G D P

全球发展展望

Global Development Perspective

中国社会科学院世界经济与政治研究所 经济发展研究中心

Working Paper No. 21.003

May 25th, 2021

Lei Hou，侯蕾

houlei@cass.org.cn

Wages, labor quality, and FDI inflows: a new non-linear approach[1]

Lei Hou, Qi Li, Yanfei Wang, Xintong Yang

Abstract Although theory predicts that locations with lower wages will attract more FDI inflows, the empirical results from traditional linear regression models are rather mixed. We consider the possible impact of labor quality, along with wages, and build a simple FDI location choice model to capture the potential nonlinear relationship among the three. Further, we propose a partially linear panel data model to investigate the wage effect on FDI location choices. Using macroeconomic province-level data from China between 1993 and 2018, we find that the marginal effect of wage is generally a decreasing function of labor quality. This implies that facing low labor quality, FDI firms prefer locations with high wages that ensure high labor quality; while facing high labor quality, their preferences for low labor costs dominate. Our nonlinear approach reconciles the mixed results in the empirical literature on how wages affect FDI inflows and the role of labor quality in attracting FDI.

Key words FDI, wage, labor quality, partially linear regression

JEL classification F21, J31

[1] Accepted for publication in *Economic Modelling*.

1 Introduction

Traditional FDI theory proposes that firms expand production abroad based on two motives: to pursue low labor costs and to access high market demand, that is, vertical and horizontal FDI, respectively (Yeaple, 2003; Zhang and Markusen, 1999). A positive effect of market demand on FDI inflows has been well-established in the empirical studies. However, the evidence on low labor costs is ambiguous.[1] The effect of "wage", a widely used measurement of labor costs in general, could be either significantly negative or positive, and even insignificant.[2] One explanation for the positive marginal effect of wage on inward FDI is that a high wage indicates high labor quality (Wei et al., 1999; UNCTAD, 1999). Because FDI firms usually enjoy a technology advantage, they may prefer a location with high-quality labor ready to adapt their technology. This research is mainly concerned with how to evaluate the wage effect on FDI inflows while reconciling the role of labor quality.

We first develop a simple model of FDI location choice in which wage is not only a measurement of local labor cost but also an important indicator of local labor quality. Local labor quality not only determines the equilibrium wage, it also influences output capacity. We show that with some regularity assumptions, there exists an optimal labor quality level such that when the labor quality falls below this level, increasing wages in a province will increase the FDI probability in that province. When the labor quality is higher than this level, increasing wages in the province will decrease the FDI probability. The intuition is that when labor quality is low, the wage can be regarded as an effective indicator of labor quality; therefore high wages attract FDI inflows. When labor quality is high enough, the labor cost effect dominates, and a high wage will impede FDI inflows. Given the roles of wage as labor cost and quality, the average marginal effect of wage on FDI inflows may be ambiguous.

We further formalize this notion by estimating the effect of wage and labor quality on FDI inflows.

[1] Nielson et al. (2017) reviewed 83 studies on FDI location choices that include wage variables. They find a negative and positive relationship between wages and FDI location choices, accounting for 49% and 17% respectively. Meanwhile, 34% of these studies find no support for a negative association as theory predicts.

[2] See Noorbakhsh and Paloni (2001), Head and Mayer (2004), among others. See Liu, Lovely and Ondrich (2010) for a review of wage and firm location choices.

In particular, we focus on the 26-year period from 1993 to 2018 across 29 Chinese provinces. The development experience of China from the early 1990s to the 2010s provides a good opportunity for examining this issue. China mainly attracted vertical FDI, especially around the 1990s, because of its large labor force, relatively low labor costs, and local governments' favorable policies toward investment. During that period, many foreign firms in developed countries either shifted their domestic and overseas production factories to China or established new factories in China, then exported their manufactures back to the developed countries. This promoted China's processing trade development and explains the simultaneous high rate of growth in its import and export levels. At the same time, China witnessed a large increase in labor quality due to the enhanced education level of its labor force, which intensifies China's advantage as an ideal FDI destination. However, China's labor cost also rose dramatically. From 2000 to 2012, China's average real wage quadrupled. As a consequence, in recent years many multinational corporations chose to shift their production lines to other Asian and Latin American developing countries with lower labor costs, or simply to close their Chinese factories. Therefore, both labor costs and labor quality may have played a role in determining China's FDI inflows over these two decades.

We first estimate a linear panel data model with province fixed effects, but this fails to capture either the effect of wages or the effect of labor quality on FDI inflows. From a geographic perspective, there are huge regional disparities between coastal and inland areas in China, in economic development levels, factor endowments, preferential policies, and the ability to attract FDI.[3] Thus, the effect of local wages and labor quality on FDI inflows may differ across regions. With our linear panel data model, we therefore perform subsample analyses across three different time periods and three geographic regions in China. These results suggest that the effects of wages and labor quality on FDI inflows are heterogeneous, both temporally and spatially. Thus, a linear regression model may not be suitable for estimating the effect of wages on FDI location choices.

In order to capture the nonlinear effects alluded above, we propose a partially linear specification to model the interdependent effect of wages and labor quality on FDI. We find that, in general, the

[3]For regional disparity in attracting FDI, see Gao (2005) and Amiti and Javorcik (2008), among others.

estimated marginal effect of wages is a decreasing function of labor quality. When labor quality is relatively low, wages have a positive effect on FDI. This implies that FDI firms prefer locations with high wages as they pursue high labor quality. As labor quality increases, the negative cost effect of the high wage on FDI kicks in. When labor quality exceeds a threshold value, the wage effect on FDI becomes significantly negative; this negative effect is intensified as labor quality increases. Moreover, we find that the marginal effect of labor quality on FDI inflows depends on the wage level. When the wage level is relatively low, labor quality will significantly attract more FDI. When the wage increases to a certain level, high labor quality will attract less FDI or even deter FDI inflows.

Recent empirical studies have investigated various novel location-specific attributes, other than the market size and labor costs, in affecting FDI, including logistic infrastructure, economic institutions, immigration, historical conflicts, local labor market flexibility, etc. (Blyde and Molina, 2015; Ascani et al., 2016; Tomohara, 2017; Gao et al., 2018; Rong et al., 2020). Given the indefinite role of labor costs in attracting FDI in existing empirical studies, our paper revisits the questions of how labor costs affect FDI and what role the local labor quality may play. From this perspective, our paper is closely related to the studies estimating the wage effect on FDI as Liu, Lovely and Ondrich (2010), and the effect of local labor quality on FDI as Gao (2005) and Iwai and Thompson (2012). [4]

Compared to the existing literature, this paper makes two contributions. First, it uses a nonlinear function to capture the potential influence of labor quality in estimating the effect of wages on FDI. In prior work, labor quality entered FDI location choices merely as a linearly additive explanatory variable, as in Gao (2005). In our study, apart from its direct impact on FDI, labor quality has an effect through wages, thus indirectly affecting FDI location choices. Our partially linear

[4] Liu, Lovely and Ondrich (2010) points out that unobserved location-specific attributes play a role in estimating the wage effect on FDI location choices. Our paper confirms their conjecture by introducing labor quality into the estimation of wage effects on FDI. Gao (2005) investigates the impact of labor quality on FDI inflows in China, and confirms labor quality as an established determinant in attracting FDI. Our paper differs from Gao (2005) by using a partially linear model estimation. We find an indirect effect of labor quality on FDI through wages, in addition to a direct effect of labor quality on FDI through a linearly additive form as in Gao (2005). In addition, our results support the hypothesis proposed by Iwai and Thompson (2012) that there exists a take-off point in labor quality for low labor cost developing countries to turn from potential FDI destination candidates into real ones. Using the Chinese data, we empirically identify such a turning point in labor quality such that above this point, a province's low labor cost is an advantage in attracting FDI; below this point, the province's low labor cost may deter instead of attracting FDI.

estimation is able to identify the nonlinear wage effect on FDI concealed by linear regressions. This nonlinearity may explain the inconclusive results on the wage effect on FDI inflows in existing empirical studies. Second, we introduce a novel nonlinear methodology. Generally, a traditional parametric method adopted to estimate the average wage effect is more suitable when wages affect FDI inflows homogeneously across the entire FDI distribution. Our study shows that when dealing with large differentiation across regions and over time, as in China, the average wage effect estimation may mask variations in the wage effect at the aggregate level. Thus, the nonlinear estimation has an advantage in making clear the nonlinear wage effects in the data. From this perspective, the nonlinear approach that we propose is useful when data have large differentiation across space and over time, which may have more general applications in economic studies.

The rest of this paper is organized as follows. Section 2 presents a simple theory to show that the wage effect on FDI location choices depends on local labor quality. Section 3 describes the data and variable constructions. Section 4 presents the estimation results of the linear panel data model with province fixed effects. We demonstrate the existence of a nonlinear relationship among FDI inflows, wages and labor quality. Section 5 presents our proposed partially linear panel data model, taking the nonlinear relationship into account, and provides our empirical results. Section 6 concludes. A numerical example that supports the theoretical model, and details about the econometric method, are relegated to the Appendices.

2 Theoretical model

The theory builds upon Liu, Lovely and Ondrich (2010). A foreign firm chooses to locate where its profits are maximized, and it seeks to invest somewhere in China. Its production technology uses local labor inputs, a vector of goods and services.[5] Note that labor input is a function of labor quality, i.e., $L(z)$. Thus, by choosing labor quality in a province, the foreign firm indirectly chooses the labor force in that province. Controllable profits for the firm if it invests in province j can be

[5]The input vector of goods and services might contain imported goods, but this does not matter for our main focus.

written as:

$$\pi_j = Q(L(z_j), x_j, z_j) - w_j(z_j)L(z_j) - p_{xj}x_j, \tag{2.1}$$

where $L(z_j)$ denotes the local labor inputs in province j, x_j denotes the input of goods and services in province j, z_j denotes the local labor quality in province j, $Q(\cdot)$ denotes the production technology, $w_j(z_j)$ denotes the labor wage level in province j, p_{xj} denotes the price level of goods and service input in province j.

Taking F.O.C. of π_j in (2.1) with respect to labor inputs $L(z_j)$, we find that the optimal labor input and wage level need to satisfy:

$$w_j(z_j) = Q_L(L(z_j), x_j, z_j). \tag{2.2}$$

Further, based on equation (2.2), we have:

$$w_j'(z_j) = Q_{LL}(L(z_j), x_j, z_j)L'(z_j) + Q_{Lz}(L(z_j), x_j, z_j)$$

Taking the derivative of π_j in (2.1) with respect to labor quality z_j, we have:

$$\begin{aligned}\frac{\partial \pi_j}{\partial z_j} &= Q_L(L(z_j), x_j, z_j) \cdot L'(z_j) + Q_z(L(z_j), x_j, z_j) - w_j(z_j)L'(z_j) - w_j'(z_j)L(z_j) \\ &= (Q_L(L(z_j), x_j, z_j) - w_j(z_j)) \cdot L'(z_j) + (Q_z(L(z_j), x_j, z_j) - w_j'(z_j)L(z_j)) \\ &= Q_z(L(z_j), x_j, z_j) - w_j'(z_j)L(z_j)\end{aligned}$$

where the third line is derived from equation (2.2). Hence, there exists a level of labor quality $\underline{z}_j$ such that $\frac{\partial \pi_j}{\partial z_j} = 0$:

$$w_j'(\underline{z}_j)L(\underline{z}_j) = Q_z(L(\underline{z}_j), x_j, \underline{z}_j).$$

Lemma 1: Assume that $Q_{zz}(L(z), x, z) < Q_{LL}(L(z), x, z) \cdot (L'(z))^2 + w''(z)L(z)$ and $Q_{LL}(L(z), x, z)L'(z) + Q_{Lz}(L(z), x, z) > 0$. When $z_j < \underline{z}_j$, $\frac{\partial \pi_j}{\partial w_j} > 0$, i.e., increasing wage level in province j will increase profits in province j. When $z_j > \underline{z}_j$, $\frac{\partial \pi_j}{\partial w_j} < 0$, i.e., increasing wage level in province j will decrease profits in province j.

Proof: Whenever $z_j = \underline{z}_j$, $\frac{\partial \pi}{\partial z_j} = 0$. To see it achieves the local maximum, we take the second derivative of π_j with respect to z_j, and have:

$$\begin{aligned}\frac{\partial^2 \pi_j}{\partial z_j^2} &= Q_{zL}(L(z_j), x_j, z_j) \cdot L'(z_j) + Q_{zz}(L(z_j), x_j, z_j) - w_j'(z_j)L'(z_j) - w_j''(z_j)L(z_j) \\ &= Q_{zz}(L(z_j), x_j, z_j) - Q_{LL}(L(z_j), x_j, z_j) \cdot (L'(z_j))^2 - w_j''(z_j)L(z_j)\end{aligned}$$

Given $Q_{zz}(L(z), x, z) < Q_{LL}(L(z), x, z) \cdot (L'(z))^2 + w''(z)L(z)$, we have $\frac{\partial^2 \pi_j}{\partial z_j^2} < 0$. This implies that when $z_j = \underline{z}_j$, firm's profit achieves a maximum, i.e., when $z_j < \underline{z}_j$, $\frac{\partial \pi_j}{\partial z_j} > 0$, and when $z_j > \underline{z}_j$, $\frac{\partial \pi_j}{\partial z_j} < 0$. Given $Q_{LL}(L(z), x, z)L'(z) + Q_{Lz}(L(z), x, z) > 0$, we have $w'(z_j) > 0$. Hence, we have when $z_j < \underline{z}_j$, $\frac{\partial \pi_j}{\partial w_j} > 0$, and when $z_j > \underline{z}_j$, $\frac{\partial \pi_j}{\partial w_j} < 0$. **Q.E.D.**

We use production function $Q(L(z), x, z)$ and do not specify its functional form in order to make our theory more general and more applicable. To be more convincing, we provide a numerical example in the Appendix to show that assumptions $Q_{zz}(L(z), x, z) < Q_{LL}(L(z), x, z) \cdot (L'(z))^2 + w''(z)L(z)$ and $Q_{LL}(L(z), x, z)L'(z) + Q_{Lz}(L(z), x, z) > 0$ hold and how to apply our general theory.

Denote the real profits of the firm in province j as $\Pi_j = \pi_j - e_j$, where e_j is an idiosyncratic cost shock in province j. Province k is chosen if the firm's real profit from investing there is maximized, i.e.,

$$\begin{aligned}Prob_k &= Pr\{\max(\Pi_1, ..., \Pi_J) = \Pi_k\} \\ &= Pr\{\Pi_k \geq \Pi_1, \Pi_k \geq \Pi_2, ..., \Pi_k \geq \Pi_J\} \\ &= Pr\{\pi_k + \epsilon_k \geq \pi_1 + \epsilon_1, \pi_k + \epsilon_k \geq \pi_2 + \epsilon_2, ..., \pi_k + \epsilon_k \geq \pi_J + \epsilon_J\} \\ &= Pr\{\epsilon_1 \leq \pi_k + \epsilon_k - \pi_1, \epsilon_2 \leq \pi_k + \epsilon_k - \pi_2, ..., \epsilon_J \leq \pi_k + \epsilon_k - \pi_J\}\end{aligned}$$

Let J denotes the set of all provinces in China. If $\{\epsilon_j\}_{j \in J}$ are independently and identically distributed according to a Type 1 Extreme Value distribution with p.d.f. $f(\epsilon) = \exp\Big(-\epsilon - \exp(-\epsilon)\Big)$ and c.d.f. $F(\epsilon) = \exp\Big(-\exp(-\epsilon)\Big)$, we have:

$$Prob_k = \frac{\exp(\pi_k)}{\sum_{j\in J}\exp(\pi_j)}. \tag{2.3}$$

Equation (2.3) shows that the probability that province k is chosen positively depends on controllable profits π_k. Then, we reach our main proposition:

Proposition 1: When $z_k < \underline{z}_k$, $\frac{\partial Prob_k}{\partial w_k} > 0$, i.e., increasing wage level in province k will increase the FDI probability in k. When $z_k > \underline{z}_k$, $\frac{\partial Prob_k}{\partial w_k} < 0$, i.e., increasing wage level in province k will decrease FDI probability in k.

We express the main idea of Proposition 1 above using the following Figure 1. The intuition is that when labor quality is low, wage can be regarded as an effective indicator of labor quality; therefore, high wages attract FDI inflows. When labor quality is high and exceeds a threshold value, the labor cost effect of wages dominates, and the high wage impedes FDI inflow.

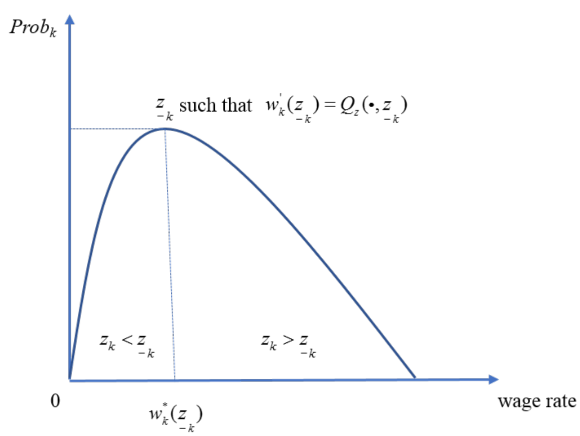

Figure 1: Illustration of how wage effect on FDI choice depends on labor quality

3 Data and variables

The theoretical model in Section 2 considers all FDI firms as a representative firm that makes the location choice about which Chinese province to invest in.[6] The model's main proposition predicts that how the wage level in province k affects the probability of the representative FDI firm investing there depends on the labor quality level in province k. This wage effect is a decreasing function of local labor quality in province k. Under the representative-firm model, the more likely a province is chosen as the FDI destination, the more FDI inflows into this province. To investigate the role of wage and labor quality in determining FDI location choices, i.e., FDI flows into different provinces, we use a panel dataset covering 29 Chinese provinces, autonomous regions and municipalities (henceforth "provinces") over a 26-year period from 1993 to 2018.[7, 8] Our sample starts in 1993 because FDI inflows expanded rapidly after Deng Xiaoping's visit to southern China in 1992. The data we use are taken from the regional database of the National Bureau of Statistics of China.

In our estimation, dependent variable is the FDI inflows into various provinces each year, measured by the logarithm of FDI firms' total investment. We include wage, labor quality, GDP, degree of openness, government expenditure, infrastructure, industry compositions, and the population structure as the main explanatory variables in our regressions. In empirical FDI studies, wage

[6]Since FDI firms in the theoretical model are homogeneous, the aggregation of homogeneous firms is equivalent to the case where the decision of a single firm is multiplied by the number of firms. Hence, the key results of the theoretical model with a representative firm will keep unchanged after aggregation. Moreover, it is important to notice that the role of the theoretical model is to inspire our empirical analysis with a nonlinear estimation approach, but not a direct test of the model, since the real data comes from the complex world that may not follow the data generating process as our theoretical model suggests. Nevertheless, we expect that "a useful model" can provide some valid explanations to the observed data. Hence, we adopt a simple theoretical setting with homogeneous firms, which does not only serve to explain the observed data to certain extents, but also enables us to focus on the novelty in empirical work.

[7]The empirical application with firm-level data is appropriate for a heterogeneous-firm model as in Liu, Lovely and Ondrich (2010). In comparison, our theory does not highlight the heterogeneous behaviors of firms, but rather the role of wage and labor quality in determining the firms' FDI flows into different provinces from the macro persepctive. Hence, a representative-firm model serves our focus well, and the provincial-level data is sufficient for our application.

[8]The 29 provinces included in our sample are Anhui, Beijing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Jilin, Liaoning, Inner Mongolia, Ningxia, Qinghai, Shandong, Shanghai, Shanxi, Shaanxi, Sichuan, Tianjin, Xinjiang, Yunnan, and Zhejiang, excluding Chongqing and Tibet.

and GDP are well-documented variables used to control for labor costs and local market demand.[9] Labor quality, calculated as the proportion of the number of people who newly graduate from high school to the total population, is included to control for not only the local human capital level, but also the local availability of human capital.[10] Degree of openness and government expenditure, two distinctive variables for China, are also included to capture their role in attracting FDI.[11] To capture different levels of infrastructure, we use the average kilometers of roads (including both highway and railway) per squared kilometer. To control for industry compositions across provinces, we use three variables: the percentage of primary sector in GDP, that of secondary sector in GDP, and the ratio of GDP in tertiary sector to that in others. To control for the aging of population that could be a potential confounding factor on labor quality, we include as a control variable the ratio of population above 65 to that between 15-64 in our regression. Due to the limitation of data availability, the time span of this variable is from 2002 to 2018. In addition to these explanatory variables, we include two time-dummy variables to control for the impacts of China's accession to the WTO in 2001 and the Global Financial Crisis in 2008. The detailed variable definitions and descriptive statistics are shown in Table 1.

[Insert Table 1]

[9] Note that in our study we use the average wage of urban employees. Urban wage is widely adopted for studying FDI-related topics (See Liu, Lovely and Ondrich (2010), Gao (2005) and Du, Lu and Tao (2008) for example). This is because on the one hand, foreign investments are located mainly in the urban areas; on the other hand, people are mostly self-employed in the rural areas, and there is a lack of definite wage statistics for rural areas.

[10] We use a flow-variable measure instead of a stock-variable measure (the share of people in population with a high school degree, for example). Both types of variables were used to measure labor quality in the existing literature. For example, Du, Lu and Tao (2008) used the flow-variable measure, the proportion of the number of students who are enrolled in higher education institutions to its total population, whereas Gao (2005) used the stock-variable measure, the proportion of the population with at least high school education. Given the relatively long sample period in our paper, we choose the flow-variable measure, the number of people who newly graduate from high school every year, because it is much longer than the stock-variable measure from our data source. Furthermore, we choose the proportion of high school graduates, not the proportion of college graduates and above, as the measure of local labor quality because it better reflects the local education level. If we considered the share of college graduates and above, we might risk underestimating labor quality in the provinces with fewer colleges and universities.

[11] Note that we use degree of openness as a policy variable to control for the policy effects. Some existing literature – see Kang and Lee (2007), Du, Lu and Tao (2008), Liu, Lovely and Ondrich (2010) among others – uses a dummy variable to indicate whether this province has Special Economic Zones (SEZs) to control for a policy effect. However, in our sample period, this dummy variable does not have time variations. Therefore, to consider the policy effect on FDI inflows, we use degree of openness, which is a highly relevant variable for the preferential policies and has enough variations across provinces and over time.

[Insert Figures 2 – 4]

Figures 2 – 4 show the time-series plots of log FDI inflow, log wage, and labor quality for every province separately. The log FDI inflow (*lfdi*) and log wage (*lwage*) for each province exhibit a clear upward trend. Moreover, the proportions of the number of people who newly graduate from high school to the total population (*lbr_qlty*) for most of the provinces are about 2‰ before year 2000, sharply increasing to a peak value around 6‰−8‰ in year 2010, and then decreasing slightly. Because most of the high school graduates between years 2000 and 2010 were born in the 1980s, this sharp upward trend in the measure of labor quality matches the occurrence of the third baby boom in China. The mild decline in the measure of labor quality may be caused by the aging population.[12,13] In addition, the percentage of total exports and imports in GDP (*open*) is less than 50%, and even close to zero, for most of the provinces except these seven: Beijing, Tianjin, Shanghai, Jiangsu, Zhejiang, Fujian, Guangdong. The economy of these seven provinces relies heavily on international trade. For all seven, the percentages increase to a different extent after China's entrance into the World Trade Organization in 2001.

4 Estimation of traditional linear panel data models

In this section, we present and discuss the estimation results of linear panel data models with province fixed effects. Then, we explore the heterogeneous effect of wage on FDI inflows over time and across regions. We aim to demonstrate the existence of a possible nonlinear relationship among FDI inflows, wages, and labor quality and thus a potential misspecification of using a linear regression model.

[12] With our measure of labor quality, we consider various dimensions of human capital changes, not only the changes in the average skill level, but also the changes in the availability of the labor force that apply their skills. Hence, the decline in our measure of labor quality is caused by the latter dimension of human capital changes, which is the aging population and hence the decline in the availability of labor force. We include an additional variable that measures the population structure to separate the population aging effect from our labor quality variable.

[13] Note that the decline in our flow-variable measure of labor quality does not necessarily imply the decline in the stock-variable measure of labor quality, but rather indicates the slower increase in the stock-variable measure of labor quality.

4.1 Estimation results of linear panel data models

Given the definitions of variables in Section 3, we first consider a traditional linear panel data model with province fixed effects and a linear time trend:

$$\begin{aligned} lfdi_{it} =& \beta_1 lwage_{it-1} + \beta_2 lbr\_qlty_{it-1} + \beta_3 lwage_{it-1} \times lbr\_qlty_{it-1} \\ &+ \mathbf{w}'_{it-1}\boldsymbol{\gamma} + \delta_1 aftr\_WTO_t + \delta_2 aftr\_GFC_t + \mu_i + \lambda t + \epsilon_{it}. \end{aligned} \tag{4.1}$$

The dependent variable $lfdi_{it}$ is log FDI inflow of province i at year t. It is important to note that FDI inflows at year t can effect the explanatory variables contemporaneously. For example, FDI inflows in year t, deemed as a part of provincial GDP in year t, could stimulate labor demand and affect other explanatory variables in the same year. To alleviate this simultaneity problem, our main explanatory variables are one-period lagged. The variables of interest, $lwage_{it-1}$ and lbr_qlty_{it-1}, are log wage and labor quality respectively. The effects of wage and labor quality on FDI may be intertwined, and we include a term interacting wage with labor quality ($lwage_{it-1} \times lbr\_qlty_{it-1}$) to capture such an interactive effect. The control variables in the vector $\mathbf{w}_{it-1}$ include log GDP, degree of openness, government expenditure, infrastructure, the percentage of primary sector in provincial GDP, that of secondary sector in provincial GDP, the ratio of GDP in tertiary sector to that in others, and the ratio of population above 65.

In equation (4.1), the fixed effects μ_i capture the impacts of province-specific time invariant factors on FDI inflows, and λt is the linear time trend that captures changes in the domestic and global macroeconomic environment. In the model, we use a linear time trend instead of year fixed effects for two reasons. First, for our panel dataset, the time span of 26 years is quite long considering there are only 29 provinces. Including both province and year fixed effects would consume too many degrees of freedom from our model and result in the invalidity of statistical inference. Second, by visual inspection of the time-series plots, we find that most of the variables exhibit similar linear upward trends. Based on these reasons, and the fact that unobserved heterogeneity across provinces is much more pronounced than over time, we make the next best choice to include province fixed effects and use the linear time trend to capture changes in the macroeconomic environment. To remedy the insufficiency of the linear time trend in capturing important events — China's accession to the WTO in 2001 and the Global Financial Crisis (GFC) in 2008 — we include two time dummy variables, the after-WTO dummy ($aftr\_WTO_t$) and the after-GFC dummy ($aftr\_GFC_t$), as additional regressors.[14]

Moreover, we assume that all of the explanatory variables are only contemporaneously exogenous. The error term ε_{it} is assumed to have an unknown serial correlation within individuals and an unknown cross-sectional correlation within time. Correspondingly, we use the within-group method to estimate the parameters in the model. For statistical inference, we use the two-way cluster-robust standard errors that are robust to both temporal and cross-sectional dependence of the error term ε_{it}.[15]

According to the classical theory, FDI prefers locations with lower wage costs, higher labor quality, larger market demand, a greater degree of openness, and better availability of infrastructure. In China, the government may supply infrastructure, finance, or preferential policies beneficial to attracting FDI, or it may deter FDI due to stringent administration and government-expenditure-induced misallocations. Thus, the coefficient of government expenditure could be either positive or negative. Considering the fact that FDI firms tend to invest in the manufacturing and service sectors in China, we would expect, ex ante, that provinces with a large primary sector are at a disadvantage for attracting FDI inflows, and therefore that the coefficient associated with the predictor percentage of primary sector in GDP ($pct\_primary$) would be negative while the coefficients associated with the predictors percentage of secondary sector in GDP ($pct\_secondary$) and ratio of GDP in tertiary sector to that in others ($rt\_tertiary$) would be positive. We would also expect, ex ante, that provinces where the population is of advanced age would see their workforce shrink and thereby be less attractive for FDI, hence the coefficient on the predictor ratio of population above 65 (rt_ab65) would be negative. In addition, China's entry into the WTO implies more openness to FDI; thus

[14] The after-WTO dummy is defined as $aftr\_WTO_t = 1$ for $t \geq 2002$, and it equals to zero otherwise. The after-GFC dummy is defined as $aftr\_GFC_t = 1$ for $t \geq 2009$, and it equals to zero otherwise.

[15] It is proposed by Cameron, Gelbach and Miller (2011) and Thompson (2011),

the sign of the time dummy for joining the WTO is positive. Moreover, the Global Financial Crisis in 2008 deterred global FDI flows; thus the sign of the time dummy for post-crisis is negative. The expected signs of all variables are summarized in the second column of Table 2.

[Insert Table 2]

Table 2 reports the estimation and related hypothesis testing results of the three linear fixed-effects panel data models. In Model 1, we only include the two variables of interest, log wage (*lwage*) and labor quality (*lbr_qlty*), as the explanatory variables in the regression. After controlling for province fixed effects and linear time trend, the coefficients of both variables are positive but statistically insignificant. After we include the interactive term between wage and labor quality ($lwage \times lbr\_qlty$) in Model 2, none of the coefficients of log wage, labor quality, and their interaction are significant. Moreover, the joint effects of log wage and labor quality on log FDI are all statistically insignificant, because the three Wald tests fail to reject the nulls that the overall and joint effects of log wage and labor quality are zero. Because the data for the ratio of population above 65 (*rt_ab65*) is only available between 2002 and 2018, we consider two models with control variables. Model 3 includes all control variables but the ratio of population above 65 and has a larger sample from 1993 to 2018. Among the three variables of interest, only the coefficients of labor quality and the interaction regressor are marginally significant; the three Wald tests still fail to reject their nulls respectively. In Model 4, we include the ratio of population above 65 to the regression with all other controls and the sample period is from 2002 to 2008. The empirical results of Model 4 are similar to those of Model 3: all the Wald tests in the two models fail to reject the nulls. Note that the estimated coefficient of the ratio of population above 65 in Model 4 is significantly positive. A possible explanation of this result may be the reverse causality from FDI to the aging of population: Locations with more FDI are usually more developed regions with higher living costs, which restrains fertility and therefore are more prone to population aging. Because we only use the ratio of population above 65 as a control variable to account for the potentially omitted factor of population aging, the estimated coefficient of the ratio of population above 65 may not reflect its causal effect on FDI.

4.2 Heterogeneity over time and across regions

One possible explanation for these insignificance results using the whole sample is that the effects of log wage and labor quality are very heterogeneous across time and/or provinces. So, in addition to using the whole sample, we investigate the effects of log wage and labor quality on FDI using subsamples, defined by different time periods or geographic locations. First we create three subsamples for three time periods: from 1993 to 1997; from 1998 to 2008; and from 2009 to 2018. The average labor quality values of the 29 provinces over the 26 years, as shown in Figure 5, exhibit drastically different patterns. During 1993 to 1997, the labor quality values are low, around 2‰ for all provinces in China. In the next 11 years, China experienced rapid growth in labor quality, rising from 2‰ to its peak value of 6.4‰ in the year 2008. From 2009 on, average labor quality values dropped a bit, becoming stable around 6‰. In addition, we create three additional subsamples corresponding to three geographic locations: the 11 provinces of East China (Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Hainan); 8 provinces of Middle China (Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan); and 10 provinces of West China (Inner Mongolia, Guangxi, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang). These three regions have different natural environment endowments and socio-economic attributes.[16] Consequently, wages and labor quality may play different roles in attracting FDI inflows in each region.

[Insert Table 3] [Insert Table 4] [Insert Figure 5]

The results in Table 3 suggest that the effects of log wage and labor quality on log FDI change over time. In the early time period from 1993 to 1997, log wage has a marginally significant positive effect on log FDI that diminishes as labor quality increases. In the first column, the *p*-value of the first Wald statistic, which tests whether the coefficients of log wage and its interaction with labor quality are both zero, is 5.7%; the coefficient of log wage (*lwage*) is positive, and that of the

[16] In addition to the heterogeneity revelation, another benefit of performing the subsample analysis across regions is to control for the bias that might arise from using variables of the urban labor force to represent the labor characteristics of a province. In our subsample analysis, within each region the provinces share similar socio-economic attributes; therefore, the bias is weakened to a certain extent.

interactive term between wage and labor quality ($lwage \times lbr\_qlty$) is negative. During the second time period from 1998 to 2008, both log wage and labor quality become statistically insignificant. In the second column, all three Wald tests fail to reject the nulls that the overall and joint effects of log wage and labor quality are zero.

Finally, in the last time period from 2009 to 2018, labor quality turns out to be a salient factor and the joint effect of log wage and labor quality on log FDI is significant, regardless of whether the ratio of population above 65 (rt_ab65) capturing the degree of population aging is included in the model or not.[17] However, the sign of the coefficients changes while remaining insignificant. In the third and fourth columns, the p-values of the second and the third Wald statistics are all around 1%; the coefficients on the predictors labor quality (lbr_qlty) and log wage ($lwage$) become negative and the coefficient on the predictor that is an interactive term between wage and labor quality ($lwage \times lbr\_qlty$) is positive.

Similarly, the results in Table 4 imply that the effects of log wage and labor quality on log FDI are quite heterogeneous across different regions of China. In the east and west regions, neither log wage nor labor quality has a significant effect on log FDI. However, in the middle region, the coefficient of log wage is significantly negative, which implies that a lower wage level generally would attract more FDI inflows into the middle region of China.

The heterogeneity of the marginal effects of log wage and labor quality on log FDI inflows over time and across regions hints at the existence of possible nonlinear relationships among the three. As the time period and region changes, the values of labor quality and log wage also change dramatically. The fact that the marginal effects of log wage and labor quality on log FDI inflow depend on the values of labor quality and log wage implies that, all other things being equal, the relationship among the three variables could be nonlinear. Therefore, the estimation results of the linear panel data models may not fully reveal the relationship among the three.

[17] Because the values of the ratio of population above 65 (rt_ab65) is available from 2002 to 2018, we include it as a control variable only in the model using subsample with time period from 2009 to 2018.

5 Estimation of a partially linear model

5.1 Specification

To account for the existence of possible nonlinear relationships among log FDI inflow, log wage and labor quality, we propose a partially linear panel data model with province fixed effects and a linear time trend:

$$lfdi_{it} = g(lwage_{it-1}, lbr\_qlty_{it-1}) + \mathbf{w}'_{it-1}\gamma + \delta_1 aftr\_WTO_t + \delta_2 aftr\_GFC_t + \mu_i + \lambda t + \epsilon_{it}. \quad (5.1)$$

The specification of our proposed empirical model (5.1) is the same as that of the traditional linear fixed effects model (4.1) except that log wage ($lwage_{it-1}$), and labor quality (lbr_qlty_{it-1}) have a joint nonlinear effect on log FDI ($lfdi_t$), characterized by an unknown smooth function $g(lwage_{it-1}, lbr\_qlty_{it-1})$. We only allow wages and labor quality to have the joint nonlinear effects on FDI. That is a trade-off between model parsimony and flexibility. It is known that the semi-parametric approach ensures the simplicity of estimation and allows for some economically meaningful flexibility of model specification, suitable for our purposes. In this study, we focus on the intertwined effect of wage and labor quality. Therefore, we only allow the effect of wage and labor quality to depend on each other, and we let other control variables keep a linear relationship, as in the existing literature.

Based on ANOVA decomposition of smooth functions, we can decompose the unknown smooth function $g(x, z)$ as follows,

$$g(lwage_{it-1}, lbr\_qlty_{it-1}) = f_1(lwage_{it-1}) + f_2(lbr\_qlty_{it-1}) + f_3(lwage_{it-1}, lbr\_qlty_{it-1}), \quad (5.2)$$

where $f_1(lwage_{it-1})$ and $f_2(lbr\_qlty_{it-1})$ are smoothed main effects of log wage and labor quality on dependent variable, and $f_3(lwage_{it-1}, lbr\_qlty_{it-1})$ captures the smooth interactive effects that include no component of the form $f_1 + f_2$. We use tensor product cubic B-splines to approximate the unknown smooth function.[18]

5.2 Empirical results

In combination with the ANOVA decomposition of smooth functions in (5.2), our proposed partially linear panel data model (5.1) can be written explicitly as

$$\begin{aligned} lfdi_{it} = {} & f_1(lwage_{it-1}) + f_2(lbr\_qlty_{it-1}) + f_3(lwage_{it-1}, lbr\_qlty_{it-1}) \\ & + \mathbf{w}'_{it-1}\gamma + \delta_1 aftr\_WTO_t + \delta_2 aftr\_GFC_t + \mu_i + \lambda t + \epsilon_{it} \end{aligned} \tag{5.3}$$

where the vector $\mathbf{w}_{it-1}$ includes the control variables, $f_1(lwage_{it-1})$ and $f_2(lbr\_qlty_{it-1})$ representing the nonlinear effects of log wage and labor quality on log FDI inflow alone respectively, and $f_3(lwage_{it-1}, lbr\_qlty_{it-1})$ captures the interactive nonlinear effects of the two variables. The construction of the three functions, f_1, f_2 and f_3, using cubic splines follows our discussion earlier in this section.

In Table 5, we report the estimation results of the four partially linear panel data models, PL1, PL2, PL3, and PL4. These specifications are similar to those of Model 1, 2, 3, and 4 shown in Table 2. In Model PL1, we consider only additively nonlinear effects of log wage and labor quality. By comparison, Model PL2 accommodates the nonlinear interactive effects. In Model PL3, we include all the control variables except the ratio of population above 65 (rt_ab65) in the linear part. In the estimation of PL1-3, we use the full sample from 1993 to 2008. In Model PL4, we incorporate the ratio of population above 65 as an additional control variable. Because the values of the ratio of population above 65 is only available from 2002 to 2018, we have to restrict our sample to the same time period. Our first finding is that the interactive nonlinear effects matter. When testing whether the nonlinear interaction term $f_3(lwage_{it-1}, lbr\_qlty_{it-1})$ is zero in Models PL2, PL3 and PL4, all

[18] The construction of the splines in the decomposition (5.2) can be found in de Boor (2001) and Wood (2017). Appendix B describes the construction of the nonlinear function and estimation method in detail. For econometric theory on partially linear panel data models with fixed effects, please see Li and Stengos (1996), Baltagi and Li (2002), and Huang, Wu and Zhou (2004) among others.

the Wald statistics strongly reject the null that the nonlinear interaction is zero.[19] In addition, the estimation results of Models PL3 and PL4 support that log wage has a statistically significant effect on log FDI inflow, because the Wald statistics used to test whether $f_1(lwage_{it-1})$ is zero has p-values of 6.4% and 3.6% respectively. Moreover, the Wald statistics testing $f_2(lbr\_qlty_{it-1}) = 0$ has p-values of 3.4% and 11.7%, showing that labor quality is at least marginally significant.

[Insert Table 5]

Comparing the estimation results of our proposed partially linear models (Models PL3 and PL4 in Table 5) to the traditional linear fixed effects model (Models 3 and 4 in Table 2), we find that the partially linear model not only fits the data better but also captures the marginal effect of wages on FDI more effectively than the FE models. Intuitively, the adjusted R-squared of Models PL3 and PL4 is 74.7% and 71.5% and that of Models 3 and 4 is 71.6% and 65.5%. Models PL3 and PL4 has twelve more regressors than Models 3 and 4 because of its construction of nonlinear functions using splines. Still, the adjusted R-squared of Models PL3 and PL4, which measures the goodness of fit after penalizing the number of regressors, is 3.1% and 6% larger than those of Models 3 and 4 respectively.

Also, we conduct a battery of linearity tests, reporting the results in Tables 6 and 7. Specifically, the bottom panels of Tables 6 and 7 display the results of testing the null hypothesis that the three nonlinear functions $f_1(lwage)$, $f_2(lbr\_qlty)$ and $f_3(lwage, lbr\_qlty)$ are jointly linear. Under this null, the specifications of both Models PL3 and PL4 coincide with those of Model 3 and Model 4 respectively, i.e. the linear fixed effects models. The Wald statistics are 43.74 and 61.28 with the p-values smaller than 0.001, strongly rejecting the null hypothesis that the linear specification of Models 3 and 4 are adequate. More interestingly, the result of linearity test for the single function $f_3(lwage, lbr\_qlty)$ also strongly rejects the null of linearity in Models PL3 and PL4. Combining it with the testing results of the null $f_3(lwage, lbr\_qlty) = 0$ in Table 5, we conclude that the interactive effect of wage and labor quality on FDI does exist and this interaction is truly nonlinear.

[19] See the test results in the third row of the middle panel in Table 5.

All of these empirical results suggest that our proposed partially linear model captures the marginal effect of wage on FDI more effectively than the linear FE models does.

[Insert Tables 6 and 7]

Using the estimation results of Model PL4, we illustrate intuitively the nonlinear effects of wage and labor quality on FDI in Figure 6 and 7. Figure 6 shows the estimated marginal effects of log wage on log FDI inflow as a function of *lwage* and *lbr_qlty*; similarly, Figure 7 shows the estimated marginal effects of labor quality on log FDI as a function of log wage (*lwage*) and labor quality (*lbr_qlty*), i.e.

$$\frac{\partial \widehat{g}(lwage, lbr\_qlty)}{\partial lwage} \quad \text{and} \quad \frac{\partial \widehat{g}(lwage, lbr\_qlty)}{\partial lbr\_qlty}.$$

The two 3-dimensional surface plots imply that the marginal effects of log wage and labor quality also depend on each other. In the 3D plots, the arrow alongside the axis indicates the direction in which the value of the corresponding variable increases; the color of the surface indicates the value of the marginal effect. In Figure 6, the marginal effect of log wage on log FDI is generally negative and decreases as labor quality increases. Specifically, when labor quality is low, the marginal effect of log wage on log FDI inflow could be positive. In Figure 7, the marginal effect of labor quality on log FDI inflow generally diminishes as log wage increases. Moreover, when log wage is low, the marginal effect of labor quality is positive.

[Insert Figures 6 and 7]

[Insert Figures 8 and 9]

Furthermore, the estimated marginal effects of log wage and labor quality shown in Figures 6 and 7 are statistically significant when log wage and labor quality are within certain ranges. Figure 8 plots the estimated marginal effect of log wage on log FDI as a function of labor quality with *lwage* fixed at its 60% percentile value of 10.130.[20] Figure 9 plots the estimated marginal effect of

[20] This confirms the key result of the empirical analysis that the estimated marginal-effect curve is downward sloping. That is to say, the marginal effects of wage on FDI are decreasing with respect to labor quality. While it is true that

labor quality on log FDI as a function of log wage ($lwage$) with labor quality (lbr_qlty) fixed at its sample median value of 4.550‰, i.e.

$$\frac{\partial \widehat{g}}{\partial lwage}(\overline{lwage}_{60\%pctl} = 10.130, lbr\_qlty) \quad \text{and} \quad \frac{\partial \widehat{g}}{\partial lbr\_qlty}(lwage, \overline{lbr\_qlty}_{50\%pctl} = 4.550).$$

Generally, both estimated curves are nonlinear and the marginal effects are statistically significant on certain sub-intervals of labor quality and log wage respectively. In Figure 8, the estimated curve is downward sloping and the marginal effects of log wage at the 60% percentile are marginally significant at the 10% level when labor quality is smaller than 3.5‰. In Figure 9, the estimated curve is hump-shaped, which shows that the marginal effect of labor quality is statistically significant with an upward slope when log wage is smaller than 9.6 and with a downward slope when log wage is larger than 9.6.

In sum, the results shown in the two 3D plots and the two 2D plots suggest that when labor quality is high, a higher wage would tend to impede FDI inflows; when the wage is moderate, higher labor quality would attract more FDI than when the wage is high.

6 Conclusion

Although in theory low-wage locations do attract FDI, the empirical evidence for this is rather mixed. Using linear panel data models with province fixed effects and a linear time trend, we find that the effects of wage and labor quality on FDI inflow are heterogeneous, not only over time from 1993 to 2018 but also across different regions of China. This finding further suggests that these effects may be nonlinear. Hence, we propose a partially linear panel data model, allowing for a nonlinear relationship between FDI inflow and the two variables of interest, and apply it to FDI location choices in China. Our main finding is that the marginal effect of wage is generally a decreasing function of labor quality. This implies that when faced with low labor quality, FDI

the marginal effects of log wage on log FDI are not significant at larger values of labor quality, it does not affect the downward sloping shape of the estimated curve.

firms would prefer locations with high wages and thus higher labor quality. When facing high labor quality, FDI firms would prioritize locations with low wages in order to reduce labor costs. Moreover, we find that the marginal effect of labor quality on FDI inflows generally diminishes as the log wage increases. This suggests that when the wage level is moderate, labor quality will attract more FDI than when the wage is high.

FDI is one of the driving forces behind the growth of transitional economies. Hence, attracting FDI inflows is the policy priority of various transitional economies. Our results help the transitional economies better understand the wage effect on FDI, i.e., low labor costs do not necessarily mean an effective comparative advantage in attracting FDI inflows. Understanding that local labor quality is vital for the low labor cost countries, the transitional economies could dedicate themselves to promoting the human capital formation and enhancing the labor quality level. Hence, the transitional economies could activate their low labor cost advantages and successfully attract FDI.

There are limitations in our current study. Although the province-level data enables us to explore heterogeneity across regions and over time, and to confirm the nonlinear relationship among wages, labor quality, and FDI location choices, it does not include the FDI industrial and bilateral information. Hence, we cannot investigate how wage effects vary across industries with different production technologies and across different FDI sourcing countries. We expect that more detailed data with the missing information can help us generate new results, which is the direction of our future research.

Appendix A Numerical example

Suppose the production function take the Cobb-Douglas form of $Q(L(z), x, z) = zL(z)^{\alpha} \cdot x^{1-\alpha}$. We assume $L(z) = \frac{L}{z}$, where L is a constant. It means that the higher the labor quality is, the smaller amount the labor demand for production is. Hence, the production function can be written as $Q(L(z), x, z) = z\left(\frac{L}{z}\right)^{\alpha} \cdot x^{1-\alpha}$ and we have:

$$w_j(z_j) = Q_L(L(z_j), x_j, z_j) = \alpha z_j \Big(\frac{L}{z_j}\Big)^{\alpha-1} x_j^{1-\alpha} = \alpha z_j^{2-\alpha} L^{\alpha-1} x_j^{1-\alpha}$$

$$Q_{LL}(L(z_j), x_j, z_j) = \alpha(\alpha-1) z_j \Big(\frac{L}{z_j}\Big)^{\alpha-2} x_j^{1-\alpha}$$

$$Q_{Lz}(L(z_j), x_j, z_j) = \alpha \Big(\frac{L}{z_j}\Big)^{\alpha-1} x_j^{1-\alpha}, \ L'(z_j) = -\frac{L}{z_j^2}$$

$$Q_z(L(z_j), x_j, z_j) = z \Big(\frac{L}{z_j}\Big)^{\alpha} \cdot x^{1-\alpha}, \ Q_{zz} = 0$$

$$w_j'(z_j) = Q_{LL}(L(z_j), x_j, z_j) L'(z_j) + Q_{Lz}(L(z_j), x_j, z_j) = \alpha(2-\alpha) z_j^{1-\alpha} L^{\alpha-1} x_j^{1-\alpha}$$

$$w_j''(z_j) = \alpha(1-\alpha)(2-\alpha) z_j^{-\alpha} L^{\alpha-1} x_j^{1-\alpha}$$

According to $w_j'(\underline{z}_j)L(\underline{z}_j) = Q_z(L(\underline{z}_j), x_j, \underline{z}_j)$, we have:

$$\underline{z}_j^{1-\alpha} L^{\alpha} \cdot x_j^{1-\alpha} = \alpha(2-\alpha)\underline{z}_j^{1-\alpha} L^{\alpha-1} x_j^{1-\alpha} L \underline{z}_j^{-1} = \alpha(2-\alpha)\underline{z}_j^{1-\alpha} L^{\alpha} x_j^{1-\alpha} \underline{z}_j^{-1}$$

So we can solve for $\underline{z}_j = \alpha(2-\alpha)$.

$$\begin{aligned}
&Q_{LL}(L(z_j), x_j, z_j)(L'(z_j))^2 + w''(z_j)L(z_j) \\
&= \alpha(\alpha-1) z_j^{3-\alpha} L^{\alpha-2} x_j^{1-\alpha} (-L z_j^{-2})^2 + \alpha(1-\alpha)(2-\alpha) z_j^{-\alpha} L^{\alpha-1} x_j^{1-\alpha} L z_j^{-1} \\
&= \alpha(\alpha-1) z_j^{-\alpha-1} L^{\alpha} x_j^{1-\alpha} + \alpha(1-\alpha)(2-\alpha) z_j^{-\alpha-1} L^{\alpha} x_j^{1-\alpha} \\
&= \alpha(\alpha-1)^2 z_j^{-\alpha-1} L^{\alpha} x_j^{1-\alpha} > 0 = Q_{zz}(L(z_j), x_j, z_j)
\end{aligned}$$

Furthermore,

$$\begin{aligned}
&Q_{LL}(L(z_j), x_j, z_j) L'(z_j) + Q_{Lz}(L(z_j), x_j, z_j) \\
&= \alpha(\alpha-1) z_j (L z_j^{-1})^{\alpha-2} \cdot x_j^{1-\alpha} \cdot (-L z_j^{-2}) + \alpha(L z_j^{-1})^{\alpha-1} \cdot x_j^{1-\alpha} \\
&= \alpha(1-\alpha) L^{\alpha-1} z_j^{1-\alpha} x_j^{1-\alpha} + \alpha L^{\alpha-1} z_j^{1-\alpha} x_j^{1-\alpha} \\
&= \alpha(2-\alpha) L^{\alpha-1} z_j^{1-\alpha} x_j^{1-\alpha} > 0
\end{aligned}$$

Hence, we have the two assumptions $Q_{zz}(L(z_j), x_j, z_j) < Q_{LL}(L(z_j), x_j, z_j) \cdot (L'(z_j))^2 + w''(z_j)L(z_j)$

and $Q_{LL}(L(z_j), x_j, z_j)L'(z_j) + Q_{Lz}(L(z_j), x_j, z_j) > 0$ hold in this numerical example.

Appendix B Estimation procedure of the partially linear model

We assume that both X_{it} and Z_{it} are distributed on compact intervals. Let L_x, $L_z \geq 1$ be the numbers of interior knots for the compact supports $[a_x, b_x]$ and $[a_z, b_z]$ associated with X_{it} and Z_{it} respectively. We divide the closed intervals $[a_x, b_x]$ and $[a_z, b_z]$ into $L_x + 1$ and $L_z + 1$ subintervals. For the interval $[a_x, b_x]$, we have $I_{j,x} = [t_{j,x}, t_{j+1,x})$, $j = 0, \ldots, L_x - 1$, $I_{L_x} = [t_{L_x,x}, b_x]$, where $\{t_{j,x}\}_{j=1}^{L_x}$ is a sequence of equally spaced interior knots, satisfying

$$t_{-2,x} = t_{-1,x} = t_{0,x} = a_x < t_{1,x} < \cdots < t_{L_x,x} < b_x = t_{L_x+1,x} = t_{L_x+2,x} = t_{L_x+3,x}.$$

Similarly, for the interval $[a_z, b_z]$, we have $I_{j,z} = [t_{j,z}, t_{j+1,z})$, $j = 0, \ldots, L_z - 1$, $I_{L_z} = [t_{L_z,z}, b_z]$, where $\{t_{j,z}\}_{j=1}^{L_z}$ is a sequence of equally spaced interior knots, satisfying

$$t_{-2,z} = t_{-1,z} = t_{0,z} = a_z < t_{1,z} < \cdots < t_{L_z,z} < b_z = t_{L_z+1,z} = t_{L_z+2,z} = t_{L_z+3,z}.$$

Based on de Boor (2001), we define the univariate cubic B-spline basis as $\mathcal{B}_{M_x} = \{B_{j,x}(x) : 1 \leq j \leq M_x\}$ and $\mathcal{B}_{M_z} = \{B_{j,z}(z) : 1 \leq j \leq M_z\}$ with the number of spline basis functions $M_x \equiv L_x + 3$ and $M_z \equiv L_z + 3$. Then we define the tensor product cubic B-spline basis by $\mathcal{B}_{M_{xz}} = \{B_{j,x}(x)B_{k,z}(z) : 1 \leq j \leq M_x,\ 1 \leq j \leq M_z\}$ with the number of spline basis functions $M_{xz} = M_x \times M_z$ for simplicity of notation. In the empirical application, we use Akaike Information Criterion to select M. The unknown smooth function $g(x, z)$ can be approximated by $\sum_{j=1}^{M_x} \sum_{k=1}^{M_z} B_{j,x}(x)B_{k,z}(z)\beta_{jk}$. Using this approximation, the partially linear model (5.1) with ANOVA decomposition (5.2) can be rewritten as

$$\begin{aligned} Y_{it} &= f_1(X_{it-1}) + f_2(Z_{it-1}) + f_3(X_{it-1}, Z_{it-1}) + W_{it-1}'\gamma + \mu_i + \lambda t + \varepsilon_{it} \\ &= \sum_{j=1}^{M_x} B_{x,j}(X_{it-1})\beta_{x,j} + \sum_{k=1}^{M_z} B_{z,k}(Z_{it-1})\beta_{z,k} + \sum_{j=1}^{M_x}\sum_{k=1}^{M_z} B_{x,j}(X_{it-1})B_{z,k}(Z_{it-1})\beta_{jk} \end{aligned}$$

$$
\begin{aligned}
&+ W'_{it-1}\boldsymbol{\gamma} + \mu_i + \lambda t + \varepsilon_{it} \\
&= B_x(X_{it-1})'\boldsymbol{\beta}_x + B_z(Z_{it-1})'\boldsymbol{\beta}_z + B'_{xz}(X_{it-1}, Z_{it-1})\boldsymbol{\beta}_{xz} + W'_{it-1}\boldsymbol{\gamma} + \mu_i + \lambda t + \varepsilon_{it} \\
&= B'(X_{it-1}, Z_{it-1})\boldsymbol{\beta} + W'_{it-1}\boldsymbol{\gamma} + \mu_i + \lambda t + \varepsilon_{it},
\end{aligned}
\tag{B.1}
$$

where

$$
\begin{aligned}
B'_x(X_{it-1}) &= \begin{pmatrix} B_{x,1}(X_{it-1}) & \cdots & B_{x,j}(X_{it-1}) & \cdots & B_{x,M_x}(X_{it-1}) \end{pmatrix} \\
B'_z(Z_{it-1}) &= \begin{pmatrix} B_{z,1}(Z_{it-1}) & \cdots & B_{z,k}(Z_{it-1}) & \cdots & B_{z,M_z}(Z_{it-1}) \end{pmatrix} \\
B'_{xz}(X_{it-1}, Z_{it-1}) &= B'_x(X_{it-1}) \otimes B'_z(Z_{it-1}) \\
B'(X_{it-1}, Z_{it-1}) &= \begin{pmatrix} B'_x(X_{it-1}) & B'_z(Z_{it-1}) & B'_{xz}(X_{it-1}, Z_{it-1}) \end{pmatrix} \\
\boldsymbol{\beta}'_x &= (\beta_{x,1}, \cdots, \beta_{x,M_x}) \\
\boldsymbol{\beta}'_z &= (\beta_{z,1}, \cdots, \beta_{z,M_z}) \\
\boldsymbol{\beta}'_{xz} &= (\beta_{11}, \cdots, \beta_{1M_z}, \cdots, \beta_{M_x 1}, \cdots, \beta_{M_x M_z}) \\
\boldsymbol{\beta}' &= (\boldsymbol{\beta}'_x,\, \boldsymbol{\beta}'_z,\, \boldsymbol{\beta}'_{xz})
\end{aligned}
$$

and by construction we restrict that $\sum_i \sum_t B_{x,j}(x_{it-1}) = 0$ and $\sum_i \sum_t B_{z,k}(z_{it-1}) = 0$ for all j and k so that the interaction terms $f_3(x_{it-1}, z_{it-1}) = B'_{xz}(x_{it-1}, z_{it-1})\boldsymbol{\beta}_{xz}$ are orthogonal to the main effects $f_1(x_{it}) = B_x(x_{it-1})'\boldsymbol{\beta}_x$ and $f_2(z_{it}) = B_z(z_{it-1})'\boldsymbol{\beta}_z$.

Stacking the equation (B.1), we obtain the following matrix form

$$
\mathbf{Y} = \boldsymbol{\mathcal{X}}\boldsymbol{\theta} + \boldsymbol{\mu} + \boldsymbol{\varepsilon} \tag{B.2}
$$

where $\boldsymbol{\mathcal{X}} = [\mathbf{B}(\mathbf{X}, \mathbf{Z}),\, \mathbf{W},\, \mathbf{R}]$, $\boldsymbol{\theta}' = (\boldsymbol{\beta}', \boldsymbol{\gamma}', \lambda)$, and other related terms are defined as follows,

$$
\begin{aligned}
\mathbf{Y} &= (Y_{12}, \ldots, Y_{1T}, \ldots, Y_{n2}, \ldots, Y_{nT})' \\
\mathbf{B}(\mathbf{X}, \mathbf{Z}) &= [B(X_{11}, Z_{11}), \ldots, B(X_{1T-1}, Z_{1T-1}), \ldots, B(X_{n1}, Z_{n1}), \ldots, B(X_{nT-1}, Z_{nT-1})]'
\end{aligned}
$$

$$\mathbf{W} = (W_{11}, \dots, W_{1T-1}, \dots, W_{n1}, \dots, W_{nT-1})'$$
$$\mathbf{R} = \iota_n \otimes (2, \cdots, T)'$$
$$\boldsymbol{\mu} = [\mu_1, \dots, \mu_n]' \otimes \iota_{T-1}$$
$$\boldsymbol{\varepsilon} = [\varepsilon_{12}, \dots, \varepsilon_{1T}, \dots, \varepsilon_{n2}, \dots, \varepsilon_{nT}]'$$

and ι_d is a d-dimensional column vector of ones.

We propose to use the within-group method to estimate the parameter vector $\boldsymbol{\theta}$ in (B.2). First, by within transformation, we remove the fixed effects by pre-multiplying the within projection matrix $\mathbf{M}$ on both sides of equation (B.2) and obtain

$$\mathbf{MY} = \mathbf{M}\boldsymbol{\mathcal{X}}\boldsymbol{\theta} + \mathbf{M}\boldsymbol{\varepsilon}. \tag{B.3}$$

where the within projection matrix $\mathbf{M}$, symmetric and idempotent, is defined by $\mathbf{M} = \mathbf{I}_n \otimes \mathbf{Q}_{T-1}$ in which $\mathbf{I}_n$ is a $T \times T$ identity matrix, $\mathbf{Q}_{T-1} = \mathbf{I}_{T-1} - \mathbf{J}_{T-1}/(T-1)$ and $\mathbf{J}_{T-1}$ is a $(T-1) \times (T-1)$ matrix with all entries equal to 1.

Since the variance-covariance structure of the error term $\boldsymbol{\varepsilon}$ is left unspecified, we simply obtain the consistent within-group OLS estimate $\widehat{\boldsymbol{\theta}}$,

$$\widehat{\boldsymbol{\theta}} = (\boldsymbol{\mathcal{X}}'\mathbf{M}\boldsymbol{\mathcal{X}})^{-1}\boldsymbol{\mathcal{X}}'\mathbf{MY}.$$

Since both serial and cross-sectional correlation are allowed for the error term, we use with two-way cluster-robust variance matrix estimate

$$\widehat{\mathrm{V}}(\widehat{\boldsymbol{\theta}}) = (\boldsymbol{\mathcal{X}}'\mathbf{M}\boldsymbol{\mathcal{X}})^{-1}\widehat{\boldsymbol{\Omega}}(\boldsymbol{\mathcal{X}}'\mathbf{M}\boldsymbol{\mathcal{X}})^{-1}$$

proposed by Cameron, Gelbach and Miller (2011) and Thompson (2011). The central matrix is

$$\widehat{\mathbf{\Omega}} = \boldsymbol{\mathcal{X}}'\mathbf{M}(\widehat{\varepsilon}\widehat{\varepsilon}' \odot \mathbf{S})\mathbf{M}\boldsymbol{\mathcal{X}}$$

where $\widehat{\varepsilon}$ is the $n(T-1) \times 1$ vector of within-group OLS residuals, $\odot$ denotes element-by-element (Hadamard) multiplication and $\mathbf{S}$ is an $n(T-1) \times n(T-1)$ indicator, or selection, matrix with ℓm entry equal to one if its corresponding (i_ℓ, t_ℓ) and (j_m, s_m) observations share at least one of the two clusters in group and time ($i_\ell = j_m$ or $t_\ell = s_m$ or both) and equal to zero otherwise. Mathematically, we have

$$\mathbf{S} = \mathbf{I}_n \otimes \mathbf{J}_{T-1} + (\mathbf{J}_n - \mathbf{I}_n) \otimes \mathbf{I}_{T-1}$$

The smooth function $g(x, z)$ is estimated by $\widehat{g}(x, z) \equiv B'(x, z)\widehat{\boldsymbol{\beta}}$ and its two-way cluster-robust variance estimate is $\widehat{\mathrm{V}}[\widehat{g}(x, z)] = B(x, z)'\widehat{\mathrm{V}}(\widehat{\boldsymbol{\beta}})B(x, z)$. Under certain smoothing condition on smooth function $g(x, z)$ and regularity conditions on matrices $\mathbf{X}$, $\mathbf{B}(\mathbf{X}, \mathbf{Z})$, and $\mathbf{W}$, we conjecture that the estimate $\widehat{\boldsymbol{\theta}}$ is consistent with typical parametric convergence rate and asymptotic normal, and the nonparametric estimator $\widehat{g}(x, z)$ has typical nonparametric convergence rate and asymptotic normality; its bias term is asymptotically negligible if certain undersmoothing condition is assumed: We use a larger number of knots (i.e. base functions) than what is needed for achieving the optimal rate of convergence.

References

Amiti, M., & Javorcik, B. S. (2008). Trade costs and location of foreign firms in China. *Journal of Development Economics*, 85(1-2), pp. 129–149.

Ascani, A., Crescenzi, R., & Iammarino, S. (2016). Economic institutions and the location strategies of European multinationals in their geographic neighborhood. *Economic Geography*, 92(4), pp. 401-429.

Baltagi, B.H., & Li, D. (2002). Series estimation of partially linear panel data models with fixed effects. *Annals of Economics and Finance*, 3, pp. 103–116.

Blyde, J., & Molina, D. (2015). Logistic infrastructure and the international location of fragmented production. *Journal of International Economics*, 95(2), pp. 319-332.

Cameron A.C., Gelbach J.B., & Miller D.L. (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2), pp. 238–249.

Cheng, L.K., & Kwan, Y.K. (2000). What are the determinants of the location of foreign direct investment? The Chinese experience. *Journal of International Economics*, 51, pp. 379–400.

de Boor, C. (2001). A Practical Guide to Splines (Applied Mathematical Sciences Vol. 27). Springer-Verlag, New York.

Du, J., Lu, Y., & Tao, Z. (2008). Economic institutions and FDI location choice: Evidence from US multinationals in China. *Journal of Comparative Economics*, 36(3), pp. 412–429.

Engle, R.F., Granger, C., Rice, J., & Weiss, A. (1986). Semiparametric estimates of the relation between weather and electricity sales. *Journal of the American Statistical Association*, 81, pp. 310–320.

Gao, T. (2005). Labor quality and the location of foreign direct investment: Evidence from China. *China Economic Review*, 16(3), pp. 274-292.

Gao, G., Wang, D., & Che, Y. (2018). Impact of historical conflict on FDI location and performance: Japanese investment in China. *Journal of International Business Studies*, 49(8), pp. 1060-1080.

Head, K., & Mayer, T.(2004). Market potential and the location of Japanese investment in the European Union. *The Review of Economics and Statistics*, 86(4), pp. 959–972.

Huang, J. Z., Wu, C. O., Zhou, L. (2004). Polynomial spline estimation and inference for varying coefficient models with longitudinal data. *Statistica Sinica*, 14, pp. 763–788.

Iwai, N., & Thompson, S. R. (2012). Foreign direct investment and labor quality in developing countries. *Review of Development Economics*, 16(2), pp. 276-290.

Kang, S. J., & Lee, H. S. (2007). The determinants of location choice of South Korean FDI in China. *Japan and the World Economy*, 19(4), pp. 441-460.

Li, Q., & Stengos, T. (1996). Semiparametric estimation of partially linear panel data models. *Journal of Econometrics*, 71, pp. 289–397.

Liu, X., Lovely, M. E., & Ondrich, J. (2010). The location decisions of foreign investors in China: Untangling the effect of wages using a control function approach. *The Review of Economics and Statistics*, 92, pp. 160–166.

Nielsen, B. B., Asmussen, C. G., & Weatherall, C. D. (2017). The location choice of foreign direct investments: Empirical evidence and methodological challenges. *Journal of World Business*, 52(1), pp. 62–82.

Noorbakhsh, F., & Paloni A. (2001). Human capital and FDI inflows to developing countries: New empirical evidence. *World Development*, 29(9), pp. 1593–1610.

Robinson, P.M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, 56, pp. 931–954.

Rong, S., Liu, K., Huang, S., & Zhang, Q. (2020). FDI, labor market flexibility and employment in China. *China Economic Review*, 101449.

Thompson S.B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics*, 99(1), pp. 1–10.

Tomohara, A. (2017). Does immigration crowd out foreign direct investment inflows? Tradeoff between contemporaneous FDI-immigration substitution and ethnic network externalities. *Economic Modelling*, 64, pp. 40-47.

UNCTAD (1999). Word Investment Report 1999: Foreign direct investment and the challenge of development. ISBN: 92-1-112440-9.

Wei, Y., Liu, X., Parker, D., & Vaidya, K. (1999). The regional distribution of foreign direct investment in China. *Regional Studies*, 33(9), pp. 857–867.

Wood, S. (2017). Generalized Additive Models. New York: Chapman and Hall/CRC.

Yeaple, S.R. (2003). The complex integration strategies of multinational and cross country dependencies in the structure of foreign direct investment. *Journal of International Economics*, 60, pp. 293–314.

You, J., & Zhou, X. (2013). Efficient estimation in panel data partially additive linear model with serially correlated errors. *Statistica Sinica*, 23, pp. 271–303.

Zhang, K., & Markusen, J.R. (1999). Vertical multinational and host-country characteristics. *Journal of Development Economics*, 59, pp. 233–252.

Table 1: Variable definitions and descriptive statistics

Variable	Definition	Obs.	Mean	SD.	Min.	Max.
fdi	The total investment of FDI firms (Million RMB Yuan in 2005)	754	462560.9	844306.3	694.0	9271464.0
wage	The average wage of urban employee (RMB Yuan in 2005)	754	23838.2	18029.7	3769.3	110831.8
lbr_qlty	The annual high school graduates over population (Unit: ‰)	754	4.395	1.951	1.093	8.666
gdp	GDP (100 Million RMB Yuan in 2005)	754	9341.9	10995.4	177.5	72801.3
open	The total export and import over GDP (Unit: 10%)	754	3.000	3.792	0.168	22.029
govern	Local government fiscal expenditure over GDP	754	0.176	0.092	0.049	0.627
infra	The density of road and railway (Unit: km/km^2)	754	0.626	0.486	0.018	2.379
pct_primary	The percentage of primary sector in GDP (Unit: %)	754	14.547	8.161	0.291	36.445
pct_secondary	The percentage of secondary sector in GDP (Unit: %)	754	44.413	7.992	16.545	59.397
rt_tertiary	The ratio of GDP in tertiary sector to that in others	754	0.757	0.454	0.381	4.914
rt_ab65	The ratio of population above 65 to that between 15 and 64 (Unit: %)	493	12.765	2.673	6.951	22.689
aftr_WTO	Indicator of China's accession to WTO after 2001	754	0.654	0.476	0.000	1.000
aftr_GFC	Indicator of Global Financial Crisis after 2008	754	0.423	0.494	0.000	1.000

All the variables in nominal term are deflated with local Consumer Price Index taking 2005 as the base year. The number of provinces is $n = 29$, and the number of years is $T = 26$. The time span of the sample is from 1993 to 2018. For the variable *rt_ab65*, the time span is from 2002 to 2018. The after WTO dummy is defined as $aftr\_WTO_t = 1$ for $t \geq 2002$, and it equals to zero otherwise. The after GFC dummy is defined as $aftr\_GFC_t = 1$ for $t \geq 2009$, and it equals to zero otherwise.

Table 2: Estimation results of fixed effects models

	Expected sign	Fixed effects Model 1	Model 2	Model 3	Model 4
log wage (*lwage*)	−	0.210 (0.306)	0.132 (0.324)	-0.298 (0.348)	-0.317 (0.451)
labor quality (*lbr_qlty*)	+	0.005 (0.023)	-0.416 (0.292)	$-0.584^{\dagger}$ (0.319)	-0.179 (0.345)
interactive term between wage and labor quality (*lwage*×*lbr_qlty*)			0.043 (0.030)	$0.060^{\dagger}$ (0.033)	0.022 (0.034)
log GDP (*lgdp*)	+			0.385 (0.349)	-0.182 (0.410)
total export and import over GDP (*open*)	+			0.023 (0.019)	0.009 (0.011)
local government fiscal expenditure over GDP (*govern*)	+/−			0.935 (0.798)	-0.060 (0.862)
density of road and railway (*infra*)	+			0.149 (0.153)	0.143 (0.146)
percentage of primary sector in GDP (*pct_primary*)	−			-0.011 (0.016)	0.010 (0.017)
percentage of secondary sector in GDP (*pct_secondary*)	+			0.008 (0.011)	0.022* (0.011)
ratio of GDP in tertiary sector to that in others (*rt_tertiary*)	+			0.325*** (0.091)	0.649*** (0.101)
ratio of population above 65 (*rt_ab65*)	−				0.035* (0.015)
after-WTO time dummy (*aftr_WTO*)	+			-0.055 (0.100)	n.id.
after-GFC time dummy (*aftr_GFC*)	−			-0.345** (0.123)	-0.285* (0.132)
Wald statistics and p-values testing exclusion of variables					
lwage, *lwage*×*lbr_qlty*=0			3.089 [0.213]	3.475 [0.176]	0.537 [0.765]
lbr_qlty, *lwage*×*lbr_qlty*=0			2.087 [0.352]	3.448 [0.178]	3.066 [0.216]
All *lwage*, *lbr_qlty* and related variables=0		0.610 [0.737]	3.232 [0.357]	3.488 [0.322]	3.274 [0.351]
Linear time trend		Yes	Yes	Yes	Yes
Province FE		Yes	Yes	Yes	Yes
Adjusted R^2		0.692	0.696	0.716	0.655
Time span		1993–2018	1993–2018	1993–2018	2002–2018
Sample size nT		754	754	754	493

In the top panel, the numbers in the parentheses are two-way cluster-robust standard errors. In the middle panel, the numbers in the brackets are the p-values of the test statistics. The number of knots (spline base functions) are selected by AIC. The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$."n.id." means the variable is not identified in the model.

Table 3: Estimation results of FE models using different time periods

	Time periods			
	1993–1997	1998–2008	2009–2018	
log wage	3.353*	0.236	-0.766	-0.847
(*lwage*)	(1.405)	(0.539)	(0.976)	(0.933)
labor quality	6.110†	0.128	-1.486	-1.549
(*lbr_qlty*)	(3.235)	(0.727)	(1.418)	(1.399)
interactive term between wage and labor quality	-0.663†	-0.015	0.153	0.159
(*lwage*×*lbr_qlty*)	(0.358)	(0.079)	(0.133)	(0.132)
log GDP	-1.687	0.971	-0.843*	-0.717†
(*lgdp*)	(1.172)	(0.823)	(0.410)	(0.434)
total export and import over GDP	-0.008	0.070*	-0.051†	-0.041
(*open*)	(0.037)	(0.030)	(0.030)	(0.031)
local government fiscal expenditure over GDP	-2.634	0.316	-0.662	-0.582
(*govern*)	(5.459)	(2.278)	(1.358)	(1.333)
density of road and railway	-3.592	-0.320†	0.860	0.826
(*infra*)	(2.398)	(0.165)	(0.561)	(0.550)
percentage of primary sector in GDP	-0.318*	-0.042	0.001	0.004
(*pct_primary*)	(0.140)	(0.034)	(0.027)	(0.029)
percentage of secondary sector in GDP	-0.285*	0.015	0.031*	0.033*
(*pct_secondary*)	(0.117)	(0.030)	(0.016)	(0.015)
ratio of GDP in tertiary sector to that in others	-7.325*	-0.047	0.469	0.510
(*rt_tertiary*)	(3.146)	(0.460)	(0.392)	(0.378)
ratio of population above 65				0.019
(*rt_ab65*)				(0.017)
after-WTO time dummy	n.id.	0.045	n.id.	n.id.
(*aftr_WTO*)		(0.128)		
after-GFC time dummy (*aftr_GFC*)	n.id.	n.id.	n.id.	n.id.
Wald statistics and p-values testing exclusion of variables				
lwage, *lwage*×*lbr_qlty*=0	5.716†	0.310	1.554	1.548
	[0.057]	[0.856]	[0.460]	[0.461]
lbr_qlty, *lwage*×*lbr_qlty*=0	3.585	0.042	8.720*	8.766*
	[0.167]	[0.9792]	[0.013]	[0.012]
All *lwage*, *lbr_qlty* and	5.810	0.310	11.094*	10.683*
their interaction=0	[0.121]	[0.958]	[0.011]	[0.014]
Linear time trend	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Adjusted R^2	0.046	0.558	0.699	0.700
Number of provinces	29	29	29	29
Sample size nT	145	319	290	290

In the top panel, the numbers in the parentheses are two-way cluster-robust standard errors. In the middle panel, the numbers in the brackets are the p-values of the test statistics. The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. "n.id." means the variable is not identified in the model.

Table 4: Estimation results of FE models using different locations

	Geographic locations		
	East	Middle	West
log wage	0.152	-0.945***	-0.251
(*lwage*)	(0.446)	(0.251)	(0.524)
labor quality	-0.383	-0.588	-0.534
(*lbr_qlty*)	(0.268)	(0.361)	(0.550)
interactive term between wage and labor quality	0.036	0.063	0.064
(*lwage*×*lbr_qlty*)	(0.026)	(0.038)	(0.059)
log GDP	-0.152	1.407**	0.493
(*lgdp*)	(0.434)	(0.518)	(0.578)
total export and import over GDP	0.024	0.153**	-0.071
(*open*)	(0.021)	(0.058)	(0.092)
local government fiscal expenditure over GDP	-2.977	2.348	-0.025
(*govern*)	(2.465)	(1.709)	(0.575)
density of road and railway	0.178	0.176	-0.110
(*infra*)	(0.287)	(0.202)	(0.299)
percentage of primary sector in GDP	-0.014	0.066	0.133
(*pct_primary*)	(0.026)	(0.060)	(0.085)
percentage of secondary sector in GDP	-0.002	0.082	0.142†
(*pct_secondary*)	(0.023)	(0.070)	(0.084)
ratio of GDP in tertiary sector to that in others	0.261	3.597†	5.398†
(*rt_tertiary*)	(0.199)	(2.000)	(3.091)
ratio of population above 65			
(*rt_ab65*)			
after-WTO time dummy	-0.069	0.098	0.024
(*aftr_WTO*)	(0.139)	(0.096)	(0.097)
after-GFC time dummy	-0.122	-0.455**	-0.442*
(*aftr_GFC*)	(0.165)	(0.151)	(0.199)
Wald statistics and p-values testing exclusion of groups of variables			
lwage, *lwage*×*lbr_qlty*=0	2.120	23.864***	1.252
	[0.347]	[0.000]	[0.535]
lbr_qlty, *lwage*×*lbr_qlty*=0	2.120	2.758	1.938
	[0.346]	[0.252]	[0.379]
All *lwage*, *lbr_qlty* and their	3.144	23.888***	1.943
interaction=0	[0.370]	[0.000]	[0.584]
Linear time trend	Yes	Yes	Yes
Province FE	Yes	Yes	Yes
Adjusted R^2	0.760	0.860	0.664
Number of provinces	11	8	10
Time span	1993–2018	1993–2018	1993–2018

In the top panel, the numbers in the parentheses are two-way cluster-robust standard errors. In the middle panel, the numbers in the parentheses are the p-values of the test statistics. The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. The variable *rt_ab65* is not included in the three models due to severely decreasing the sample sizes if otherwise.

Table 5: Estimation results of partially linear models

	Partially linear models			
	Model PL1	Model PL2	Model PL3	Model PL4
$f_1(lwage)$	Included	Included	Included	Included
$f_2(lbr\_qlty)$	Included	Included	Included	Included
$f_3(lwage,\ lbr\_qlty)$		Included	Included	Included
log GDP			0.345	-0.122
($lgdp$)			(0.303)	(0.360)
total export and import over GDP			0.027	0.025†
($open$)			(0.020)	(0.014)
local government fiscal expenditure over GDP			0.961	0.685
($govern$)			(0.792)	(0.695)
density of road and railway			0.068	0.147
($infra$)			(0.148)	(0.144)
percentage of primary sector in GDP			-0.032*	-0.004
($pct\_primary$)			(0.014)	(0.012)
percentage of secondary sector in GDP			-0.006	0.012
($pct\_secondary$)			(0.011)	(0.009)
ratio of GDP in tertiary sector to that in others			-0.174	0.132
($rt\_tertiary$)			(0.181)	(0.172)
ratio of population above 65				0.045**
(rt_ab65)				(0.015)
after-WTO time dummy			-0.100	n.id.
($aftr\_WTO$)			(0.078)	
after-GFC time dummy			-0.196**	-0.126†
($aftr\_GFC$)			(0.071)	(0.068)
Wald statistics and p-values testing exclusion of groups of variables				
$f_1(lwage)$=0	9.69*	2.78	7.246†	8.564*
	[0.021]	[0.428]	[0.064]	[0.036]
$f_2(lbr\_qlty)$=0	25.3***	11.9**	8.656*	5.900
	[0.000]	[0.008]	[0.034]	[0.117]
$f_3(lwage,\ lbr\_qlty)$=0		365.2***	29.99***	17.140*
		[0.000]	[0.000]	[0.047]
Linear time trend	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
# of splines	6	15	15	15
Adjusted R^2	0.734	0.742	0.747	0.715
Time span	1993–2018	1993–2018	1993–2018	2002–2018
Sample size nT	754	754	754	493

In the top panel, the numbers in the parentheses are two-way cluster-robust standard errors. In the middle panel, the numbers in the parentheses are the p-values of the test statistics. The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. "n.id." means the variable is not identified in the model.

Table 6: Linearity tests of nonlinear functions in Model PL3

Null hypothesis	Wald statistic	p-value
Linearity tests for single functions		
$f_1(lwage)$ is linear.	3.322	[0.190]
$f_2(lbr\_qlty)$ is linear.	$5.804^{\dagger}$	[0.055]
$f_3(lwage, lbr\_qlty)$ is linear.	27.707^{***}	[0.000]
Joint linearity test for all three functions		
f_1, f_2 and f_3 are all linear.	43.735^{***}	[0.000]

The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 7: Linearity tests of nonlinear functions in Model PL4

Null hypothesis	Wald statistic	p-value
Linearity tests for single functions		
$f_1(lwage)$ is linear.	3.437	[0.179]
$f_2(lbr\_qlty)$ is linear.	$4.806^{\dagger}$	[0.090]
$f_3(lwage, lbr\_qlty)$ is linear.	46.86^{***}	[0.000]
Joint linearity test for all three functions		
f_1, f_2 and f_3 are all linear.	61.28^{***}	[0.000]

The asterisks in the superscript denote the significance level, † $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

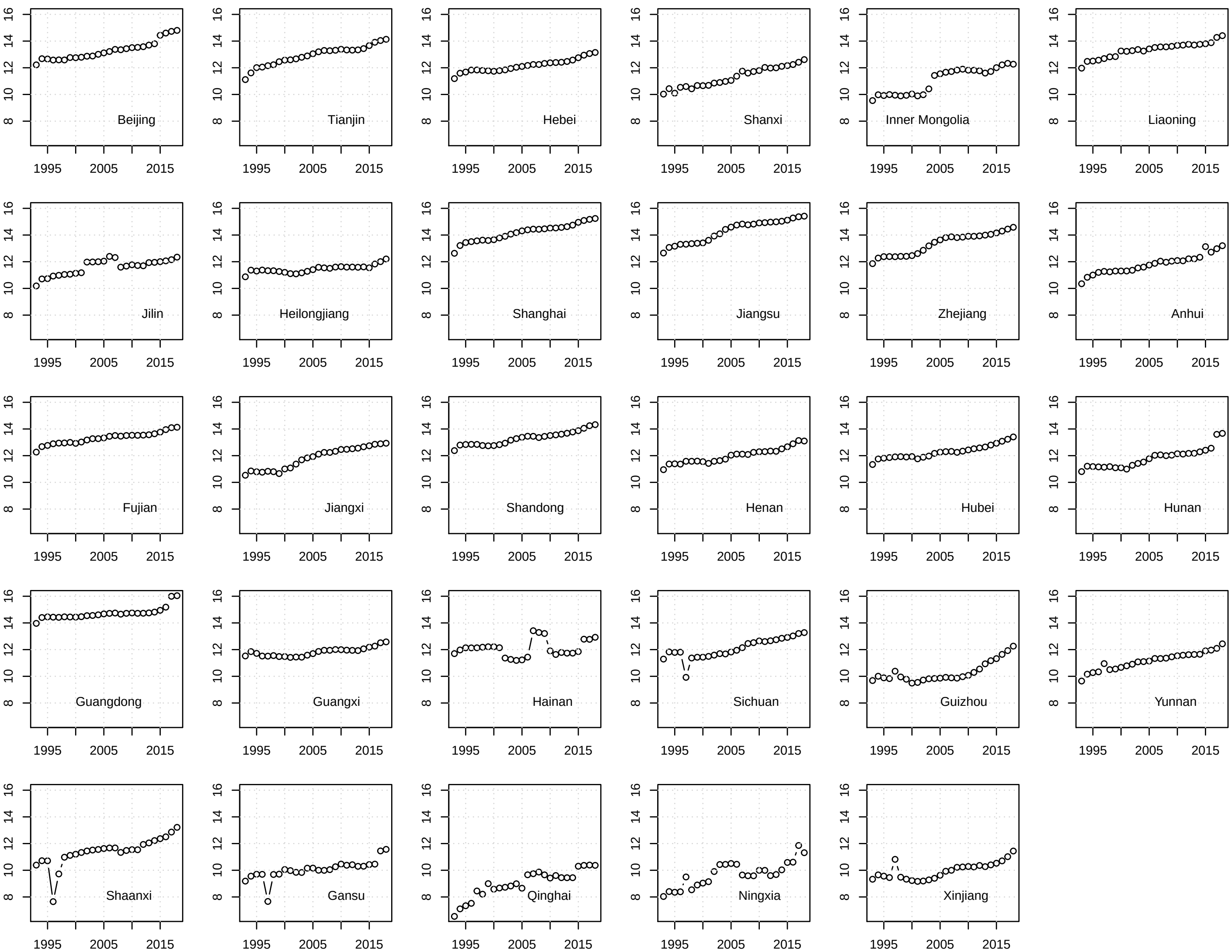

Figure 2: Log FDI (*lfdi*) series of 29 provinces in China (1993–2018)

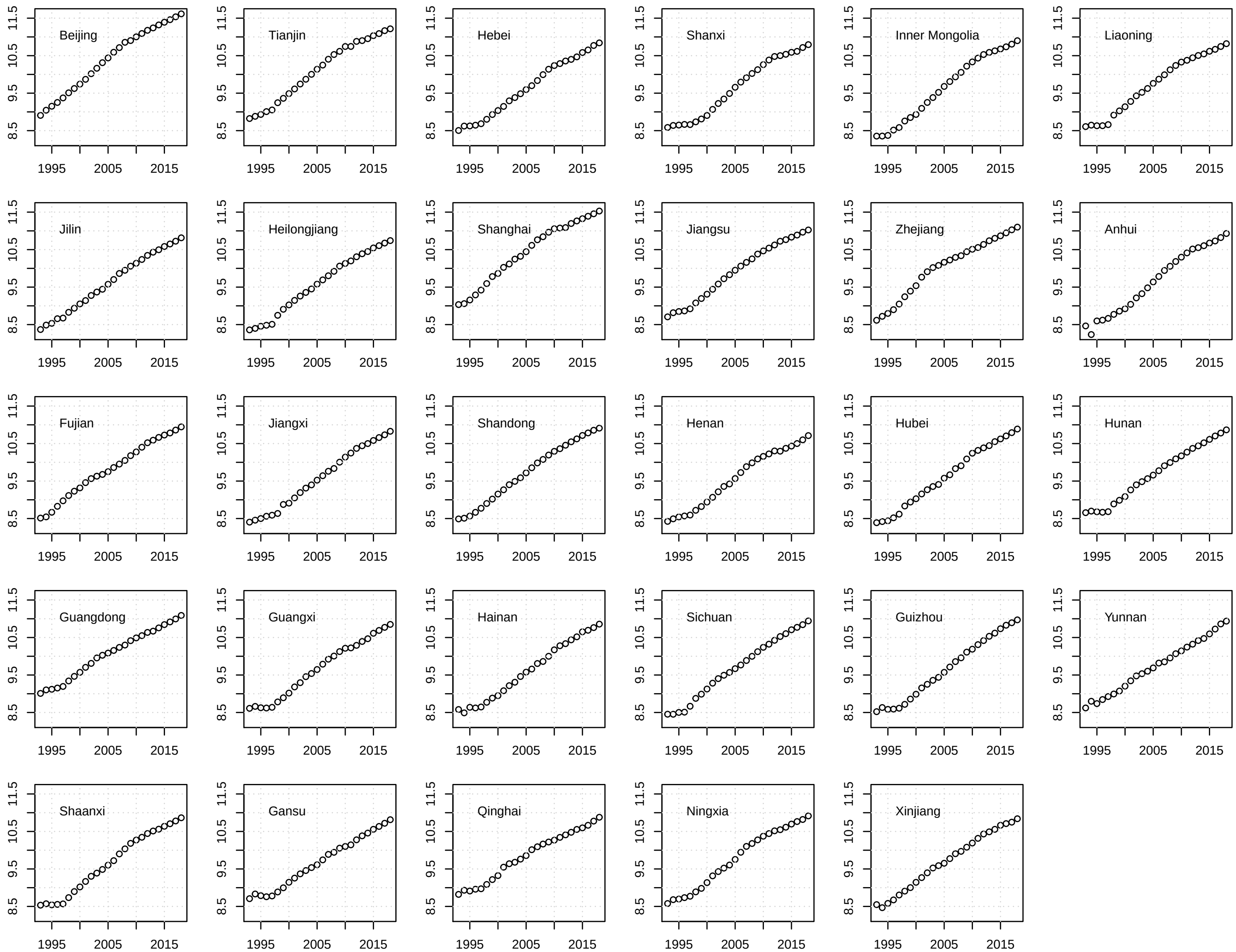

Figure 3: Log wage (*lwage*) series of 29 provinces in China (1993–2018)

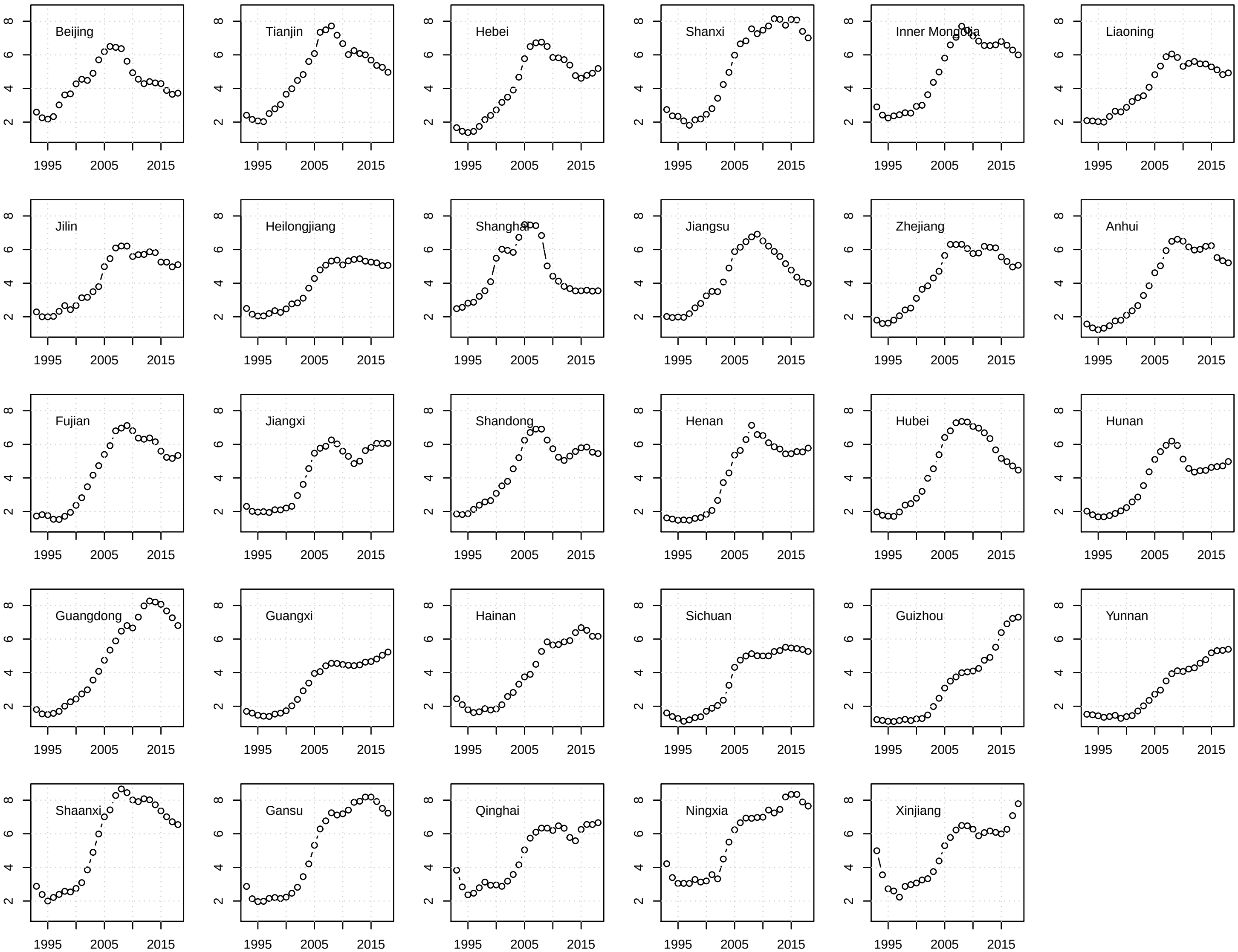

Figure 4: Labor quality (*lbr_qlty*) series of 29 provinces in China (1993–2018)

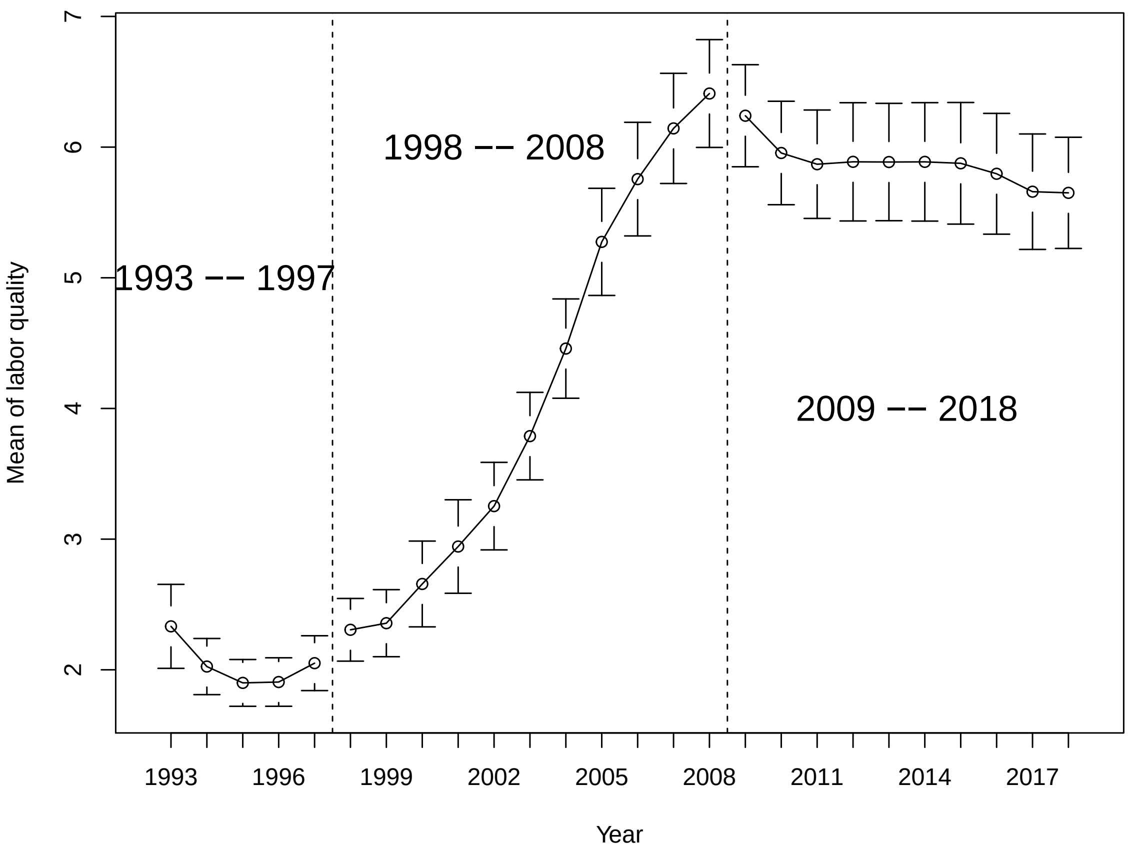

Figure 5: Average labor quality of 29 provinces in China (1993 – 2018)

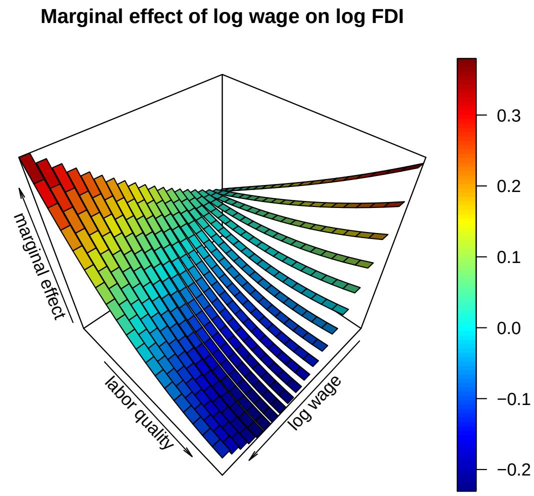

Figure 6: Estimated marginal effect of log wage on log FDI as a function of log wage and labor quality: $\dfrac{\partial \widehat{g}}{\partial lwage}(lwage, lbr\_qlty)$

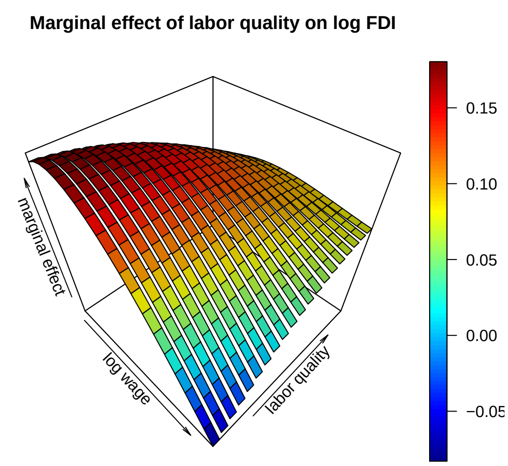

Figure 7: Estimated marginal effect of labor quality on log FDI as a function of log wage and labor quality: $\dfrac{\partial \widehat{g}}{\partial lbr\_qlty}(lwage, lbr\_qlty)$

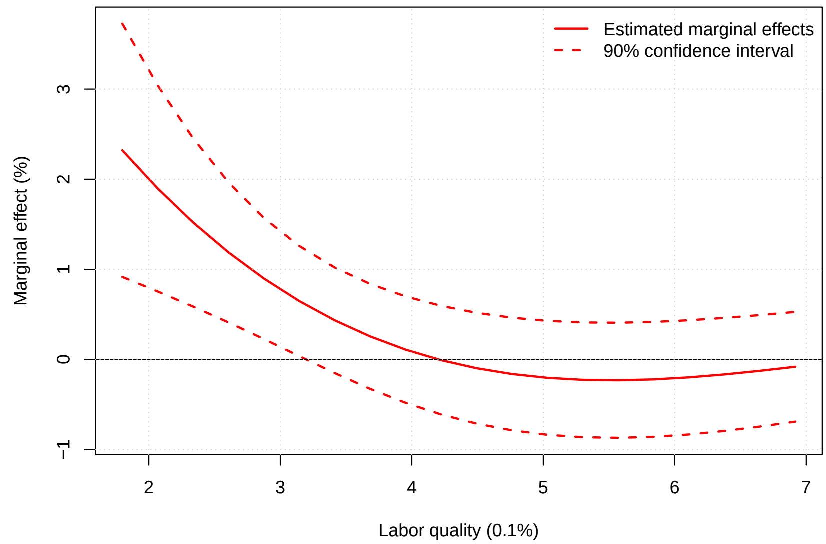

Figure 8: Estimated marginal effect of log wage on log FDI as a function of labor quality: $\dfrac{\partial \widehat{g}}{\partial lwage}(\overline{lwage}_{60\%pctl}, lbr\_qlty)$ where $lwage$ is fixed as its 60% percentile value.

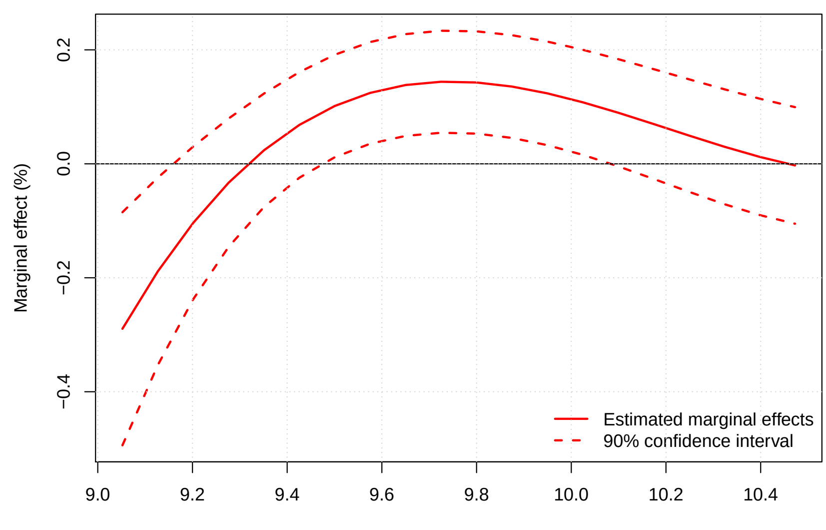

Figure 9: Estimated marginal effect of labor quality on log FDI as a function of log wage: $\dfrac{\partial \widehat{g}}{\partial lbr\_qlty}(lwage, \overline{lbr\_qlty}_{50\%pctl})$ where lbr_qlty is fixed as its sample median value.

联系邮箱：xuqiy@163.com