Dances with Chinese data: are the reform period Chinese provincial panel data reliable?

He, Qichun

December 2011

Dances with Chinese Data: Are the Reform Period Chinese Provincial Panel Data Reliable?

Qichun He*

December, 2011

Abstract

There is a debate over the reliability of the Chinese data (e.g., Young, 2003; Holz, 2003, 2006). In this paper we test the Chinese provincial panel data for the period 1978-2002 against the predictions from the technology diffusion model. We find that the estimated coefficient on initial real GDP per worker is negative and significant, showing strong evidence of conditional convergence; the estimated coefficients on secondary school human capital investment rate and labor force growth are positive and negative respectively, significant at the 5% level, in both LSDV (Least squares dummy variables) estimation and system GMM (Generalized method of moments) estimation that overcomes the endogeneity of these variables. The test accepts that the estimate coefficients on physical capital and human capital investment rates are equal, with absolute magnitudes about half of that on labor force growth in LSDV estimation. The estimated coefficient on the FDI to GDP ratio that captures technology diffusion is insignificant in LSDV estimation but becomes significant in system GMM estimation. All these are consistent with the technology diffusion model (and the augmented Solow model). Therefore, the reform period Chinese provincial panel data may be reliable for growth regressions.

JEL Classification: C23

Keywords: Technology Diffusion, Convergence Equation, Panel Data, System GMM (Generelized method of moments)

*China Economics and Management Academy, Central University of Finance and Economics, No. 39 South College Road, Haidian District, Beijing, China. 100081. Email: qichunhe@gmail.com, heqichun@cufe.edu.cn.

1 Introduction

The People's Republic of China (hereafter China) has achieved impressive growth in the past three decades. The Chinese experience may help to solve many theoretical debates and offer useful lessons for other countries. Unfortunately, there is a debate over the reliability of the Chinese data. Some authors emphasize the statistical discrepancies of the Chinese data (see e.g., Hsueh and Li, 1999; Young, 2003), while others argue that the Chinese data are reliable and the criticism is due to misunderstanding (see e.g., Chow, 1993; Holz, 2003, 2006).[1] Some authors have used the Chinese macro data to test theories (see e.g., Zhang and Zou, 1998; Démurger, 2001; Narayan et al., 2008). Therefore, it is appealing to investigate whether the controversial Chinese macro data are reliable for researches on its growth and business fluctuations. In this paper we take an indirect approach by testing the Chinese macro panel data against the augmented Solow model and a technological diffusion model to see whether they are reliable for growth regressions.

Why does such an indirect approach using growth regressions in the line of Barro (1991) and Mankiw et al. (1992) seem meaningful? There is already a large literature that estimates the production function for China (see e.g., Chow and Li, 2002; Li, 2003). But as Li (2003) summarizes, this approach cannot deal with the potential endogeneity of the explanatory variables. Moreover, it cannot isolate the contributions of human capital and technological change to output growth. Our approach complements and improves over this approach in the following sense. First, it is similar to production function estimation with a particular production function form, namely the augmented Solow model. Although this is a strong assumption, it is widely used in the empirical growth literature (e.g., Mankiw et al.). In so doing, we can produce results on the determinants of Chinese economic growth that are comparable to the existing large literature on empirical growth pioneered by Barro and Mankiw et al.. If it turns out the Chinese provincial panel data perform well in such framework, the potential benefit is substantial. As mentioned, China has achieved impressive growth after its reform and opening-up in 1978. The Chinese experience should provide useful lessons for other developing and transitional economies. Moreover, it may provide a natural experiment that may help to solve many theoretical and empirical debates. One example would be whether Chinese reform merely follows growth (i.e., growth causes reform) or it has a causal effect on growth (i.e., reform causes growth). The reform could be any kind of reform such as financial reform, fiscal decentralization, trade liberalization, and even political decentralization. Each of these topics has a large literature that is full of hot debate. The growth regressions following Barro and Mankiw et al. may help more to solve these debates as explained below.

Second, our approach can avoid or address the aforementioned drawbacks of the production function estimation approach. Specifically, the empirical specification in our

[1] See Young (2003) for a thorough survey of the literature on the Chinese data. Because of his excellent reference to it, we shall omit detailed discussion of the literature.

approach naturally involves dynamic panel data, which allows us to use the system GMM (Generalized method of moments) estimation (see Arellano and Bover, 1995; Blundell and Bond, 1998; Roodman, 2006) to overcome the potential endogeneity of important explanatory variables of output growth. Arellano and Bover (1995) and Blundell and Bond (1998) show that system GMM estimator can dramatically improve efficiency and avoid the weak instruments problem in the first-difference GMM estimator. At the macro level, almost all explanatory variables of output growth may be endogenous to the growth process. For instance, human capital accumulation, technological progress, and all kinds of policies are all possibly endogenous due to the feedback effect from growth. The aforementioned debates are all partly related to the direction of causality between growth and its determinants. Therefore, our regression approach allows us to establish a causal relationship between growth and its determinants. Moreover, it can isolate the contributions of human capital and technological change to output growth.

We combine the technological diffusion model based on Acemoglu (2009, ch. 18) with the augmented Solow model from Mankiw et al. (1992) to describe the reform and opening-up period Chinese economy. The production function depends on technology, physical capital, human capital and raw labor. The technological progress of the reform and opening-up period Chinese economy depends on the absorption of world frontier technologies and its own domestic technological innovations. Following previous studies (see e.g., Findlay, 1978; Borensztein et al., 1998; Keller and Yeaple, 2003), foreign direct investment (FDI) is assumed to be the main channel for advanced technologies to be transferred to the backward Chinese economy. Solving the model and then approximating around the steady state, we derive an empirical formulation similar to the augmented Solow model (see Mankiw et al., 1992), with two additional independent variables capturing China's absorption of world frontier technologies and its domestic innovations.

Then we build the necessary macroeconomic series by choosing the most consistent data provided by the Statistical Yearbook of China (hereafter SYC). Specifically, we choose the period 1978-2002 to avoid the statistical adjustment on GDP (Gross Domestic Product) in 2005 – detailed later – by the National Bureau of Statistics of China (hereafter NBSC) (see Holz, 2008). We take five-year averages of the data to avoid the influence from business cycle fluctuations, which matches the political cycle in China. We use the nominal GDP and GDP indexes from SYC to calculate the provincial real GDP. We use the labor force data from SYC rather than the provincial statistical yearbooks to avoid the large upward adjustment in labor force in 1990 by some provinces (detailed in Young, 2003, section IV). Then we calculate the growth rate of real GDP per worker and initial real GDP per worker. Following Mankiw et al. (1992), we use both the primary and the secondary school enrollment rates to measure human capital investment rate. Following Borensztein et al. (1998), we use the ratio of nominal FDI to nominal GDP to measure technological diffusion. To avoid the deflators' problem on physical capital investment (see Young, section VI), we use nominal physical investment to nominal GDP to measure physical capital investment rate.

The reform period Chinese data provide a balanced panel data of 27 provinces and 5 time periods, which allows us to control for fixed time and province effects. We test the data against the predictions of the model. Both LSDV (Least squares dummy variables) and system GMM estimations yield similar results. The estimated coefficient on initial real GDP per worker is negative and significant at the 1% level, showing strong evidence of conditional convergence. The estimated coefficients on secondary school human capital investment rates and labor force growth are positive and negative respectively, significant at least at the 5% level. The test accepts that the estimate coefficients on physical capital and secondary school human capital investment rates are equal, with absolute magnitudes about half of that on labor force growth in LSDV estimation. In LSDV estimations, the estimated coefficient on physical capital investment rate is significant without the FDI to GDP ratio in the regression and insignificant with FDI/GDP in the regression, while it is always insignificant in system GMM estimation. The estimated coefficient on the FDI to GDP ratio is insignificant in LSDV estimation but becomes significant in system GMM estimation. Therefore, except for physical capital investment, the Chinese macro panel data fit the model well.

The paper is organized as follows. In section 2 we derive the empirical formulation and construct the variables. Section 3 presents the estimation results. Section 4 concludes.

2 The Data

2.1 Deriving the Empirical Specification

China has undertaken the market-oriented reform and opening-up in 1978. That is, China has not only made continuous efforts to reform its economic institutions, but also opened its borders to foreign investors and trade.[2] Therefore, the Chinese provinces can be treated as backward small open economies that rely on the absorption of technological expertises from abroad to achieve technological progress. Barro and Sala-i-Martin (2004, p. 350), for instance, have stated that the absorption of technological expertises from Hong Kong has been important for China's technological progress. Therefore, we use the technology diffusion and absorption model based on Acemoglu (2009, ch. 18) to derive the empirical formulation. Following previous works (e.g., Findlay, 1978; Keller and Yeaple, 2003), FDI is assumed to be the main channel for advanced technology to be transferred to the backward Chinese economy.

[2]Technological imitation from leading countries is emphasized by Deng (1975), the designer of the reform and opening-up and the leader of China after 1978.

For a Chinese province i in year t, its aggregate production function for a unique final good is

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

where K, H, and L are physical capital, human capital, and raw labor respectively. A_{it} is its level of technology, and $g_{it} = \frac{\dot{A}_{it}}{A_{it}}$ is the growth rate of technology. The output per effective labor is $y_{it} = k_{it}^{\alpha} h_{it}^{\beta}$, where the effective capital-labor ratio, k_{it}, and the effective human capital-labor ratio, h_{it}, evolve according to

$$\dot{k}_{it} = s_k y_{it} - \left(n + g_{it} + \delta\right) k_{it} \tag{2}$$

$$\dot{h}_{it} = s_h y_{it} - \left(n + g_{it} + \delta\right) h_{it}, \tag{3}$$

where s_k, s_h are exogenous physical and human capital investment rates respectively. n and δ are exogenous population growth rate and depreciation rate respectively.

The world technological frontier A_t^w is assumed to grow at an exogenous rate g^w. However, not all world frontier technologies are available for imitating. We assume that, at any time, the available pool of technology for imitating depends on how many foreign firms conduct direct investment in the backward economy, which is measured as inward FDI to GDP ratio (denoted as FDI_{it}). Therefore, following Acemoglu, we posit the following law of motion for technology:

$$\dot{A}_{it} = \sigma_{it} \cdot \left(A_t^w \cdot FDI_{it} - A_{it}\right) + \gamma_i A_{it}, \tag{4}$$

where the first term on the right-hand-side (RHS) of equation (4) measures the absorption/imitation of world technology and the second term, γ, measures domestic innovations. Technology absorption depends on the product of the absorptive capability (σ) and the technology gap between world technology frontier available for absorption and the domestic level of technology, $(FDI_{it} \cdot A_t^w - A_{it})$.

As in Acemoglu, we define the inverse of the distance to the world frontier, $a_{it} < 1$, as $a_{it} = \frac{A_{it}}{A_t^w}$. Using equation (4), we have

$$\dot{a}_{it} = \sigma_{it} \cdot FDI_{it} - \left(\sigma_{it} + g^w - \gamma_i\right) a_{it}. \tag{5}$$

We begin with the steady state. In the steady state, the technological progress rate of the small economy, g, is equal to g^w. And in steady state, $\dot{k}_{it} = 0$ and $\dot{h}_{it} = 0$. Then steady state output per effective labor can be solve as

$$y_i^* = \left(s_k\right)^{\frac{\alpha}{1-\alpha-\beta}} \left(s_h\right)^{\frac{\beta}{1-\alpha-\beta}} \left(n + g^w + \delta\right)^{-\frac{\alpha+\beta}{1-\alpha-\beta}}. \tag{6}$$

Approximating around the steady state, the speed of convergence is $\lambda = (1 - \alpha - \beta)\left(n + g^w + \delta\right)$.

Following the steps in Mankiw et al. (1992, p. 423), we end up with

$$\ln(y_{it}) - \ln(y_{it-1}) = -\left(1-e^{-\lambda}\right)\ln(y_{it-1}) + \left(1-e^{-\lambda}\right)\ln(y_i^*), \tag{7}$$

where $\ln(y_i^*)$ can be expressed as exogenous parameters as in equations (6). Since the above equation is output per effective labor, we transform it into output per labor. Output per labor is $\frac{Y}{L}$, which is equal to yA. Hence we have

$$\ln\left(\frac{Y}{L}\right)_{it} - \ln\left(\frac{Y}{L}\right)_{it-1} = \left[\ln(y_{it}) - \ln(y_{it-1})\right] + \left[\ln(A_{it}) - \ln(A_{it-1})\right]. \tag{8}$$

Combining equations (7) and (8) yields

$$\ln\left(\frac{Y}{L}\right)_{it} - \ln\left(\frac{Y}{L}\right)_{it-1} = -\left(1-e^{-\lambda}\right)\ln(y_{it-1}) + \left(1-e^{-\lambda}\right)\ln(y_i^*) + g_{it}. \tag{9}$$

The technological growth rate of the small economy, g_{it}, is

$$g_{it} = \frac{\dot{A}_{it}}{A_{it}} = \frac{\dot{a}_{it}}{a_{it}} + g^w = \frac{FDI_{it}}{a_{it}}\sigma_{it} - \sigma_{it} + \gamma_i. \tag{10}$$

According to equation (10), higher inflow of FDI will increase the technological growth rate of the backward economy. Substituting out g_{it} using equation (10) and $\ln(y_i^*)$ using equation (6) from equation (9), we have our final empirical specification as

$$\begin{aligned}\ln\left(\frac{Y}{L}\right)_{it} - \ln\left(\frac{Y}{L}\right)_{it-1} &= \frac{FDI_{it}}{a_{it}}\sigma_{it} - \sigma_{it} + \gamma_i - \left(1-e^{-\lambda}\right)\ln(y_{it-1}) \\ &+ \left(1-e^{-\lambda}\right)\frac{\alpha}{1-\alpha-\beta}\ln(s_k)_{it} + \left(1-e^{-\lambda}\right)\frac{\beta}{1-\alpha-\beta}\ln(s_h)_{it} \\ &- \left(1-e^{-\lambda}\right)\frac{\alpha+\beta}{1-\alpha-\beta}\ln(n_{it}+g^w+\delta). \end{aligned} \tag{11}$$

The last four terms in equation (11) are exactly the same as those in the augmented Solow model (see Mankiw et al., 1992). The first two terms on the RHS of equation (11) are new and capture the technological progress of the backward economy. Higher inflow of FDI would raise its growth rate. The domestic technological advances of the backward economy (γ) would also speed up the growth of the backward economy.

Specifically, we use the following formulation for empirical assessment:

$$\begin{aligned} growth_{it} &= \beta_0 \ln(\frac{FDI}{GDP})_{it} - \beta_1 \ln\left(\frac{GDP}{L}\right)_{i,t-1} + \beta_2 \ln(\frac{I}{GDP})_{it} \\ &+\beta_3 \ln(School)_{it} - \beta_4 \ln(n_{it}+g^w+\delta) + u_i + T_t + \varepsilon_{it} \end{aligned} \tag{12}$$

where $growth_{it}$ is the average annual growth of real GDP per worker for i^{th} province at period t; As in Borensztein et al. (1998), $\frac{FDI}{GDP}$ is the ratio of nominal FDI to nominal GDP, which measures technological diffusion/absorption from abroad. $\ln\left(\frac{GDP}{L}\right)_{i,t-1}$, real GDP per worker at the beginning of period t, controls for conditional convergence. $\frac{I}{GDP}$ and $School$ measure physical capital investment rate and human capital investment rate respectively. $(n_{it}+g^{w}+\delta)$ measures labor force growth. u_i and T_t stand for fixed province and time effects respectively. The fixed province effects could capture the time-invariant provincial technological advances.

According to equation (11), the coefficients on physical and human capital investment rates should be positive and equal, which are half of the absolute magnitude of the expected negative coefficient on labor force growth. Moreover, the coefficient on initial real GDP per worker is negative and that on $\frac{FDI}{GDP}$ is positive.

2.2 Constructing the Variables

To test the reliability of the Chinese data, we want to use as many data as possible. And we rely on the SYC for our data source. The reason is two-fold. First, many researchers would turn to SYC for macro data on China when they need them. This is because SYC is the most authoritative in providing the macro-data on China. Some (e.g., Chow, 1993; Holz, 2003) argue that the Chinese data are intrinsically consistent. Second, if we find out that the Chinese data from SYC are not far from truth, one may still be able to use them, provided that one deals with the statistical adjustments made by the NBSC.

2.2.1 The Data Sample

We try to include as many years as possible for the reform period. This is actually consistent with our model that studies a backward open economy. China has become an open economy in 1978. Moreover, we can avoid additional issues on the data of the pre-reform period as argued by Young. We will follow the common practice in the empirical growth literature to take five-year average of the reform period data. This yields six time periods, with the last one covering 2003-2007. However, the NBSC has made a significant statistical adjustment at the end of 2005 concerning GDP accounting (see Holz, 2008). Resultantly, the Chinese GDP has been revised upward around 16.8% due to the large adjustment in the service sector. Therefore, the data on GDP after year 2004 differ a lot from previous years. Although NBSC has provided a revised series of GDP data back to 1993, we use the old series simply for consistency. Therefore, we only include five time periods covering 1978-2002. The five-year averaging of the data matches exactly with the political cycle in China. The National People's Congress (NPC) and the Chinese People's Political Consultative Conference (CPPCC) are held in the same year every five years starting from 1978, setting up all the important economic policies in China.

Before 1998, among the 31 provincial governments in China, four are municipalities and four are autonomous regions. We delegate the usage 'province' to all. Before 1997, Chongqing was a city of Sichuan province, hence both of them are excluded from the sample. Hainan was part of Guangdong before it became an independent province. Since there is a complete set of data for Guangdong, it is kept in the data sample while Hainan is dropped. Tibet is excluded because there are many missing data. In summary, the data sample comprises panel data of 27 provinces and 25 years (1978-2002), which produces a balanced panel with 135 observations.

2.2.2 Data on FDI

The provincial FDI inflow data and the nominal GDP data are available from SYC. The FDI data are in US dollars, we multiply them by the fixed exchange rate of the Chinese currency (yuan) against the US dollar in each year to get the FDI data in Chinese currency. China has adopted the fixed exchange rate regime until year 2005 in which the government allows its currency to appreciate gradually each year. We then calculate the ratios of FDI over nominal GDP in each year as our measure of FDI, denoted by FDI/GDP. Although China opened its borders to foreign investors in 1978, many Chinese provinces did not receive FDI until the early 1980s. Only three Chinese provinces (Liaoning, Fujian and Guangdong) received any FDI during the period 1978-82. The datum for Liaoning for 1978-82 is zero in the Appendix because it is too small. The FDI data for eary years are from the China Center for Economic Research at Peking University.

2.2.3 Data on Real GDP

The SYC only provides nominal GDP and GDP indexes for each province. The provincial GDP data for early years are also from the China Center for Economic Research at Peking University. The problem with the implicit GDP deflator has been analyzed by Young (2003). Young decomposes the Chinese output into sectors and uses available price indices. He finds out that this would lower the aggregate GDP growth by 1.7%. Although we do not make systematic adjustments, we argue that this problem may apply to all Chinese provinces. In our cross province comparisons, this problem can be treated as measurement error on the dependent variable. With the nominal GDP and the GDP indexes and 1978 as our base year, the real GDP can be calculated as follows. We multiply the nominal GDP in 1978 by the GDP index in that year then divide the result by 100. The GDP deflator, which is not needed in our analysis, can then be backed out.

2.2.4 Data on Labor Force

To calculate our dependent variable, we need data on the labor force. However, there is a large statistical adjustment in 1990 on labor force. This has been analyzed in Young (1233-1234). For instance, the provincial statistical bureau of Jiangsu reported its labor force by using a new measurement detailed in Young. Resultantly, its labor force jumps from 35.19 million in 1989 to 42.25 million in 1990, while the SYC lists its labor force at

35.69 million in 1990. The provincial statistical bureau reports 6.56 million more workers. The provincial statistical bureau of Jiangsu should revise its labor force data before 1989 accordingly, but it did not. The same happened to Hubei province. It revised, using the new accounting method, its labor force from 24.33 million in 1989 to 30.40 million in 1990, while the SYC lists its labor force at 24.79 million in 1990. Around half of Chinese provinces made the changed in 1990. One can infer that it is not the case that the provincial statistical bureau has made up the numbers. Instead, it is just the change in statistical caliber as detailed in Young. Fortunately, SYC has maintained the original statistical caliber and provided the data on provincial labor force.[3] Therefore, this relative more consistent series provided by SYC allow us to cover the periods before and after 1990 to avoid "spurious labor force growth" (Young, p. 1234).

Now with the labor force data and the real GDP data, we can calculate two needed variables: the growth rate of real GDP per worker (our dependent variable), and the labor force growth rate. Labor force growth, $(n + g^w + \delta)$, is measured as labor force growth rate (n) plus 0.08 that is assumed for $(g^w + \delta)$. That is, we assume a 2% world annual growth and a 6% depreciation rate for China. As in Mankiw et al. (1992), our result is insensitive to the assumed number for $(g^w + \delta)$. The labor force data are also used to measure the human capital investment rate.

2.2.5 Data on Human Capital

Based on the convergence formulation, human capital investment rate is an important determinant of growth, which has been overlooked in the studies on China in Weeks and Yao (2003). Mankiw et al. (1992) use the secondary school enrollment rate – the ratio of the secondary school enrollment to the working age population – to measure human capital investment rate. They pointed out that they ignore education at the primary and higher levels. The SYC provides complete data on the student enrollments for all levels of education in China: the primary, the secondary and the higher education levels. Since China is a backward country, we, following Mankiw et al., have two measures of human capital investment rate: SCHOOL01 is the ratio of primary school enrollment to the labor force; SCHOOL02 is the ratio of secondary school enrollment (grade 7 to 12) to labor force. There are studies that group all levels of education together with a quality measure (e.g., years of schooling and student performance) as weight (see e.g., Wößmann, 2002). We do not follow this approach because the quality of education has its own measurement problem detailed in Young (section V). The student enrollment in China may be one of the most accurately documented series in SYC.

[3] For the majority of the years and provinces, the labor force data provided by SYC seem reasonable. However, we also find out some rare anomaly in it. For instance, the labor force datum for Beijing jumps to 7.99 million in 2002 from 6.29 million in 2001, while the provincial statistical yearbook lists the numbers in 2002 and 2001 as 6.79 and 6.29 million respectively.

2.2.6 Data on Physical Capital

China's physical capital investment generates the largest controversy in previous literature (see Chow, 1993; Hsueh and Li, 1999; Young, section VI; Perkins and Rawski, 2008). According to Young, the deflator of physical capital investment (the gross capital formation in SYC) has been downwardly reported by the Chinese provincial statistical bureaus. Therefore, if one uses the gross capital formation and its indexes to calculate real investment, some provinces would have unbelievably high real investment rates. For instance, the rates in some years of Tianjin and Gansu are over 100%. Without systematic adjustment, we resort to use the nominal invest rate, which is the ratio of nominal physical capital investment to the nominal GDP. This would avoid the deflators' problem and the over 100% invest rates. The average nominal investment rate for all the provinces in our data sample is 39% comparing to 41% for the average real investment rates. Table 1 lists the summary statistics of our data. The detailed data are listed in the Appendix.

[Table 1 Here]

3 Estimation Results

3.1 Regressions Based on Augmented Solow Model

In this paper, to fully test the reliability of the Chinese macro panel data, we first report the results by dropping the FDI/GDP from our empirical formulation (11). In so doing, it becomes the empirical formulation derived from the standard augmented Solow model (see Mankiw et al., 1992). It has been widely used in cross-country growth regressions. Therefore, it is meaningful to see how well the Chinese model fits it.

3.1.1 LSDV Regressions

We first use LSDV estimation. That is, we use OLS (Ordinary least squares) estimation that includes 27 province dummies and 5 time dummies. Table 2 summarizes the results.

Column 2.1 in Table 2 reports the LSDV results with secondary school enrollment rate, ln(SCHOOL02), as the measure of human capital investment rate. One can see that the estimated coefficient on initial real output per worker is negative and significant at the 1% level, showing strong evidence of conditional convergence. This is consistent with previous studies such as Weeks and Yao (2003). That is, after controlling for other factors, richer provinces tend to grow slower, consistent with augmented Solow model. The speed of convergence is 5.8% per year, similar to the 4-6% per year predicted by the augmented Solow model but larger than the 2% per year in empirical studies (see Weeks and Yao, 2003). The estimated coefficient on $ln\left(\frac{I}{GDP}\right)$ is positive and significant at the 5% level. That is, a higher domestic physical capital investment rate is associated with a higher rate of ecoomic growth, consistent with augmented Solow model. The estimated coefficient on ln(SCHOOL02) is positive and significant at the 5% level, while that on $ln\left(n + g^w + \delta\right)$ is negative and significant at the 1% level. In other words, a higher secondary school human capital investment rate is associated with a higher rate of ecoomic growth, consistent with augmented Solow model. A higher rate of labor force growth is associated with a lower rate of economic growth, consistent with augmented Solow model.

According to the prediction of the theory, we test whether the estimated coefficients on $ln\left(\frac{I}{GDP}\right)$ and ln(SCHOOL02) are equal. The F-test yields a p-value of 0.81, meaning we accept the null that the coefficients on $ln\left(\frac{I}{GDP}\right)$ and ln(SCHOOL02) are equal. Similarly, we test whether the estimated coefficient on $ln\left(\frac{I}{GDP}\right)$ is half of the absolute magnitude of that on $ln\left(n + g^w + \delta\right)$. The F-test yields a p-value of 0.28, supporting the theoretical prediction. According to the F-test, we also accept the null hypothesis that the estimated coefficient on ln(SCHOOL02) is half of the absolute magnitude of that on $ln\left(n + g^w + \delta\right)$. The adjusted R-squared is 0.81, meaning our regression explains more than 80% of the variation in the growth of real GDP per worker. Taken together, the often-suspected Chinese aggregate data fit the model quite well.

Column 2.2 in Table 2 reports the LSDV results with primary school enrollment rate, ln(SCHOOL01), as the measure of human capital investment rate. One can see that the estimated coefficient on ln(SCHOOL01) is negative and insignificant, while those on the other variables are still significant at the 5% level with quite similar magnitudes to those in column 2.1. Column 2.3 in Table 2 reports the LSDV results with both ln(SCHOOL01) and ln(SCHOOL02) in the regression. One can observe that the results on the main variables are also very similar to those in column 2.1.

From the LSDV results in Table 2, one can observe that the results on initial real GDP per worker, physical capital investment rate, and labor force growth are very stable. Using secondary school enrollment rate instead of the primary school enrollment rate to measure human capital investment rate improves the model fit.

[Table 2 Here]

3.1.2 System GMM Estimation

Our model has the characteristics listed in Roodman (2006), namely, "small T, large N; a linear functional relationship with a single left-hand-side variable that is dynamic, depending on its own past realizations; independent variables that are not strictly exogenous; fixed individual effects; heteroskedasticity and autocorrelation within individuals, but not across them." The dynamic structure of the model makes the LSDV estimators biased and inconsistent. Arellano and Bover (1995) and Blundell and Bond (1998) show that system GMM estimator can dramatically improve efficiency and avoid the weak instruments problem in the first-difference GMM estimator. Therefore, we re-estimate our model with system GMM estimator. In using the system GMM, we treat initial real GDP per worker as predetermined, and all the other main independent variables as endogenous. Following Roodman (2006), the province dummies are excluded, while the time dummies are used as exogenous instruments to make sure that the number of instruments is smaller than the number of groups (i.e., 27).[4] The results are reported in Table 3.

Column 3.1 in Table 3 reports the system GMM estimation results with secondary school enrollment rate as the measure of human capital investment rate. Both the Sargan and the Hansen tests for over-identifying restrictions confirm that the instrument set can be considered valid. The F-test shows that the overall regression is significant. The Arellano-Bond test accepts the hypothesis of no autocorrelation of the second order, supporting the model specification. Comparing with the LSDV estimation results in Table 2, one can see that the estimated coefficient on initial real GDP per worker remains negative and significant at the 1% level, with a larger magnitude. The estimated coefficient on $ln\left(\frac{I}{GDP}\right)$ becomes negative and insignificant at the 10% level. The estimated coefficient on $ln\left(n+g^{w}+\delta\right)$ remains negative but becomes significant at the 5% level, with a similar magnitude. The estimated coefficient on ln(SCHOOL02) remains positive but becomes significant at the 1% level, with a much larger magnitude.

Column 3.2 in Table 3 reports the system GMM results with primary school enrollment rate as the measure of human capital investment rate. One can see that the estimated coefficient on ln(SCHOOL01) is negative and insignificant, while those on the other variables are still significant at the 5% level. Moreover, the Hansen test shows that the instruments are invalid. This may be due to omitting other important variables such as the secondary school enrollment rate. Column 3.3 in Table 3 presents the results with both ln(SCHOOL01) and ln(SCHOOL02) in the regression. One can observe that the results on the main variables are also very similar to those in column 3.1.

Comparing the LSDV and system GMM estimation results, one can observe that the estimated coefficient on initial real GDP per worker is always negative and significant at the 5% level, showing strong evidence of conditional convergence. The estimated coefficient on labor force growth is always negative and significant at the 5% level. The estimated coefficient on human capital investment rate measured by secondary school enrollment rate is always positive and significant at the 5% level. All these are consistent with the augmented Solow model. The estimated coefficient on physical capital investment rate, however, is significant in LSDV estimation but insignificant in system GMM estimation. Nevertheless, even with real physical capital investment rate, Weeks and Yao (2003) still show that its estimated coefficient is insignificant in system GMM estimation.

[Table 3 Here]

[4] Our command in STATA10 is "xtabond2 growth lninigdp lninvest lnlabor lnschool2 t1-t5, gmm(L.(growth lninvest lnlabor lnschool2) lninigdp, lag(2 2) eq(diff)) iv(t1-t5) robust small", where t1-t5 are time dummies.

3.2 Regressions Based on Technological Diffusion Model

Now we report the results by including FDI/GDP in our empirical formulation. That is, we are using equation (12). It is derived from a combination of the technological absorption model and the augmented Solow model. However, as discussed, because of the lack of FDI to the Chinese provinces for the period 1978-82, we now have 110 observations. Moreover, given the results in the previous section, we use the secondary school enrollment rate to measure human capital investment rate.

3.2.1 LSDV Regressions

The LSDV estimation results are presented in column 2.4 in Table 2. One can observe that the estimated coefficients on initial real GDP per worker and $ln\,(n + g^w + \delta)$ remain negative and significant at the 1% level, with magnitudes similar to those in LSDV estimation. The estimated coefficient on $ln\left(\frac{I}{GDP}\right)$ remains positive but becomes insignificant at the 10% level, with a similar magnitude. The estimated coefficient on ln(SCHOOL02) remains positive and significant at the 5% level, with a much larger magnitude. Therefore, the results are robust when we further include FDI/GDP to control for technological absorption. However, although the estimated coefficient on $\ln\left(\frac{FDI}{GDP}\right)$ is positive, it is insignificant at the 10% level. This is consistent with previous studies (see e.g., Borensztein et al., 1998, and their discussion and explanation on this issue).

3.2.2 System GMM Estimation

The system GMM estimation results with $\ln\left(\frac{FDI}{GDP}\right)$ as an additional regressor are presented in column 3.4 in Table 3. Both the Sargan and the Hansen tests for over-identifying restrictions confirm that the instruments are valid. Moreover, the p-values of both Sargan and Hansen over-identification tests become much larger comparing to column 3.1. Although it is well-known that over-identification tests have very little statistical power, the much larger p-values of these tests show that there is little evidence of omitted variable bias. That is, by further including $\ln\left(\frac{FDI}{GDP}\right)$ in the regression, our model specification is more suitable for the reform period Chinese economy. The F-test shows that the overall regression is significant. The Arellano-Bond test rejects the hypothesis of no autocorrelation of the first order at the 5% level, and accepts the null hypothesis of no autocorrelation of the second order. All these support our model specification.

One can observe that the results on other variables are pretty similar to those in column 3.1 without $\ln\left(\frac{FDI}{GDP}\right)$ in the regression. The estimated coefficient on $\ln\left(\frac{FDI}{GDP}\right)$ remains positive, which becomes significant at the 1% level. That is, after overcoming its potential endogeneity problem, FDI is conducive to the economic growth in China. This confirms the technological diffusion model in section 2.1.

4 Conclusions

The importance to study the Chinese economy goes without emphasizing. The statistical data are vital for any serious research on the Chinese economy. In this paper we focus the macro-data provided by SYC. As Young argues: "The data of most economies are filled with apparently inconsistent series...Consequently, the only value added in a paper of this sort lies in its treatment and exposition of the data." We build necessary variables by choosing the consistent series and using alternative measures to avoid the detected inconsistency. We end up with a Chinese macro panel data on 27 provinces for the reform and opening-up period 1978-2002.

Using a technological diffusion/absorption model to describe the Chinese economy during our sample period, we derive the empirical specification. It is similar to the augmented Solow model with one additional independent variable capturing the technological absorption from abroad. We then use the Chinese macro panel data to test how well they fit the model. We find that the results on conditional convergence (measured by initial real GDP per worker), human capital investment rate (proxied by secondary school enrollment rate), and labor force growth confirm the theoretical predictions. The results are robust to the inclusion/exclusion of technological progress from absorbing world frontier technologies. Moreover, both LSDV estimation and system GMM estimation that overcomes the potential endogeneity of growth determinants yield these results, despite that the estimated magnitudes are somewhat different between the two estimation methods. Last but not least, the estimated coefficient on the FDI to GDP ratio is significant in system GMM estimation, meaning the technological diffusion model is more suitable to describe the reform period Chinese economy.

Therefore, except for physical capital investment, the Chinese data may be reliable for growth regressions. The results are not surprising as Holz (2003) finds: "What is left is problems in understanding the meaning and coverage of various Chinese statistics,..., but no evidence of data falsification at the national level." He examines the recent criticism of statistics on China's GDP and shows that the criticism is unfounded as it is based on misunderstanding. Nevertheless, the problem with physical capital (extensively discussed in Chow, 1993, and Young, 2003, section VI) may pose serious problem for business cycle studies on China that rely on the volatile fluctuations in investment to explain business fluctuations (see e.g., King, Plosser and Rebelo, 1988). We leave this to future research.

Appendix

Province	Annual Growth	initial real GDP per worker	School02	$n+g^w+\delta$	I/GDP	FDI/GDP
Beijing78-82	1.48	2451	15.7	12.8	28.4	0
Beijing83-87	7.24	2952	9.1	10.3	51.9	0.60
Beijing88-92	3.12	4900	7.2	10.3	60.2	3.56
Beijing93-97	8.41	6515	9.1	7.8	76.4	7.77
Beijing98-02	3.20	11529	10.3	12.4	65.2	6.33
Tianjin78-82	2.93	2254	14.4	11.5	31.2	0
Tianjin83-87	6.52	2731	8.1	10.3	42.3	0.78
Tianjin88-92	3.59	4060	7.6	8.1	49.9	1.24
Tianjin93-97	9.85	5380	9.2	8.8	57.4	13.20
Tianjin98-02	9.73	11076	13.7	4.2	51.3	9.00
Hebei78-82	2.13	868	14.6	10.7	30.1	0
Hebei83-87	6.42	1015	9.4	11.1	35.4	0.03
Hebei88-92	4.82	1541	7.4	11.1	36.2	0.28
Hebei93-97	9.22	2264	9.4	9.4	43.0	1.95
Hebei98-02	7.04	4013	14.0	7.8	46.2	1.56
Shanxi78-82	3.41	912	17.2	10.4	31.4	0
Shanxi83-87	5.07	1211	13.5	10.9	44.9	0.01
Shanxi88-92	3.20	1637	11.3	10.8	42.5	0.19
Shanxi93-97	7.03	2138	10.5	9.1	41.1	0.83
Shanxi98-02	6.49	3419	14.2	7.1	46.6	1.30
Inner Mongolia78-82	4.51	889	19.3	12.0	30.5	0
Inner Mongolia83-87	6.80	1168	13.1	11.2	34.4	0.01
Inner Mongolia88-92	4.03	1769	11.5	9.9	40.8	0.07
Inner Mongolia93-97	7.01	2347	10.5	9.4	47.2	0.69
Inner Mongolia98-02	7.39	3810	13.0	7.2	43.9	0.62
Liaoning78-82	-1.09	1828	19.5	13.8	22.8	0
Liaoning83-87	7.57	1881	10.9	11.2	31.6	0.14
Liaoning88-92	3.26	3029	9.9	9.3	37.3	1.17
Liaoning93-97	5.74	4100	9.7	9.1	38.0	4.80
Liaoning98-02	6.69	6715	11.8	5.9	32.2	3.98
Jilin78-82	-0.56	1270	25.4	15.2	30.6	0
Jilin83-87	4.58	1507	14.9	12.0	35.0	0.01
Jilin88-92	1.59	2050	10.7	11.5	39.4	0.26
Jilin93-97	8.73	2492	10.3	8.2	40.8	2.57
Jilin98-02	7.49	4614	13.0	5.7	38.9	1.49
Heilongjiang78-82	2.17	1736	21.9	11.0	25.1	0
Heilongjiang83-87	3.44	1989	16.1	11.3	37.5	0.06
Heilongjiang88-92	3.21	2518	12.9	10.1	36.6	0.26
Heilongjiang93-97	5.11	3143	11.5	10.4	35.6	1.88
Heilongjiang98-02	7.91	4228	14.6	7.6	33.7	0.95
Shanghai78-82	3.72	3907	9.6	10.3	20.9	0
Shanghai83-87	6.52	5041	6.3	8.6	35.5	0.60
Shanghai88-92	6.13	7652	6.3	7.4	45.3	2.15
Shanghai93-97	9.35	12335	9.1	8.2	59.3	11.81
Shanghai98-02	5.94	24914	11.4	7.5	47.8	6.67

Appendix (Continued)

Province	Annual Growth	initial real GDP per worker	School02	$n+g^w+\delta$	I/GDP	FDI/GDP
Jiangsu78-82	5.65	897	11.5	9.9	30.4	0
Jiangsu83-87	8.31	1309	8.8	10.8	40.6	0.08
Jiangsu88-92	6.37	2321	7.9	9.7	43.7	1.30
Jiangsu93-97	10.48	3809	8.3	8.1	48.5	7.59
Jiangsu98-02	8.80	7357	10.7	6.7	45.6	6.86
Zhejiang78-82	7.50	689	9.3	11.0	25.9	0
Zhejiang83-87	9.72	1023	7.5	11.9	33.1	0.09
Zhejiang88-92	6.46	1808	6.8	9.4	34.5	0.47
Zhejiang93-97	10.92	3009	7.7	8.6	47.9	3.35
Zhejiang98-02	7.01	5825	9.0	9.0	45.7	2.46
Anhui78-82	4.76	608	12.4	11.6	19.8	0
Anhui83-87	6.38	825	8.8	11.5	31.7	0.03
Anhui88-92	2.20	1149	7.9	11.1	31.9	0.13
Anhui93-97	9.81	1517	8.6	10.2	40.3	1.85
Anhui98-02	5.96	2697	10.7	8.5	37.0	0.83
Fujian78-82	6.99	718	10.7	10.7	30.3	0.01
Fujian83-87	7.04	1051	9.4	11.8	31.9	0.89
Fujian88-92	6.28	1651	8.0	11.8	31.2	4.66
Fujian93-97	10.79	2748	10.3	9.6	44.9	15.98
Fujian98-02	6.37	5210	14.0	9.2	46.4	8.30
Jiangxi78-82	3.97	694	11.4	11.4	33.4	0
Jiangxi83-87	6.36	865	10.0	11.1	33.1	0.04
Jiangxi88-92	4.69	1282	10.0	10.3	34.8	0.31
Jiangxi93-97	5.42	1880	9.8	10.1	39.2	2.11
Jiangxi98-02	6.82	2781	13.4	6.8	38.9	1.92
Shandong78-82	4.92	759	12.9	10.4	32.0	0
Shandong83-87	9.32	952	9.6	11.1	36.1	0.05
Shandong88-92	5.15	1655	9.1	11.2	43.5	0.92
Shandong93-97	9.08	2570	10.2	9.3	47.0	4.40
Shandong98-02	7.63	4534	13.9	8.2	48.0	2.93
Henan78-82	4.60	580	15.6	10.9	30.3	0
Henan83-87	5.36	865	10.2	11.8	36.4	0.02
Henan88-92	3.37	1199	8.7	11.0	41.1	0.24
Henan93-97	7.67	1609	8.9	10.7	41.0	1.37
Henan98-02	4.91	2575	11.7	10.0	42.1	0.85
Hubei78-82	5.70	790	15.8	10.5	24.7	0
Hubei83-87	7.70	1094	11.0	10.2	31.2	0.04
Hubei88-92	4.43	1694	8.8	9.8	31.8	0.40
Hubei93-97	9.62	2391	9.4	9.1	40.3	2.35
Hubei98-02	8.03	4418	14.0	6.2	44.2	2.10
Hunan78-82	3.45	645	12.0	10.7	22.9	0
Hunan83-87	5.16	819	9.1	10.7	25.5	0.06
Hunan88-92	3.16	1111	8.0	10.6	28.1	0.23
Hunan93-97	6.99	1442	8.2	9.6	34.9	2.19
Hunan98-02	6.94	2278	11.4	7.3	35.7	1.75

Appendix (Continued)

Province	Annual Growth	initial real GDP per worker	School02	$n+g^w+\delta$	I/GDP	FDI/GDP
Guangdong78-82	6.64	817	10.6	10.6	28.7	0.39
Guangdong83-87	9.69	1198	8.5	10.9	33.5	2.63
Guangdong88-92	8.30	2188	7.6	11.1	34.6	5.53
Guangdong93-97	8.63	3958	9.3	10.2	41.9	15.67
Guangdong98-02	6.72	6799	12.1	9.0	36.9	10.18
Guangxi78-82	3.79	521	10.3	11.5	29.7	0
Guangxi83-87	3.72	633	6.4	11.3	30.1	0.26
Guangxi88-92	5.87	777	6.7	10.5	29.6	0.76
Guangxi93-97	6.64	1200	8.3	10.0	38.5	4.97
Guangxi98-02	5.68	1811	11.0	8.9	34.1	2.30
Guizhou78-82	4.43	442	10.1	11.5	36.2	0
Guizhou83-87	5.25	608	6.9	11.5	32.3	0.02
Guizhou88-92	1.83	834	6.0	11.9	32.4	0.17
Guizhou93-97	4.79	989	5.9	10.0	35.7	0.70
Guizhou98-02	5.35	1349	8.1	9.6	50.2	0.30
Yunnan78-82	3.41	526	7.2	12.1	35.5	0
Yunnan83-87	6.00	659	6.1	10.9	32.5	0.04
Yunnan88-92	3.68	1004	6.4	11.0	33.8	0.09
Yunnan93-97	6.79	1309	6.0	9.7	43.4	0.69
Yunnan98-02	4.98	1966	8.0	8.8	40.9	0.51
Shaanxi78-82	2.52	752	15.2	11.8	34.5	0
Shaanxi83-87	7.49	890	12.3	11.0	42.7	0.39
Shaanxi88-92	1.96	1519	8.9	11.2	43.2	0.91
Shaanxi93-97	5.74	1857	8.3	9.3	47.0	2.87
Shaanxi98-02	6.16	2714	12.7	8.7	49.8	1.52
Gansu78-82	-2.56	933	12.1	13.9	37.9	0
Gansu83-87	6.24	820	9.2	13.6	37.7	0
Gansu88-92	3.85	1224	7.9	10.8	41.6	0.02
Gansu93-97	6.15	1526	6.3	11.3	40.2	0.88
Gansu98-02	5.91	2252	8.5	9.1	43.8	0.45
Qinghai78-82	-0.04	1074	12.8	12.3	54.6	0
Qinghai83-87	5.17	1164	12.3	10.5	54.1	0
Qinghai88-92	1.55	1589	10.7	10.2	45.0	0.01
Qinghai93-97	4.86	1874	8.6	9.7	48.7	0.12
Qinghai98-02	6.49	2661	9.5	9.0	63.9	0.49
Ningxia78-82	2.13	959	15.1	11.5	54.0	0
Ningxia83-87	5.75	1178	13.5	12.2	58.1	0
Ningxia88-92	2.23	1702	13.4	11.3	57.8	0.11
Ningxia93-97	4.28	2045	9.1	11.0	51.8	0.42
Ningxia98-02	5.77	2757	9.5	9.6	64.4	0.80
Xinjiang78-82	5.50	794	16.0	10.1	45.0	0
Xinjiang83-87	8.37	1161	17.0	9.8	47.3	0.16
Xinjiang88-92	7.15	1906	14.1	9.7	53.8	0.12
Xinjiang93-97	5.77	2958	11.6	9.7	63.5	0.56
Xinjiang98-02	5.40	4311	16.8	8.3	52.1	0.13

Note: Except for initial real GDP, all variables are five-year averages and in percent per year.
School02 is the percentage of labor force in secondary school.
I/GDP and FDI/GDP are nominal investment and FDI to nominal GDP ratios respectively.

References

[1] Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth.* Princeton University Press.

[2] Arellano, M., and O. Bover. 1995. "Another look at the instrumental variables estimation of errorcomponents models." *Journal of Econometrics* 68: 29-51.

[3] Barro, R. 1991. "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics* 106 (May): 407-433.

[4] Barro, R., and X. Sala-i-Martin. 2004. *Economic Growth (2^{nd} edition).* Cambridge, MA: MIT Press.

[5] Blundell, R., and S. Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics* 87: 11–143.

[6] Borensztein, E., J. De Gregorio, and J-W. Lee, 1998. "How does foreign direct investment affect economic growth?" *Journal of International Economics* 45, 115-135.

[7] Chow, G. 1993. "Capital Formation and Economic Growth in China," *Quarterly Journal of Economics* 108 (August), 809-42.

[8] Chow, G. and Kui-Wai Li. 2002. "China's Economic Growth: 1952-2010." *Economic Development and Cultural Change* 51 (1), 247-56

[9] Démurger, Sylvie. 2001. "Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China?" *Journal of Comparative Economics* 29(1): 95-117.

[10] Deng, Xiaoping, 1975. "Several Points on Developing Industry" (in Chinese) [Guan Yu Fa Zhan Gong Ye de Ji Dian Yi Jian].

[11] Findlay, R., 1978. "Relative backwardness, direct foreign investment, and the transfer of technology: a simple dynamic model." *Quarterly Journal of Economics* 92, 1-16.

[12] Harrison, A. E. and M. S. McMillan, 2003. "Does direct foreign investment affect domestic credit constraints?" *Journal of International Economics* 61, 73-100.

[13] Holz, Carsten A. 2003. "'Fast, Clear and Accurate:' How Reliable Are Chinese Output and Economic Growth Statistics?" *The China Quarterly*, 173: 122-63.

[14] ____. 2006. "China's Reform Period Economic Growth: How Reliable Are Angus Maddison's Estimates." R*eview of Income and Wealth* 52 (1): 85-119.

[15] ____. 2008. "China's 2004 Economic Census and 2006 Benchmark Revision of GDP Statistics: More Questions Than Answers." *The China Quarterly*, 193: 150-63.

[16] Hsueh, T. and Q. Li (eds). 1999. *China's National Income, 1952-1995.* Boulder, Colo.: Westview.

[17] Li, Kui-Wai. 2003. "China's Capital and Productivity Measurement Using Financial Resources." Yale University Economic Growth Center Discussion Paper No. 851.

[18] Keller, W. and S. Yeaple. 2003. "Multinational Enterprises, International Trade, and Productivity Growth: Firm-level Evidence from the United States," *NBER Working Paper* No. 9504.

[19] King, R.G., C. Plosser, and S. Rebelo. 1988. "Production, growth and business cycles: I. The basic neoclassical model," *Journal of Monetary Economics*, 21 (2-3): 195-232.

[20] Mankiw, G., D. Romer and D. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107: 407-37.

[21] Narayan, P.K., I. Nielsen, and R. Smyth. 2008. "Panel data, cointegration, causality and Wagner's law: Empirical evidence from Chinese provinces," *China Economic Review*, 19 (2): 297-307.

[22] Perkins, D. and T. Rawski. 2008. "Forecasting China's Economic Growth to 2025." in *China's Great Economic Transformation* (Edited by Loren Brandt and Thomas Rawski). Cambridge: Cambridge University Press.

[23] Roodman, D. 2006. *How to do xtabond2: an introduction to "Difference" and "System" GMM in Stata*. Center for Global Development Working paper no. 103.

[24] Weeks, M. and J. Yao. 2003. "Provincial Conditional Income Convergence in China, 1953-1997: A Panel Data Approach," *Econometric Reviews* 22: 59-77.

[25] Wößmann, L. 2002. "Cross-Country Evidence on Human Capital and the Level of Economic Development: The Role of Measurement Issues in Education," *Historical Social Research*, 27(4): 47-76.

[26] Young, A. 2003. "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period," *Journal of Political Economy* 111: 1220-61.

[27] Zhang, T. and H. Zou. 1998. "Fiscal decentralization, public spending, and economic growth in China," *Journal of Public Economics* 67(2): 221-240.

[28] *Statistical Yearbook of China* [Zhong Guo Tong Ji Nian Jian]. Beijing: China Statistical Press 1981, 1983-2007 (annual).

Table 1: Descriptive statistics

	Mean	Standard deviation	Minimum	Maximum
growth (annual, %)	5.65	2.45	−2.56	10.92
ln(GDP/L)$_{t-1}$	7.50	0.75	6.09	10.12
ln(School01)	3.19	0.31	2.39	3.87
ln(School02)	2.34	0.29	1.77	3.23
ln(FDI/GDP)	−0.67	2.06	−6.49	2.77
ln(I/GDP)	3.66	0.24	2.99	4.34
ln($n + g^w + \delta$)	2.30	0.18	1.44	2.72

Observations: 135 (110 for ln(FDI/GDP)). The data are five-year averages for 27 provinces. Except for growth, and $ln(\frac{GDP}{L})_{t-1}$, all other variables are multiplied by 100 and then taken logarithms.

Table 2. LSDV Regressions Results. Dep. Var.: Average Annual Growth Rate of Real GDP per worker 1978-82, 1983-87, 1988-92, 1993-97, 1998-2002.

	Regression number			
Independent Variable	2.1	2.2	2.3	2.4
$\ln\left(\frac{GDP}{L}\right)_{i,t-1}$	-5.02^{***}	-4.08^{***}	-4.87^{***}	-5.77^{***}
	(0.82)	(0.77)	(0.83)	(1.01)
$\ln\left(\frac{I}{GDP}\right)$	2.11^{**}	2.11^{**}	2.29^{**}	1.98
	(0.90)	(0.93)	(0.92)	(1.33)
$\ln(n+g^w+\delta)$	-6.52^{***}	-6.44^{***}	-6.61^{***}	-5.89^{***}
	(0.94)	(0.96)	(0.94)	(1.01)
$\ln(School02)$	1.81^{**}		2.06^{**}	2.96^{**}
	(0.90)		(0.93)	(1.11)
$\ln(School01)$		-0.54	-1.08	
		(0.98)	(1.00)	
$\ln\left(\frac{FDI}{GDP}\right)$				0.19
				(0.18)
Time Fixed Effects	Yes	Yes	Yes	Yes
Province Fixed Effects	Yes	Yes	Yes	Yes
test $\ln\frac{I}{GDP}$=ln(School)	F(1,100)=0.06	F(1,100)=3.3	F(1,99)=0.04	F(1,74)=0.25
(p-value)	(0.81)	(0.07)	(0.85)	(0.62)
test $2\ln\frac{I}{GDP}+\ln(n+g^w+\delta)=0$	F(1,100)=1.18	F(1,100)=1.04	F(1,99)=0.90	F(1,74)=0.47
(p-value)	(0.28)	(0.31)	(0.35)	(0.50)
test $2\ln(School)+\ln(n+g^w+\delta)=0$	F(1,100)=2.14	F(1,100)=11.2	F(1,99)=1.53	F(1,74)=0.00
(p-value)	(0.15)	(0.001)	(0.22)	(0.99)
R^2	0.86	0.85	0.86	0.86
R^2(adjusted)	0.81	0.80	0.81	0.79
Observations	135	135	135	110

***Significant at the 0.01 level, ** at the 0.05 level, * at the 0.10 level
(standard errors in parentheses)

Table 3. System GMM Regressions Results. Dep. Var.: Average Annual Growth Rate of Real GDP per worker 1978-82, 1983-87, 1988-92, 1993-97, 1998-2002.

Dynamic panel-data estimation, one-step system GMM

	Regression number			
Independent Variable	3.1	3.2	3.3	3.4
$\ln\left(\frac{GDP}{L}\right)_{i,t-1}$	-8.37^{***}	-3.49^{**}	-8.20^{***}	-7.65^{***}
	(2.18)	(1.34)	(2.42)	(2.09)
$\ln\left(\frac{I}{GDP}\right)$	-0.17	3.90^{**}	-0.02	-1.73
	(2.24)	(1.89)	(2.65)	(2.25)
$\ln(n + g^w + \delta)$	-5.89^{**}	-6.86^{***}	-6.15^{**}	-6.14^{**}
	(2.45)	(2.42)	(2.40)	(2.28)
$\ln(School02)$	5.92^{***}		5.59^{**}	8.36^{***}
	(2.38)		(2.47)	(2.61)
$\ln(School01)$		-1.25	0.18	
		(3.17)	(2.44)	
$\ln\left(\frac{FDI}{GDP}\right)$				1.48^{***}
				(0.34)
Time Fixed Effects	Yes	Yes	Yes	Yes
Sargan overID test p-value	0.188	0.113	0.251	0.789
Hansen overID test p-value	0.181	0.061	0.218	0.406
Arellano-Bond test for AR(1)	Pr>z = 0.137	Pr>z = 0.007	Pr>z = 0.114	Pr>z = 0.034
Arellano-Bond test for AR(2)	Pr>z = 0.932	Pr>z = 0.298	Pr>z = 0.895	Pr>z = 0.369
number of instruments	16	16	18	18
Effective observations	135	135	135	110
F-test	33.20^{***}	24.12^{***}	28.39^{***}	31.23^{***}

Note: $\ln\left(\frac{GDP}{L}\right)_{i,t-1}$ is treated as predetermined. All other independent variables except the time dummies are treated as endogenous. Time dummies are used as instruments.

***Significant at the 0.01 level, ** at the 0.05 level, * at the 0.10 level

(standard errors in parentheses)