CENTRE D'ÉTUDES
ET DE RECHERCHES
SUR LE DÉVELOPPEMENT
INTERNATIONAL

Document de travail de la série

Etudes et Documents

E 2007.23

Sources of income differences across Chinese provinces during the reform period: A development accounting exercise

Rui HAO

CERDI, Université d'Auvergne, 63000 Clermont-Ferrand, France

E-mail address: rui.hao@u-clermont1.fr

novembre 2007
36 p.

Abstract

China's rapid and uneven growth since 1978 has not eliminated but even reinforced the persistent income inequality across provinces. While existing literature focuses mainly on the provincial variation in growth performance using cross-province growth regressions or growth accounting, few efforts has been made to directly study the differences in income levels across provinces. This paper explores the proximate causes of cross-province income differences in the framework of development accounting. Rather than assuming a priori values for output elasticities of capital and labor, we estimate them from an aggregate production function using panel data. The accounting results show that differences in total factor productivity (TFP) and in physical capital intensity are both important sources of cross-province income differences, each accounting for roughly half of the variation of income levels. Differences in human capital accumulation explain only a small amount of income differences across provinces. The results are robust to whether or not the assumption of constant returns to scale is imposed, and are valid in the long run. We do not exclude the possibility that interaction between factor accumulation and TFP plays an important role in determining cross-province income differences.

Keywords: Income differences; Development Accounting; China
JEL classification: O18, O40, O53

Sources of income differences across Chinese provinces during the reform period: A development accounting exercise

1. Introduction

One central problem in the growth literature is the need to explain the large and persistent differences in both levels and growth rates of output per worker across countries. The bulk of the literature focuses on either growth accounting initiated by Solow (1957), Denison (1962, 1967) and Maddison (1982), or growth regression motivated by Barro (1991), Mankiw *et al.* (1992) and Barro and Sala-i-Martin (1992). The former intends to ask to what extent can differences in factor inputs and productivity contribute to the cross-country differences in growth rates, while the latter aims to explain the transitory differences in growth rates across countries, assuming that long-run growth rates are determined by the identical exogenous technological progress. These two strands of research have provided ample insights into understanding the question like "why do some countries grow faster than others?"

More recent studies, however, focus on cross-country differences in income levels instead of differences in growth rates (e.g., Hall and Jones, 1999; Acemoglu *et al.*, 2001; Caselli, 2005) for at least three reasons. First, there are huge differences across countries in income levels and this enormously dispersed distribution is roughly stable over time. In contrast, growth rates over decades are only weakly correlated (Easterly *et al.*, 1993; Caselli, 2001). In particular, if the neoclassical models best describe the real world, the income gap between the rich countries and the poor should eventually stabilize thanks to mechanisms of diminishing returns of capital and international technology transfer, and thereby there is no difference in growth rates in the long-run. So long-run differences in levels are the most relevant variable to explain. Second, income levels capture the enormous contemporary differences across countries that have important welfare implications (Bourguignon and Morrisson, 2002). Growth rates are important only to the extent that they are a determining factor of levels. Third, the convergence literature provides relatively little evidence in terms of poor countries catching up with rich countries, i.e., statistical tests for convergence have failed to address the notion of convergence in an economically interesting sense (Quah, 1993; Durlauf, 2003). Therefore, it is more informative to directly analyze the cross-section relation in levels, which help answer the question "why are some countries so much richer than others?"

As a first step to answer this question, the development accounting (also acknowledged as level accounting) seeks to assess the relative importance of the proximate causes of economic performance, namely factor inputs and total factor productivity (TFP), to the differences in income levels across countries at one point in time. While it does not address the more fundamental determinants, i.e., why factor inputs and/or TFP differ across countries, the development accounting is nonetheless a useful diagnostic tool in understanding the vast cross-country income differences. For instance, if factor inputs are found to account for most of the income differences, then one could focus on explaining low rates of factor accumulation; instead, if one found that variation in TFP explains a large fraction of the variation in income, then one could point focus toward the sources of the differences in TFP levels. Obviously, the development accounting has the same idea with the growth accounting, but differs in the sense of cross-country differences replacing cross-time differences. The development of this method dates back to the pioneering work of Denison (1967). Standard development accounting exercises are generally based on the one-sector neoclassical growth model and employ an aggregate Cobb-Douglas production function with the assumption of constant returns to scale (e.g., King and Levine, 1994; Klenow and Rodriguez-Clare, 1997; Prescott, 1998; Hall and Jones, 1999; Parente and Prescott, 2000; Caselli, 2005).

The development accounting framework studying income differences across countries can be easily extended to analyzing income differences across regions within a country, especially for a large country like China. China's market-oriented and open-door reforms since 1978 has gradually transformed the Chinese economy. While every province has achieved a rapid economic growth there is considerable variation between provinces. The poor provinces have not displayed a consistent tendency to catch up with the rich. As a result, the long-existing income gap among provinces has expanded rather than narrowed. For example, the proportional gap between the richest province Shanghai and the poorest Guizhou grew from 5.8 in 1985 to 13.4 in 2000. Explaining such differences is one of the most important concerns in contemporary China and has spurred a growing literature on the subject. However, most studies continue to focus on cross-province differences in growth rates using growth regressions and convergence tests (e.g., Chen and Fleisher, 1996; Raiser, 1998; Brun *et al.*, 2002). Recently a few authors have attempted, in the framework of growth accounting, to examine the economic growth inequality between coastal and interior provinces due to inequality in capital and in technology (Liu and Li, 2006)[1]. Instead of estimating the TFP levels, they constructed the investment in innovation as a measure for technology, by arguing expenditure on technology in China is largely state-driven. Hence they can estimate the production function in first-difference with capital and technology data. Their main focus is still the differences in growth rates and the comparison is limited within estimated parameters among two regions of provinces.

This paper implements the development accounting procedure to investigate the enormous disparity of income levels across provinces in post-reform China. Since a number of studies on this subject already exist, it is useful to clarify how our study differs. First, it aims at studying the sources of cross-province differences in income levels, i.e., "why some provinces are much richer than others", rather than differences in economic growth, i.e., "why do some provinces grow faster than others". The main objective is thus to estimate what part of income differences across provinces is attributed to differences in capital intensities and what part to differences in TFP. As mentioned above, the accounting analysis does not attempt to uncover the ultimate reasons of income differences across provinces, but only the proximate ones. Second, by allowing the data to choose the parameters, we estimate econometrically the key parameters, namely the output elasticities of capital and labor, instead of assuming a value of experience. In view of the existence of heterogeneity for both cross-section and time dimensions across Chinese provinces we conduct panel data estimations of the production function. In addition, theory does not provide clear guidance about whether estimation of the

[1] Most studies in the framework of growth accounting focus on the Chinese economy as a whole rather than the regional growth inequality. See, for instance, Borensztein and Ostry, 1996; Hu and Khan, 1997; Chow and Li, 2002; Wang and Yao, 2003.

production function in levels or in growth rates (first-difference) would be preferable (Romer, 2001)[2]. Since the growth rate of output per worker varies much more than does the growth rate of (physical and human) capital per worker, the link between output per worker growth and inputs growth is likely to be very weak. In light of this, the production function will be estimated in levels in this paper. Finally, the usual assumption of constant returns to scale (CRS) is relaxed. While most related studies insist on this hypothesis and some provide evidence based on statistical tests in support of it, our results of estimation clearly show that CRS is not supported by the data used in this paper. Nonetheless, we also consider the CRS cases for comparison purposes.

Our analysis evolves as follows. Section 2 describes the data employed. Section 3 reports the estimations results of the production function under four specifications: imposing and relaxing the assumption of CRS, and with and without the stock of human capital as an input. Based on the estimated parameters, Section 4 conducts the development accounting using the traditional Denison approach and the Hall and Jones (1999) approach. Section 5 makes some further discussions. Finally, Section 6 concludes.

2. Data issues

The objective of development accounting is to decompose differences in income levels into contributions from differences in inputs levels and differences in TFP levels. Operationally, the key step lies in estimating the production function from which we derive the measures of TFP levels. This primarily depends on the proper measurement of income and factors. We use data on output, labor input, physical capital, and human capital for 28 Chinese provinces over the period 1982-2005[3]. The data come largely from the official dataset in *Comprehensive Statistical Data and Materials on 55 Years of New China* (2005) and *China Statistical Yearbook* (various years) published by the National Bureau of Statistics of China (NBS). The output measure (Y) is real GDP computed using the nominal GDP and the implicit deflators provided by the NBS. All current values are expressed in local currency (RMB) in 1995 constant prices. The labor force (L) is measured by the number of employed workers. Data on the stock of physical and human capital are not directly found in NBS dataset. Such data need to be calculated from the related NBS statistics, and this indirect measurement has been the object of a long-lasting dispute. The remainder of this section thus describes their construction.

2.1 Provincial stock of physical capital

Like previous studies, we generate estimates of the physical capital stock for each province over 1982-2005 using the perpetual inventory method, that is, with data on real investment ($I_{i,t}$), initial capital stock ($K_{i,0}$), and a common depreciation rate (δ), we can construct a sequence of physical capital

[2] Romer (2001, pp.66) states that if the change in the capital stock has less variation than the level, and is less correlated with the saving rate, the bias of the estimates in growth rates tends to rise. If the change in the residual is likely to have less variation than the level, this tends to increase the precision of estimates in growth rates. In addition, the bias of the estimates depends also on how the covariances of the instruments with the capital stock variable and the residual change.

[3] In this paper, China refers to mainland China and Hong Kong, Macao and Taiwan are excluded from analysis. Hainan and Tibet are also excluded from analysis because of incomplete data on educational attainment for construction of human capital. Data for Chongqing, which became a municipality in 1997, are included in the data for Sichuan.

stock using the following equation:

$$K_{i,t} = (1-\delta)K_{i,t-1} + I_{i,t}$$

However, there exist debates over each term of this equation in empirical work: the selection of the nominal investment series and the corresponding investment deflators, the estimates of the initial physical capital stock as well as the depreciation rate. Early studies like Chow (1993) used accumulation data in official statistics as the measurement of investment. Beginning from 1994, Chinese official national income statistics changed from the former “national income available” (consumption + accumulation) under the Material Product Balances System (MPS) to the “GDP” (final consumption expenditure + gross capital formation + net export of goods and services) under the System of National Accounts (SNA). From then on the accumulation data has not been available. Researchers have also attempted to use gross capital formation to measure investment (e.g., Chow and Li, 2002; Liu and Li, 2006)[4]. However, gross capital formation is calculated as the sum of gross fixed capital formation (GFCF) and inventory changes. Whether or not the inventory changes should be included in the capital stock is controversial. In a recent study, Young (2000) argues

> “*…the ‘changes in stocks’ figures reported in the national accounts of developing countries are frequently a residual, fabricated, item used to conceal large discrepancies between the production and expenditure sides of the accounts. In addition, the proper measurement of inventory changes, including the adjustment for differences between current valuations and accounting conventions, is technically more challenging than the measurement of the flow value of investment in fixed capital. Finally, in the context of the People's Republic (China), considering the unsold inventories of state enterprises as a productive element of the capital stock would seem to be an egregious error.*”

Following Young (2000), we exclude inventories from measure of the physical capital stock, and focus on GFCF alone. Data for some provinces in some years are missing from our data source. To maintain a complete sample, we estimate the GFCF series for these provinces on an important component of GFCF, namely Investment in Capital Construction (ICC), for the years where both series are available. With the estimated coefficients and the data on ICC we can thus obtain the GFCF using linear interpolation for these provinces for the years without data.

With a measure of nominal investment in hand, it is necessary to derive an appropriate fixed capital formation deflator. The Price Index of Investment (PII) is an appropriate choice, but this index is not available until 1991. However, *The Gross Domestic Product of China 1952-1995* published in 1997 by the NBS provides the index of GFCF for all provinces over the period 1952-1995, from which we can derive the implicit investment deflator in a similar way as deriving the implicit GDP deflator. As a check, we compare the calculated implicit investment deflators for the period 1991-1995 with the published PII for the same period, and find that they are extremely similar. It is therefore reasonable to believe that these two series are internally consistent. Hence, we use the calculated implicit investment deflators for 1952-1995 and the PII for 1996-2005, both converted to 1995 constant prices, to deflate the nominal investment (GFCF) to real value.

---

[4] In addition, some early authors used “total social investment in fixed assets (TSIFA)” collected under MPS as measure of investment to estimate the physical capital stock. The TSIFA series provide figures quite close to those in the GFCF series. This is not surprising, since the data on GFCF have been compiled based on some small adjustments upon TSIFA (Xu, 2000, who is in charge of the National Accounts Division in the NBS).

Alternatively, Young (2000) constructed an implicit deflator for GFCF as a residual between the GDP deflator and the deflators for other components of GDP, including private consumption, government consumption, inventories and export and import[5]. As compared to Young (2000, Figure IV), the overall GFCF deflators we calculated are highly similar, as shown in Figure 1 which graphs the averages of the GFCF deflators of the 30 provinces during the period 1952-2005. However, as Wang and Yao (2003) point out, the inherent risk of Young's complex method is that a measurement error in any of these deflators would be passed onto the "residual" investment deflator.

Figure 1 Here

Turning to depreciation rate, a large number of studies assume a value equal to 6% and apply it to all countries for all years (McGrattan and Schmitz, 1999; Hall and Jones, 1999; Young, 2000; Bernanke and Gürkaynak, 2001). Others either take a value of 3% (Klenow and Rodriguez-Clare, 1997), or 4% (Chow and Li, 2002; Liu and Li, 2006), or 5% (Perkins, 1988; Woo, 1998; Wang and Yao, 2003), or even 10% (Gong and Xie, 2004). Theoretically, the effect of varying the depreciation rate in the perpetual inventory calculation is to change the relative weight of old and new investment. A higher rate of depreciation may increase the relative capital stock in countries that have experienced high investment rates towards the end of the sample period. In practice, however, such an effect is showed to be minimal and the accounting analysis is not sensitive to a higher or lower depreciation rate (Klenow and Rodriguez-Clare, 1997; Caselli, 2005). In this paper we adopt an overall depreciation rate of 9.6% advocated by Zhang *et al.* (2004)[6], who conducted a detailed estimation on depreciation rate in the case of China.

With regard to initial physical capital stock, standard practice in the literature initiate the capital stock K_0 as $I_0/(g+\delta)$, where I_0 is the value of the investment series in the first year with available data, and g is the geometric average growth rate of the investment between the first five or ten years. The underlying rationale is that $I/(g+\delta)$ is the expression for the capital stock in the steady state of the Solow model. For instance, Hall and Jones (1999) take 1960 as the first year and estimate the 1960 capital stock as $I_{60}/(g+\delta)$, where g is calculated as the geometric average growth rate of the real investment over 1960-1970. Young (2000) initiate the capital stock of China in 1952 as real investment in 1952 divided by the depreciation rate of 6% plus the average annual growth of real investment between 1952 and 1957 that equals roughly to 4%. This way of estimating initial capital stock, by assuming that the economy is roughly on a balanced growth path and that the real investment growth recorded in the first several years of the data extends to the infinite past, leads to a good approximation. Although one may suspect that most of the poorer economies certainly do not satisfy the steady state condition that motivates the assumption $K_0=I_0/(g+\delta)$, the importance of the estimate of the initial capital stock tends to diminish over time rapidly due to depreciation (Hall and Jones, 1999; Easterly and Levine, 2001). In other words, "*the initial guess has very small persistence*" (Caselli, 2005). In

[5] In a similar manner, Liu and Li (2006) employed the national income identity rather than the investment deflator to obtain the real investment series, that is, $GDP=C+I+(E\text{-}X)$. Real investment (I) is obtained by subtracting real consumption (C) and real net exports ($E\text{-}X$) from real GDP figures (GDP). Real GDP and real net exports are calculated through nominal figures deflated by the implicit GDP deflators, while real consumption is nominal consumption deflated by the consumption price index. However, as mentioned above, their measure of investment includes inventory change, in that it is difficult to distinguish between real GFCF and real inventories.

[6] As a robustness check, we tried others values of depreciation rate like 4% and 6% and found that the final results are not sensitive to the choices.

particular, in the case of China, given that the period of interest is the reform period (i.e., after 1978), so "*with 26 years of data before the beginning of the analysis, the initial capital stock is fairly irrelevant, and any assumption will do equally well*" (Young, 2000). Due to lack of more convictive alternatives, our estimations of the initial capital stocks in 1952 for each province are calculated by their respective real investment in 1952 divided by 10% as in Young (2000).

2.2 Provincial stock of human capital

We follow the standard practice of transforming education attainment (i.e., years of schooling) into our measure of human capital stock (h) using estimates of private returns to education[7], namely,

$$(1) \qquad h = e^{\varphi(E)}$$

where E is the average years of schooling in the total population, and $\varphi(.)$ is a continuous, piecewise linear function constructed to match rates of return to schooling reported in Psacharopoulos (1994). Specifically, for schooling years between 0 and 4, the return to schooling $\varphi'(.)$ is assumed to be 13.4% which is an average for sub-Saharan Africa. For schooling years between 4 and 8, the return to schooling is assumed to be 10.1%, which is the world average. With 9 or more years, the return is assumed to be 6.8%, which is the average for the OECD countries. Despite the doubt about the applicability of these rates of return to schooling to the China's case, there is hardly any alternative[8].

The challenging problem is thus to construct the educational attainment data (E). There have been a number of attempts to measure educational attainment for individual countries and across countries. For instance, Mankiw *et al.* (1992) use the proportion of adult population enrolled in secondary school to capture educational attainment. However, only the secondary education does not adequately measure the aggregate stock of human capital available contemporaneously as an input for production process. In addition, it is problematic to use a flow variable, the enrollment rate, to represent the stock of human capital. Using various levels of enrollment ratios, Barro and Lee (2001) derive the percentage of population that has successfully completed a given level of schooling from a perpetual inventory approach, and hence construct the average years of schooling.

In this paper, we also use a perpetual inventory procedure close in spirit to the work of Démurger (2001), Wang and Yao (2003) and Liu and Li (2006) to construct our educational attainment data. The new school graduates are used as flows that are added to the human capital stock annually. The number of graduates is a more accurate measure to reflect the addition to the existing educated human capital stock than the enrollment ratios used by Barro and Lee (2001). As in Démurger (2001), we take account of three schooling levels: primary, secondary (comprising junior secondary, senior secondary and specialized secondary), and high education, as follows:

$$(2) \qquad H_{j,i,t} = (1-\delta_{i,t})H_{j,i,t-1} + Grad_{j,i,t}$$

where $H_{j,i,t}$ is the number of accumulated graduates who have completed at least j schooling level in

[7] Many studies use the educational attainment directly as a proxy for human capital stock (e.g., Mankiw *et al.*, 1992; Barro and Lee, 2001). The present paper, however, adopts the conventional approach of constructing human capital stock in the literature of development accounting along the lines of Klenow and Rodriguez-Clare (1997), Hall and Jones (1999), and Bils and Klenow (2001), among others.

[8] Nonetheless, we address this concern by conducting a sensibility analysis, i.e., assuming a constant rate of return to schooling equal to 10% and 15%, respectively. The final results are robust to these changes.

province i, $Grad_{j,i,1}$ is the annual number of net graduates at j schooling level, j=1 for primary, 2 for secondary, and 3 for high education, and $\delta_{i,t}$ is the mortality rate of the population used as a proxy of the depreciation rate.

Following Démurger (2001), we derive the initial educational attainment from the 1982 Chinese census which provides the 1‰ sample population surveys by schooling levels for 28 provinces (without Hainan and Tibet). As such, the initial values for these three schooling levels ($H_{j,i,0}$) are given as follows:

$$H_{1,i,0} = Prim_{i,0} + Sec_{i,0} + High_{i,0}$$
$$H_{2,i,0} = Sec_{i,0} + High_{i,0}$$
$$H_{3,i,0} = High_{i,0}$$

where $Prim_{i,0}$, $Sec_{i,0}$ and $High_{i,0}$ are referred as the 1‰ sampling number of people in province i that have completed primary, secondary and high education, respectively, multiplied by the population of the province i in 1982, $Pop_{i,0}$.

Using these initial values for each level of schooling, combined with mortality rates and annual number of net graduates, the three series $H_{j,i,t}$, which are the number of accumulated graduates at different schooling levels for province i at the end of the period t, can be obtained from a perpetual inventory procedure in equation (2). We then assume that the length of schooling cycles for primary, secondary and high education is 5, 10 and 14.5 years, respectively. Hence, the average years of schooling per capita is given by:

$$E_{i,t} = (5H_{1,i,t} + 10H_{2,i,t} + 14.5H_{3,i,t}) \,/\, Pop_{i,t}$$

where $Pop_{i,t}$ is the total population of province i in year t.

We should note, however, the educational attainment figures we calculated here are somewhat crude, as compared to some previous studies. For instance, Wang and Yao (2003) construct a time series of educational attainment of China as a whole from 1952 to 1999, while Liu and Li (2006) construct a panel data covering 30 provinces over 1985-2001, both using the perpetual inventory method. Our construction of schooling differs from theirs mainly in two ways. First, we assemble junior, senior and specialized secondary in a global category for secondary education, whereas they deal with them separately and accord them respective weight (length of schooling cycles). The simplicity in our work is primarily due to the fact that there are no initial figures for the three secondary schooling levels separately in the 1982 census. Second, we use the provincial data on total population instead of population aging 15-64. The various *China Population Statistical Yearbook* provide only the 1% or 10% samplings from the population aging 15-64, and even this is not available for all post-reform years. Though Liu and Li (2006) have attempted to make some adjustments on these figures by using the national figures in the World Bank's *World Development Indicators* and then decomposing these national figures by the respective provincial employment figures, we wonder if this complicated procedure could undertake more measurement errors from each step of calculation. Therefore, we rely on the relative consistent series namely the total population.

The final step of measuring stock of human capital is to use estimates of the return to education from wage regressions of log wages on years of schooling, as mentioned above. Hence, applying the educational attainment ($E_{i,t}$) figures to equation (1), we obtain the stock of human capital ($h_{i,t}$) for 28 provinces over 1982-2005.

It may be argued that the human capital stock thus obtained is only the quantity-based measure of human capital which does not incorporate any adjustment for variations in the quality of education. Indeed, one year of education in province A may generate more human capital than in province B. One way of taking into account this possibility is to rewrite equation (1) as follows (Caselli, 2005):

$$h = A_h e^{\varphi(E)}$$

Apparently, the measures we have adopted so far are built on the assumption that A_h is constant across provinces. One can use some indicators to augment the quantity-based measure of human capital. For example, consider the possibility that A_h is variable,

$$A_h = p^{\varphi(p)} m^{\varphi(m)} k^{\varphi(k)} t^{\varphi(t)}$$

where p is the teacher-pupil ratio, m is the amount of teaching materials per student (textbooks, etc.), k is the amount of structures per student (classrooms, gyms, labs, etc.), t is the human capital of teachers or represent externalities in the process of acquiring human capital, and $\varphi(p)$, $\varphi(m)$, $\varphi(k)$ and $\varphi(t)$ represent the corresponding elasticities. Other extensions that augment quantity-based human capital involve allowing for differences across economies in experience levels (Klenow and Rodriguez-Clare, 1997; Bils and Klenow, 2001), in nutrition and health status (Shastry and Weil, 2003), and in international tests of academic performance in mathematics and science (Hanushek and Kimko, 2000), or considering social return to schooling instead of private return used in this paper (Prichett, 2003; Cordoba and Ripoll, 2005). However, taking into account these issues is extremely difficult in our framework. This is not only because the relevant data are not available but also because these issues *per se* are the object of intense controversy when entering human capital. Hence, we regard our constructed dataset as a reasonable measurement of human capital stock upon available information. It is noted that our data on output, labor, and physical capital stock contain 30 provinces over the period 1978-2005, but the addition of the human capital stock reduce the sample to 28 provinces over 1982-2005. The descriptive statistics of the data are given in Table 1.

Table 1 Here

3. Panel estimation of the production function

3.1 Specification

Development accounting essentially asks the following question: is the persistent inequality across provinces explained mainly by differences in factor accumulation or by differences in TFP levels? The starting point is an aggregate production function assumed to be the same across provinces. Specifically, consider the following Cobb-Douglas aggregate production function:

$$(3) \qquad Y = A\, K^{\alpha}\, (hL)^{\beta} \qquad 0<\alpha<1 \text{ and } 0<\beta<1$$

where Y is real GDP, A is an index of TFP, k is physical capital stock, h is per capita human capital stock, L is total employment, and hL is skill-adjusted labor input. α and β are output elasticities with respect to physical capital and skill-adjusted labor, respectively. Since we want to understand why income varies across provinces, income defined to be output per worker (y=Y/L), is our object of interest rather than Y. Hence, dividing the equations (3) by the labor force L yields:

$$(4) \qquad y = A\, k^{\alpha}\, h^{\beta}\, L^{\alpha+\beta\text{-}1}$$

where y is income that equals to real GDP per worker, k is the stock of physical capital per worker k=K/L.

Given the data on income and factor inputs, it is crucial to obtain plausible and robust estimates of parameters. This in turn raises several relevant issues that deserve mention. The first involves the choice of Cobb-Douglas specification. Due to its simplicity, the Cobb-Douglas has been the most widely used form of production function, though other forms are equally valid such as the translog or the CES (Kim and Lau, 1994). The Cobb-Douglas has an additional advantage of sidestepping the arbitrage between Harrod neutral and Hicks neutral technical change. As Felipe and McCombie (2004) state, in the context of the Cobb-Douglas production function, there is no conceptual difference between Harrod neutral and Hicks neutral technical change.

How human capital enters the production function is controversial. There are two principal ways in the literature, one involving Mankiw *et al.* (1992) and one involving Hall and Jones (1999). In the former case, human capital is modeled precisely as a second kind of capital good, the production of human capital sharing the same way as that of physical capital. It is actually an entry in the vector of capital stocks, and thus is better viewed as analogous to physical capital. In the latter case, however, the benefits of education are assumed to be labor-augmenting. Human capital is therefore embodied in the labor force, i.e., human capital multiplies the labor input to produce effective labor input. Hall and Jones' production is substantially more restrictive than the one used by Mankiw *et al.* (1992). As argued by Caselli (2005), the great advantage of the Hall and Jones' formulation is that it generates the log-linear relation between wages and years of schooling. Since our construction of human capital is in line with Hall and Jones (1999), we adopt the labor-augmenting human capital formulation.

It is also worth noting that no consensus has emerged about whether human capital should be included as an input in the production function. Whereas Mankiw *et al.* (1992) obtain a better fit when including human capital in their regressions, Benhabib and Spiegel (1994) conclude that human capital does not enter the production function as an input, but rather affects growth via its effect on TFP. Therefore, we consider both cases that exclude and incorporate human capital in the production function in order to look at the resulting effect on estimated elasticities with respect to capital and labor.

Another important issue comes from the usual assumption of constant returns to scale (CRS), that is, $\alpha+\beta=1$. While most researchers argue that the CRS is a reasonable assumption in the context of China, based on statistical tests after regressions (e.g., Chow, 1993; Démurger, 1999; Wu, 2000; Liu and Li, 2006), we believe that this is a rather strong hypothesis that needs to be carefully treated. For instance, Gong and Xie (2004) find that the parameters estimates are not sensitive to the specifications of production function but more sensitive when the assumption of CRS is imposed. Wang and Yao (2003), while not exploring the possibility of variable returns to scale (VRS), admit that CRS may not be appropriate for a transition economy like China. However, allowing for the possibility of VRS does come at a cost of less tractability. We start by estimating the production functions under CRS and then extend the analysis by relaxing this assumption to check the robustness of our results.

The estimation of the production function also raises the issue of whether to estimate it in levels or in first differences. Unlike previous studies, we estimate the production function in levels because the first difference operator removes all the long-run information and thus emphasizes short-run fluctuations in the data, as documented by the cointegration literature. By differencing, we risk disregarding the most valuable part of information in the data, for example, the pure cross-section variation (Blundell and Bond, 1997). Differencing may also decrease the signal-to-noise ratio, thereby exacerbating measurement error bias (Griliches and Hausman, 1986). In addition, our data show that the growth rate of real GDP is not strongly linked with the growth rate of capital and labor inputs. This implies that level regressions should yield more accurate estimates of the parameters. Hence, taking logs on equation (4), we have

$$\text{(5)} \qquad \ln y = \ln A + \alpha \ln k + \beta \ln h + (\alpha+\beta-1)\ln L$$

3.2 Endogeneity bias

As is well-known, the estimation of an aggregate production function confronts the researchers with the problem of endogeneity bias. The parameters estimates of the production function by OLS are biased and inconsistent if there is contemporaneous correlation between the error term and the factor inputs. Three related reasons for endogeneity of right hand side (RHS) variables are usually recognized in the econometric literature: omitted variables, measurement error in a RHS variable and simultaneity of an explanatory variable (Wooldridge, 2002). Classical cross-section regressions assume that the omitted variables are independent of the included RHS variables and are independently, identically distributed. In the case of China, however, these omitted variables may contain some unobservable or unmeasurable heterogeneity across provinces and/or across time. Unless these heterogeneities are controlled for, the estimation will generate misleading results due to omitted variable bias. We use panel data estimation to address this problem (Knight *et al.*, 1993; Loayza, 1994; Islam, 1995). In particular, we restrict our regressions to fixed-effect estimators since the random-effect model requires that the individual effects are uncorrelated with the other regressors, an assumption that is necessarily violated in the context of our model[9]. Hence, rewriting equation (5) in the two-way fixed-effect form yields the following

$$\text{(6)} \qquad \ln y_{i,t} = \alpha \ln k_{i,t} + \mu_{i,t} + \eta_{i,t} + \varepsilon_{i,t} \qquad \text{(CRS without human capital)}$$

$$\text{(7)} \qquad \ln \tilde{y}_{i,t} = \alpha \ln \tilde{k}_{i,t} + \mu_{i,t} + \eta_{i,t} + \varepsilon_{i,t} \qquad \text{(CRS with human capital)}$$

$$\text{(8)} \qquad \ln y_{i,t} = \alpha \ln k_{i,t} + (\alpha+\beta-1)\ln L_{i,t} + \mu_{i,t} + \eta_{i,t} + \varepsilon_{i,t} \qquad \text{(VRS without human capital)}$$

$$\text{(9)} \qquad \ln \tilde{y}_{i,t} = \alpha \ln \tilde{k}_{i,t} + (\alpha+\beta-1)\ln(hL)_{i,t} + \mu_{i,t} + \eta_{i,t} + \varepsilon_{i,t} \qquad \text{(VRS with human capital)}$$

where $\ln\tilde{y}=\ln(Y/hL)$, $\ln\tilde{k}=\ln(K/hL)$, μ is an unobservable province-specific effect, η is a time-specific effect, ε is the error term, and i and t represent province and year, respectively. As argued above, we

[9] The Hausman specification tests do confirm that fixed-effect estimators are preferred to random-effect estimators in the context of our model.

consider the specifications that impose and relax the assumption of CRS, and both without and with human capital as an input. Clearly, equation (6)-(8) are special cases of equation (9). In particular, we can test the plausibility of CRS by testing whether the coefficient of ln*L* or ln(*hL*) equals zero.

Our panel data fixed-effect estimators involve a consistent treatment of the individual effect but still require a strong assumption of strict exogeneity of the factor inputs. Measurement error and simultaneity are also possible candidates to cause regressors and error term to be correlated. We rely heavily on official Chinese data, while there exists a dispute on their accuracy (Rawski, 2001). Our construction of the stock of physical capital and of human capital is based on many assumptions. Some missing values were obtained using linear interpolation. Therefore these series are likely to be measured with error. Furthermore, if there is causality between income and factor inputs, simultaneity problems emerge. Income and inputs measures are also likely to be simultaneously influenced by certain omitted factors. In such circumstances, the instrumental variables (IV) approach is called for, in the sense that it may alleviate the endogeneity bias and achieve consistency of estimates by instrumenting the factor inputs with regressors that are correlated with them and at the same time orthogonal to the errors[10].

However, applied researchers often must confront several hard choices in this context. As Wooldridge states (2003) states, an important cost of performing IV estimation when regressors and errors are actually uncorrelated is the much larger asymptotic variance of the IV estimator as compared to that of the OLS estimator. Consequently, although there may be good reason to suspect non-orthogonality between regressors and errors, the use of IV estimation for the sake of consistency must balanced be balanced against the inevitable loss of efficiency vis-à-vis OLS (Baum *et al.*, 2003). Thus a test of the appropriateness of OLS and the necessity to resort to IV is needed. The Wu-Hausman test (Wu, 1973; Hausman, 1978) can be used to diagnose whether the regressors are endogenous variables and if it is the case, OLS is inconsistent and IV is required[11].

But even when IV is judged to be the appropriate estimation technique, we may still question whether our instruments are relevant and valid instruments. In reality, to find instrumental variables that satisfy both conditions has proved very difficult (Wooldridge, 2002). The relevance of instruments may be assessed by examining the explanatory power (or the significance) of the excluded instruments in the first stage IV regressions. In the case of a single endogenous variable, we use a test statistic recommended by Bound *et al.* (1995), namely partial R^2 which is the R^2 of the first-stage regression with the included instruments partialled-out. For models with multiple endogenous regressors we employ additionally Shea (1997) partial R^2 statistic which takes the intercorrelations among the instruments into account[12]. If the relevance of excluded instruments is weak, the bias in the estimated

[10] It is argued that the generalized method of moments (GMM) approach deals consistently and efficiently with both the correlated individual effects and endogenous explanatory variables problems in the framework of panel data. Although appealing, this procedure critically hinges on the assumption of lack of second-order serial correlation in the errors of the equations in levels. However, the AR(2) tests suggest none of the four specifications satisfies such a necessary condition in the context of our data.

[11] The null hypothesis of the Wu-Hausman test is that OLS estimator would yield consistent estimates. The test proceeds in two steps. The first step involves estimating the first-stage regression for each of the endogenous variables to generate their residual series. The original model is augmented with these residuals and reestimated with OLS in the second step. An *F*-test (a *t*-test in the case of a single endogenous variable) of the significance of the residuals in this auxiliary regression is then a direct test of the null hypothesis.

[12] For a model containing a single endogenous regressor, the two R^2 measures are equivalent. As a rule of thumb, a

IV coefficients increase and conventional asymptotics fail. In the extreme case of zero relevance of instruments the model is essentially unidentified with respect to that endogenous variable and IV becomes inconsistent, hence nothing is gained from instrumenting.

To be a valid instrument, the instrument must be independent from the error process. Such a condition seems even more difficult to satisfy. When estimating production functions, some authors propose to instrument the regression using investment (Olley and Pakes, 1996) or intermediate materials (Levinsohn and Petrin, 2003). However, investment (or intermediate materials) may be to some extent demand-induced, in that they may follow as well as lead output growth, i.e., output and investment (or intermediate materials) may still be simultaneously determined. Therefore, the validity of investment (or intermediate materials) as instruments would be suspectable. Econometrically, the validity of instruments can be diagnosed via a Hansen (1982) *J*-test of overidentifying restrictions[13]. A rejection of the null hypothesis implies that the instruments are not satisfying the orthogonality conditions required for their employment. Note that this type of tests can be performed only when the order condition for identification is satisfied in inequality, that is, the number of excluded instruments exceeds the number of included endogenous variables. In this paper we follow Liu and Li (2006) to use twice-lagged endogenous regressors as instruments. Due to the fact that there are hardly any alternative *a priori* plausible instruments, our IV models are exactly identified. Therefore, the two periods lags of the endogenous regressors we employed as instruments would not be guaranteed for their validity *a posteriori*. Furthermore, as Nakamura and Nakamura (1998) point out, if some variables are chosen as instruments that are actually endogenous then there will exist an endogeneity problem even after instrumentation which may be worse than the original endogeneity problem. Therefore, pursuing a Wu-Hausman pre-test strategy and IV estimation when the instruments are weak becomes questionable. In such circumstances the OLS results would be preferred despite the significant Hausman test results (Harvey *et al.*, 1998). Given these caveats, both OLS and IV results are reported.

3.3 Estimation

Table 2 reports the estimates of Equation (6) and (7). All the estimated coefficients are significant at the 1% level. Column 1 and 2 give the estimates of the specification without human capital. The Wu-Hausman test rejects that $\ln k$ is exogenous at the 1% level, suggesting that OLS results may be inconsistent. A high value (0.880) of Bound *et al.* partial R^2 indicates a strong relevance of the used instrument, namely twice-lagged $\ln k$, albeit lacking test for its validity. However, the OLS and IV estimations yield extremely similar estimates for output elasticity of physical capital (α), 0.481 and 0.490 respectively. Moreover, the hypothesis that the estimated elasticity α obtained by IV technique is not statistically different from that using OLS can not be rejected, an F-statistic being only 0.12.

The similar pattern arises when human capital enters the production function, as shown in column 3 and 4 of Table2. The estimates for α increase slightly to 0.503 and 0.519, using OLS and IV respectively. The null hypothesis that OLS estimator is consistent is rejected at the 5% level, and a high

large value of the Bound *et al.* R^2 combined with a small value of the Shea R^2 indicate a lack of sufficient relevance of instruments to explain all the endogenous regressors, and the model may be in effect unidentified (Baum *et al.*, 2003).

[13] The commonly used tests of overidentifying restrictions include also Sargan (1958) statistic and Basmann (1960) statistic. However, neither of them is valid in the presence of conditional heteroskedasticity. If robust inference is sought in an IV model, one can resort to the standard *J*-statistic (Baum *et al.*, 2003).

value (0.885) of partial R^2 statistic suggests that the instrument twice-lagged $\ln\tilde{k}$ satisfies the relevance condition that it should be correlated with the endogenous variable $\ln\tilde{k}$. Nonetheless, since we can not provide statistic evidence for the orthogonality condition of the instrument, and the equality of coefficients between OLS and IV estimators is clearly not rejected, we focus on the OLS results only. To summarize, the estimates of α range from 0.481 to 0.503 under the assumption of CRS, implying a value for output elasticity with respect to an agglomerate of raw labor and human capital (β) ranging from 0.519 to 0.497.

Table 2 Here

Turning to the specifications where the assumption of CRS is relaxed. Table 3 reports the estimates of Equation (8) and (9). Column 1 and 2 ignore the human capital. The Wu-Hausman test rejects the overall exogeneity of $\ln k$ and $\ln L$, but only at the 10% level. The coefficients equality tests do not reject that there is no statistical difference between OLS and IV estimators for α, β, or α and β jointly. As before, we discuss the OLS results only. The OLS estimate of α equals 0.430, which is significant at the 1% level. The most striking result is that the coefficient of $\ln L$ equals -0.371, indicating high decreasing returns to scale. The F-statistic for testing CRS is 31.67, strongly rejecting that CRS is a reasonable assumption in the context of our data.

The main results essentially do not alter when human capital is considered as shown in column 3 and 4 of Table 3, except that the Wu-Hausman test fails to provide support for the endogeneity of regressors. Hence the OLS estimations should yield appropriate estimates[14]. As can be seen from column 3, the estimated α is 0.433, a value only slightly higher as compared to the case without human capital. The coefficient of $\ln(hL)$ is now -0.377, indicating even larger decreasing returns to scale. The implied output elasticity with respect to effective labor (β) falls in the range from 0.200 (without human capital) to 0.190 (with human capital), suggesting that the higher β of 0.519-0.497 in the specifications under CRS capture much of the influence of returns to scale[15]. Therefore, imposing the assumption of CRS may potentially generate misleading results. In contrast with the estimates of β that are volatile relative to whether the assumption of CRS is imposed or relaxed, the estimates of α are relatively stable: when CRS relaxed, the estimated α range from 0.430 to 0.433, only slightly lower than the estimates under CRS of 0.481-0.503. In particular, as shown in Table 4, our estimates of α are in line with the usual values used in the literature on China and thus provide a plausible and even more precise range of values.

Table 3 and 4 Here

[14] We have also conducted the IV estimation treating $\ln k$ alone as an endogenous variable for equation (8), and the IV estimations treating $\ln\tilde{k}$ alone as an endogenous variable for Equation (9). The messages regarding the estimated coefficients and the tests statistics are comparable to those reported in column 2 and 4 of Table 4.

[15] Under the assumption of CRS, Liu and Li (2006) find that the estimates of the marginal productivity of labor range from 0.027 for the interior provinces to 0.159 for the coastal provinces. In addition, studies of labor's share have often found lower values in developing countries than in industrial countries (Elias, 1993). The possible reasons involve the distortion of labor market in developing countries, causing a divergence between the price per unit of each employed labor and its marginal value product. This may be most likely the case of China in transition.

4. Development accounting

4.1 Denison level accounting

Once we get the estimates of the production function parameters, we are able to perform development accounting for identifying the "sources of development", that is, what part of cross-province income inequality is accounted for by differences in factor accumulation and what part by differences in TFP levels. We first employ the traditional Denison approach (Denison, 1967; King and Levine, 1994; Easterly and Levine, 2001). Rewrite the terms of equation (4) as ratios to the values of the reference province *j*, we have:

$$y_i/y_j = A_i/A_j\ (k_i/k_j)^{\alpha}\ (h_i/h_j)^{\beta}\ (L_i/L_j)^{\alpha+\beta-1}$$

To describe the relative level of development, we calculate the percentage shortfall in income for province *i* relative to the reference province *j*:

$$P_i = 100(y_j\text{-}y_i) / y_j$$

Then to describe the extent to which physical capital accounts for cross-province differences in income, we construct the ratio:

$$\varphi^k{}_i = \alpha \ln(k_i/k_j) / \ln(y_i/y_j)$$

which is the fraction of income difference attributive to physical capital. The contribution of physical capital to the level of development is thus given as $P_i\varphi^k{}_i$. Likewise, the fractions due to human capital and scale efficiency are given, respectively, as follows:

$$\varphi^h{}_i = \beta \ln(h_i/h_j) / \ln(y_i/y_j)$$
$$\varphi^L{}_i = (\alpha+\beta-1) \ln(L_i/L_j) / \ln(y_i/y_j)$$

The contributions of human capital and scale efficiency to the level of development equal $P_i\varphi^h{}_i$ and $P_i\varphi^L{}_i$, respectively. Finally, the contribution of differences in the residual A is calculated as:

$$P_i\varphi^A{}_i = P_i(1\text{-}\varphi^k{}_i\text{-}\varphi^h{}_i\text{-}\varphi^L{}_i)$$

Hence, for each province the differences in income levels can be apportioned to differences in physical capital, in human capital, in scale efficiency and in the residual A.

Notice that in contrast with the standard development accounting exercises that impose the assumption of CRS (i.e., $Y=Ak^{\alpha}h^{\beta}$, $\alpha+\beta=1$), our decomposition in equation (4) adds the scale efficiency term $L^{\alpha+\beta-1}$ ($\alpha+\beta\neq1$) to capture the effect of returns to scale. It is not clear whether the contribution of the scale efficiency term should be fully assigned to TFP, or to factors, or if it should be treated as a separate component. In principle, the conventional notion of TFP, which is "a measure of our ignorance" (Abramovitz, 1986), involves many factors like technology innovation, technology imitation and adoption, efficiency of resource allocation, institutional change, omitted variables and measurement errors. If we broadly view the TFP as the melange of factors other than inputs, then the scale efficiency and the residual A can be integrated to produce the TFP in the general sense[16]. In what follows, we loosely use the term TFP to refer to the combination of the scale efficiency and the residual A, while the term X is used to capture "factor intensities", that is

$$(10) \qquad y = X \cdot TFP$$

where $X=k^{\alpha}h^{\beta}$ and $TFP=AL^{\alpha+\beta-1}$.

Since development accounting typically assess the relative contribution of differences in factors intensities and differences in TFP to cross-province income differences at a particular time, we must choose the year under study. In this paper, we provide the accounting results for 2000 since it is in this year that the dispersion in output per worker across provinces reached its peak. The richest Shanghai is the reference province. The results are plotted into visual figures in order to reassure ourselves that the results are not being driven by a few outliers.

Figure 2 and 3 summarize the level accounting results under the assumption of CRS. Each figure includes 27 provinces plotted by descending order from left to right (i.e., from poorest to richest relative to Shanghai). From Figure 2 where human capital is ignored, TFP differences account for a large fraction of differences in income levels: on average a 43% of the income shortfall relative to Shanghai is due to the residual (TFP). Once human capital is considered, as shown in Figure 3, the contribution of TFP differences falls to 36%, suggesting that the addition of human capital chips away a fraction from the residual and thus reduces the measure of our ignorance. Nonetheless, a still high proportion (36%) indicates that TFP does play a very large role in explaining cross-province income differences, which accords in spirit with the consensus view in the development accounting literature. On the other hand, it is clear that the differences in physical capital accumulation account for a bulk of cross-province differences in income levels, with a contribution ranging from 57% (without human capital) to 60% (with human capital). Therefore, the differences in physical capital accumulation and in TFP are both important to understand the vast cross-province income differences, while the human capital differences play a less crucial role (less than 4%).

Figure 2 and 3 Here

We now turn to the more general specifications where the assumption of CRS is relaxed (Figure 4 and 5). In such circumstances, our TFP measure bundles together the scale efficiency and the residual A. As shown in Figure 4, the differences in TFP (comprising scale efficiency and residual A) contribute significantly to the cross-province differences in income levels, accounting for 49% of the income shortfall relative to Shanghai. With human capital being incorporated in Figure 5, the contribution of TFP still attains 47%. The contribution of physical capital accumulation equals 51.0% (without human capital) and 51.3% (with human capital) respectively. Once again, the contribution of human capital is rather small (less than 2%). These results are consistent with our earlier finding that most of the cross-province variation in income levels is due to both differences in physical capital accumulation and differences in TFP.

[16] In the literature of production frontier analysis the change in TFP is often decomposed into technological progress, technical efficiency change and scale efficiency (e.g., Färe *et al.*, 1994).

In particular, a careful investigation of Figure 4 and 5 makes it clear that the scale efficiency is important as a part of TFP, accounting for more than 50% of the variation of TFP. Of course, the consideration of scale efficiency can also be regarded as a research strategy to chip away at the measure of our ignorance, like the treatment of human capital. Nonetheless, this issue has not yet adequately been studied and should be a topic for the future research work on China. On the other hand, given that the basic results remain unchanged qualitatively when CRS is relaxed, this assumption broadly used in existing literature may be justified to some extent.

Figure 4 and 5 Here

4.2 Hall and Jones (1999) level accounting

Some recent authors perform a variance decomposition exercise to assess the contributions of TFP and X to world income dispersion (e.g., Klenow and Rodriguez-Clare, 1997; Caselli, 2005). To see this, for instance, from equation (10) we have:

$$var[\ln y] = var[\ln X] + var[\ln \mathrm{TFP}] + 2cov[\ln X, \ln TFP]$$

A difficulty arises from how to handle the covariance between X and TFP, a term that accounts for 35% of income dispersion, and for which exogenous theories of TFP have no predictions (Cordoba and Ripoll, 2005). Indeed, this large covariance suggests that the TFP level may be actually endogenous. In order to capture this possibility, Klenow and Rodriguez-Clare (1997) assign half of the covariance term as part of the contribution to X and the other half to TFP, that is, they define the contributions of factors (VD_X) and TFP (VD_{TFP}) as:

$$VD_X = var[\ln X]+cov[\ln X, \ln TFP] \; / \; var[\ln y]$$
$$VD_{TFP} = var[\ln TFP]+cov[\ln X, \ln TFP] \; / \; var[\ln y]$$

Rather than using the variance decomposition as Klenow and Rodriguez-Clare (1997) do, Hall and Jones (1999) decompose output per worker in each country into the three multiplicative terms (i.e., equation (4) under the assumption of CRS): the contribution of physical capital intensity, the contribution of human capital, and the contribution of TFP[17]. Then they calculate the inter-percentile differential (e.g., the five richest countries to the five poorest, or the 90th percentile to the 10th percentile) for each term:

$$y_{\mathrm{rich}}/y_{\mathrm{poor}} = (k^{\alpha}_{\mathrm{rich}}/k^{\alpha}_{\mathrm{poor}})\ (h^{\beta}_{\mathrm{rich}}/k^{\beta}_{\mathrm{poor}})\ (TFP_{\mathrm{rich}}/TFP_{\mathrm{poor}})$$

An important advantage of Hall and Jones' approach is that it does not have to deal with the distribution of the covariance term. In addition, the inter-percentile differential is less sensitive to outliers compared to variances.

[17] It is worth noting that we work with the expression for output per worker in terms of *K/L* ratio instead of the *K/Y* ratio used by Mankiw *et al.* (1992), Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999). Hall and Jones (1999, pp.88) argue that there are two advantages in writing the decomposition in terms of the *K/Y* ratio. The first is that the *K/Y* ratio is proportional to the investment rate along a balanced growth path. The second is that writing *y* in terms of the *K/L* ratio tends to increase falsely the explanatory power of factor intensities. However, as argued by Caselli (2005), the *K/Y* is not invariant to differences in TFP and is less appropriate for answering the development accounting question, while the *K/L* ratio writing is more "*intuitive and cleaner*". We have chosen to use the *K/L* formulation primarily because the assumption of a constant *K/Y* ratio seems unreasonable for our sample of Chinese provinces in transition, most of which should be very far from steady-state conditions. See also Bosworth and Collins (2003, pp.131-133) for a good discussion about this issue.

The accounting results for the four specifications using the Hall and Jones approach, given in Table 5 and 6, tell a consistent story. The variations in physical capital intensity and in TFP both explain a large amount of the variation of income levels across provinces, while differences in human capital accounts for a very small fraction. Look at, for instance, the results obtained from the specification under VRS with human capital (the right part of Table 6). All numbers are levels relative to Shanghai in 2000. The first observation is the less variation in human capital, with a standard deviation of only 0.021. Most provinces have about the same human capital as Shanghai. In contrast, both physical capital and TFP (comprising scale efficiency and residual A) are important for the rich provinces and for the poor provinces as well. For example, output per worker in Jiangsu is about 39% of that in Shanghai and this income gap is simultaneously due to lower physical capital intensity (58% of Shanghai value) and to lower TFP level (69% of Shanghai value). The poorest Guizhou has only 31% of Shanghai's physical capital intensity and only 26% of Shanghai's TFP level. Shortage in both these factors explains the relative poverty of Guizhou. On average the three poorest provinces in our sample have just 10.3% of Shanghai's output per worker, i.e., output per worker in Shanghai is about 9.72 times higher. This 9.72-fold difference can be decomposed as the multiplication of a 2.95-fold difference in physical capital, a 1.07-fold difference in human capital and a 3.08-fold difference in TFP level. Obviously, physical capital and TFP play large and almost equally important roles in explaining cross-province income differences.

Table 5 and 6 Here

5. Further discussion

5.1 Robustness check over time

It is useful to check the robustness of our results through time, i.e., whether the contribution of factor inputs and TFP to cross-province income differences would change greatly over time? One way to think about this is to look at what has happened to the standard deviations of y, k^{α}, h^{β} and *TFP* (log) across time, namely the presence or not of σ-convergence. Table 7 reports the standard deviations of these variables throughout the 1982 to 2005 period. The standard deviation of the ln$y$ increases during this period, implying an increasing income inequality across provinces. This increasing income inequality stems from an increasing inequality in TFP combined with an increasing inequality in physical capital intensity. The standard deviation of human capital is consistently small and even experiences a slight decrease. Therefore, our results that differences in both TFP and physical capital intensity account for a large portion of the cross-province income differences are not sensitive to the choice of the year under concern.

Table 7 Here

Another way to check the robustness of our results in the long run is to examine the cross-province distributions dynamics. Figure 6 plots the kernel density estimates of y, k^{α}, h^{β} and *TFP* (log) for 1985, 1990, 1995, 2000, and 2005. Obviously, the cross-province distribution of income has become increasingly twin-peaked. Turning to the components of income, we see that both the physical capital and TFP show movements similar to the income distribution and therefore appear to be the proximate causes of cross-province differences in the long run behavior of income levels. The distribution of the human capital, on the other hand, shows some tendency toward convergence to a high level. Combined with the overall lack of variation in human capital, it is reasonable to conclude that human capital accumulation alone can not explain the long-term cross-province income differences. The apparent bimodality in long-run distribution of income levels is the product of bimodality in the long-run distributions of both the physical capital and TFP.

Figure 6 Here

5.2 Interaction between factors and TFP

An important caveat on our results arises because we have assumed implicitly thus far that TFP is exogenous. If instead there is interaction between factors and TFP, then relying mono-causally on factors and/or TFP to explain cross-province income differences may be inadequate. For instance, recall that the inter-percentile differential measures used by Hall and Jones (1999) bypass the treatment of covariance term between factor inputs and TFP. However, we do observe that provinces with higher physical capital intensity tend to exhibit higher TFP. Performing variance decompositions analogous to those of Klenow and Rodriguez-Clare (1997) yields a contribution of (physical and human capital) capital intensities VD_X=48.1% and a contribution of TFP VD_{TFP}=50.3% to the variation of income across provinces, respectively (the right part of Table 6). Given the small variation of human capital in X, we reach the same conclusion that cross-province income differences are almost equally attributive to differences in physical capital intensity and in TFP levels. But recall that we have assigned half of the covariance term as part of the contribution to X and the other part half to TFP. In fact the covariance term *per se* account for around 44% of the cross-province income dispersion. To examine further this large covariance Table 8 reports the correlations among income, physical and human capital intensity, and TFP[18]. There is a very strong positive correlation between physical capital intensity and TFP. The correlation is 0.842. Despite lack of variation across provinces, human capital is also highly positively correlated with TFP levels (0.604).

Table 8 Here

How to interpret these quite positive correlations between factors and TFP? A possible reason is that there may exist common driving forces beneath factors and TFP such as institutions, geography, language and climate. For instance, Hall and Jones (1999) document that differences in "social infrastructure" drive cross-country variation in both factors accumulation and TFP, and are therefore the fundamental determinant of income differences across countries[19].

However, these fundamental factors are themselves highly persistent and tend to change slowly over

[18] The correlation matrix for the other three specifications are omitted to save space, the results of which are consistent with Table 8.

[19] The social infrastructure involves the institutions and government policies that provide the incentives for individuals and firms: either positive incentives that can encourage productive activities such as skills accumulation and new production techniques, or negative incentives that can encourage predatory behavior such as rent-seeking and corruption (Hall and Jones, 1999, pp.95).

time. It is more likely to think about TFP changes for most provinces as a process of adopting technologies developed at the technological frontier. In such a case, the high correlation between capital intensities and TFP reflects the possibility of positive externalities, suggesting that TFP is not exogenous as invoked by the Solow model. The TFP level could be affected by the factors abundance. Recent authors (e.g., Griffith *et al.*, 2000; Howitt, 2002; Aghion *et al.*, 2005) argue that if technology transfer is costly, the receiving country can not keep up with the frontier just by copying technologies developed in leading countries. The country must make technology investments of its own to acquire foreign technologies and adapt them to the local environment because technological knowledge is often circumstantially sensitive. There may also be linkages between human capital accumulation and TFP. High human capital also facilitates technology adoption. In fact, productive externalities generated by human capital accumulation are the basis for many early endogenous growth models. Therefore, provinces with more factors endowments can undertake more adoption activities, accumulate a larger stock of knowledge, and become more efficient. In particular, though human capital exhibits a very limited exogenous contribution to cross-province income differences, it may still play an important role through reinforcing the externalities between human capital and TFP.

This interpretation of the interaction between capital intensities and TFP is also consistent with the multiple equilibria literature. For instance, Feyrer (2003) examines the joint distributions of capital (physical and human) and TFP using a Markov chain approach. He finds that there is some threshold level of human capital above which TFP growth is driven by spillovers from the R&D producing countries, whereas countries below this level lack the skills to utilize overseas technologies and therefore stagnate. In any case, the intricate interplay between factors and TFP is crucial for understanding the ultimate causes of cross-province income differences and remains a matter for future inquiry.

6. Conclusion

This paper performs the development accounting exercises to investigate the sources of income differences across Chinese provinces during the reform period. Unlike the growth accounting and growth regressions that focus on the sources or the differences in growth rates, the development accounting focuses on the proximate causes of differences in income levels.

Using panel data from a sample of 28 provinces over the period 1982-2005, we estimate the parameters from a Cobb-Douglas aggregate production function involving output per worker and physical capital, both with and without the stock of human capital as an input, and with and without the assumption of constant returns to scale. The parameters estimates are not sensitive to the addition of human capital and to the estimation techniques. In particular, the output elasticity with respect to physical capital is relatively stable within a range of around 0.4-0.5, which is in line with the usual values used in the existing literature. The assumption of constant returns to scale is strongly rejected regardless of the presence of human capital. The production function exhibits high decreasing returns to scale.

With the estimated parameters we conduct the level accounting for the year 2000 using both the traditional Denison approach and the Hall and Jones (1999) approach. We find differences in TFP and in physical capital intensity are both important sources of cross-province income differences, each accounting for roughly half of the variation in income levels. At the same time, the bulk of income differences across provinces has little to do with differences in stock of human capital. The accounting results are robust to whether or not the assumption of constant returns to scale is imposed, and are valid in the long run. Despite small fluctuations, the relative contributions of TFP, physical and human capital intensities are rather stable through time. Our finding is supportive of the consensus view in the development accounting literature that TFP differences can explain a substantially large share of income differences. Meanwhile, as emphasized by Jones (1997) and Easterly and Levine (2001), we should avoid an all-or-nothing view to ignore the equally important role played by physical capital accumulation.

References

Acemoglu, D., Johnson S. and Robinson J. A. (2001). The colonial origins of comparative development: An empirical investigation, *American Economic Review*, 91(5), pp.1369-1401.

Aghion, P., Howitt, P. and Mayer-Foulkes, D. (2005). The effect of financial development on convergence: Theory and evidence, *Quarterly Journal of Economics*, 120(1), pp.173-222.

Barro, R. (1991). Economic growth in a cross section of countries, *Quarterly Journal of Economics*, 106(2), pp.407-443.

Barro, R. and Lee, J-W. (2001). International data on educational attainment: Updates and implications, *Oxford Economic Papers*, 53(3), pp.541-563.

Barro, R. and Sala-i-Martin, X. (1992), Convergence, *Journal of Political Economy*, 100, pp.223-251.

Basmann, R. (1960). On finite sample distributions of generalized classical linear identifiability test statistics. *Journal of the American Statistical Association*, 55(292), pp.650-659.

Baum, C. F., Schaffer, M. E. and Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *Working Paper No.545*, Department of Economics, Boston College.

Benhabib, J. and Spiegel, M. M. (1994). The role of human capital in economic development: Evidence from aggregate cross-country data. *Journal of Monetary Economics*, 34, pp.143-173.

Bernanke, B. S. and Gürkaynak, R. S. (2001). Is growth exogenous? Taking Mankiw, Romer, and Weil seriously, *NBER Macroeconomics Annual*, 16, pp.11-57.

Bils, M. and Klenow, P. J. (2001). Quantifying quality growth, *American Economic Review*, 91(4), pp.1006-1030.

Blundell, R. and Bond, S. (1997). Initial conditions and moment restrictions in dynamic panel data models, *Unpublished discussion paper No.97-07*, University College London.

Borensztein, E. and Ostry, J. D. (1996). Accounting for China's growth performance, *American Economic Review*, 86(2), pp.224-228.

Bosworth, B. P. and Collins, S. M. (2003). The empirics of growth: An update, *Brookings Papers on Economic Activity*, 2, pp.113-179.

Bound, J., Jaeger, D. A. and Baker, R. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90, pp.443-450.

Bourguignon F. and Morrisson C. (2002). Inequality among world citizens: 1820-1992, *American Economic*

Review, 92(4), pp. 727-744.

Brun, J. F., Combes, J. L. and Renard, M. F. (2002). Are there spillover effects between coastal and noncoastal regions in China? *China Economic Review*, 13(2-3), pp.161-169.

Caselli, F. (2001). Is Growth Exogenous? Taking Mankiw, Romer, and Weil Seriously: Comment, *NBER Macroeconomics Annual*, 16, pp.58-62.

Caselli, F. (2005). Accounting for cross-country income differences, Chapter 9 in: Aghion P. and Durlauf, S. (ed.), *Handbook of Economic Growth*, pp.679-741.

Chen, J. and Fleisher B. M. (1996). Regional income inequality and economic growth in China, *Journal of Comparative Economics*, 22(2), pp.141-164.

Chow, G. C. (1993). Capital Formation and Economic Growth in China, *Quarterly Journal of Economics*, 108(3), pp.809-842.

Chow, G. C. and Li, K-W. (2002). China's economic growth: 1952-2010, *Economic Development and Cultural Change*, 51(1), pp.247-256.

Cordoba, J. C. and Ripoll, M. (2005). Endogenous TFP and cross-country income differences, *Working Paper No.247*, Department of Economics, University of Pittsburgh.

Démurger, S. (1999). Elément de comptabilité de la croissance chinoise, *CERDI Working Paper No.13*, Université d'Auvergne.

Démurger, S. (2001). Infrastructure development and economic growth: An explanation for regional disparities in China? *Journal of Comparative Economics*, 29(1), 95-117.

Denison, E. F. (1962). *Sources of Economic Growth in the United States and the Alternatives Before Us*, New York: Committee for Economic Development.

Denison, E. F. (1967). *Why Growth Rates Differ*, Washington DC: The Brookings Institution.

Durlauf, S. N. (2003). The convergence hypothesis after 10 years, *Working Paper No.6*, University of Wisconsin at Madison.

Easterly, W. and Levine, R. (2001). It's not factor accumulation: Stylized facts and growth models, *World Bank Economic Review*, 15(2), pp.177-219.

Easterly, W., Kremer, M., Pritchett, L. and Summers, L. H. (1993). Good policy or good luck? Country growth performance and temporary shocks, *Journal of Monetary Economics*, 32(3), pp.459-483.

Elias, V. J. (1993). Sources of Growth: A study of seven Latin American economies, *Journal of Latin American Studies*, 25(3), pp.675-677.

Färe, R., Grosskopf, S., Norris, M. and Zhang, Z. (1994). Productivity growth, technical progress, and efficiency change in industrialized countries, *American Economic Review*, 84(1), pp.66-83.

Felipe, J. and McCombie, J. S. L. (2004). Is a theory of Total Factor Productivity really needed? *CAMA Working Paper No.12*, Centre for Applied Macroeconomic Analysis, Australian National University.

Feyrer, J. (2003). Convergence by parts, *mimeo*, Dartmouth College.

Gong, L. and Xie, D. (2004). Factor mobility and dispersion in marginal products: A case on China, *Economic Research Journal*, 1, pp.45-53.

Griffith, R., Redding, S. J. and Van Reenen, J. (2000). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries, *CEPR Discussion Papers No.2457*.

Griliches, Z. and Hausman, J.A. (1986). Errors in variables in panel data. *Journal of Econometrics*, 31, pp.93–118.

Hall, R. E. and Jones C. I. (1999). Why do some countries produce so much more output per worker than others? *Quarterly Journal of Economics*, 114(1), pp.83-116.

Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, 50(3), pp.1029-1054.

Hanushek, E. A. and Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American*

Economic Review, 90(5), pp.1184-1208.

Harvey, D. I., Leybourne, S.J. and Newbold, P. (1998). Tests for forecast encompassing, *Journal of Business and Economic Statistics*, 16(2), pp.283-298.

Hausman, J. A. (1978). Specification tests in econometrics, *Econometrica*, 46(6), pp.1251-1271.

Howitt, P. (2002). R&D, implementation and stagnation: A Schumpeterian theory of convergence clubs, *NBER Working Paper No.9104*.

Hu , Z. and Khan, M. S. (1997). Why is China growing so fast? *IMF Staff Papers*, 44(1).

Islam, N. (1995). Growth empirics: A panel data approach. *Quarterly Journal of Economics*, 110(4), pp.1127-1170.

Jones, C. I. (1997). The neoclassical revival in growth economics: Has it gone too far?: Comment, *NBER Macroeconomics Annual*, 12, pp.107-113.

Kim, J-I. and Lau, L. J. (1994). The sources of economic growth of the East Asian Newly Industrialized Countries, *Journal of the Japanese and International Economies*, 8(3), pp.235-271.

King, R. G. and Levine, R. (1994). Capital fundamentalism, economic development, and economic growth, *Carnegie-Rochester Conference Series on Public Policy*, 40, pp.259-292.

Klenow, P. J. and Rodriguez-Clare, A. (1997). The neoclassical revival in growth economics: Has it gone too far? *NBER Macroeconomics Annual*, 12, pp.73-103.

Knight, M., Loayza, N. and Villaneuva, D. (1993). Testing the neoclassical growth model. *IMF Staff Papers*, 40, pp.512-541.

Levinsohn, J. and Petrin, A. (2003). Estimating production functions using inputs to control for unobservables, *Review of Economic Studies*, 70(2), pp.317-341.

Li, J., Gong, F. and Zheng, Y. (1995). Productivity and China's economic growth, 1953-1990, in Tsui, K., Hsueh, T. T. and Rawski, T. G. (eds.), *Productivity, efficiency and reform in China's economy*, Hong Kong Institutes of Asia-Pacific Studies, Chinese University of Hong Kong.

Liu, T. and Li, K-W. (2006). Disparity in factor contributions between coastal and inner provinces in post-reform China, *China Economic Review*, 17(4), pp.449-470.

Loayza, N. (1994). A test of the international convergence hypothesis using panel data. *Policy Research Working Paper No.1333*, The World Bank.

Maddison, A. (1982), *Phases of Capitalist Development*, Oxford: Oxford University Press.

Maddison, A. (1998). *L'Economie Chinoise: Une Perspective Historique*, Etudes du Centre de Développement, OCDE, Paris.

Mankiw, N., Romer, D. and Weil, D. N. (1992). A contribution to the empirics of economic growth, *Quarterly Journal of Economics*, 107, pp.407-437.

McGrattan, E. R. and Schmitz, J. Jr. (1999). Explaining cross-country income differences, chapter 10 in Taylor, J. B. and Woodford, M. (ed.), *Handbook of Macroeconomics*, pp.669-737.

Nakamura, A. and Nakamura, M. (1998). Model Specification and Endogeneity, *Journal of Econometrics*, 83(1-2), pp.213-237.

National Bureau of Statistics of China. (2005). *Comprehensive Statistical Data and Materials on 55 years of New China*, Beijing: China Statistics Press.

National Bureau of Statistics of China. (Various years). *China Statistical Yearbook*, Beijing: China Statistics Press.

Olley, G. S. and Pakes, A. (1996). The dynamics of productivity in the telecommunications equipment industry, *Econometrica*, 64(6), pp.1263-1297.

Parente, S. and Prescott, E. C. (2000). *Barriers to Riches*, Cambridge, MA: MIT press.

Perkins, D. H. (1988). Reforming China's economic system, *Journal of Economic Literature*, 26(2), pp.601-645.

Prescott, E. C. (1998). Needed: A theory of Total Factor Productivity, *International Economic Review*, 39, pp.525-552.

Pritchett, L. (2003). Does learning to add up add up? The returns to schooling in aggregate data, Chapter 11 in Hanushek, E. and Welch, F. (eds.), *Handbook of the Economics of Education*, pp.635-695.

Psacharopoulos, G. (1994). Returns to investment in education: A global update, *World Development*, 22(9), pp.1325-1343.

Quah, D. (1993). Galton's fallacy and tests of the convergence hypothesis, *CEPR Discussion Papers No.820*.

Raiser, M. (1998). Subsidizing inequality: Economic reforms, fiscal transfer and convergence across Chinese provinces, *Journal of Development Studies*, 34(3), pp.1-26.

Rawski, T. (2001). What's happening to China's GDP statistics? *China Economic Review*, 12, pp.347-354.

Romer, D. (2001). Is growth exogenous? Taking Mankiw, Romer, and Weil seriously: Comment, *NBER Macroeconomics Annual*, 16, pp.62-70.

Sargan, J. (1958). The estimation of economic relationships using instrumental variables. *Econometrica*, 26(3), pp.393-415.

Shastry, G. K. and Weil, D. N. (2003). How much of cross-country income variation is explained by health? *Journal of the European Economic Association*, 1(2-3), pp.387-396.

Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *Review of Economics & Statistics*, 79(2), pp.348-352.

Solow, R. (1957). Technical change and the aggregate production function, *Review of Economics and Statistics*, 39(3), pp.312-320.

Wang Y. and Yao Y. (2003). Sources of China's economic growth 1952–1999: Incorporating human capital accumulation, *China Economic Review*, 14(1), pp.32-52.

Woo, W. T. (1998). Chinese economic growth: Sources and prospects, in Fouquin, M. and Lemoine, F. (eds.), *The Chinese Economy*, Economica.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. 1st ed. Cambridge, MA: MIT Press.

Wooldridge, J. M. (2003). *Introductory Econometrics: A Modern approach*. 2nd ed. New York: Thomson Learning.

Word Bank. (1996). *The Chinese economy: Fighting inflation, deepening reforms*, Washington. D.C: World Bank.

Word Bank. (1997). *China 2020: Development challenges in the new century*, Washington. D.C: World Bank.

Wu, D-M. (1973). Alternative tests of independence between stochastic regressors and disturbances, *Econometrica*, 41(4), pp.733-750.

Wu, Y. (2000). Is China's economic growth sustainable? A productivity analysis, *China Economic Review*, 11(3), pp.278-296.

Xu, X. (2002). China's gross domestic product estimation, *China Economic Quarterly*, 2(1), pp.23-36.

Young, A. (2000). Gold into base metals: Productivity growth in the People's Republic of China during the reform period, *NBER Working Paper No.7856*.

Zhang, J., Wu, G. and Zhang, J. (2004). Capital stock estimation of Chinese provinces, *Economic Research Journal*, 10, pp.35-44.

Table 1: Descriptive statistics

Variable		Mean	Std. Dev.	Minimum	Maximum	Observations
Y	overall	2181.39	2481.73	50.84	17025.30	N = 672
	between		1566.91	182.56	5866.03	n = 28
	within		1946.26	-2812.39	13340.66	T = 24
K	overall	3872.83	4619.80	163.80	31561.27	N = 672
	between		2559.24	495.74	9467.30	n = 28
	within		3875.22	-4738.42	25966.80	T = 24
L	overall	2069.94	1403.48	155.61	6335.30	N = 672
	between		1392.91	222.39	5853.53	n = 28
	within		309.94	733.03	3344.66	T = 24
h	overall	2.27	0.34	1.57	3.15	N = 672
	between		0.29	1.75	2.90	n = 28
	within		0.19	1.94	2.82	T = 24

Table 2: Production function estimates (Constant returns to scale)

	Dependent variable: ln$y$		Dependent variable: ln$\tilde{y}$	
	(1)	(2)	(3)	(4)
	OLS	IV	OLS	IV
ln$k$	0.481***	0.490***		
	(21.21)	(17.92)		
ln$\tilde{k}$			0.503***	0.519***
			(22.65)	(19.62)
Implied β	0.519***	0.510***	0.497***	0.481***
	(22.85)	(18.63)	(22.42)	(18.19)
R^2	0.990		0.980	
Number of obs.	672	616	672	616
Heteroskedasticity test				
	258.10	203.07	229.38	200.24
	[0.000]	[0.000]	[0.000]	[0.000]
Instrument used		L2.ln$k$		L2.ln$\tilde{k}$
Instrument relevance test				
Bound *et al*. partial R^2		0.880		0.885
Wu-Hausman test				
		3.22***		2.42***
		[0.001]		[0.016]
Coefficients equality test				
α		0.12		0.44
		[0.727]		[0.506]

Notes: OLS is ordinary least squares (fixed-effect) estimation. IV is two-stage least squares (fixed-effect) estimation with two periods lags of the regressors as instrument variables. All regressions contain province-specific and time-specific fixed effects that are statistically significant at the conventional levels and are not reported. Numbers in parentheses under the estimated coefficients are *t*-statistics calculated from heteroskedasticity-robust standard errors.

Numbers in heteroskedasticity test are White/Koenker statistic for OLS and Pagan-Hall statistic for IV. Numbers in Wu-Hausman endogeneity test are *t*-statistic of the residuals generated from OLS estimation of the first-stage regression and added in the regression of the original model. All numbers in the square brackets under the test statistics are *p*-values.

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table 3: Production function estimates (Variable returns to scale)

	Dependent variable: ln$y$		Dependent variable: ln$\tilde{y}$	
	(1) OLS	(2) IV	(3) OLS	(4) IV
ln$k$	0.430***	0.438***		
	(16.65)	(14.79)		
ln$L$	-0.371***	-0.406***		
	(-5.63)	(-5.20)		
ln$\tilde{k}$			0.433***	0.440***
			(17.73)	(15.48)
ln(hL)			-0.377***	-0.432***
			(-7.29)	(-6.88)
Implied β	0.200***	0.156**	0.190***	0.128***
	(3.60)	(2.32)	(4.14)	(2.28)
R^2	0.991		0.988	
Number of obs.	672	616	672	616
F-Test for CRS				
	31.67	27.04	53.10	47.37
	[0.000]	[0.000]	[0.000]	[0.000]
Heteroskedasticity test				
	304.82	216.56	297.70	207.83
	[0.000]	[0.000]	[0.000]	[0.000]
Instrument used		L2.ln$k$ and L2.ln$L$		L2.ln$\tilde{k}$ and L2.ln(hL)
Instrument relevance test				
Bound *et al*. partial R^2		0.880 and 0.672		0.886 and 0.717
Shea partial R^2		0.900 and 0.687		0.897 and 0.727
Wu-Hausman test				
		2.79*		1.90
		[0.063]		[0.150]
Coefficients equality test				
α		0.10		0.08
		[0.756]		[0.781]
β		0.29		0.91
		[0.589]		[0.340]
α and β jointly		0.22		0.55
		[0.804]		[0.579]

Notes: OLS is ordinary least squares (fixed-effect) estimation. IV is two-stage least squares (fixed-effect) estimation with two periods lags of the regressors as instrument variables. All regressions contain province-specific and time-specific fixed effects that are statistically significant at the conventional levels and are not reported. Numbers in parentheses under the estimated coefficients are t-statistics calculated from heteroskedasticity-robust standard errors.

Numbers in heteroskedasticity test are White/Koenker statistic for OLS and Pagan-Hall statistic for IV. Numbers in Wu-Hausman endogeneity test are F-statistic of the residuals generated from OLS estimation of the first-stage regression and added in the regression of the original model. All numbers in the square brackets under the test statistics are p-values.

*** significant at the 1% level, ** significant at the 5% level, * significant at the 10% level.

Table 4: Summary of output elasticities of capital used in the studies on China

Author	Period under study	Output elasticity with respect to capital
Perkins (1988)	1953-1985	0.4
Li, Gong and Zheng (1995)	1953-1990	0.4-0.5
Chow (1993)	1952-1985	0.6 (Total)
		0.25 (Agriculture)
		0.68 (Industry)
Borensztein and Ostry (1996)	1953-1978	0.47
	1978-1994	0.63
World Bank (1996)	1985-1994	0.5
Hu and Khan (1997)	1953-1994	0.411
	1952-1978	0.386
	1979-1994	0.453
World Bank (1997)	1978-1995	0.4
Maddison (1998)	1952-1995	0.3
Woo (1998)	1979-1993	0.4-0.6
Démurger (1999)	1978-1996	0.548-0.590
Chow and Li (2002)	1952-1998	0.558-0.628
Wang and Yao (2003)	1953-1999	0.33-0.67
Liu and Li (2006)	1986-1998	0.612 (Coastal)
		0.582 (Interior)
Our study	1982-2005	0.481-0.503 (CRS)
		0.430-0.433 (VRS)

Table 5: Hall and Jones level accounting (specifications under CRS)

		Without human capital Contribution from		With human capital Contribution from		
Province	y	k^{α}	A	k^{α}	h^{β}	A
Shanghai	1.000	1.000	1.000	1.000	1.000	1.000
Tianjin	0.615	0.782	0.786	0.774	1.008	0.789
Beijing	0.575	0.811	0.710	0.803	1.017	0.704
Jiangsu	0.393	0.545	0.721	0.531	0.938	0.789
Guangdong	0.388	0.525	0.740	0.510	0.876	0.869
Liaoning	0.372	0.526	0.706	0.512	1.014	0.716
Fujian	0.365	0.502	0.728	0.487	0.897	0.837
Zhejiang	0.351	0.529	0.663	0.515	0.937	0.728
Heilongjiang	0.301	0.471	0.639	0.456	1.010	0.654
Shandong	0.288	0.455	0.634	0.439	0.956	0.687
Xinjiang	0.288	0.569	0.506	0.555	0.927	0.559
Jilin	0.267	0.437	0.611	0.421	0.993	0.638
Hubei	0.254	0.434	0.586	0.418	0.938	0.648
Hebei	0.223	0.428	0.521	0.413	0.942	0.575
I. Mongolia	0.210	0.390	0.538	0.374	0.958	0.585
Shanxi	0.186	0.384	0.485	0.368	0.973	0.520
Qinghai	0.168	0.447	0.376	0.432	0.871	0.448
Jiangxi	0.164	0.321	0.509	0.306	0.926	0.578
Hunan	0.163	0.322	0.506	0.306	0.953	0.558
Anhui	0.155	0.327	0.476	0.311	0.903	0.554
Ningxia	0.152	0.421	0.361	0.405	0.902	0.416
Sichuan	0.143	0.335	0.426	0.319	0.903	0.495
Guangxi	0.139	0.323	0.431	0.307	0.923	0.491
Henan	0.139	0.320	0.435	0.304	0.939	0.487
Shaanxi	0.135	0.367	0.369	0.351	0.956	0.404
Yunnan	0.126	0.332	0.379	0.316	0.822	0.484
Gansu	0.116	0.299	0.389	0.283	0.854	0.480
Guizhou	0.075	0.272	0.274	0.257	0.845	0.343
Average, 27 prov.	0.219	0.423	0.519	0.407	0.931	0.579
Standard deviation	0.136	0.133	0.141	0.135	0.052	0.138
Average, 3 poorest prov.	0.103	0.300	0.343	0.285	0.840	0.431
Variance decomposition		0.509	0.475	0.595		0.390

Notes: All numbers are measured as ratios to the Shanghai values.

Table 6: Hall and Jones level accounting (specifications under VRS)

		Without human capital Contribution from			With human capital Contribution from			
Province	y	k^{α}	$L^{\alpha+\beta-1}$	A	k^{α}	h^{β}	$L^{\alpha+\beta-1}$	A
Shanghai	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Tianjin	0.615	0.803	1.205	0.635	0.802	1.003	1.209	0.632
Beijing	0.575	0.829	1.030	0.674	0.828	1.007	1.030	0.670
Jiangsu	0.393	0.582	0.539	1.252	0.580	0.976	0.534	1.302
Guangdong	0.388	0.563	0.523	1.319	0.560	0.950	0.518	1.409
Liaoning	0.372	0.564	0.693	0.951	0.562	1.005	0.688	0.956
Fujian	0.365	0.540	0.716	0.945	0.538	0.959	0.711	0.995
Zhejiang	0.351	0.567	0.597	1.037	0.564	0.975	0.592	1.076
Heilongjiang	0.301	0.511	0.720	0.819	0.509	1.004	0.716	0.824
Shandong	0.288	0.495	0.488	1.194	0.492	0.983	0.482	1.236
Xinjiang	0.288	0.604	1.000	0.476	0.602	0.972	1.000	0.492
Jilin	0.267	0.478	0.840	0.666	0.475	0.997	0.837	0.673
Hubei	0.254	0.475	0.614	0.873	0.472	0.976	0.609	0.906
Hebei	0.223	0.469	0.546	0.871	0.467	0.977	0.540	0.906
I. Mongolia	0.210	0.431	0.858	0.566	0.429	0.984	0.856	0.581
Shanxi	0.186	0.426	0.758	0.577	0.423	0.990	0.755	0.589
Qinghai	0.168	0.488	1.469	0.235	0.485	0.948	1.479	0.247
Jiangxi	0.164	0.363	0.676	0.667	0.361	0.971	0.672	0.696
Hunan	0.163	0.363	0.545	0.822	0.361	0.982	0.539	0.852
Anhui	0.155	0.368	0.550	0.767	0.366	0.962	0.545	0.811
Ningxia	0.152	0.462	1.395	0.236	0.459	0.961	1.403	0.245
Sichuan	0.143	0.377	0.442	0.855	0.374	0.962	0.436	0.908
Guangxi	0.139	0.365	0.612	0.624	0.362	0.970	0.607	0.653
Henan	0.139	0.361	0.457	0.842	0.359	0.976	0.451	0.881
Shaanxi	0.135	0.408	0.693	0.479	0.406	0.983	0.688	0.493
Yunnan	0.126	0.374	0.635	0.530	0.371	0.928	0.630	0.580
Gansu	0.116	0.340	0.812	0.421	0.338	0.941	0.809	0.452
Guizhou	0.075	0.313	0.662	0.360	0.310	0.937	0.658	0.389
Average, 27 prov.	0.219	0.464	0.705	0.670	0.462	0.973	0.701	0.697
Standard deviation	0.136	0.128	0.268	0.283	0.128	0.021	0.272	0.296
Average, 3 poorest prov.	0.103	0.341	0.699	0.431	0.339	0.935	0.694	0.467
Variance decomposition		0.454	0.530		0.481		0.503	

Notes: All numbers are measured as ratios to the Shanghai values.

Table 7: Standard deviations (specification under VRS with human capital)

Year	ln$y$	ln$X$			lnTFP
			ln(k^{α})	ln(h^{β})	
1982	0.422	0.217	0.205	0.028	0.299
1983	0.429	0.218	0.204	0.027	0.289
1984	0.432	0.221	0.205	0.028	0.281
1985	0.428	0.225	0.207	0.027	0.266
1986	0.429	0.228	0.210	0.027	0.257
1987	0.435	0.236	0.217	0.026	0.254
1988	0.446	0.247	0.228	0.026	0.251
1989	0.448	0.253	0.233	0.026	0.243
1990	0.456	0.259	0.239	0.026	0.242
1991	0.463	0.262	0.242	0.026	0.247
1992	0.472	0.263	0.243	0.026	0.255
1993	0.484	0.265	0.245	0.025	0.262
1994	0.496	0.270	0.249	0.025	0.267
1995	0.511	0.280	0.259	0.026	0.269
1996	0.515	0.286	0.265	0.026	0.264
1997	0.526	0.288	0.268	0.025	0.269
1998	0.538	0.293	0.273	0.026	0.272
1999	0.569	0.300	0.282	0.024	0.294
2000	0.583	0.297	0.284	0.022	0.309
2001	0.584	0.296	0.281	0.022	0.311
2002	0.576	0.287	0.272	0.022	0.315
2003	0.577	0.281	0.267	0.021	0.323
2004	0.577	0.277	0.264	0.021	0.327
2005	0.574	0.270	0.260	0.020	0.331
Average	0.499	0.263	0.246	0.025	0.279

Table 8: Correlation matrix (specification under VRS with human capital)

	y	k^{α}	h^{β}
k^{α}	0.955		
h^{β}	0.639	0.576	
TFP (=$AL^{\alpha+\beta\text{-}1}$)	0.963	0.842	0.604

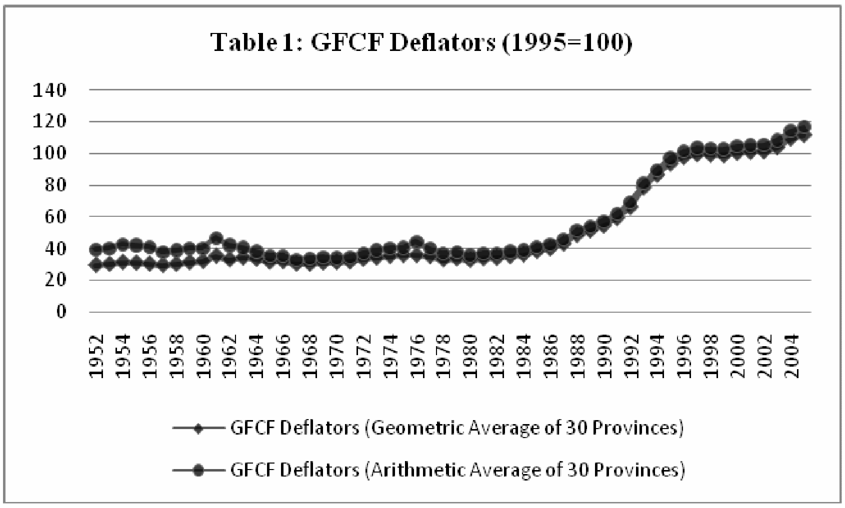
Table 1: GFCF Deflators (1995=100)
140
120
100
80
60
40
20
0
1952 1954 1956 1958 1960 1962 1964 1966 1968 1970 1972 1974 1976 1978 1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004
GFCF Deflators (Geometric Average of 30 Provinces)
GFCF Deflators (Arithmetic Average of 30 Provinces)

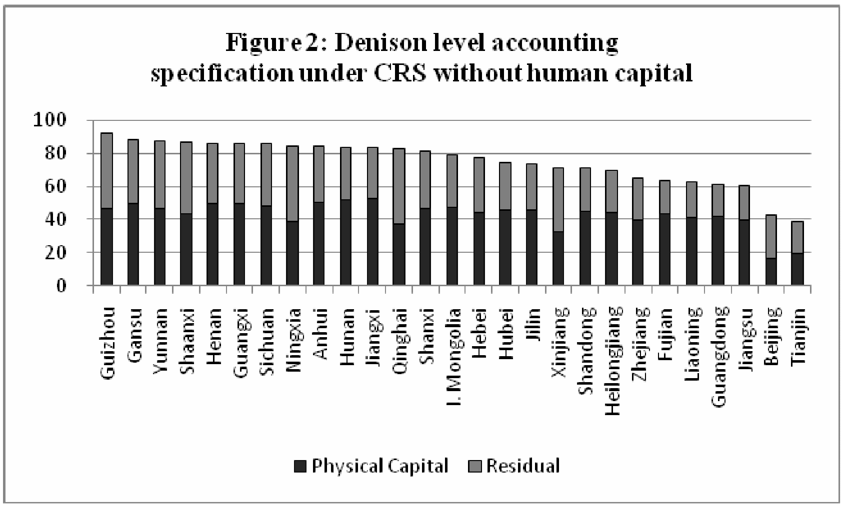
Figure 2: Denison level accounting
specification under CRS without human capital
100
80
60
40
20
0
Guizhou
Gansu
Yunnan
Shaanxi
Henan
Guangxi
Sichuan
Ningxia
Anhui
Hunan
Jiangxi
Qinghai
Shanxi
I. Mongolia
Hebei
Hubei
Jilin
Xinjiang
Shandong
Heilongjiang
Zhejiang
Fujian
Liaoning
Guangdong
Jiangsu
Beijing
Tianjin
Physical Capital
Residual

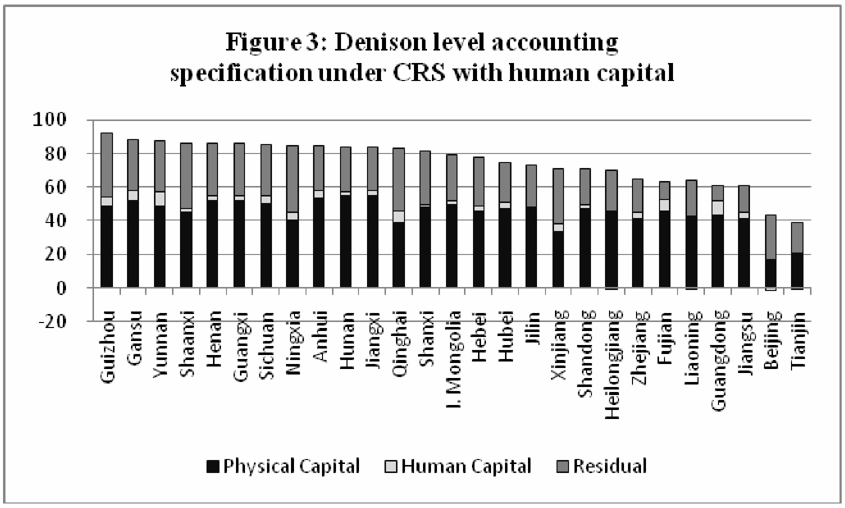
Figure 3: Denison level accounting
specification under CRS with human capital
100
80
60
40
20
0
-20
Guizhou
Gansu
Yunnan
Shaanxi
Henan
Guangxi
Sichuan
Ningxia
Anhui
Hunan
Jiangxi
Qinghai
Shanxi
I. Mongolia
Hebei
Hubei
Jilin
Xinjiang
Shandong
Heilongjiang
Zhejiang
Fujian
Liaoning
Guangdong
Jiangsu
Beijing
Tianjin
Physical Capital
Human Capital
Residual

Figure 4: Denison level accounting specification under VRS without human capital

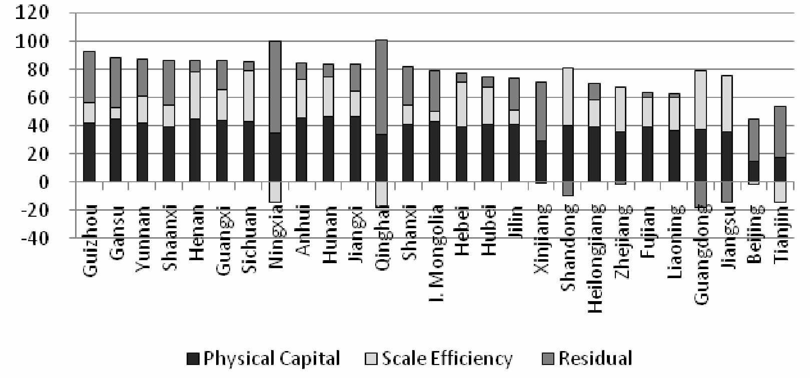

Figure 5: Denison level accounting specification under VRS with human capital

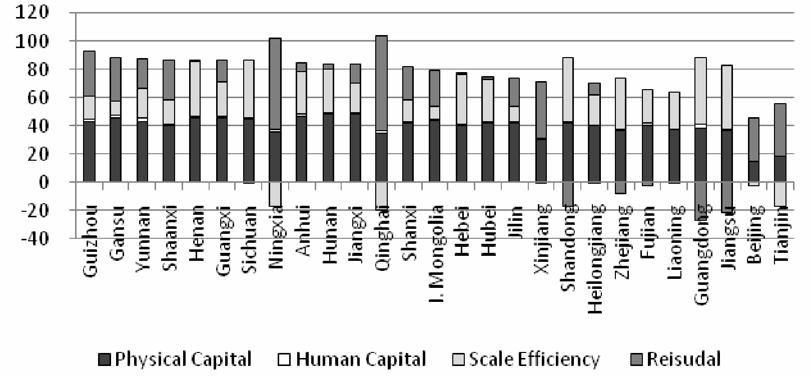

Figure 6: Evolution of the distributions of income, factors inputs and TFP specification under VRS with human capital