathbf{g} - \bar{g}\mathbf{1}_N)$ is dominated by $(\mathbf{I}_N - \rho\mathbf{W})'\bar{g}\mathbf{1}_N$, i.e., $(\mathbf{I}_N - \rho\mathbf{W})^{-1}(\mathbf{g} - \bar{g}\mathbf{1}_N) \approx \frac{1}{1-\rho}(\mathbf{g} - \bar{g}\mathbf{1}_N)$, with a similar approximation for $\mathbf{d}_t$ vector, we obtain

$$\ln \mathbf{A}_t = (\mathbf{I}_N - \rho\mathbf{W})^{-1} \ln \mathbf{A}_0 + \frac{1}{1-\rho}\left(\mathbf{g}t + \xi\mathbf{d}_t\right), \tag{3}$$

and

$$\frac{\dot{\mathbf{A}}_t}{\mathbf{A}_t} = \frac{1}{1-\rho}\left(\mathbf{g} + \xi\dot{\mathbf{d}}_t\right). \tag{4}$$

Eq. (4) indicates that that if $0 < (1 - \rho) < 1$, i.e., the interregional technology spillovers are present, the magnitude of the marginal impact of FDI intensity on TFP growth rate will be $\xi/(1-\rho)$, the absolute value of which is greater than $|\xi|$ (if $|\xi| > 0$), the magnitude of the otherwise standard intra-regional spillover effect of FDI on TFP.

To incorporate factor mobility we recognize that the mobility of labour and capital is typically a ‘gravity-driven’ flow, moving from economies with low return rates to those with high return rates. Denote capital per effective labourer as $\hat{\mathbf{k}}_t = (\hat{k}_{1t}, \hat{k}_{2t}, \cdots, \hat{k}_{Nt})'$. We postulate a labour migration function $m_i^L(\mathbf{k}_t) = b^L(\ln \hat{k}_{it} - \sum_{j\neq i}^N m_{ij} \ln \hat{k}_{jt})$ for province i, with $b^L > 0$ to indicate that labour moves from an economy with lower capital-labour ratio to one with higher capital-labour ratio, where, m_{ij} is the (i, j) entry of a spatial weights matrix $\mathbf{M}_N$ and $\sum_{j\neq i}^N m_{ij} \ln \hat{k}_{jt}$ is the weighted average of $\ln \hat{k}_{jt}$, which can be regarded as a proxy for wage rate, across neighbouring provinces. Thus we have

$$\frac{\dot{L}_{it}}{L_{it}} = n_i + m_i^L(\mathbf{k}_t). \tag{5}$$

Eq. (5) indicates that the change in labour supply in province i comes from two sources: the exogenous growth rate of population n_i and migration from (to) other provinces as specified by $m_i^L(\mathbf{k}_t)$.[3]

[3] What we need is a postulation of a migration function and thus the optimization problem for migrants is not considered here.

We postulate a capital movement function $m_i^K(\mathbf{k}_t) = b^K(\ln \hat{k}_{it} - \sum_{j\neq i}^{N} m_{ij} \ln \hat{k}_{jt})$ in a parallel fashion. The change in capital stock is then given by

$$\dot{K}_{it} = s_i Y_{it} - \delta_i K_{it} + m_i^K(\hat{\mathbf{k}}_t) \cdot K_{it}, \tag{6}$$

where s_i is the saving rate and δ_i the depreciation rate. Eq. (6) adds a capital movement term $m_i^K(\hat{\mathbf{k}}_t) \cdot K_{it}$ to the otherwise standard Solow-Swan capital formation model.

The dynamics of $\hat{k}_{it}$ can be derived from Eq. (6) as follows.

$$\begin{aligned} \dot{\hat{k}}_{it} &= s_i \hat{y}_{it} + m_i^K(\mathbf{k}_t) \cdot \hat{k}_{it} - \left(\frac{\dot{L}_{it}}{L_{it}} + \frac{\dot{A}_{it}}{A_{it}} + \delta_i\right) \cdot \hat{k}_{it} \\ &= s_i \hat{k}_{it}^{\alpha} - \hat{k}_{it}\left(n_i + \frac{g_i}{1-\rho} + \delta_i + \xi \dot{\mathbf{d}}_t + m_i^L(\mathbf{k}_t) - m_i^K(\mathbf{k}_t)\right) \end{aligned} \tag{7}$$

In the steady state, $\dot{\hat{k}}_{it} = 0$, $\dot{\mathbf{d}}_t = 0$, and $m_i^L(\mathbf{k}_t) = m_i^K(\mathbf{k}_t) = 0$, which implies that

$$\hat{k}_i^* = \left(\frac{s_i}{n_i + g_i/(1-\rho) + \delta_i}\right)^{1/(1-\alpha)} \text{ and thus } \hat{y}_i^* = \left(\frac{s_i}{n_i + g_i/(1-\rho) + \delta_i}\right)^{\alpha/(1-\alpha)}. \tag{8}$$

Using $\dot{\hat{k}}_{it} / \hat{k}_{it} = \frac{d}{dt}(\ln \hat{k}_{it} - \ln \hat{k}_i^*)$, assuming that $\dot{\mathbf{d}}_t \approx 0$ in the neighbourhood of the steady state, and approximating around the steady state based on Eq. (7), the speed of convergence is given by

$$\frac{d}{dt}(\ln \hat{k}_{it} - \ln \hat{k}_i^*) = -\lambda_i (\ln \hat{k}_{it} - \ln \hat{k}_i^*), \tag{9}$$

where λ_i is the rate of convergence and given by

$$\lambda_i = (1-\alpha)(n_i + g_i/(1-\rho) + \delta_i) + b^L - b^K. \tag{10}$$

Eq. (11) indicates that if 0 < (1 – ρ) < 1, i.e., the interregional technology spillovers are present, and $b^L - b^K > 0$, e.g., in the case that labour moves from an economy with lower capital-labour ratio to one with higher capital-labour ratio and capital moves in the opposite direction, the convergence rate will be greater than the standard rate of $(1-\alpha)(n_i + g_i + \delta_i)$ as suggested in the convergence literature (Barro and Sala-i-Martin, 1992; Islam, 1995). Nevertheless, if $b^L - b^K < 0$, e.g., in the case that both labour and capital move towards higher capital-labour ratio to pursue economies of agglomeration, and ρ is sufficiently small, the convergence rate could become less than the standard rate of $(1-\alpha)(n_i + g_i + \delta_i)$. For the feasibility of empirical estimation, we assume that λ_i is same for all *i* and denoted by λ.

Solving the first-order differential equation (9) and noting that the path of $\ln \hat{k}_{it}$ is the same as that of $\ln \hat{y}_{it}$ because $\hat{y}_{it} = \hat{k}_{it}^{\alpha}$, we obtain

$$\ln \hat{y}_{it_2} = e^{-\lambda\tau} \ln \hat{y}_{it_1} + (1 - e^{-\lambda\tau}) \ln \hat{y}_i^*, \tag{11}$$

where $\hat{y}_i^*$ is given by (8) and $\tau = t_2 - t_1$. Placing $\ln y_{it} = \ln \hat{y}_{it} + \ln A_{it}$ into (11) and using vector form we have the following equation

$$\ln \mathbf{y}_{t_2} = e^{-\lambda\tau} \ln \mathbf{y}_{t_1} + \ln \mathbf{A}_{t_2} - e^{-\lambda\tau} \ln \mathbf{A}_{t_1} + (1 - e^{-\lambda\tau}) \ln \hat{\mathbf{y}}^*. \tag{12}$$

A combination of (12) and (3) gives

$$\begin{aligned}(\mathbf{I}_N - \rho\mathbf{W}) \ln \mathbf{y}_{t_2} = {} & (\mathbf{I}_N - \rho\mathbf{W}) e^{-\lambda\tau} \ln \mathbf{y}_{t_1} + (1 - e^{-\lambda\tau}) \ln \mathbf{A}_0 + (t_2 - e^{-\lambda\tau} t_1)\mathbf{g} \\ & + \xi(\mathbf{d}_{t_2} - e^{-\lambda\tau}\mathbf{d}_{t_1}) + (\mathbf{I}_N - \rho\mathbf{W})(1 - e^{-\lambda\tau}) \ln \hat{\mathbf{y}}^*.\end{aligned} \tag{13}$$

Eqs. (13) and (8) lead to our basic estimation equation as follows.

$$\begin{aligned}\ln \mathbf{y}_{t_2} = {} & \rho\mathbf{W} \ln \mathbf{y}_{t_2} + e^{-\lambda\tau} \ln \mathbf{y}_{t_1} - \rho e^{-\lambda\tau} \mathbf{W} \ln \mathbf{y}_{t_1} + (t_2 - e^{-\lambda\tau} t_1)\mathbf{g} + \xi(\mathbf{d}_{t_2} - e^{-\lambda\tau}\mathbf{d}_{t_1}) \\ & + (1 - e^{-\lambda\tau}) \ln \mathbf{A}_0 + \frac{\alpha(1 - e^{-\lambda\tau})}{1 - \alpha} (\mathbf{I}_N - \rho\mathbf{W}) \ln \mathbf{x_t} + \mathbf{v}_t .\end{aligned} \tag{14}$$

where $\mathbf{x} = \left(\frac{s_1}{n_1 + g_1/(1-\rho) + \delta_1}, \frac{s_2}{n_2 + g_2/(1-\rho) + \delta_2}, \cdots, \frac{s_N}{n_N + g_N/(1-\rho) + \delta_N}\right)'$ and $\mathbf{v}_t = (v_{1t}, v_{2t}, \cdots, v_{Nt})'$ is the transitory error terms that are assumed to be *i.i.d.* across *i* and *t*. It can be seen that the spillover effect ρ is crucial in the derivation of Eq. (14). If $\rho = 0$, Eq. (14) will be reduced to one similar to Eq. (11) in Islam (1995), with different convergence rate due to the presence of factor mobility as indicated in Eq. (10).

In our model, the exogenous rate of a province's own technology progress, g_i, varies across provinces and is free for a further augmentation. While it is difficult to have a complete account for g_i, Temple and Wöβmann (2006) suggest a partial account which quantifies the direct contribution of labour reallocation to aggregate TFP growth in economies with sizeable differentials in the marginal product of labour across different sectors. The intuition is that if the marginal product of labour is lower in agriculture, the movement of agricultural workers to sectors where the marginal product is higher will raise the total output. Because this additional output is produced without additional input of capital and labour, the reallocation of labour raises aggregate productivity. In a large developing economy like China where regional differentials on industrialization has been substantial (cf. Sun at al. 2009, Section 3), their account is particularly relevant and would be able to capture a substantial part of aggregate TFP growth. Their final regression specification of the relationship between the aggregate TFP growth and structural change is as follows.

$$\frac{\dot{z}}{z} = \beta'\mathbf{V} + (\kappa - 1)\phi MGRORTH + \kappa\phi\frac{1}{\psi}DISEQ\,, \tag{15}$$

where z is the aggregate TFP, $\mathbf{V}$ is a vector of determinants of aggregate TFP growth including regional dummies and $ln\mathbf{A}_0$, and the structural change terms are

$$MGROWTH = (1-a)\frac{\dot{m}}{m} \approx \Delta m\,, \tag{16a}$$

$$DISEQ = \frac{p}{1-p}(1-a)\frac{\dot{m}}{m} \approx \frac{p}{1-p}\Delta m\,. \tag{16b}$$

where a is the share of agricultural employment in total employment, $m = 1 - a$ is the share of non-agricultural employment in the total employment, $\Delta m = m_{t_2} - m_{t_1}$, $p = -\Delta a / a$ ($\Delta a = a_{t_2} - a_{t_1}$) is the migration propensity, ϕ is proximately equal to the labour share in total output, and the parameters κ and ψ capture the intersectoral wage ratio at and the speed of adjustment to the long-run migration equilibrium respectively.

Intuitively speaking, *MGROWTH* captures the effect of labour reallocation on TFP growth for a fixed marginal product ratio. This effect is essentially that examined in Kuznets (1961) and Denison (1967). *DISEQ* implies that the growth impact of a given extent of structural change will be greater in those regions experiencing more rapid structural change, because at least on average the intersectoral differential in those regions is greater. Because of this distinction, the former is termed as the ‘linear’ effect and the latter ‘nonlinear’ effect. A combination of Eqs (14) and (15) give the final estimation equation.

$$\begin{aligned} ln\mathbf{y}_{t_2} &= \rho\mathbf{W}\,ln\mathbf{y}_{t_2} + e^{-\lambda\tau}\,ln\mathbf{y}_{t_1} - \rho e^{-\lambda\tau}\mathbf{W}\,ln\mathbf{y}_{t_1} + \frac{(1-e^{-\lambda\tau})(1-\alpha)+1}{1-\alpha}ln\,\mathbf{A}_0 \\ &+ (t_2 - e^{-\lambda\tau}t_1)\phi\left(\frac{\kappa-1}{1-\alpha}MGROWTH + \frac{\kappa}{(1-\alpha)\psi}DISEQ\right) \\ &+ \xi(\mathbf{d}_{t_2} - e^{-\lambda\tau}\mathbf{d}_{t_1}) + \frac{\alpha(1-e^{-\lambda\tau})}{1-\alpha}(\mathbf{I}_N - \rho\mathbf{W})ln\,\mathbf{x}_\mathbf{t} + \mathbf{v}_t\,. \end{aligned} \tag{17}$$

Eq. (17) presents a consistent theoretical setting to comprehensively control for model specification errors and influential variables which may alter the relationship between investment intensity d_{it} and TFP. A complete mechanical adoption of Eq. (17) for real regression would be problematic due to nonlinear parameter restrictions and high correlation between some variables. Our first estimation will not impose parameter restrictions and the second one will impose the linear restrictions only. We will exclude some variables that are of secondary importance and highly correlated with key variables. Another pragmatic approximation is necessary for $\ln(n_i + g_i/(1-\rho) + \delta_i)$ because it contains the spatial effect parameter ρ. We will focus on the variation of population growth n_i and take the average value of technological progress and depreciation for $g_i/(1-\rho) + \delta_i$. This is to assume that variations in $g_i/(1-\rho) + \delta_i$ is likely to be modest in relation to the inter-provincial variations in population growth in a large developing country like China.

2.2. Spillovers from FDI

Spillovers from FDI carry the features of the typical notion of technology spillovers/externalities and can be divided into pecuniary (also called vertical or rent) and non-pecuniary (also called horizontal or knowledge) spillovers. Pecuniary or vertical spillovers flow in two directions of suppliers and buyers, i.e., forward and backward linkages, respectively. It occurs when quality improvement in inputs and outputs are not entirely reflected in the price of such goods and services. Recipients of these welfare enhancing externalities experience a cost reduction and a subsequent rent gain (Harris, 2009). It is also often the case that foreign companies require higher standards for input quality and on-time delivery which confer incentives on domestic suppliers to upgrade their production technology and management.

In contrast, non-pecuniary or knowledge spillovers are disembodied from goods and services and instead arise when firms share a general pool of knowledge, which can shift their production possibility frontier (Harris, 2009). For example, FDI firms usually possess better technology in comparison with that of local firms and local firms can become more productive by watching and imitating these FDI firms. non-pecuniary spillovers result from the nature of knowledge as a local or partial public good. The codified part of an invention (innovation) is likely to be a fully public good. The embodied and tacit part of the inventive (innovative) knowledge is linked to the experiences of the inventors (innovators) and attached to people. This stock of knowledge grows in a region as technology transfer from FDI occurs as well as local inventors and innovators discover new ideas. It diffuses mostly via personal interactions and labour mobility. It is local public good because it mainly benefits inventors and innovators within the region or its neighbourhoods but fades farther away as interactions and labour mobility decrease (Botazi and Peri, 2003).

Apart from the potential positive contribution of FDI spillover to local TFP and income growth as discussed above, it is also worth noting that a spillover effect of FDI can be zero or negative. Local firms may experience difficulties in absorption of new technology due to a lack of complementarities or due to the "gap" problem in terms of technology and human capital. Foreign investors can pick the best local firm, allowing that company to dominate the market and crowd out other firms. A significant increase of market share of FDI firms in product, labour, and capital markets may push domestic firms up their average cost curves. For example, FDI firms may "poach" for better staff by offering higher pay and attractive career development and thus exert upward pressure on wage cost (Harris, 2009; Hanousek et al., 2010).

While the above discussion on FDI spillover indicates a clear-cut direction for causality running from FDI to TFP and growth and leaves the sign and extent of the comprehensive spillover effect to be determined by empirical estimation, an unbiased and consistent estimation of the sign and scale of the effect must control for the following two lines of reverse causality. First, FDI may gravitate to regions where productivity is higher and this type of agglomeration would lead to positive association between FDI and TFP. A lack of control for this line of reverse causality may generate an upward bias. Second, FDI may also be attracted to regions where wage is low to reap the low-cost advantage. Because low-wage is typically associated with low productivity, this type of FDI presence would bring in negative association between FDI and TFP. A lack of control for this second line of reverse causality would create a downward bias. For these reasons, in our system GMM estimations, we will treat FDI intensity variable d_{it} as endogenous.

(See *Table 1*)

3. Data, Variables, and Estimation Methods

Data used in our estimations is a panel of 29 provinces and municipalities for the period 1980-2005. Among all 31 provinces and municipalities in China, Tibet is excluded mainly because of lack of data. For data consistency Chongqing is included in Sichuan, since Chongqing became a central municipality out of Sichuan province in 1997. We collected data from SSB (1999) for the period 1980-1998 and SSB (2000-2006) for the period 1999-2005, respectively. All the value variables are deflated using 1980 prices.

Following the tradition of growth empirics (Islam, 1995), we opt for five-year time intervals in order to reduce the influence of business cycle fluctuations and to alleviate the problem of parameter heterogeneity. For the period 1980-2005 we have five data point for each province: 2005, 2000, 1995, 1990, and 1985. When $t = t_2 = 1985$, $t - 1 = t_1 = 1980$, saving and population growth variables are averages over 1980-1985, and the structural change is the change between 1980 and 1985. The analogy holds for other intervals.

Because $\mathbf{W}\ln\mathbf{y}_{t-1}$, $\mathbf{W}\ln\mathbf{s}_t$ and $\mathbf{W}\ln(\mathbf{n}_t + \bar{g}/(1-\bar{\rho}) + \bar{\delta})$ are highly correlated with $\mathbf{W}\ln\mathbf{y}_t$ (r = 0.99, 0.96, and 0.79 respectively), in our first set of regressions, which focuses on the endogenous impact of $\mathbf{W}\ln\mathbf{y}_t$, we remove them from the regression equation. Following the conventional notion of the panel data literature and noting that $\ln\mathbf{A}_0$ is time invariant, we can reformulate Eq. (17) as follows.

$$\begin{aligned} ln\, y_{it} &= \gamma\, ln\, y_{i,t-1} + \rho(\mathbf{W}\, ln\,\mathbf{y})_{it} + \beta_1 ln\, s_{it} + \beta_2 ln(n_{it} + \bar{g}/(1-\bar{\rho}) + \bar{\delta}) \\ &+ (\beta_3 + \beta_4 t)\cdot MGROWTH_{it} + (\beta_5 + \beta_6 t)\cdot DISEQ_{it} + \beta_7 d_{it} \\ &+ \beta_8 d_{i,t-1} + \eta_i + \mu_t + \varepsilon_{it}, \end{aligned} \tag{18}$$

where $t = t_2$, $t - 1 = t_1$, $\gamma = e^{-\lambda\tau}$, and $\tau = t_2 - t_1$. In the unrestricted setting, we expect that $\beta_1 > 0$, $\beta_2 < 0$, β_3 to β_6 are nonnegative, β_3 or $\beta_4 > 0$, and β_5 or $\beta_6 > 0$. $\beta_7 \geq 0$ is associated with $\beta_8 \leq 0$ and $\beta_7 < 0$ with $\beta_8 > 0$. In the restricted setting, we expect $\beta_1 = -\beta_2$ in addition. Moreover, because *MGROWTH* is highly correlated with $t\cdot MGROWTH$ and *DISEQ* is highly correlated with $t\cdot DISEQ$ (r > 0.99 in both case), we have to drop one of the paired two variables and examine the effect of the other.

It is worth noting that the omission of $\mathbf{W}\ln\mathbf{x}_t$ may result in spurious significance of $\mathbf{W}\ln\mathbf{y}_t$ and d_{it} in Eq. (18). To check this, we conduct second set of regressions based on Eq. (19):

$$\begin{aligned} ln\, y_{it} &= \gamma\, ln\, y_{i,t-1} + \rho(\mathbf{W}\, ln\,\mathbf{y})_{it} + \beta_{1a} ln\, s_{it} + \beta_{2a} ln(n_{it} + \bar{g}/(1-\bar{\rho}) + \bar{\delta}) \\ &+ \beta_{1b}(\mathbf{W}\, ln\, s)_{it} + \beta_{2b}(\mathbf{W}\, ln(n + \bar{g}/(1-\bar{\rho}) + \bar{\delta}))_{it} \\ &+ (\beta_3 + \beta_4 t)\cdot MGROWTH_{it} + (\beta_5 + \beta_6 t)\cdot DISEQ_{it} + \beta_7 d_{it} \\ &+ \beta_8 d_{i,t-1} + \eta_i + \mu_t + \varepsilon_{it}, \end{aligned} \tag{19}$$

In the unrestricted setting, we expect that $\beta_{1a} > 0$, $\beta_{2a} < 0$, $\beta_{1b} < 0$, and $\beta_{2b} > 0$. Expectations on others are same as in Eq. (18). In the restricted setting, we expect

$\beta_{1a} = -\beta_{2a}$ and $\beta_{1b} = -\beta_{2b}$ in addition. Given that Eq. (19) mainly serves the purpose of robustness check, unless specially mentioned we will focus on Eq. (18) in the following.[4]

Table 1 presents the definition of the variables in Eq. (18) and the corresponding summary statistics. It is worth highlighting that while the saving rate incorporates the capital accumulation contribution of FDI as an integral part of fixed capital formation, FDI intensity further measures the relatively position of FDI in total capital formation. The disturbance term consists of the unobserved provincial fixed effect that is constant over time (η_i), the unobserved time effect that is common across provinces (μ_t) and the transitory errors (ε_{it}) that varies across provinces and time periods and has mean equal to zero.

The spatial weight matrix **W** describes the spatial arrangement of the N regions concerned. Let w_{jk} denote the (j, k)-th element of **W**, where j and $k = 1, \cdots, N$. It is assumed that all w_{jk} are known constants, all diagonal elements of **W** are zero, and the characteristic roots of **W** are known. The first assumption excludes the possibility that the spatial weight matrix is parametric. The second one implies that no region can be viewed as its own neighbour. The third presupposes that the characteristic roots of **W** can be computed accurately using the computing technology typically available to empirical researchers and ensures that the log-likelihood function of the spatial regression models we distinguish can be computed (Elhorst, 2003). In this research the off-diagonal elements of **W** are first defined as $w_{jk} = 1/d_{jk}$ where d_{jk} is the distance between the capital city of province j and that of province k, with $k \neq j$.[5] This **W** is then row-normalized so that each row sums to unity. This row-normalized **W** being multiplied by the vector of $\ln \mathbf{y_t}$ leads to the vector of the spatial lagged dependent variable.

There is a host of methods to estimate Eq. (18). The most representative ones include pooling regression with OLS, fixed effects estimator, spatial panel regression with maximum likelihood (ML) estimator, and the combined spatial and dynamic panel regression with system GMM. In the literature on growth empirics which does not use spatial econometric technique it is well-known that the simple pooled OLS estimate of the

[4] Please note that the number of observation in our database is small only 145, which makes it infeasible to include nonlinear parametric constraints in the regressions. Once this inclusion becomes feasible, it would lead to more accurate estimations of the structural parameters. We leave this for the attention of future work.

[5] We experimented with alternative weighting schemes such as the binary contiguity matrix and got broadly similar results to those reported in this paper.

coefficient on the initial income term, $\hat{\gamma}$, is likely to be inconsistent and biased upwards due to the positive correlation between $\ln y_{i,t-1}$ and η_i (Hsiao, 2003). The fixed effects estimator, although the within groups transformation wipes out the time-invariant province-specific effects (η_i), produces the opposite, a downward bias with the extent of attenuation increasing when exogenous covariates are added (Nickell, 1981). Bond et al. (2001) and Caselli et al. (1996) suggest a bound for $\hat{\gamma}$: the observed biases in the OLS and within group estimators are used as references to define upper and lower bounds for this serial autoregressive parameter.

The inclusion of a spatially lagged dependent variable $(\mathbf{W}\ln\mathbf{y})_{it}$ on the right-hand side of Eq. (18) further causes a simultaneity problem and renders both OLS and fixed effects estimators inconsistent. Although this simultaneity problem can be solved by employing ML estimator established in spatial econometrics, the existing spatial ML estimators are not designed to solve the endogeneity problem caused by the inclusion of a serially lagged dependent variable $\ln y_{i,t-1}$ (Abreu et al., 2005a; Elhorst, 2003). An exception is Elhorst (2005) who uses a first-differenced model to eliminate fixed effects and then derives an unconditional likelihood function. However, his ML method does not allow for instrumental treatment to control the potential endogeneity of other explanatory variables than the serially and spatially lagged dependent variable, and thus is not applicable to our case because we must consider FDI intensity, d_{it}, as endogenous.[6]

Badinger et al. (2004) propose a two-step procedure to take care of spatial interdependence in the estimation of a dynamic panel-data model. The variables are first filtered to remove spatial autocorrelation and then a standard GMM estimator (Arellano and Bover, 1995; Blundell and Bond, 1998) is applied to the filtered data. Although this procedure avoids addressing the joint problem of serial and spatial endogeneity in one equation, it is incapable of making explicit inference for spatially lagged variables and removes some of the variation that could potentially explain differences in growth rates. In addition, the two-step procedure implies that the coefficients of the spatial filtering and the other coefficients in the model are not determined jointly but sequentially, and the properties of such a sequential estimator are not known (Abreu et al., 2005a).

[6] Kukenova and Monteiro (2008) provide a survey on ML, QML, S2SLS, and partial GMM, and point to the same limitation: a lack of instrumental treatment to other potential endogenous variables.

In comparison, a direct application of the system GMM to Eq. (18) appears to be the best estimator available as it deals with the joint problem of serial and spatial endogeneity and corrects for the potential endogeneity of other explanatory variables. The basic idea of the system GMM is to estimate Eq. (18) as a system of two equations. One is in first difference, which allows the removal of the fixed effects, and the other is in levels, which brings in the technical gains of additional level moment conditions and increased efficiency. Lagged first differences and lagged levels are used as instruments for equations in levels and for equations in first differences, respectively. The use of instrumental variables allows consistent estimation of parameters even in the presence of measurement error and endogenous right-hand-side variables. On a practical ground, the system GMM also avoids the inversion of high dimension spatial weights matrix **W** and the computation of its eigenvalues as in the case of ML and QML, which involves accuracy problems when **W** is large. Furthermore, the Monte Carlo investigation in Kukenova and Monteiro (2008) also recommends the application of system GMM to the joint problem of serial and spatial endogeneity.

Considering that the consistency of the system GMM estimator depends on whether a selected set of lagged level and first-differenced values of the explanatory variables are valid instruments in the regression, three sets of specification tests are employed. First, the overall validity of the instruments is tested by the standard Hansen's *J* test of over-identifying restrictions, which analyses the sample analogue of the moment conditions used in the estimation process. Second, following the recommendations in Roodman (2009), Difference-in-Hansen tests for the full set of instruments for the levels equation as well as for the subset based on the dependent variable are conducted. The number of instruments generated for the regressions are reported. Third, because significant second-order serial correlation of the first-differenced residuals indicates serial correlation in the original error terms and therefore misspecification of the instruments, we also test for first-order and second-order serial correlation in the first-differenced residuals. If the original error terms are not serially correlated, there should be evidence of a significant negative first-order serial correlation in differenced residuals and no evidence of second-order serial correlation in the first-differenced residual. In addition to the validity tests, a finite-sample correction to the two-step covariance matrix as suggested in Windmeijer (2005) is implemented.

(See *Tables 2-4*)

4. Empirical Results

Eq. (18) is estimated by the OLS, fixed effects, spatial ML and system GMM estimator, respectively. The estimation results are presented in Table 2. Let us first look at the coefficients on d_{it} and $d_{i,\,t-1}$, the key concern of this research. All four estimations produce significantly positive coefficients on d_{it} , ranging between 0.814 and 1.117. In addition, coefficients on $d_{i,\,t-1}$ are significantly negative in three of the four estimations and particularly in the system GMM estimation. This consistence on coefficient estimation for d_{it} across four estimators is supervising given the absence of control for the endogeneity of FDI intensity in OLS, fixed effects, and spatial ML estimations (cf. Section 3). It suggests that the empirical value of ξ in Eq. (2) is about unity. Following Eq. (4), the multiplier from the marginal increment of FDI intensity, $\dot{d}_{it}$, to the percent growth rate of TFP is $\hat{\xi}/(1-\hat{\rho})$. Let us take $\hat{\rho}$ from the system GMM estimation because it addresses the joint endogeneity of serially and spatially lagged income and performs better than other three estimators, this gives $\hat{\xi}/(1-\hat{\rho})$ = 1.698, meaning that one percentage point increase in FDI intensity, say an increase from the mean value of 6.1 percent to 7.1 percent (cf. Table 1), would potentially lead to about 1.7 percent growth in the level of TFP according to Eq. (4) and in the level of per capita income according to Eq. (17).

In terms of growth empirics, each estimator produces qualitatively similar results as its counterpart in Sun et al (2009, Table 2) does. However, the introduction of FDI intensity in the TFP specification reduces the size of the coefficient on serially lagged income consistently in corresponding to each of the four estimators and hence leads to a higher convergence rate. With OLS the $\hat{\gamma}$ estimation, although upward biased, decreases from 0.968 to 0.936, which implies an increase in annual convergence rate, λ, from 0.65% to 1.3%. With the fixed effects $\hat{\gamma}$ estimation, which is known to be downward biased, falls from 0.669 to 0.624, suggesting an increase of annual convergence rate from 8.0% to 9.4%. With the better behaved system GMM $\hat{\gamma}$ decreases from 0.902 to 0.862, indicating an increase of annual convergence rate from 2.1% to 3.0%. With regard to the spatially lagged per capita income, the coefficient estimates for it are significantly positive for the estimators of OLS, spatial ML, and system GMM, respectively, and this spatial autoregressive parameter lies between 0 and 1, as assumed in Eq. (2). These results are consistent with those in Sun et al. (2009) and confirm the importance of including spatially lagged income in growth convergence theory and empirics.

Let us now to have a close look at the results of system GMM estimator, which addresses the joint endogeneity of serially and spatially lagged income, the endogeneity of FDI intensity and possible endogeneity of other explanatory variables, and possible measurement errors, and passes the specification tests of Hansen's *J*, Difference-in-Hansen, AR(1) and AR(2). All coefficients in the systems GMM estimation are statistically significant and have the expected sign. The significantly positive $\hat{\xi}$ and $\hat{\rho}$ coefficients produced under the full control of various endogeneity provide strong evidence in favour of our theoretical specification, which introduce FDI spillovers and further general interregional technology spillovers into the TFP specification of a standard Solow model. Empirically, significantly positive $\hat{\xi}$ and $\hat{\rho}$ indicate that the overall effect of FDI spillovers, as well as the general interregional technology spillovers, on TFP and income growth are positive and highly significant. The significantly positive $\hat{\rho}$ also provides empirical support to the strong presence of interregional technological diffusion and factor mobility across provinces and suggests that this presence has generated a complementary effect between a given province and its neighbours although the effect decays in distance. In addition, both the linear and nonlinear components of structural change make significantly positive contribution to the growth of TFP and further per capita income. The results also indicate that the implied convergence rate is about 3% per annum and capital investment exerts significantly positive contribution to the growth of per capita income. On the other hand, population growth plays a significantly negative role as expected.

Table 3 reports that the imposition of the linear parameter restriction of $\beta_1 = -\beta_2$ on all above regressions generates qualitatively unchanged and quantitatively similar results to those presented in Table 2. To test the possible spurious significance of the coefficient on the spatial autoregressive term and on FDI intensity, we estimate Eq. (19) and the results from the unrestricted regressions are presented in Table 4.[7] The results indicate that the incorporation of $\mathbf{W}\ln\mathbf{x}_t$ into the regressions does not crowd out the significance of either $\mathbf{W}\ln\mathbf{y}_t$, d_{it} or $d_{i,t-1}$ in the ML and system GMM regressions.

[7] The corresponding results for the restricted regressions are qualitatively unchanged and quantitatively similar. These and other results on robustness tests are available upon request.

Moreover, with the well-behaved system GMM estimator, the coefficients of both $\mathbf{W}\cdot \ln s_{it}$ and $\mathbf{W}\cdot \ln(n_{it} + g + \delta)$ become statistically insignificant. These findings suggest that the significance of the coefficient on either the spatial autoregressive variable or FDI intensity variable is very unlikely to be spurious.

Our findings are consistent with the emerging empirical evidence at the firm level in China. Wei and Liu (2006) assess the productivity enhancing effects of technology spillovers across firms from R&D, exports and the very presence of FDI in China's manufacturing sector. Their estimations based on a panel of more than 10,000 indigenous and foreign-invested firms for 1998–2001 indicate strong productivity enhancing effects of inter-industrial spillovers from R&D and exports, and of both intra- and inter-industrial spillovers from foreign presence to indigenous Chinese firms. While their study is confined to the impact of knowledge spillovers on the productivity of indigenous Chinese firms only and does not pay attention to the geographical scale of proximity, Wei et al. (2008), using the same panel dataset, show that mutual productivity spillovers between foreign and local firms are significantly positive and both national and regional in scale. Hale and Long (2006) use the World Bank's 2001 survey of 1500 firms in five Chinese cities to study whether the presence of foreign firms produces technology spillovers on domestic firms operating in the same city and industry. They find that the transfer of technology occurs through movement of high-skilled workers from FDI firms to domestic firms as well as through network externalities among high-skilled workers. Brambilla et al. (2009) propose a new channel of FDI spillovers on domestic firms, which operates through imitation of original products. Using firm-level panel data for China from the World Bank's 2001 and 2003 Investment Climate Surveys they find that the presence of foreign firms generates incentives for imitation because they introduce original products that are vertically differentiated from domestic products.

In comparison with Madariaga and Poncet (2007), our results confirm qualitatively the presence of strong positive relationship between FDI and income growth as they have put forward. In addition, our approach corrects the upward "aggregation bias" presented in their setting. It is worth noting that our results are not in line with the literature on FDI-income growth "paradox" in the context of China as reviewed in Laurenceson and Tang (2007), which indicates that the actual evidence for positive spillovers from FDI is surprisingly weak. However, the latter literature heavily depends on growth accounting approaches, which require the imposition of many strong assumptions. For example, a key assumption in growth-accounting exercises is that factor prices coincide with social marginal products. If this assumption is violated, the estimated value of TFP calculated from standard Solow's formulae deviates from the true contribution of TFP to economic growth (Barro and Sala-i-Martin, 2004). The calculation of TFP typically pre-excludes technological externalities within and across regions. The accounting decomposition can easily attribute to capital accumulation something that should be attributed to technological progress. The key criticism is that an accounting relationship is not the same thing as a causal relationship. "Even though there is evidence that somewhere between 30 and 70 per cent of the growth of output per worker in OECD countries can be 'accounted for' by capital accumulation," ··· "it might still be that all of the growth is caused by technological progress" (Aghion and Howitt, 2007). In sharp contrast, our approach is largely free from these limitations and thus capable of producing more convincing results.

5 Concluding Remarks

In this paper, we examine the effect of foreign direct investment (FDI) on regional economic performance of China from a spatial economic perspective. We disentangle the capital accumulation effects of FDI from the total-factor-productivity (TFP) effects of externalities generated by FDI. By augmenting the specification of the neoclassical growth model in Yu (2007) and Sun at al. (2009), we explicitly incorporate the effects of intraregional spillover of FDI, inter-provincial technological spillover, factor mobility, and structural change into the specification of the TFP function. The process of mathematical derivation naturally leads to a spatial dynamic model capable of providing comprehensive and consistent control for influential variables and specification errors which may alter the relationship between FDI intensity and TFP.

We apply this augmented spatial dynamic model to the regression of cross-province growth performance in China over the period 1980-2005. The estimations have to address the joint endogeneity problem of the serially and spatially lagged dependent variable, as indicated in the theoretical model, and further the endogeneity of the FDI intensity variable, the key concern of this research. We employ system GMM estimator to deal with this join endogeneity problem. For robustness, we also make use of alternative estimators and model structures. The estimations show that the impact of FDI externalities on TFP is significant and positive. Numerically speaking, one percentage point increase in FDI intensity, say, from the mean value of 6.1 percent to 7.1 percent, would lead to about 1.7

percent growth in the level of TFP and per capita income. In comparison with the results in Sun et al. (2009), we show that the incorporation of FDI externalities, as an addition to structural change and spatial interdependence, provides an improved account of interprovincial variations in income levels and economic growth. The spatial interdependence is strong mainly due to interregional technological spillover and factor mobility. The speed of convergence increases from about 2.1% to about 3%. These results indicate that capital investments, FDI intensity in capital investment, changes in the structure of employment, force of conditional convergence, and population growth are the main sources of the income and growth difference across Chinese provinces.

Our results are consistent with the emerging empirical evidence based on panel data at the firm level in China (e.g., Hale and Long, 2006; Wei and Liu, 2006; Wei et al., 2008; Brambilla et al., 2009) but not in line with the growth accounting based literature as reviewed in Laurenceson and Tang (2007), which suggests a FDI-income growth "paradox" in the context of China. This contrast implies that a theory-rooted research design is crucial for a proper analysis of productivity spillovers generated by FDI. In terms of policy implication, our findings indicate that FDI can make important contributions to economic growth, more considerably by way of productivity enhancement than capital accumulation. Local conditions matter and may limit the extent to which the benefits of FDI materialize. Local FDI promotion should pay great attention to those sectors where the payoffs from technological spillovers are likely to be highly significant. Policy efforts should encourage foreign capital that is complementary to local capital stock because this type of FDI is more likely to enhance local productivity and competitiveness and less likely to simply crowd out local investment. Moreover, policy and regulatory efforts should promote technology diffusion and movement of labour and capital across regions and industries because such diffusion and mobility will not only enhance TFP and income growth of individual regions but also the convergence forces across regions in the long run, leading to reduction in regional income disparity.

Reference

Abreu, Maria, Henri L F de Groot, and Raymond J G M Florax. 2005a. “Space and Growth: A Survey of Empirical Evidence and Methods,” *Région et Développement*, 21, 13-44.

Abreu, Maria, Henri L F de Groot, and Raymond J G M Florax. 2005b. “A Meta-analysis of *β*-convergence: The Legendary 2%,” *Journal of Economic Surveys*, 19 (3), 389-420.

Adams, James D. and Adam B. Jaffe. 1996. “Bounding the Effects of R&D: An Investigation Using Matched Establishment-firm Data,” *RAND Journal of Economics*, 27(4), 700-721.

Aghion, Philippe and Peter Howitt. 2007. “Capital, innovation, and growth accounting.” *Oxford Review of Economic Policy*, 23 (1), 79–93.

Aitken, B. and Ann E. Harrison. 1999. “Do Domestic Firms Benefit from Direct Foreign Investment? Evidence from Venezuela.” *American Economic Review*, 89, 3, 605-617.

Arellano, Manuel and Olympia Bover. 1995. “Another Look at Instrumental Variable Estimation of Error-Component Models,” *Journal of Econometrics*, 68 (1), 29-52.

Badinger, Harald, Werner G Muller, and Gabriele Tondl. 2004. “Regional Convergence in the European Union, 1985-1999: A Spatial Dynamic Panel Analysis,” *Regional Studies*, 38(3), 241-253.

Barro, Robert J. and Xavier Sala-i-Martin. 2004. *Economic Growth*, second edition. Cambridge, MA: The MIT Press.

Barro, Robert J. and Xavier Sala-i-Martin. 1992. “Convergence,” *Journal of Political Economy*, 100 (2), 223-251.

Blundell, Richard and Stephen Bond R. 1998. “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics*, 87(1), 115-143.

Bond, Stephen R., Anke Hoeffler, and Jonathan Temple. 2001. “GMM Estimation of Empirical Growth Models,” CEPR Discussion Paper No. 3048, London: Centre for Economic Policy Research.

Bottazzi, Laura and Giovanni Peri. 2003. “Innovation and Spillovers in Regions: Evidence from European Patent Data,” *European Economic Review*, 47(4), 687-710.

Brambilla, Irene, Galina Hale and Cheryl Long. 2009. “Foreign Direct Investment and the Incentives to Innovate and Imitate.” *Scandinavian Journal of Economics*, 111 (4), 835-861.

Caselli, Francesco, Gerardo Esquivel and Fernando Lefort. 1996. “Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics,” *Journal of Economic Growth*, 1(3), 363-389.

Ciccone, A., and R. E. Hall, 1996. “Productivity and the Density of Economic Activity.” *American Economic Review* 86, 1, pp. 54-70.

Denison, Edward F. 1967. *Why Growth Rates Differ: Postwar Experiences in Nine Western Countries*. Washington, D.C.: Brookings Institution.

Elhorst, J. Paul. 2003. "Specification and Estimation of Spatial Panel Data Models," *International Regional Science Review*, 26(3), 244-268.

Elhorst, J. Paul. 2005. "Unconditional Maximum Likelihood Estimation of Dynamic Models for Spatial Models," *Geographical Analysis*, 37(1), 85-106.

Fagerberg, J. 1994. "Technology and International Differences in Growth Rates." *Journal of Economic Literature*, 32, 3, pp. 1147-1175.

Gilboy, G. F., 2004. The myth behind China's miracle. Foreign Affairs 83 (4), 33-48.

Görg, Holger and David Greenaway, 2004. "Much Ado about Nothing? Do Domestic Firms Really Benefit from Foreign Direct Investment?" *World Bank Research Observer* 19, 171-197.

Hale, G., Long, C., 2006. "Are there productivity spillovers from foreign direct investment in China?" Working Paper Series 2006-13, Federal Reserve Bank of San Francisco.

Hanousek, Jan, Evzen Kocenda and Mathilde Maurel, 2010. "Direct and Indirect Effects of FDI in Emerging European Markets: A Survey and Meta-analysis." William Davidson Institute Working Paper Number 976, University of Michigan.

Harris, Richard, 2009. "Spillover and Backward Linkage Effects of FDI: Empirical Evidence for the UK," Spatial Economics Research Centre (SERC) Discussion Paper 0016, University of Glasgow.

Hsiao, Cheng. 2003. *Analysis of Panel Data*. 2nd edition. Cambridge: Cambridge University Press.

Islam, Nazrul. 1995. "Growth Empirics: A Panel Data Approach," *Quarterly Journal of Economics*, 110(4), 1127-1170.

Keller, Wolfgang. 2002. "Geographic Localization of International Technology Diffusion," *American Economic Review*, 92(1), 120-142.

Koo, Jun. 2005. "Technology Spillovers, Agglomeration, and Regional Economic Development," *Journal of Planning Literature*, 20 (2), 99-115.

Kukenova, Madina and Jose-Antonio Monteiro. 2008. "Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation." MPRA Paper No. 13405. Online at http://mpra.ub.uni-muenchen.de/13405/.

Kuznets, Simon. 1961. "Economic Growth and the Contribution of Agriculture: Notes on Measurement," *International Journal of Agrarian Affairs*, 3 (1), 56-75.

Laurenceson, James and Kam Tang, 2007. "The FDI-Income Growth Nexus: A Review of the Chinese Experience." East Asia Economic Research Group Discussion Paper No. 9, School of Economics, University of Queensland.

Madariaga, Nicole and Sandra Poncet, 2007. "FDI in Chinese Cities: Spillovers and Impact on Growth." *The World Economy*, 30, pp. 837-862.

Nickell, Stephen J. 1981. "Biases in Dynamic Models with Fixed Effects," *Econometrica*, 49(6), 1417-1426.

Paas, Tiiu and Friso Schlitte. 2008. "Regional Income Inequality and Convergence Processes in the EU-25," *Italian Journal of Regional Science*, 7(2), 29-49.

Roodman, David. 2009. "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.

State Statistical Bureau (SSB). 2000-2009. *China Statistical Yearbook.* Beijing: China Statistical Publishing House.

State Statistical Bureau (SSB). 1999. *Statistical Data on 50 Years of New China, 1949-1998*. Beijing: China Statistical Publishing House.

Sun, Laixiang, Eunsuk Hong and Tao Li. 2009. "Incorporating technology diffusion, factor mobility and structural change into cross-region growth regression: An application to China." *Journal of Regional Science*, published online 5 August 2009.

Temple, J. and L. Wöβmann. 2006. "Dualism and Cross-Country Growth Regressions." *Journal of Economic Growth*, 11, 187–228.

Wei, Yingqi and Xiaming Liu. 2006. "Productivity Spillovers from R&D, Exports and FDI in China's Manufacturing Sector," *Journal of International Business Studies*, 37 (4), 544-557.

Wei, Yingqi, Xiaming Liu and Chengang Wang. 2008. "Mutual Productivity Spillovers between Foreign and Local Firms in China," *Cambridge Journal of Economics*, 32 (4), 609-631.

Windmeijer, Frank. 2005. "A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators," *Journal of Econometrics*, 126(1), 25-51.

Yu, Jihai. 2007. "Essays on Spatial Dynamic Panel Data Model: Theories and Applications," Ph.D. dissertation, Department of Economcis, The Ohio State University.

TABLE 1. Definition of Variables and Summary Statistics

Variable	Label	Definition	Obs.	Mean	Std. Dev.	Min	Max
Income level	$\ln y_{it}$	Logarithm of real GDP per capita in 1985, 1990, 1995, 2000, and 2005.	145	7.393	0.829	5.900	9.874
Initial income level	$\ln y_{i,t-1}$	Logarithm of real GDP per capita in 1980, 1985, 1990, 1995, and 2000.	145	6.956	0.804	5.380	9.625
Saving rate	$\ln s_{it}$	Logarithm of the average share of fixed asset investment in real GDP. The 6-year average is over 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05, respectively.	145	-1.136	0.281	-1.860	-0.322
Population growth rate	$\ln(n_{it} + \overline{g}/(1-\overline{\rho})+\overline{\delta})$	Logarithm of the sum of average population growth rate (*n*), exogenous technology progress rate (*g*) and capital depreciation rate (δ), where $\overline{g}/(1-\overline{\rho})+\overline{\delta}$ = 0.07. The 6-year average of population growth is over 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05, respectively.	145	-2.498	0.097	-2.791	-2.054
Structural change (linear part)	*MGROWTH*	The change in the share of non-agricultural population in total population over each of the five-year intervals of 1980-85, 1985-90, 1990-95, 1995-2000, and 2000-05.	145	0.049	0.062	-0.016	0.339
Structural change (non-linear part)	*DISEQ*	$DISEQ = \Delta MGROTH \cdot p\,/\,(1 - p)$, where *p* is the migration propensity, defined by $p = -\Delta a\,/a$, where *a* is the share of agricultural population in total population.	145	0.035	0.200	-0.588	1.797
FDI intensity	d_{it}	Share of utilized FDI in total fixed asset investment in 1985, 1990, 1995, 2000, and 2005.	145	0.061	0.083	0	0.493
Initial FDI intensity	$d_{i,t-1}$	Share of utilized FDI in total fixed asset investment in 1980, 1985, 1990, 1995, and 2000.	145	0.049	0.084	0	0.493
Income level of neighbouring provinces	$(W \cdot \ln \mathbf{y})_{it}$	Spatially lagged dependent variable	145	7.420	0.642	6.448	8.582

Note: The inconsistence of data on employment by sector and province is well-known and the use of them, in fact, generates many negative values for *MGROWTH* and *DISEQ*, which are in contradiction to observed structural changes. In comparison, data on agricultural versus non-agricultural population at the provincial level, although under-representing the extent of rural non-farming activities, are regarded as basically consistent and the measurements of *MGROWTH* and *DISEQ* based on these data are largely in line with the observed structural changes.

TABLE 2. Estimation Results (unrestricted regression, number of observation = 145)

	OLS	Within Groups	Spatial ML	System GMM
$\ln y_{i,t-1}$	0.936***	0.624***	0.494***	0.862***
	[0.019]	[0.065]	[0.051]	[0.041]
$\ln s_{it}$	-0.010	0.037***	-0.030	0.060*
	[0.039]	[0.059]	[0.052]	[0.031]
$\ln(n_{it} + g + \delta)$	-0.453***	-0.429***	-0.496***	-0.395***
	[0.101]	[0.083]	[0.074]	[0.063]
$t \cdot MGROWTH_{it}$	0.087**	0.096*	0.165***	0.077*
	[0.041]	[0.057]	[0.031]	[0.039]
$t \cdot DISEQ_{it}$	0.009	0.004	-0.002	0.009***
	[0.010]	[0.006]	[0.006]	[0.003]
d_{it}	1.103***	0.814***	0.992***	1.117***
	[0.128]	[0.150]	[0.108]	[0.176]
$d_{i,t-1}$	-0.361**	0.152	0.305***	-0.233**
	[0.140]	[0.152]	[0.119]	[0.114]
$W \cdot \ln \mathbf{y}_{it}$	0.051*	0.173	0.400***	0.411*
	[0.028]	[0.292]	[0.046]	[0.225]
Constant	-0.707***	0.782***		-2.930
	[0.253]	[2.438]		[1.769]
R^2	0.987	0.992	0.995	
Number of Instruments				22
Hansen J test (*p*-value)				(0.887)
Difference-Hansen tests (*p*-value)				
All system GMM instrument				(0.910)
Those based on lagged income only				(0.829)
AR(1) test in differences (*p*-value)				(0.003)
AR(2) test in differences (*p*-value)				(0.630)
Implied λ	0.013	0.094	0.141	0.030

Notes. Numbers in [ ] and ( ) are standardized errors and *p*-values respectively. *, ** and *** denotes significance at the 10%, 5% and 1% level respectively. Province dummy in Spatial ML and Period dummy in the Within Groups and System GMM are not reported.

TABLE 3. Estimation Results (restricted regression, number of observation = 145)

	OLS	Within Groups	Spatial ML	System GMM
$\ln y_{i,t-1}$	0.917*** [0.019]	0.617*** [0.068]	0.449*** [0.056]	0.856*** [0.057]
$\ln s_{it} - \ln(n_{it} + g + \delta)$	0.129 [0.042]	0.182*** [0.044]	0.160*** [0.040]	0.106** [0.043]
$t \cdot MGROWTH_{it}$	0.097** [0.043]	0.073 [0.060]	0.141*** [0.033]	0.044 [0.080]
$t \cdot DISEQ_{it}$	-0.001 [0.010]	0.002 [0.008]	-0.002 [0.007]	0.007 [0.011]
d_{it}	1.078*** [0.137]	0.681*** [0.153]	0.808*** [0.112]	0935*** [0.216]
$d_{i,t-1}$	-0.462*** [0.150]	0.075 [0.158]	0.236* [0.131]	-0.282** [0.127]
$W \cdot \ln \mathbf{y}_{it}$	0.090*** [0.029]	0.017 [0.303]	0.423*** [0.051]	0.708*** [0.190]
Constant	0.267* [0.158]	2.936 [2.478]		-4.548*** [1.488]
R^2	0.984	0.991	0.994	
Number of Instruments				22
Hansen J test (*p*-value)				(0.454)
Difference-Hansen tests (*p*-value)				
All system GMM instrument				(0.172)
Those based on lagged income only				(0.589)
AR(1) test in differences (*p*-value)				(0.006)
AR(2) test in differences (*p*-value)				(0.982)
Implied λ	0.017	0.097	0.160	0.031

Notes. The same as in Table 2.

TABLE 4. Robustness of the Results on the Spatial Autoregressive Term (unrestricted regressions, number of observation = 145)

	OLS	Within Groups	Spatial ML	System GMM
$\ln y_{i,t-1}$	0.937***	0.659***	0.560***	0.869***
	[0.022]	[0.065]	[0.053]	[0.044]
$\ln s_{it}$	0.105**	0.205***	-0.021	0.054*
	[0.050]	[0.061]	[0.051]	[0.029]
$\ln(n_{it} + g + \delta)$	-0.366***	-0.316***	-0.478***	-0.383***
	[0.113]	[0.090]	[0.073]	[0.076]
$t \cdot MGROWTH_{it}$	0.086*	-0.037*	0.155***	0.050
	[0.050]	[0.065]	[0.032]	[0.058]
$t \cdot DISEQ_{it}$	0.004	0.007	0.002	0.009**
	[0.012]	[0.008]	[0.007]	[0.004]
d_{it}	0.007	-0.004	1.033***	1.058***
	[0.004]	[0.003]	[0.108]	[0.197]
$d_{i,t-1}$	0.000	-0.000	0.190	-0.219*
	[0.003]	[0.003]	[0.124]	[0.110]
$W \cdot \ln s_{it}$	-0.934***	0.811**	-0.274	-0.042
	[0.217]	[0.406]	[0.232]	[0.287]
$W \cdot \ln(n_{it} + g + \delta)$	-0.324	2.085***	-0.546*	0.956
	[0.529]	[0.687]	[0.307]	[0.873]
$W \cdot \ln \mathbf{y}_{it}$	0.272***	0.184	0.398***	0.451*
	[0.080]	[0.367]	[0.077]	[0.237]
Constant	-3.832***	7.048*		-0.868
	[1.107]	[3.782]		[1.980]
R^2	0.983	0.991	0.995	
Number of Instruments				24
Hansen J test (p-value)				(0.927)
Difference-Hansen tests (p-value)				
All system GMM instrument				(0.878)
Those based on lagged income only				(0.882)
AR(1) test in differences (p-value)				(0.002)
AR(2) test in differences (p-value)				(0.348)
Implied λ	0.013	0.083	0.116	0.028

Notes. The same as in Table 2.