Review of Income and Wealth
Series 53, Number 1, March 2007

FORCES SHAPING CHINA'S INTERPROVINCIAL INEQUALITY

BY KAI-YUEN TSUI*

The Chinese University of Hong Kong

This paper explores the forces that shaped China's interprovincial inequality in the last five decades of Communist rule. In so far as the change in interprovincial inequality is the result of differential growth in provincial GDP per capita and provincial economic growth, it may be decomposed into contributions by total factor productivity (TFP) and other factor inputs. A new method is introduced to make this decomposition. Care is exercised in taking into account problems of Chinese official data when implementing the decomposition analysis. The findings suggest that TFP and factor inputs exerted different and sometimes opposing effects on interprovincial inequality in the Maoist and reform era. The increase in inequality from the mid 1960s to the mid 1970s is due to the contribution of TFP overwhelming that of physical capital. The opposite is true for the 1980s. The increase in the 1990s is mainly driven by the skewed distribution of investments in favor of the richer coastal provinces reinforced by the increasing contribution of TFP.

1. INTRODUCTION

The unveiling of the Western Development Program in 1999 by China's former Party Secretary, Jiang Zemin, is a reminder that, in spite of its spectacular economic success, China is still a country full of stark contrasts, conjuring up an image of the western periphery lagging far behind the eastern core. Uneven regional development remains a hotly debated issue half a century after Chairman Mao declared the contradictions between the coastal and inland provinces as one of the so-called "Ten Cardinal Relations" (*shida guanxi*). Often in the limelight is whether regional inequality increased or not. The debate between Lardy (1976) and Donnithorne (1976) in the mid 1970s is a case in point, though this early attempt to tackle the issue was unresolved due to the lack of data at that time. With the release of provincial data in the reform era, subsequent works (see below) have rekindled interest in the subject. Focusing on the historical trend of interprovincial inequality, these studies have tried to loosely tie the oscillation in the trend to policy changes. However, in so far as the change in interprovincial inequality is the end result of different forces at work, reinforcing or counteracting each other, existing studies have left unanswered how important each of those forces are in shaping the trajectory of interprovincial inequality. This paper is an attempt to fill this gap.

The paper introduces a novel framework for identifying quantitatively the factors behind the oscillation in interprovincial inequality since the 1960s. In view

Note: The first draft of this paper was presented at the conference on "Inequality and Poverty in China" organized by WIDER, United Nations University and the Institute of Population and Labour Economics, Chinese Academy of Social Sciences, held in Beijing on April 23–24, 2005. I thank Professor Shorrocks and participants of the conference as well as two referees and the editor of this journal for their valuable comments.

*Correspondence to: Kai-yuen Tsui, Department of Economics, The Chinese University of Hong Kong, Shatin, N.T. (kaiyuentsui@cuhk.edu.hk).

of the concern with the quality plaguing Chinese statistics (e.g. Young, 2000b; Rawski, 2001; Holz, 2003, 2004), this paper also tries to address seriously the problems with Chinese data when analyzing the trend in interprovincial inequality and its contributing factors.

To motivate the subsequent empirical exercise, Section 2 reviews the existing literature to single out those forces affecting regional inequality mentioned in previous studies. In Section 3, a novel framework decomposing overall interprovincial inequality into contributions by factor inputs and total factor productivity is introduced. Addressing the data problems is the subject of Section 4. The next two sections present the empirical results. To motivate the decomposition analysis, the trends in interprovincial inequality are first summarized in Section 5. Next, Section 6 presents the empirical results of the decomposition exercise from which Section 7 extracts the salient findings. Section 8 discusses limitations of our theoretical framework and empirical findings. The concluding section explores possible directions of extensions.

2. Background and Literature Review

There is by now a large literature on China's interprovincial inequality, e.g. Bhalla *et al.* (2003), Fujita and Hu (2001), Jian *et al.* (1996), Lardy (1978), Lyons (1991), Naughton (2002), Lin *et al.* (1999), Lin and Cai (2003), Long (1999), Raiser (1998), Tsui (1991, 1996) and Wang and Hu (1999); the list is by no means exhaustive. An early attempt to seriously study this issue is Lardy (1978). Based on fragmentary statistical information available in the 1970s, he contended that there was a decrease in interprovincial inequality in the prereform era, a conclusion challenged by Donnithorne (1976). Subsequent works since the 1980s have benefited from the release of provincial national income data. Starting with Lyons (1991) and Tsui (1991), these studies have invariably found a trend in interprovincial disparity for the prereform era to be oscillatory rather than a decreasing inequality conjectured by Lardy (1978). After 1978, most studies have found a decreasing trend in the 1980s to be reversed since the early 1990s.

With the issue of coastal-inland disparity being in the limelight, Lin *et al.* (1999), Lin and Cai (2003), and Bhalla *et al.* (2003), among others, have gone one step further by partitioning China into inland and coastal regions and decomposed overall inequality indices into within- and between-region contributions to inequality. Their findings confirm that the gap between the inland and coastal provinces is widening after 1978.

What are the forces driving the changes in inequality? Previous studies have invariably tried to tie the changes in inequality loosely to different forces induced by policy zigzags before and after 1978. Gleaning from the existing literature are two sets of factors driving interprovincial inequality. The first affects provincial economic growth through changing spatial allocation of factor inputs induced by shifts in regional development strategies. A second set of factors may be thought of as changes impinging on the overall efficiency or total factor productivity of the provinces. Setting the stage for our subsequent empirical analysis, the remainder of this section will review these two clusters of factors in turn.

How shifts in regional development strategies before and after 1978 have changed the spatial allocation of resources and ultimately interprovincial inequality is a recurrent theme in previous studies trying to shed light on the trend in interprovincial inequality. For the prereform era, Lardy (1978) and later Naughton (2002) have pointed out that the apparatus of central planning gave the Chinese government a handle to push ahead with its egalitarian regional development strategy, resulting in a spatial allocation of budgetary resources in favor of inland provinces. A case in point is the Third Front Campaign, which was a defense-related program to relocate industries to inland provinces between the mid 1960s and the early 1970s (Naughton, 1988). Lardy (1978) came to the conclusion that regional inequality decreased before 1978.

In addition to investments to boost industrialization and to build up infrastructure in poor provinces, Lardy (1978) also argued that poor provinces in the prereform era appropriated increasing shares of social expenditures on such public goods as health and education, presumably accelerating the accumulation of human capital in less developed regions. The spread of basic education to less-developed provinces was a legacy of the Maoist period that has prompted some scholars, among them Bramall (2000), to argue that investment in education in the Maoist era undergirds the rapid economic growth in the reform era. At the dawn of the reform era, China had a population that was more educated than those of countries with comparable levels of development. In so far as education boosts human capital formation and thus output, the spread of education to less developed provinces may eventually help reduce gaps between rich and poor provinces.[1]

The reform era has witnessed a policy break with the past. New forces unleashed by economic reform have fundamentally altered the spatial distribution of budgetary funds. Raiser (1998) and Naughton (2002) emphasize that the decline in interprovincial inequality in the 1980s was, among other things, the result of a retreat from central planning and redistributive investment policies favoring less developed regions. Fiscal decentralization has allowed local governments, administrative agencies and state-owned enterprises to retain more revenues generated within their jurisdictions, opening up more opportunities to boost their fiscal intakes (e.g. through township and village enterprises). The result has been an explosion in self-raised funds, the distribution of which is highly skewed in favor of the richer coastal provinces.

Changing spatial allocation of resources is not the only channel through which policy shifts affect regional inequality. Some studies have implicitly and explicitly pointed out that switching from central planning to a more market-driven development strategy may impact on the overall efficiency or total factor productivity of provincial economies. In the prereform era, the spatial allocation of funds in favor of inland provinces often did not have such efficiency considerations as comparative advantages and agglomeration economies in mind. The Third Front Campaign, a massive plan to transfer industrial capacities to less-developed regions in complete disregard of their comparative advantages, is

[1]However, the rapid expansion of basic education raises questions about the quality of prereform education.

often perceived as an extreme manifestation of such an apparently egalitarian but highly inefficient development strategy leading to enormous economic waste (e.g. Naughton, 1988; Tsui, 1991). Inefficient regional development policies were allegedly reinforced by the call for self reliance and the formation of a cellular economy (Donnithorne, 1972) that militated against specialization and the emergence of efficient economic structures. The effect of increases in investment favoring some inland provinces was thus offset by other forces that undercut overall productivity.

Many studies have contended that the reversal of these inefficient policies during the reform period has induced productivity-enhancing but spatially differentiated restructuring of provincial economies in response to market incentives and comparative advantages. Rural reforms are often cited as an important reason behind the decrease in regional inequality in the early phase of economic reform (e.g. Raiser, 1998; Naughton, 2002; Bhalla *et al.*, 2003). On the other hand, the emergence of efficient township and village enterprises mainly in coastal provinces since the mid 1980s may have had the opposite effect (e.g. Lin and Cai, 2003).

It is also frequently noted that spatial flows of industrial investments have fallen more in line with comparative advantages, enhancing overall productivity of provincial economies. The reorientation of the industrial strategy from an excessive focus on heavy industries across the board to a judicious exploitation of China's comparative advantage in labor-intensive industries has upset the status quo, and boosted the competitiveness and productivity of such provinces as Guangdong and Fujian on the one hand and set off declines in such industrial powerhouses as Liaoning on the other (Naughton, 2002). Fujita and Hu (2001) have highlighted the industrial agglomeration effect in the coastal region that has become stronger in the reform era when market forces have played an increasingly important role in industrial restructuring. The emergence of such manufacturing hubs as the Pearl River Delta has helped reap the benefits of agglomeration economies. However, how important such effects may be remains a debatable issue. Local protectionism in China, by blocking the free flow of resources between jurisdictions, may have weakened the effect of agglomeration and comparative advantage on regional inequality. How serious market fragmentation induced by local protectionism has been in the reform era remains an unsettled issue (see, e.g. Young, 2000a; Naughton, 2003; Poncet, 2003; Bai *et al.*, 2004). Empirical evidence on Chinese cities seems to suggest that China has yet to fully exploit the effect of agglomeration on productivity (Au and Henderson, 2006). Putting together all the above complicated dynamics, it is *a priori* difficult to ascertain how industrial restructuring impinges on regional inequality without resorting to empirical analysis.

A discussion of economic reform's impact on provincial economic restructuring is not complete without mentioning China's open-door policy. Increasing openness has played a critical role in shaping regional inequality as noted, among others, by Fujita and Hu (2001). Coastal provinces such as Guangdong and Fujian were one step ahead in opening to the outside world. Richer coastal provinces attract preponderant shares of foreign direct investment that have not only increased physical capital for production but also have allegedly brought with them technology and management know-how, boosting productivity in richer provinces.

What emerges from this review of the literature is that different policy regimes emerging in the last four decades unleashed different forces with differential and, at times, opposing impacts on interprovincial inequality. One therefore expects that interprovincial inequality fluctuated in response to the complicated interactions of those forces discussed above. What is missing in previous studies is, however, a sharper analysis to disentangle those forces and to assess their impact on interprovincial inequality. To fill this lacuna, this paper introduces a new framework to assess quantitatively their relative importance by decomposing changes in China's interprovincial inequality into their contributions.

There is one last point regarding the existing literature that is worth mentioning. The empirical studies cited above more often than not measure interprovincial inequality in terms of real GDP per capita with the data derived from official sources. However, there are concerns that official GDP growth rates may have been inflated and that population data based on household registration fail to take proper account of interprovincial migration (e.g. Young, 2000b; Rawski, 2001; Holz, 2003, 2004). The concern with the inflation of provincial growth rates suggests that care must be exercised in using official statistics in general. As explained below, this paper aims to address the data problems seriously.

3. Conceptual Framework and Empirical Strategy

In light of the literature review above, this section introduces a new framework to disentangle the major forces shaping interprovincial inequality. First, an introduction of our notation is in order. The basic unit of our analysis is a province. The set of provinces may, following the Chinese convention, be further partitioned into three regions—East, Central and West.[2] In what follows, the subscript g is used to denote a region (e.g. the East) and gm to denote the m-th province in the g-th region (e.g. Jiangsu in the East). The nominal provincial GDP of the m-th province in the g-th region is $\hat{Y}_{gm}$ (e.g. the nominal GDP of Jiangsu province in the East); its real counterpart is $Y_{gm} = \hat{Y}_{gm}/\Pi_{gm}$, where Π_{gm} is the GDP deflator. The real provincial GDP per capita, y_{gm}, is then defined as Y_{gm}/P_{gm}, where P_{gm} is the total population and the corresponding vector of provincial GDP per capita is $\mathbf{y}$.

Before turning to the nuts and bolts of our new framework, an intuitive explanation may help clarify its logic. Let $I(\mathbf{y})$ be some measure of interprovincial inequality. We try to decompose the change in interprovincial inequality, $dI(\mathbf{y})/dt$,

[2]The directly administered municipalities are incorporated into their neighboring provinces. Hainan and Xizang are excluded. The older classification including Guangxi within the coastal region was first officially used in documents of the 7th Five Year Plan. The new classification began with the unveiling of the Western Development Program. We follow the recent practice of assigning it to the western region. Formerly classified as a province in the central region, Neimenggu has officially been included in the western region after the introduction of the Western Development Program. The partition reflects a mental framework often invoked by officials in their debate regarding and as a basis for the formulation of regional policies. Such a partition therefore provides an interface with and helps better understanding of issues related to policy shifts and regional development.

which depends on the growth rates of provincial GDP per capita, i.e. $\dot{y}_{gm}/y_{gm} = (dy_{gm}/dt)/y_{gm}$. Since $\dot{y}_{gm}/y_{gm}$ is obviously a function of provincial GDP growth rates $\dot{Y}/Y_{gm}$ (see equation (3)), which then depends on policy-induced forces included in standard growth accounting framework, $dI(\mathbf{y})/dt$ is ultimately driven to the interprovincial differential growth in physical and human capital as well as the growth in total factor productivity (TFP).

The remainder of this section is an exposition of the decomposition framework involving two steps. First, we specify and explain how we estimate the provincial production functions. Then, we plug the estimated parameters into the expressions for the contributions of TFP and factor inputs to be derived below.

With regard to the first step, it is convenient to think in terms of provincial production functions:

$$(1) \qquad Y_{gm} = A_{gm} Q_{gm}(K_{gm}, H_{gm}), \quad H_{gm} = E_{gm} L_{gm}$$ [3]

where K_{gm} is the capital stock, H_{gm} is labor adjusted for education attainment being a product of an index of schooling received, E_{gm}, and the labor force, L_{gm}. Q_{gm} is an increasing function of K_{gm} and H_{gm}. A_{gm} is the term capturing total factor productivity (TFP). Recalling our discussion in Section 2, the change in overall provincial efficiency A_{gm} may depend on an R-vector $z_{gm} = (z_{gm1}, \ldots, z_{gmR})$ of factors affecting the overall efficiency of a provincial economy.

Provincial economic growth may then be decomposed into the contribution of the growth of TFP and factor inputs:

$$(2) \qquad \frac{\dot{Y}_{gm}}{Y_{gm}} = \frac{\dot{A}_{gm}}{A_{gm}} + \left(\frac{\partial Q_{gm}/\partial K_{gm}}{Y_{gm}}\right)\frac{\dot{K}_{gm}}{K_{gm}} + \left(\frac{\partial Q_{gm}/\partial H_{gm}}{Y_{gm}}\right)\frac{\dot{H}_{gm}}{H_{gm}},$$

where $\dot{X}/X$ denotes the growth of X. Interprovincial differences in the growth of A_{gm}, K_{gm}, and H_{gm} result in provincial outputs expanding at different paces, contributing ultimately to the change in interprovincial inequality. What is often of interest is the growth in GDP per capita in previous studies on interprovincial inequality. In this connection, the next equation connects the growth in GDP per capita with GDP growth:

$$(3) \qquad \dot{y}_{gm}/y_{gm} = (\dot{Y}_{gm}/Y_{gm}) - (\dot{P}_{gm}/P_{gm}).$$

The growth in provincial GDP per capita is in turn linked to the input factors and total factor productivity via the first term which is growth in GDP. Substituting (3) into (2) results in

[3]There is a question as to why the average number of years of schooling, E_{gm}, is not treated as a separate factor. I try to have a parsimonious model with fewer parameters to be estimated. Adopting more flexible functional forms easily runs into estimation problems such as multicollinearity. As shown below, the present specification results are reasonable.

$$\frac{\dot{y}_{gm}}{y_{gm}} = \frac{\dot{A}_{gm}}{A_{gm}} + \alpha^K_{gm}\frac{\dot{K}_{gm}}{K_{gm}} + \alpha^H_{gm}\frac{\dot{H}_{gm}}{H_{gm}} - \frac{\dot{P}_{gm}}{P_{gm}}, \tag{4}$$

where $\alpha^K_{gm} = (\partial Q_{gm}/\partial K_{gm})/Y_{gm}$ and $\alpha^H_{gm} = (\partial Q_{gm}/\partial H_{gm})/Y_{gm}$. Differential growth rates in TFP, physical capital, human capital, and population result in different growth rates in provincial GDP per capita.

Next, $dI(\mathbf{y})/dt$ hinges on the variation in $\dot{y}_{gm}/y_{gm}$, which in turn depends on changes in A_{gm}, K_{gm}, H_{gm} and P_{gm}. To measure interprovincial inequality, we resort to a population-weighted version of the mean logarithm deviation (also referred to as Theil's entropy measure):[4]

$$I(\mathbf{y}) = \sum_{g=1}^{G}\sum_{m=1}^{M_g} f_{gm}\ln\left(\frac{\bar{y}}{y_{gm}}\right), \quad \bar{y} = \sum\nolimits_{g=1}^{G}\sum\nolimits_{m=1}^{M_g} f_{gm}y_{gm}, \tag{5}$$

where G is the number of regions (eastern, central and western regions in the present context), M_g is the number of provinces in the g-th region; $\mathbf{y} = (\mathbf{y}_1, \ldots, \mathbf{y}_G)$ where $\mathbf{y}_g = (y_{g1}, \ldots, y_{gM_g})$; $f_{gm} = P_{gm}/P$, $P = \sum_{g=1}^{G}\sum_{m=1}^{M_g} P_{gm}$, is population share of a province. First, we differentiate equation (5) with respect to time:

$$\frac{dI(\mathbf{y})}{dt} = \sum_{g=1}^{G}\sum_{m=1}^{M_g}(s_{gm} - f_{gm})\frac{\dot{y}_{gm}}{y_{gm}} + \sum_{g=1}^{G}\sum_{m=1}^{M_g}(s_{gm} - f_{gm}\ln(y_{gm}))\frac{\dot{f}_{gm}}{f_{gm}}, \tag{6}$$

where t is time and $s_{gm} = Y_{gm}/Y$, $Y = \sum_{g=1}^{G}\sum_{m=1}^{M_g} Y_{gm}$.[5] The first term on the right hand side of equation (6) captures the impact on inter-provincial inequality of differential growth rates across provinces, while the second term summarizes the impact of changing population shares. It is interesting to note that the impact of $\dot{y}_{gm}/y_{gm}$ on inequality hinges on the sign of the term $(s_{gm} - f_{gm})$. With the income share of a province, s_{gm}, falling below its population share, f_{gm}, transferring more output to that province reduces inequality.[6]

Substituting (4) into (6), the change in inequality depends on the growth of TFP and factor inputs, i.e.,

[4]Measuring interprovincial inequality in terms of provincial GDP per capita is a common practice in previous studies, not least because provincial GDP data are most readily available, though some studies (e.g. Lyons, 1991; Bhalla *et al.*, 2003) do also use consumption per capita in addition to GDP per capita. Following this practice thus facilitates comparing our results with those in previous studies. Furthermore, GDP is the appropriate aggregate output measure for the estimation of provincial production functions, i.e. equation (1), in the present context. It is well known that GDP is just the output produced by a province, while what local residents actually receive should be measured by provincial GNP, which includes transfers into and out of a province. However, provincial GNP data are not available. While GDP per capita is often used as a measure for the level of economic development, GNP better reflects income at the disposal of residents in a province.

[5]Details for deriving equation (6) are provided in Appendix A.

[6]This is analogous to the Pigou–Dalton transfer principle whereby an inequality measure satisfies this principle if transfer of a dollar from a richer to a poorer person without changing their total income leads to a fall in inequality (see, e.g. Sen, 1997).

$$\frac{dI(\mathbf{y})}{dt} = CA + CK + CH + CP + CF \tag{7}$$

where the mathematical expressions for the right hand side terms may be found in Appendix A. The terms *CA*, *CK*, *CH* may be interpreted as the contributions of spatial variations in the growth in TFP, physical capital and human capital to the *change* in interprovincial inequality respectively. Capturing the effect of population growth on inequality is the term *CP*: faster population growth in a poor region *ceteris paribus* results in an increase in interprovincial inequality. Finally, *CF* summarizes the effect of changing population shares on inequality. Whenever any of the right-hand-side components is positive, it contributes to an increase in interprovincial inequality. An example with respect to *CK* helps illustrate the mechanism linking the contributions to the change in inequality. Since *CK* is equal to $\sum_{g=1}^{G}\sum_{m=1}^{M_g}(s_{gm} - f_{gm})\alpha_{gm}^{K}(\dot{K}_{gm}/K_{gm})$ (see Appendix A) and a "poor" province's share of GDP is less than its population share, a 1 percent increase in K_{gm} leads to a marginal increase in its output equal to α_{gm}^{K} and thus a decrease in inequality because $(s_{gm} - f_{gm}) < 0$. Holding all other things constant, faster growth in GDP of a poor province due to more physical investment reduces interprovincial disparity. The same logic applies to the other components.

An important dimension of interprovincial inequality that has generated a lot of attention is the gap between the coastal provinces as opposed to the central and western provinces. Indeed, as shown above, hidden behind the oscillation in *overall* interprovincial inequality may be divergent changes in between-region and within-region inequalities. To gain a richer picture of interprovincial inequality, each term in equation (7) may be further decomposed into between- and within-region contributions. To see this, equation (5) is first decomposed into within-region and between-region inequality as follows:[7]

$$I(\mathbf{y}) = WG(\mathbf{y}) + BG(\mathbf{y}), \tag{8}$$

where within-region inequality is defined as

$$WG(\mathbf{y}) = \sum_{g=1}^{G} f_g I(\mathbf{y}_g), \quad f_g = \sum\nolimits_{m=1}^{M_g} P_{gm}/P,$$

while between-region inequality assumes the following form:

$$BG(\mathbf{y}) = \sum_{g=1}^{G} f_g \ln\left(\frac{\bar{y}}{\bar{y}_g}\right), \quad \bar{y}_g = \sum\nolimits_{m=1}^{M_g} Y_{gm} \Big/ \sum\nolimits_{m=1}^{M_g} P_{gm}.$$

Differentiating $WG(\mathbf{y})$ and $BG(\mathbf{y})$ with respect to time results in the following two expressions:

[7]See, e.g. Shorrocks (1984).

$$(9)\qquad \frac{dW(\mathbf{y})}{dt} = WCA + WCK + WCH + WCP + WCF,$$

$$(10)\qquad \frac{dBG(\mathbf{y})}{dt} = BCA + BCK + BCH + BCP + BCF,$$

where the terms on the right hand side of equations (9) and (10) are respectively the within-region and between-region contributions corresponding to the components in equation (7) (see Appendix A for the exact expressions). For example, CK is the sum of within-region (WCK) and between-region (BCK) contributions.

Finally, to facilitate our discussion in subsequent sections, it is helpful to derive *cumulative* changes in inequality induced by the different components above. For example, in the case of TFP, its cumulative contribution is defined as follows:

$$(11)\quad CCA_T = \sum\nolimits_{s=T_0}^{T} CA_s,\quad WCCA_T = \sum\nolimits_{s=T_0}^{T} WCA_s,\quad BCCA_T = \sum\nolimits_{s=T_0}^{T} BCA_s$$

where CCA_T, $WCCA_T$ and $BCCA_T$ are the *cumulative* contribution of TFP to the *change* in overall, within-region and between-region inequality from T_0 up to the period T. Other cumulative changes in equations (7), (9) and (10) may be defined accordingly. The cumulative contributions are presented instead of their one-period (annual) counterparts because it is *visually* easier to detect longer-term trends which are often hard to discern from the one-period contributions due to their short-run volatilities. If one-period increases in contributions outweigh their one-period decreases, then cumulative contributions will exhibit an upward trajectory. This formulation also renders it easier to relate to the level of inequality $I(\mathbf{y})$ in so far as it is the cumulative counterpart of $dI(\mathbf{y})/dt$.

Implementing the decomposition exercise above involves estimating equation (1). To render the estimation manageable, the provincial production functions are assumed to be log-linear:

$$(12)\qquad Y_{gm} = A_{gm}(\mathbf{z}_{gm}) K_{gm}^{\alpha_K} H_{hm}^{\alpha_H},\quad A_{gm}(\mathbf{z}_{gm}) = e^{\lambda_{0gm} + \sum_{r=1}^{R} \lambda_r z_{gmr}}.$$ [8]

The estimated parameters are then substituted into the equations for the contributions, i.e. the right-hand-side terms in equations (7), (9) and (10) to arrive at contributions of the different factors and TFP (see Appendix A). We assume that, with the possible exception of λ_{0gm} corresponding to provincial fixed effects, all other coefficients are assumed to be stable across provinces so that provincial data are pooled for panel estimation. Given our specification above, $\dot{A}_{gm}/A_{gm} = \sum_{r=1}^{R} \lambda_r dz_{gmr}/dt$. Using the expression of CA from Appendix A, the contribution of TFP may further be expressed as the sum of the contributions by z_r:

[8]Similar specifications appear in, for example, Bosworth and Collins (2003).

(13) $$CA = \sum_{g=1}^{G} \sum_{m=1}^{M_g} (s_{gm} - f_{gm})(\dot{A}_{gm} / A_{gm}) = \sum_{r=1}^{R} Cz_r$$

where $Cz_r = \sum_{g=1}^{G} \sum_{m=1}^{M_g} (s_{gm} - f_{gm}) \lambda_r dz_{gmr} / dt$, $r = 1, 2, \ldots, R$, is the contribution of z_r. The corresponding between- and within-region contributions may be derived accordingly.

Total factor productivity as summarized by the term A_{gm} is often estimated as a residual after deducting the contribution of capital and labor and is thus often a black box containing possibly a smorgasbord of factors ranging from technological advances to policy changes that may ultimately impact on productivity. Motivated by our discussion in Section 2, we try below to explore some policy-related factors that may impinge on TFP, though caution must be exercised regarding the interpretation of TFP and our results below are tentative. We incorporate into $\mathbf{z}_{gm}$ the share of foreign direct investment actually utilized (*shiji liyong waizi*) to GDP (FDI) as a proxy for openness.[9] Since FDI measures the importance of foreign direct investment in the provincial economy and *not* the level of investment, the variable is not already part of K. Its inclusion is intended to capture productivity-boosting effects induced by the transfer of technology and management know-how. Another factor that possibly affects TFP pertains to the spatial reshuffling of industries alluded to in Section 2. In the present context, we treat the effects induced by industrial restructuring as among the factors in $\mathbf{z}_{gm}$ that may affect productivity. The effect is measured by a province's output from the secondary sector as a share of the national share, i.e. $Y_{2k} / \Sigma_m Y_{2m}$, where Y_{2k} is the k-th province's output of the secondary sector and the denominator is the national total. This is a measure of industrial agglomeration for a province.[10] As an illustration, Figure 1 summarizes the shares for Guangdong and Liaoning for the period under study. Underlying the increase in the share of Guangdong since 1978 may be an increase in the productivity of the province due to the exploitation of comparative advantages and economies of scale producing for the world market.[11]

In addition to these two key variables, also included is a dummy variable for the initial years of the Cultural Revolution (CRV) to capture productivity shocks

[9]The data are mainly from the National Bureau of Statistics, Department of Comprehensive Statistics (1999), supplemented by data from provincial yearbooks. The choice of FDI as a proxy was decided after much deliberation and data searching. For one thing, a complete and consistent time-series and cross-section dataset for imports and exports for the period 1964–99 is hard to come by. Many provinces switched from reporting trade data collected by the former Ministry of Foreign Trade to reporting customs statistics, rendering provincial time series intertemporally inconsistent. Indeed, many provinces only report trade data for the reform period (see, e.g. National Bureau of Statistics, Department of Comprehensive Statistics, 1999). In addition, changing definitions for certain categories of trade also render time series incomparable. A case in point is Guangdong's processing trade. There was a jump in Guangdong's exports from US$42.51 million in 1986 to US$101.4 million in 1987 because value added instead of total value of processing trade was recorded before 1987.

[10]In the literature on agglomeration, the regional share of industry is often used as a measure for agglomeration (see, e.g. Brakman *et al.*, 2001).

[11]We have also experimented with the share of primary output. The rationale is that TFP growth in many developing countries is due to structural transformation. However, in the present context, this variable does not seem to be significant.

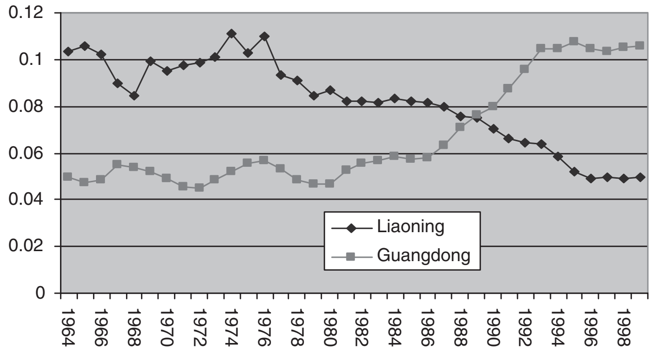

Figure 1. Shares of Secondary Sectors of Guangdong and Liaoning

Notes: For each province, the value for each year is equal to the secondary sector of the province divided by the national total output of the secondary sector.

induced by political turmoil. A time trend for the reform era (TD) is also included in $\mathbf{z}_{gm}$ to take account of any time-dependent factors, technological and otherwise, that may be left out. The variables included far from exhaust all the factors driving TFP growth. Data limitations, however, prevent us from embarking on a more comprehensive investigation.

Including all the above explanatory variables and taking the logarithm of equation (12), the equation to be estimated is as follows:

$$(14)\qquad \ln(Y_{gm}) = \lambda_{0gm} + \lambda_{FDI}FDI + \lambda_{IND}IND + \lambda_{CRV}CRV + \lambda_{TD}TD + \alpha_K \ln(K) + \alpha_H \ln(H).$$

If there is constant returns to scale, i.e. $\alpha_H = 1 - \alpha_K$, the above equation becomes:

$$(15)\qquad \ln(Y_{gm}/H_{gm}) = \lambda_{0gm} + \lambda_{FDI}FDI + \lambda_{IND}IND + \lambda_{CRV}CRV + \lambda_{TD}TD + \alpha_K \ln(K/H_{gm}).$$

We estimate both of these equations and choose the one that best fits the data. The estimated coefficients are then substituted into equations (7)–(10) (see also Appendix A).

4. Data Issues

In view of the concerns over China's official data alluded to above, let us first explain how we tackle the problem to arrive at the adjusted data used for our subsequent empirical exercise. The section also explains how we have constructed data for provincial capital stocks and schooling.[12]

[12]A longer document explaining in greater detail the data adjustments is available from the author on request.

Real Provincial GDP

Instead of using the official implicit GDP deflators, we follow the suggestion of Keidel (2001), deflating the expenditure components making up GDP by their respective price indices.[13] Specifically, nominal GDP is the sum of the following expenditure categories:

$$\hat{Y}_{gm} = CR_{gm} + CU_{gm} + CG_{gm} + CF_{gm} + NX_{gm}$$

where $\hat{Y}_{gm}$ = nominal GDP, CR_{gm} = rural consumption, CU_{gm} = urban consumption CG_{gm} = government consumption, CF_{gm} = gross capital formation (gross fixed capital formation plus changes in inventories), and NX_{gm} = net export. The data are from the National Bureau of Statistics, Department of Comprehensive Statistics (1999) up to 1998 and from provincial yearbooks for 1999.[14] For the period from the mid 1980s onwards, provincial rural and urban consumer price indices (CPIs) are used to deflate CR_{gm} and CU_{gm} respectively. Provincial CPI indices are used to deflate CG_{gm}. For the preceding period, only provincial cost-of-living indices for staff and workers are available for deflating CU_{gm}. We have no choice but to resort to provincial retail price indices (RPIs) for deflating CR_{gm}. Comparing RPIs and CPIs for a few provinces for which both indices are available, their movements are very similar. In any case, prices in the prereform period were by and large fixed administratively.

In the case of gross capital formation CF_{gm}, provincial price indices for fixed asset investment (*guding zichan touzi jiage zhishu*) are available for the period from 1992 to 1999 (National Bureau of Statistics, Department of Fixed Assets Investment Statistics (2002)). Between 1985 and 1991, the price indices for capital goods recently released by the National Bureau of Statistics, Urban Socio-Economic Survey Team (2001) are used to arrive at provincial price indices for capital investment. From 1964 to 1984, deflators based on the official implicit deflators for gross capital formation are used.

Finally, provincial retail price indices are used to deflate net exports.

Real Capital Stock

To estimate the production function in the text, data for real capital stock are required. We first derive provincial initial capital stocks for 1952 following a procedure proposed by Nehru and Dhareshwar (1993). Then, real provincial capital stock series are derived using the perpetual inventory approach:

[13]Our construction of the provincial GDP deflators implicitly assumes that there is no spatial variation in prices. At the time of doing research for this paper, we had no choice but to make this assumption because provincial prices for *all the provinces* for constructing provincial GDP deflators, taking into account spatial variation in prices, were hard to come by. Recently, after completion of this paper, an interesting and useful paper by Brandt and Holz (forthcoming) has filled this gap.

[14]Jiangxi and Guangdong do not have data for the prereform period. However, these provinces have information on consumption (*xiaofei*) and accumulation (*jilei*) under the socialist material product system (MPS) available in National Bureau of Statistics, Department of National Economic Balance (1987). To derive deflators for these provinces, the best we can do is to deflate consumption and accumulation as defined in the MPS by their respective price indices. In the case of accumulation, the national implicit deflator for capital formation is used to deflate accumulation. The implicit deflators so derived are then used to deflate the prereform GDP series for these two provinces.

$$K_{gmt} = (1-\delta)\, K_{gm,t-1} + FA_{gmt}.$$

where FA_{gmt} is gross fixed capital formation (*guding ziben xingchen zonge*) and δ is the rate of depreciation assumed to be 5 percent. Since the sample period for our econometric estimation below is 1964–99, biases embedded in the initial capital stocks may be ameliorated after 12 years. The perpetual inventory approach requires nominal data for fixed asset investment and their corresponding price deflators. The nominal data used are those of gross fixed capital formation from the national income account. Up to 1998, the data are from the National Bureau of Statistics, Department of Comprehensive Statistics (1999) and the 1999 data are again from provincial yearbooks. The same method for the deflation of gross capital formation explained above is also applied to arrive at provincial deflators for FA_{gmt}. For the period before 1985, the official implicit deflators for gross fixed capital formation are derived using data from the National Bureau of Statistics, Department of National Accounts (1997). It is to be noted that this deflator is not the same as those for gross capital formation in the previous section, which includes changes in inventories. Between 1985 and 1991, the same deflator alluded to above is constructed from prices of capital goods in the National Bureau of Statistics, Urban Socio-Economic Survey Team (2001). After 1992, the official price indices for fixed asset investment from the National Bureau of Statistics, Department of Fixed Assets Investment Statistics (2002) are used. Finally, missing prereform data for Guangdong and Jiangxi are estimated with the help of fixed-asset accumulation figures under the socialist Material Product System from the National Bureau of Statistics, Department of National Economic Balance (1987).[15]

Population

Using population data based on household registration records (*huji renkou*) in the reform era to derive provincial GDP per capita is not without problems due to the failure to take proper account of interprovincial population mobility. In so far as the data problems created by migration are not so serious for the prereform period, the official population figures from the National Bureau of Statistics, Department of Comprehensive Statistics (1999) and from provincial yearbooks are used. From 1982 onwards, we extrapolate the population figures made up of age-specific data from the 1982, 1990 and 2000 population censuses. We have collected age-specific populations for all the provinces from the 1982, 1990 and 2000 censuses. The "residual method" is used (e.g. Bogue *et al.*, 1982), whereby survival rates from the provincial life tables are used to project age-specific population cohorts forward from one census to the next. Discrepancies between the projected and actual figures from the next census are assumed to be equal to net in-migration (i.e. in-migration minus out-migration). Such differences are then allocated evenly to the inter-census years as in the case of schooling.

[15]Details may be found in a longer document available from the author on request.

Labor Force Adjusted for Education Attainments

Provincial labor force adjusted for education attainments is the product of the labor force, L_{gm} (*congye renyuan*) multiplied by average years of schooling E_{gm}. Provincial labor force statistics up to 1998 may be found in the National Bureau of Statistics, Department of Comprehensive Statistics (1999) and the 1999 data are again from provincial yearbooks. Workers on furlough (*xiagang*) were subtracted from total labor force for the period from 1994 to render the series intertemporally consistent.

To arrive at E_{gm}, we follow previous studies by first dividing the total labor force into categories with respect to a characteristic capturing labor quality (e.g. Denison, 1967; Griliches, 1970; Jorgenson and Fraumeni, 1992; Barro and Lee, 1996). Our choice of education attainment as such a characteristic is due to data limitations. Essentially, shares of adult population with different levels of education are multiplied by the respective durations of schooling and are then aggregated to arrive at average years of schooling, an approach analogous to Barro and Lee (1996). A similar approach has been adopted by Wang and Yao (2003) as well as Qian and Smyth (2006) to study the effect of education on growth for China at the national level.

We arrive at provincial estimates of average years of schooling by resorting to age-specific data from the 1982, 1990 and 2000 censuses on working-age populations with different levels of education, viz. no formal education (*wenmang*), primary (*xiaoxue*), junior high (*chuzhong*), senior high (*gaozhong*) and university (*daxue*) collected from *all* the provincial publications on the censuses. A method similar to the estimation of population above is used, taking into account interprovincial mobility. Interprovincial mobility is also taken into consideration as above. E_{gm} is then the weighted sum of provincial population shares with different education levels, with weights being their respective durations of schooling (e.g. six years for primary school, etc).

5. The Trend in Interprovincial Inequality

The present and the next section show the empirical results of this study. Before presenting the results on the contributions of different factors to interprovincial inequality, we first turn to the historical trend in China's inter-provincial inequality for the period 1952–99 to provide a bird's-eye view of the changes in interprovincial inequality with respect to provincial GDP per capita, paving the way for our decomposition exercise.[16]

As the data adjustments described above make a difference, we compare the trends in interprovincial inequality based on official and adjusted data. Figure 2 reports three set of results using the mean logarithm deviation (or Theil's entropy)

[16]The period under study stops at 1999. While it is in theory possible to extend the study using more recent data, there are problems with changing definitions, especially for provincial populations after 2000, rendering some data not intertemporally comparable. Before the 2000 census, permanent population (*changzhu renkou*) includes temporary residents (*zanzhu renkou*) staying within the jurisdiction for more than one year. The new definition changes the cutoff point to half a year. As a result, provincial populations in 2000 are much larger.

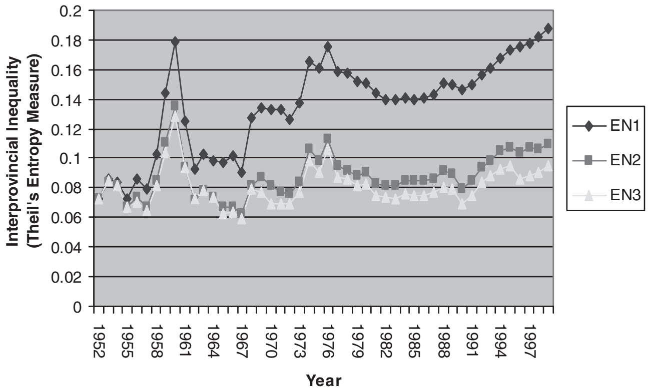

Figure 2. Overall Interprovincial Inequality

Notes: Trends in interprovincial inequality based on mean logarithm deviation.

EN1 is the trend in interprovincial inequality based on official data; EN2 and EN3 are trends based on adjusted data as detailed in Section 4.

index. EN1 is based on unadjusted official data. EN2 and EN3 are derived using our estimates of provincial GDP deflators and population figures as detailed in the last section. EN3 is different from EN2 in the deflation of gross capital formation. While EN3 adopts the official implicit deflators for capital formation, we adjust for some anomalous figures of the official deflators in arriving at EN2. It turns out that EN2 and EN3 are not that different. In what follows, we thus focus only on a discussion of EN1 and EN2.

A striking feature is that the magnitude of EN1 is distinctly higher than EN2. What is even more interesting is that the two trajectories diverge in some sub-periods even though their overall patterns of oscillations share certain salient features. Up to 1967 and except for the anomalous years of the Great Leap Forward, EN1 ratchets upwards while EN2 edges downwards. Inequality shoots up in 1968, possibly due to production disruptions when the chaos induced by the Cultural Revolution was in full swing. For the period between 1968 and the early 1970s, EN2 moves downwards while EN1 is almost stationary. This period coincides with the heydays of the Third Front campaign with investment funds pouring into inland provinces. But since 1973 all series were climbing upward, reaching a peak at 1976. The dawn of the reform era set off a conspicuous decline in interprovincial inequality up to the mid 1980s for both EN1 and EN2. From then onwards, the two trends initially crawl upwards to be followed by sharp increases in the first half of the 1990s. Unlike EN2 whose trend remains by and large stable since 1995, the trend using official data continues to climb upwards. This finding seems to suggest that the richer provinces underestimate their rates of inflation, thereby exaggerating the increase in inequality.

Since much attention has been focused on the gap between the coastal and inland provinces, we decompose $I(\mathbf{y})$ into within- and between-region inequality as in equation (8) and reports in Figures 3 and 4 the trends using different deflators, with WGi and BGi corresponding to ENi, i = 1, 2, 3. Within-region inequality oscillates up to the mid 1970s and then moves downward. The decrease tapers off

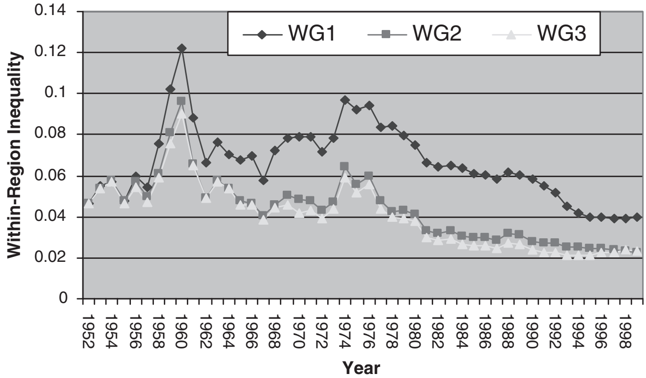

Figure 3. Within-Region Inequality

Notes: WG1 is based on official data; WG2 and WG3 are based on adjusted data. WGi corresponds to ENi, i = 1,2,3, in Figure 2.

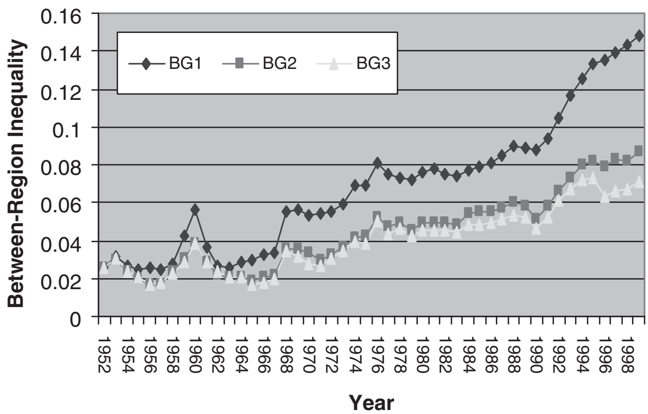

Figure 4. Between-Region Inequality

Notes: BG1 is based on official data; BG2 and BG3 are based on adjusted data. BGi corresponds to ENi, i = 1,2,3, in Figure 2.

in the second half of the 1990s. With regard to BGi, i = 1, 2, 3, their trends oscillate but are essentially increasing. The jumps around the late 1950s and in 1968 capture the policy shocks of those chaotic years. There is a short interlude coinciding with the Third Front Campaign when downward trends are discernible. From then onwards, inequality increases for all the three cases, reaching a temporary peak at 1976. The reform era ushers in a period of mildly declining or almost stationary between-region inequality. Since the mid 1980s, between-region inequality for the three trends inches upwards first and then the increase accelerates. The three trajectories however diverge in the second half of the 1990s with the trend based on official data still exhibiting a distinctly increasing trend while the trends for the other two series are much less pronounced.

TABLE 1

Panel Unit Tests for Variable Used in Regressions

Lag p	t-bar Test Statistics	Lower Tail Area
	LNY	
2	6.65	0.99971
3	6.47	0.99994
	LNK	
2	12.62	0.99971
3	13.07	0.99994
	LNH	
2	−3.27	0.99971
3	−2.48	0.99994
	FDI	
2	3.44	0.99971
3	3.87	0.99994
	IND	
2	0.67	0.74793
3	1.39	0.91804

Source: Author's calculations.

Notes:

The t-bar statistics are derived from the panel unit root tests with serially correlated errors recommended by Im *et al.* (2003). The t-bar statistics are asymptotically normal. Critical value for the test depends on the time series dimension T and p, the lag with respect to the autogressive process assumed for the time series in question (see table 3 in Im *et al.*, 2003, p. 66). The results without time trend and for p equal to 2 and 3 are presented.

The null is that the series is non-stationary. The null is not rejected.

6. Decomposition Analysis

The decomposition exercise involves two steps as detailed in Section 3. With regard to the first step, we present the results from the estimation of the production function, i.e. equations (14) and (15). The second step involves plugging the estimates for λ_r, α_K and α_H into equations (7), (9) and (10) to arrive at the contributions of TFP, K and H.

We briefly summarize how we estimate equation (14). Covering the period from 1964 to 1999, provincial data are pooled to increase the degrees of freedom. The data for the regression are derived as explained in Section 4, which gave a set of real GDP per capita series as indicated as EN2 above. Panel unit-root tests proposed by Im *et al.* (2003) reject the time series of the variables included in the above equation as stationary (Table 1). To avoid the potential problem of spurious correlation based on level data, we estimate the first-difference form of equation (14) (Temple, 1999):

$$(16)\quad \Delta \ln(Y_{gm}) = \lambda_{FDI}\Delta FDI + \lambda_{IND}\Delta IND + \lambda_{CRV}\Delta CRV + \lambda_{TD}\Delta TD + \alpha_K \Delta \ln(K) + \alpha_H \Delta \ln(H)$$

where $\Delta X = X - X_{-1}$, X being any of the variables in the level equation. First differencing is also applied to equation (15). The provincial fixed effects, if any,

TABLE 2
Regression Results from Production Function Estimation

	Log-Linear	Constant Returns to Scale
Ln(K)	0.4765***	
	(5.1743)	
Ln(H)	0.5989***	
	(2.5311)	
Ln(K/H)		0.4513***
		(3.6700)
FDI	0.4525**	0.4606**
	(0.2457)	(1.8581)
IND	7.1664***	7.2270***
	(12.2763)	(12.8106)
CRV	–0.1030***	–0.1041***
	(–6.8251)	(–9.5102)
TD	0.0117	0.0164**
	(0.9156)	(1.6503)
Adjusted R^2	0.315	0.325
No. of observations	900	900

Notation:

Dependent variable: logarithm of real provincial GDP.

Explanatory variables: ln(K) = logarithm of capital stock; ln(H) = logarithm of labor force adjusted for education attainments; FDI = share of foreign direct investment actually utilized to GDP; IND = provincial share of secondary sector to national total; CRV = dummy for the Cultural Revolution; TD = time trend for the reform era.

Notes:

The estimates for the log-linear specification is based on equation (14) with *no* restriction imposed on the sum of α_K and α_H. Assuming constant returns to scale (i.e. $\alpha_K + \alpha_H = 1$) results in equation (15) in the text from which the above estimated coefficients under constant returns to scale are derived.

The estimates are based on pooling the provincial data and estimating the first-difference form of the log-linear production, i.e. equation (16).

The variance–covariance matrix is robust to autocorrelation and heteroskedasticity following Arellano (2003, pp. 18–19). Figures in brackets are robust t statistics.

***Significant at 1%, **significant at 10%.

captured by λ_{0gm} are removed after first differencing.[17] The results from the pooled time-series and cross-section regressions are summarized in Table 2. Two sets of estimates are presented, one for regressions with a constant-returns-to-scale restriction (i.e. $\alpha_H + \alpha_K = 1$) and one without. Robust t statistics are based on variance–covariance matrices, taking into account autocorrelation and heteroskedasticity (Arellano, 2003, pp. 18–19).

The estimated models are subject to a series of tests and the results are summarized in Table 3. On the whole, the estimates are with expected signs to a reasonable extent. To test whether the right-hand-side variables are correlated with the error term and thus inconsistent, we invoke an F test and a robust Wald test for strict exogeneity proposed by Wooldridge (2002, p. 285). The F test assumes no heterogeneity and autocorrelation for the error term, while the Wald test is based on a robust variance–covariance matrix suggested by Arellano (2003). As reported in Table 3, neither of them detects endogeneity of the explanatory variables, implying that the estimates are consistent. We test the constant-returns-

[17]The time trend is among one of the variables z_{gmr} in equation (12). Such a formulation is also used by Chow and Lin (2002).

TABLE 3
Summary of Tests for the Estimated Production Function

A. *Exogeneity test*

	Log-Linear		Constant Returns to Scale	
	Test Statistics	p-value	Test Statistics	p-value
F test	0.98	0.42	1.59	0.17
Wald test	2.97	0.7	1.84	0.77

B. *Testing the linear restriction $\alpha_K + \alpha_H = 1$ (log-linear vs. constant returns to scale)*

	Test Statistics	Upper Tail Area
F test	2.29	0.13
Wald test	0.139	0.7091

Notes:
The F tests assume no heteroskedasticity and autocorrelation for the error term. The Wald tests, which are chi-square statistics, are robust test using the robust variance–covariance matrix suggested by Arellano (2003).
The exogeneity tests are for each of the log-linear and the constant-returns-to-scale specification.

to-scale restriction using the F test as well as the robust Wald test depending on whether heteroskedasticity and autocorrelation are assumed. As reported in Table 3, both tests do not reject the linear restriction $\alpha_K + \alpha_H = 1$. We therefore use the set of parameters from the constant-returns specification in the decomposition exercise in the next section.

The second step involves plugging the coefficients derived above (based on the regression assuming constant returns to scale) into the expressions for the cumulative contributions of different factors (see Appendix A). The cumulative contributions of the different factors are all summarized in Figure 5 to facilitate a comparison of the relative magnitudes of these contributions for the period 1965–99. The vertical axis of the figure represents the cumulative contribution to interprovincial inequality. CCK and CCH are the cumulative contributions of physical capital and labor adjusted for education attainments. Decomposing the contribution of TFP growth (CCA), CCFDI pertains to the contribution of openness and CCIND that of spatial industrial restructuring which captures the agglomeration effect. What remains after deducting CCFDI and CCIND is CCOTH, the remaining contribution of TFP growth (i.e. CCA – CCFDI – CCIND). In what follows, we further decompose each of these components into their respective within- and between-region counterparts (recall equations (9) and (10)).

Being the focus of many previous studies, e.g. Naughton (2002), the cumulative contribution of capital (CCK) unambiguously declined up to 1972 and then heads upwards all the way as shown in Figure 6. The initial decline is attributable to a fall in both within-region contribution of capital (WCCK) and between-region contribution of capital (BCCK), though the magnitude of the latter is larger.[18] The

[18]Such a partition does not inexorably result in an increase in BG and a decrease in WG because an initially poorer province may catch up with their richer neighbors, in which case BG decreases. For example, during the third front campaign, BG decreases because, in the second half of the 1960s and early 1970s, western provinces grew faster than their eastern counterparts because of more investment in the west. The same logic applies to WG.

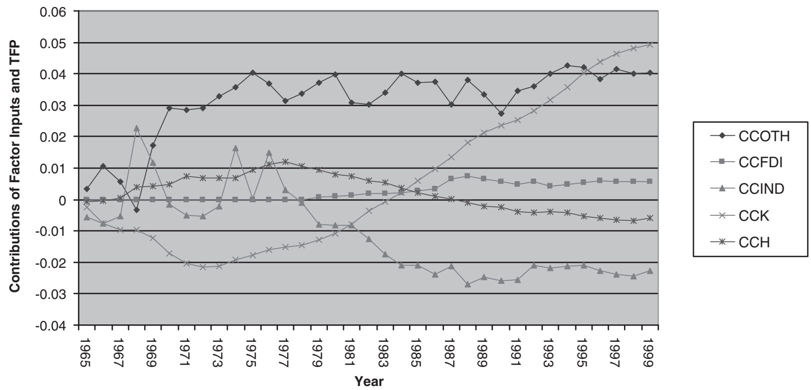

Figure 5. Contributions to Interprovincial Inequality

Notes: CCFDI = cumulative contribution of openness, CCIND = cumulative contribution of industrial restructuring, CCK = cumulative contribution of physical capital, CCH = cumulative contribution of labor adjusted for education attainment, CCOTH = cumulative contribution of TFP after deducting CCFDI and CCIND.

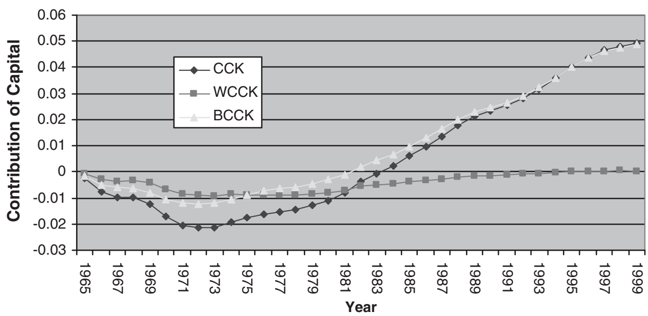

Figure 6. Overall, Between-Region and Within-Region Contributions of Capital

Notes: CCK = cumulative contribution of physical capital, WCCK = within-region cumulative contribution of physical capital, BCCK = between-region cumulative contribution of physical capital.

decline in BCCK coincides with the Third Front campaign, a period with massive state investment directed to the inland provinces.

The increasing trajectory of CCK since 1973 is largely explained by the widening gap among the coastal, central and western regions, so much so that by the 1990s, the upward trend in CCK is entirely propelled by BCCK. The heydays of the Third Front Campaign was already over in the early 1970s (Naughton, 1988, 2002) and this policy reversal is captured in Figure 6 by the sharp increase in the contribution of capital progressively driven by the between-region contribution of capital (BCCK) and is the major force in the reform era driving interprovincial inequality upwards. As our literature review has suggested, reform has ushered in a period with rapid growth in new sources of investment funds such as self-raised funds and foreign capital, the allocation of which is not in the hands of the central government. The spatial distribution of investment funds in the reform era is

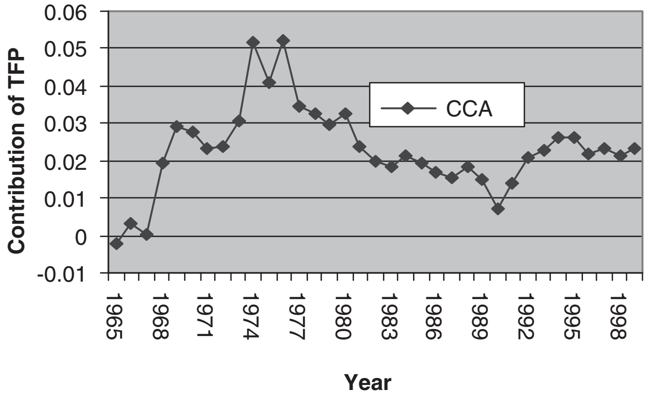

Figure 7. Cumulative Total Factor Productivity

Notes: CCA = cumulative contribution of total factor productivity.

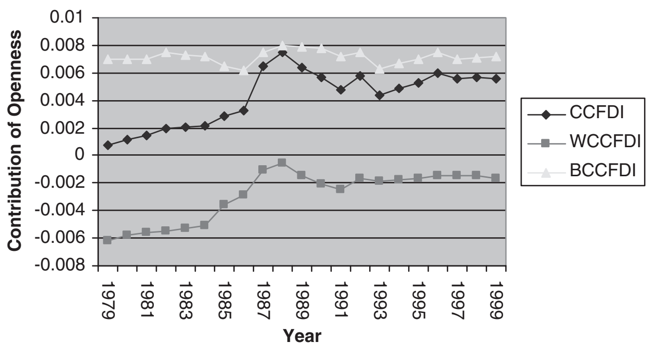

Figure 8. Contribution of Openness

Notes: CCFDI = cumulative contribution of openness, WCCFDI = within-region cumulative contribution of openness, BCCFDI = between-region cumulative contribution of openness.

increasingly skewed in favor of richer provinces, explaining the sustained increase in BCCK.

Next, Figure 7 summarizes the contribution of TFP (CCA) (equal to the sum of CCOTH, CCFDI, and CCIND in Figure 5). CCA is equally if not more important in certain sub-periods and it has a trajectory that is at times divergent from that of CCK. Figure 7 summarizes the contribution of TFP (CCA). During the period of the Cultural Revolution (from 1967 to 1977), the magnitude of the upward trend in CCA overwhelms CCK so that the overall inequality ratchets upwards.

Using equation (13), the contributions of openness (FDI) and spatial industrial restructuring (IND) may be extracted from that of TFP (CCA). Figure 8 shows the trends for the contribution of openness (CCFDI) together with its between- and within-region components, BCCFDI and WCCFDI. FDI, our proxy for openness, initially contributes to an increase in inequality in the reform era but

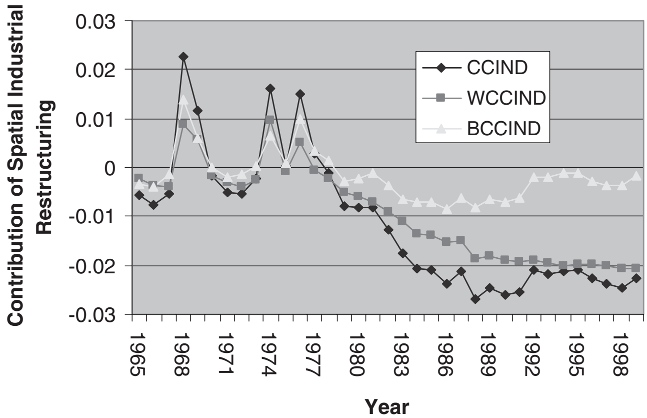

Figure 9. Contribution of Industrial Restructuring

Notes: CCFDI = cumulative contribution of industrial restructuring, WCCFDI = within-region cumulative contribution of industrial restructuring, BCCFDI = between-region cumulative contribution of industrial restructuring.

its effect tapers off from the mid 1980s onwards. After a jump induced by initial flows of FDI into such coastal provinces as Guangdong, the between-region contribution of FDI (BCCFDI) oscillates and does not exhibit a discernible trend. The trajectory of the within-region contribution (WCCFDI) exhibits a one-shot decrease to be followed by an upward trend, consistent with the fact that the initial beneficiaries of FDI are such poorer coastal provinces as Guangdong and Fujian endowed with the Special Economic Zones, but foreign capital subsequently spread to other coastal provinces. On the whole, these contributions are relatively small in magnitude (see Figure 5).

Figure 9 summarizes the overall contribution of spatial industrial restructuring (CCIND), as well as the corresponding between- (BCCIND) and within-region contributions (WCCIND). In the prereform era, there are a number of inequality-increasing spikes but no discernible trends are detectable. Since the late 1970s, CCIND by and large contributes to a reduction in overall interprovincial inequality, though its inequality-reducing effect is tapering off in the 1990s. Further decomposing CCIND into its between- and within-region contributions, its declining trend is largely attributable to within-region contribution and to a smaller extent to between-region contribution. In the latter case, the trend seems to have reversed in the 1990s. Further splitting WCCIND into the contributions by the eastern, central and western regions, the decline is largely within the eastern region.[19] The decrease in within-region contribution in the late 1970s and the first half of the 1980s is consistent with the demise of such industrial powerhouses as Liaoning heavily relying on heavy industries and the rapid ascendance of new industrial growth poles such as Guangdong, Fujian and Zhejiang within the eastern region exploiting their comparative advantages and agglomeration econo-

[19]Each term on the right-hand-side of equation (9) is the sum of the individual contributions of the three regions; the figures are not reported but available on request.

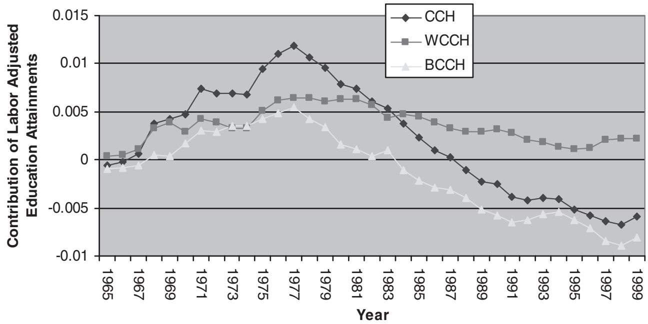

Figure 10. Contribution of Labor Adjusted for Education Attainment

Notes: CCFDI = cumulative contribution of labor adjusted for education attainment, WCCFDI = within-region cumulative contribution of labor adjusted for education attainment, BCCFDI = between-region cumulative contribution of labor adjusted for education attainment.

mies in labor-intensive industries, thereby boosting their productivity (more discussion below). With these newly industrializing provinces being less developed at the dawn of the reform era, such a TFP-enhancing industrial restructuring has thus set off a decline in interprovincial inequality.

Next, as depicted in Figure 10, CCH contributes to an increase in interprovincial inequality up to the mid 1970s but the upward trend has been reversed. The trajectories of between-region (BCCH) and within-region contributions (WCCH) are similar, with the inequality-reducing effect of the between-region contributions being more prominent. There are two forces at work. First, the growth in the labor force is faster for poor provinces. Second, there is a spread of education to less developed regions. With better-educated school age children in poor provinces gradually entering the labor force in the 1960s and 1970s, the rapid improvement in labor quality E in poor provinces also leads to a faster growth in H. Our empirical findings suggest that this is translated into a decline in interprovincial inequality in the reform era. The impact is, however, small relative to the contributions of TFP and physical capital (see Figure 5).

7. Salient Findings

This section highlights three important key results from our empirical findings presented above. One of them pertains to how the contribution of physical capital relative to that of TFP growth shapes the overall trend of interprovincial inequality. The second set of results looks into the coastal–inland divide. These results also paint a more complex picture of regional inequality than can be captured by the coastal–inland dichotomy. A final set of results sheds light on the effect of the prereform spread of education on growth in the reform era.

First, our empirical results show that changing investment allocation across space far from fully explains the trend of interprovincial inequality. The contribution of physical capital has a trajectory that is at times divergent from that of TFP. As pointed out in Section 3, one has to be cautious in the interpretation of the contribution of TFP which may reflect the effects of many complex factors.

Keeping this in mind, we present a tentative explanation of the changing contribution of TFP. In the prereform era and especially during the Cultural Revolution, the contribution of TFP growth (CCA) overwhelms that of physical investment (CCK). Jumps in the contribution of TFP coincide with shocks induced by political turmoil as well as the policy environment in the prereform era. This explains why interprovincial inequality during the Cultural Revolution did not decrease contrary to the conjecture of Lardy (1978). However, the roles of TFP growth and investment have reversed since the late 1970s with a decreasing contribution of TFP more than offsetting the increasing contribution of capital, accounting for the decline in interprovincial inequality in the 1980s. As alluded to above, these changes may be due to changing spatial distribution of budgetary resources and market-driven industrial restructuring. The inequality-decreasing contribution of TFP peters out in the 1990s so that the contribution of capital becomes dominant.

Spatial industrial restructuring turns out to be important in explaining much of the decline in the contribution of TFP in the post-Mao era. Relinquishing the policy of self reliance, the spatial reconfiguration of industries has become more in line with comparative and agglomeration advantages. As summarized by our empirical results, industrial reshuffling in the 1980s had the effect of rectifying a distorted spatial industrial structure and induced both a decline in within-region and between-region inequality. This effect is so powerful that it overwhelms other inequality-increasing forces, leading to an overall decrease in interprovincial inequality. With the completion of the spatial restructuring of industries, the spatial distribution of TFP growth contributes to an increase in between-region inequality in the 1990s. Institutional reforms seem to be moving at much faster paces in the coastal provinces, probably reinforced by agglomeration economies.

The next set of results sheds light on the coastal-inland divide. In the prereform era, the between-region inequality-reducing contribution of capital is more than offset by the increase in the between-region contribution of TFP. The lesson from this finding is that simply pumping more resources into poorer provinces does not necessarily narrow the coastal–inland gap. The introduction of efficiency-enhancing institutions is also important. Indeed, the retreat from the distortionary policies in the Maoist era and the embrace of market institutions after 1978 did lead to a reduction in regional inequality in the 1980s. But market-driven industrial restructuring also sets off such forces as agglomeration economies enhancing efficiencies in richer provinces, thereby increasing inequality. Indeed, the increase in interprovincial inequality in the 1990s is due to the mutually reinforcing contributions of TFP and capital.

While the focus on regional inequality has often been on the gap between the inland as opposed to inland provinces, there are also significant changes *within* each region. This is particularly the case within the eastern regions. As shown in Figure 3, there is a sustained decline in within-region inequality from the mid 1970s to the end of the 1990s. Particularly worth mentioning is the industrial restructuring among such coastal provinces as Guangdong, Fujian and Zhejiang, triggering a fall in interprovincial inequality in the 1980s. The emergence of new growth poles in the eastern region has capitalized on the development of non-state industrial enterprises (e.g. township and village enterprises) and the open-door policy. In sharp contrast is the decline in the old industrial centers (especially those in the northeast) as their moribund state-owned enterprises have failed to fend off competition brought about by reform. The above transformation is translated into a decline in overall interprovincial inequality in the 1980s, with the within-region contribution of industrial restructuring exhibiting a downward trend.

Finally, the third set of our empirical results addresses the effect of schooling on China's interprovincial inequality. This paper is a preliminary attempt to incorporate this issue into the analysis of interprovincial inequality. Much remains to be done to improve the measure of schooling. While communism inflicted political calamities on the Chinese people, there is no denying that much had been done before 1978 in spreading education to less developed provinces. Some scholars, e.g. Bramall (2000), have even gone so far as to argue that the prereform investment in education laid the foundation for the spectacular growth in the reform era. Our empirical results suggest that contribution of schooling first increases and then reduces interprovincial inequality. In so far as school-age children in less developed provinces only gradually entered the labor force in the 1970s, our interpretation is that the inequality-reducing effect of schooling is a consequence of the prereform expansion of basic education. The contribution of labor adjusted for education attainment to growth is however small relative to capital and TFP.

8. Limitations of the Approach and Interpretation of Empirical Results

This paper introduces a new tool for understanding the forces driving interprovincial inequality. Furthermore, unlike many previous studies on interprovincial inequality simply using official data, this paper tries seriously to take into account the problems plaguing China's provincial GDP growth rates and population data. These departures from common practices naturally raise questions that we try to address in the rest of this section.

On tackling the data problems, there are various routes possible, each with its own limitations.[20] One concern regarding the approach adopted here to derive provincial GDP deflators is the appropriateness of the price indices applying to the different expenditure components. Much of our effort was expended looking up all possible data sources in search of appropriate indices. As far as this approach is concerned, any improvement can only come about if more and better data are released. Another question is whether the adjustments make a difference. As shown in Figure 2, interprovincial inequality using deflators we derive (i.e. EN2) not only differs in magnitude from that using official deflators (i.e. EN1), their trends also differ over certain sub-periods, e.g. the period 1995–99.

As far as the scope of this paper is concerned, the focus on interprovincial inequality sheds light on the important role of political jurisdictions in shaping regional development policies. There is no denying that interprovincial inequality is just one, albeit important, dimension of China's economic inequality. For one thing, within-province inequality is left out of our discussion though it may be large. To some extent captured by the coastal–inland divide, rural–urban inequality is another important issue that has attracted much attention. Each of these

[20]For a survey of these approaches, see, e.g. Wu (2000).

dimensions is a research project on its own. Interesting and important as all these issues may be, it is perhaps too ambitious to incorporate all of them in one paper, not least because branching out into these topics quickly encounters almost insurmountable data problems.

There is also the difficult decision regarding the range of factors we may bring into account for changes in interprovincial inequality. While TFP often remained a black box in previous studies, we go a step further by introducing FDI and IND. Admittedly, within the Chinese context, there is a whole host of institutional and technological factors other than FDI and IND that one may consider incorporating into $\mathbf{z}_{gm}$. The availability of intertemporally consistent data, however, limits what we can actually include. Our framework is nonetheless set up to accommodate any institutional factors that are deemed important.

The study of interprovincial inequality often brings within its purview the role of factor mobility. How do flows in labor and capital impinge on interprovincial inequality?[21] In this paper, growth in provincial factor inputs ($\dot{K}_{gm}/K_{gm}$ or $\dot{H}_{gm}/H_{gm}$) may be due to *own* accumulation of physical and human capital as well as *cross-province* flows of factors. For example, as pointed out above and by Naughton (2002), the retreat of central planning results in a reconfiguration of the cross-province distribution of investment growth, $\dot{K}_{gm}/K_{gm}$, impinging on provincial economic growth ($\dot{Y}_{gm}/Y_{gm}$) and thus the change in interprovincial inequality ($dI(\mathbf{y})/dt$). The same logic applies to labor mobility. While the impact of factor mobility has been implicitly taken into account by our empirical results, to isolate its impact requires detailed breakdowns of $\dot{K}_{gm}/K_{gm}$ and $\dot{H}_{gm}/H_{gm}$ into different sources. Such time-series and cross-section data are presently not available, but once the information emerges, K_{gm} or H_{gm} may be decomposed into different sources, and the cumulative contributions CCK or CCH may be split into their respective sub-components, one of which may be the contribution of cross-province net factor inflows.[22]

With regard to the estimation of the production function as a prerequisite to the decomposition exercise, we adopt the simple log-linear specification following such precedents as Chow and Lin (2002) in the case of China as well as the many studies in the general literature on growth accounting (for surveys, see Temple, 1999; Bosworth and Collins, 2003). The reason for such a choice is twofold. First, such a form fits perfectly into our framework if we are to have a neat decomposition of $dI(\mathbf{y})/dt$ so that it is just equal to the sum of the contributions of the factors, i.e. equation (7). Second, the choice of flexible functional forms such as the translog often runs into the problem of multicollinearity, a problem that can be alleviated with more parsimonious specifications.

[21]This question is closely related to a growing literature on whether interregional resource mobility has increased in the reform era; see, e.g. Naughton (2003), Young (2000a), Poncet (2003) and Bai *et al.* (2004).

[22]Mathematically, $CK = \sum_{g=1}^{G}\sum_{m=1}^{M_g}(s_{gm} - f_{gm})\alpha_{gm}^{K}\frac{\dot{K}_{gm}}{K_{gm}}$. If different sources including net capital inflow are available, then, $\dot{K}_{gm} = \sum_j \dot{K}_{gmj}$, where $\dot{K}_{gmj}$ is the j-th source of investment. Thee contribution of the j-th source is then $CK_j = \sum_{g=1}^{G}\sum_{m=1}^{M_g}(s_{gm} - f_{gm})\alpha_{gm}^{K}\frac{\dot{K}_{gmj}}{K_{gm}}$.

9. Concluding Remarks

This paper attempts to extend the research on China's regional inequality by first resorting to a more careful use of China's data and then putting forward a new framework that enhances our understanding of the forces behind the changes in China's interprovincial inequality. In particular, quantitative estimates of the contributions of factors impinging on inequality are presented and furnish empirical evidence for a more informed discussion of China's ongoing debate on regional disparity and regional policies. In so far as reducing regional disparity is high on the agenda of the Chinese leadership, there are pressures on the central government of China to pump more resources into the inland provinces through intergovernmental transfers. However, the findings in this paper suggest that redistribution alone may not be enough in bringing about a decrease in regional inequality without at the same time improving the quality of institutions. A case in point is China's prereform experience. Despite massive investments in inland provinces, interprovincial inequality actually increased for the period straddling the Cultural Revolution. Removing institutions and policies impeding efficient spatial allocation of resources may on the other hand help the poor more than the rich provinces as attested by the decline of regional inequality in the early 1980s.

There is room to extend the current study in a number of directions. Much remains to be done to improve the understanding of the contribution of TFP growth if data permit. There are certainly other important variables driving TFP growth other than those explored in this study. Among them are the effects of the role of factor mobility involving capital and labor, the varying importance of state owned enterprises, the growth of township and village enterprises, the role of labor-intensive industries, local protectionism and improvement in transportation network, etc. The list is by no means exhaustive.

Concerns with the quality of Chinese data have presented major problems for those doing research on the Chinese economy. The method used to adjust the data in this paper is not the only one conceivable. For example, Young (2000b) tackles the problem by first coming up with deflators for the different sectors of the national economy, while others such as Wu (2000) derive growth indices based on physical output. There are pros and cons with these approaches and how well they perform when applied to provincial data remains to be seen. Another potential extension of our study is to experiment with different ways to construct provincial indices for human capital, a subject that is too complicated to be elaborated in this paper.

All the above extensions may be subjects for future research. As explained above, the decomposition framework proposed in this paper can easily handle these extensions in so far as data are available.

Appendix A

This appendix sketches how the equations in the text are derived. The derivations for the expressions on the right hand sides of equations (7), (9) and (10) are presented in detail. The population-weighted version of Theil's entropy measure is as follows:

$$\text{(A1)} \qquad I(\mathbf{y}) = \sum_{g=1}^{G}\sum_{m=1}^{M_g} f_{gm} \ln\left(\frac{\bar{y}}{y_{gm}}\right) = \ln(\bar{y}) - \sum_{g=1}^{G}\sum_{m=1}^{M_g} f_{gm} \ln(y_{gm})$$

where $\bar{y} = \sum_{g=1}^{G}\sum_{m=1}^{M_g} f_{gm} y_{gm}$. Differentiate equation (A1):

$$\text{(A2)} \qquad \frac{dI}{dt} = \frac{\dot{\bar{y}}}{\bar{y}} - \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(f_{gm}\frac{\dot{y}_{gm}}{y_{gm}} + \dot{f}_{gm}\ln(y_{gm})\right)$$

where, for any variable x, $\dot{x} = dx/dt$. It is to be noted that

$$\text{(A3)} \qquad \frac{\dot{\bar{y}}}{\bar{y}} = \sum_{g}^{G}\sum_{m}^{M_g}\left(\frac{\dot{f}_{gm}y_{gm} + f_{gm}\dot{y}_{gm}}{\bar{y}}\right) = \sum_{g}^{G}\sum_{m}^{M_g}\left(\frac{f_{gm}y_{gm}}{\bar{y}}\frac{\dot{f}_{gm}y_{gm} + f_{gm}\dot{y}_{gm}}{f_{gm}y_{gm}}\right).$$

With P defined as the total population, the term

$$s_{gm} = \frac{f_{gm}y_{gm}}{\bar{y}} = \frac{Pf_{gm}y_{gm}}{P\bar{y}},$$

is the share of GDP accruing to the m province in the g-th region and equation (A3) becomes:

$$\frac{\dot{\bar{y}}}{\bar{y}} = \sum_{g}^{G}\sum_{m}^{M_g}\left(\frac{f_{gm}y_{gm}}{\bar{y}}\frac{\dot{f}_{gm}y_{gm} + f_{gm}\dot{y}_{gm}}{f_{gm}y_{gm}}\right) = \sum_{g}^{G}\sum_{m}^{M_g} s_{gm}\left(\frac{\dot{f}_{gm}}{f_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}}\right).$$

Thus, equation (A2) becomes:

$$\frac{dI}{dt} = \sum_{g}^{G}\sum_{m}^{M_g} s_{gm}\left(\frac{\dot{f}_{gm}}{f_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}}\right) - \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(f_{gm}\frac{\dot{y}_{gm}}{y_{gm}} + \dot{f}_{gm}\ln(y_{gm})\right).$$

Regrouping the terms:

$$\text{(A4)} \qquad \frac{dI}{dt} = \sum_{g}^{G}\sum_{m}^{M_g}\left((s_{gm} - f_{gm})\frac{\dot{y}_{gm}}{y_{gm}}\right) - \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left((s_{gm} - f_{gm}\ln(y_{gm}))\frac{\dot{f}_{gm}}{f_{gm}}\right).$$

Since

$$\text{(A5)} \qquad \dot{y}_{gm}/y_{gm} = (\dot{Y}_{gm}/Y_{gm}) - (\dot{P}_{gm}/P_{gm})$$

and

$$\text{(A6)} \qquad \frac{\dot{Y}_{gm}}{Y_{gm}} = \frac{\dot{A}_{gm}}{A_{gm}} + \left(\frac{\partial Q_{gm}/\partial K_{gm}}{Y_{gm}}\right)\frac{\dot{K}_{gm}}{K_{gm}} + \left(\frac{\partial Q_{gm}/\partial H_{gm}}{Y_{gm}}\right)\frac{\dot{H}_{gm}}{H_{gm}}$$

and plugging (A5) and (A6) into (A4) results in

$$\frac{dI(\mathbf{y})}{dt} = CA + CK + CH + CP + CF$$

where

$$CA = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_{gm}\right)\frac{\dot{A}_{gm}}{A_{gm}}, \quad CK = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_{gm}\right)\alpha_{gm}^{K}\frac{\dot{K}_{gm}}{K_{gm}},$$

$$CH = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gn} - f_{gm}\right)\alpha_{gm}^{H}\frac{\dot{H}_{gm}}{H_{gm}}, \quad CP = -\sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_{gm}\right)\frac{\dot{P}_{gm}}{P_{gm}},$$

$$CF = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_{gm}\ln\left(y_{gm}\right)\right)\frac{\dot{f}_{gm}}{f_{gm}},$$

the within-region component is

$$WG(\mathbf{y}) = \sum\nolimits_{g}^{G} f_g I\left(\mathbf{y}_g\right), \quad f_g = \sum\nolimits_{m}^{M_g} f_{gm}$$

$$\frac{dWG(\mathbf{y})}{dt} = \sum\nolimits_{g}^{G}\left(f_g\frac{dI_g}{dt} + I_g\dot{f}_g\right), \quad I_g = I\left(\mathbf{y}_g\right).$$

In so far as I_g corresponds to the set of provinces in region g, the form of the expression of dI_g/dt is similar to dI/dt, i.e. (A4):

$$\frac{dI_g}{dt} = \sum\nolimits_{m}^{M_g}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\right)\frac{\dot{y}_{gm}}{y_{gm}}\right] + \sum\nolimits_{m}^{M_g}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\ln\left(y_{gm}\right)\right)\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}}\right]$$

where $\hat{s}_{gm} = Y_{gm}/Y_g$, $\hat{f}_{gm} = P_{gm}/P_g$, where Y_g and P_g are total GDP and total population of region g respectively. Variables with "^" denote shares or averages with respect to a region. Thus,

$$\begin{aligned}\frac{dWG(\mathbf{y})}{dt} &= \sum\nolimits_{g} f_g \sum\nolimits_{m}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\right)\frac{\dot{y}_{gm}}{y_{gm}}\right] + \\ &\quad \sum\nolimits_{g}\left\{I_g\dot{f}_g + f_g\sum\nolimits_{m}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\right)\ln\left(y_{gm}\right)\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}}\right]\right\} \\ &= \sum\nolimits_{g} f_g \sum\nolimits_{m}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\right)\frac{\dot{y}_{gm}}{y_{gm}}\right] + \\ &\quad \sum\nolimits_{g} f_g\left\{I_g\frac{\dot{f}_g}{f_g} + \sum\nolimits_{m}\left[\left(\hat{s}_{gm} - \hat{f}_{gm}\right)\ln\left(y_{gm}\right)\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}}\right]\right\}.\end{aligned}$$

Again, using (A5) and (A6), the above expression expands to equation (9) in the text, where

$$WCA = \sum_{g=1}^{G} f_g \left[\sum_{m=1}^{M_g} \left(\hat{s}_{gm} - \hat{f}_{gm} \right) \frac{\dot{A}_{gm}}{A_{gm}} \right], \quad WCK = \sum_{g=1}^{G} f_g \left[\sum_{m=1}^{M_g} \left(\hat{s}_{gm} - \hat{f}_{gm} \right) \alpha_{gm}^{K} \frac{\dot{K}_{gm}}{K_{gm}} \right],$$

$$WCH = \sum\nolimits_{g=1}^{G} f_g \left[\sum_{m=1}^{M_g} \left(\hat{s}_{gm} - \hat{f}_{gm} \right) \alpha_{gm}^{H} \frac{\dot{H}_{gm}}{H_{gm}} \right], \quad WCP = -\sum_{g=1}^{G} f_g \left[\sum_{m=1}^{M_g} \left(\hat{s}_{gm} - \hat{f}_{gm} \right) \frac{\dot{P}_{gm}}{P_{gm}} \right],$$

$$WCF = \sum_{g=1}^{G} f_g \left[\sum_{m=1}^{M_g} \left(\hat{s}_{gm} - \hat{f}_{gm} \ln\left(y_{gm} \right) \right) \frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}} + I\left(\mathbf{y}_g \right) \frac{\dot{f}_g}{f_g} \right].$$

The term $(\hat{s}_{gm} - \hat{f}_{gm})$ is the difference between the income share and population share of the m-th province *within the* g-th region and its interpretation is similar to $(s_{gm} - f_{gm})$ discussed above. In the case of the between-region contribution, since $(s_{gm} - f_g\hat{s}_{gm}) = \hat{s}_{gm}(s_{gm}/\hat{s}_{gm} - f_g)$ and $s_{gm}/\hat{s}_{gm}$ is in fact the share of income accruing to region g, so that $(s_{gm}/\hat{s}_{gm} - f_g)$ turns out to be the difference between the income share and the population share of the g-th region. The between-region component is:

$$BG(\mathbf{y}) = \sum\nolimits_{g}^{G} f_g \left(\ln(\bar{y}) - \ln\left(\bar{y}_g \right) \right)$$

where $\bar{y}_g = \sum_{m=1}^{M_g} f_{gm} y_{gm}$ is the GDP per capita of region g. Thus,

$$\begin{aligned} \frac{dBG(\mathbf{y})}{dt} &= \sum\nolimits_{g}^{G} \left[f_g \left(\frac{\dot{\bar{y}}}{\bar{y}} - \frac{\dot{\bar{y}}_g}{\bar{y}_g} \right) + \dot{f}_g \left(\ln(\bar{y}) - \ln\left(\bar{y}_g \right) \right) \right] \\ &= \frac{\dot{\bar{y}}}{\bar{y}} - \sum\nolimits_{g}^{G} f_g \frac{\dot{\bar{y}}_g}{\bar{y}_g} + \sum\nolimits_{g}^{G} \dot{f}_g \left(\ln(\bar{y}) - \ln\left(\bar{y}_g \right) \right) \\ &= \frac{\dot{\bar{y}}}{\bar{y}} - \sum\nolimits_{g}^{G} f_g \frac{\dot{\bar{y}}_g}{\bar{y}_g} - \sum\nolimits_{g}^{G} \dot{f}_g \ln\left(\bar{y}_g \right), \end{aligned}$$

where $\sum\nolimits_{g}^{G} \dot{f}_g = 0$ since $\sum\nolimits_{g}^{G} f_g$, the sum of the population shares of the G regions, must be one so that an increase in the share of one region must be offset by decreases in the shares of the other regions.

By substituting the following terms

$$\frac{\dot{\bar{y}}}{\bar{y}} = \sum\nolimits_{g}^{G} \sum\nolimits_{m}^{M_g} s_{gm} \left(\frac{\dot{f}_{gm}}{f_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}} \right), \quad \frac{\dot{\bar{y}}_g}{\bar{y}_g} = \sum\nolimits_{m}^{M_g} \hat{s}_{gm} \left(\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}} \right),$$

into the expression for $dBG(\mathbf{y})/dt$ above, we have the following expression:

$$\frac{dBG(\mathbf{y})}{dt} = \sum_g^G \sum_m^{M_g} s_{gm}\left(\frac{\dot{f}_{gm}}{f_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}}\right) - \sum_g^G f_g \sum_m^{M_g} \hat{s}_{gm}\left(\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}} + \frac{\dot{y}_{gm}}{y_{gm}}\right) - \left(\sum_g^G \dot{f}_g \ln(\bar{y}_g)\right)$$

$$= \sum_g^G \sum_m^{M_g}\left[\left(s_{gm} - f_g\hat{s}_{gm}\right)\frac{\dot{y}_{gm}}{y_{gm}}\right] + \sum_g^G \left\{\sum_m^{M_g}\left[s_{gm}\frac{\dot{f}_{gm}}{f_{gm}} - f_g\hat{s}_{gm}\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}}\right] - \dot{f}_g \ln(\bar{y}_g)\right\}$$

where the notation is the same as in the text. Using (A5) and (A6), the above expression expands to equation (10), where

$$BCA = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_g\hat{s}_{gm}\right)\frac{\dot{A}_{gm}}{A_{gm}}, \quad BCK = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_g\hat{s}_{gm}\right)\frac{\dot{K}_{gm}}{K_{gm}},$$

$$BCL = \sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_g\hat{s}_{gm}\right)\frac{\dot{H}_{gm}}{H_{gm}}, \quad BCP = -\sum_{g=1}^{G}\sum_{m=1}^{M_g}\left(s_{gm} - f_g\hat{s}_{gm}\right)\frac{\dot{P}_{gm}}{P_{gm}},$$

$$BCF = \sum_{g=1}^{G}\left[\sum_{m=1}^{M_g}\left(s_{gm}\frac{\dot{f}_{gm}}{f_{gm}} - f_g\hat{s}_{gm}\frac{\dot{\hat{f}}_{gm}}{\hat{f}_{gm}}\right) - f_g \ln(\bar{y}_g)\frac{\dot{f}_g}{f_g}\right], \quad \bar{y}_g = \sum_{m=1}^{M_g}\hat{f}_{gm} y_{gm}.$$

References

Arellano, Manuel, *Panel Data Econometrics*, Oxford University Press, Oxford, 2003.

Au, C. C. and V. Henderson, "Estimating Net Urban Agglomeration Economies with an Application to China," *Journal of Development Economics*, 80(2), 350–88, 2006.

Bai, Chong-En, Yingjuan Du, Zhigang Tao, and Sarah Y. Tong, "Local Protectionism and Regional Specialization: Evidence from China's Industries," *Journal of International Economics*, 63(2), 397–417, 2004.

Barro, R. and J. W. Lee, "International Measures of Schooling Years and Schooling Quality," *American Economic Review, Papers and Proceedings*, 86(2), 218–23, 1996.

Bhalla, Ajit S., Shujie Yao, and Zhang Zongyi, "Causes of Inequalities in China, 1952 to 1999," *Journal of International Development*, 15(8), 939–95, 2003.

Bogue, D. J., K. Hinze, and M. White, *Techniques of Estimating Net Migration,* Community and Family Study Center, University of Chicago, Chicago, 1982.

Bosworth, Barry P. and Susan M. Collins, "The Empirics of Growth: An Update," *Brookings Papers on Economic Activities*, Issue 2, 113–206, 2003.

Brakman, Steven, Harry Garretsen, and Charles van Marrewijk, *An Introduction to Geographical Economics: Trade, Location and Growth*, Cambridge University Press, Cambridge, 2001.

Bramall, Chris, *Sources of Chinese Economic Growth 1978–1996*, Oxford University Press, Oxford, 2000.

Brandt, L. and C. Holz, "Spatial Price Differences in China: Estimates and Implications," *Economic Development and Cultural Change*, 55(1), 43–86, 2006.

Chow, G. and A. Lin, "Accounting for Economic Growth in Taiwan and Mainland China: A Comparative Analysis," *Journal of Comparative Economics*, 30(3), 507–30, 2002.

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

Donnithorne, Audrey, "China's Cellular Economy: Some Economic Trends since the Cultural Revolution," *China Quarterly*, 52, 605–19, 1972.

———, "Comment," *The China Quarterly*, 66, 328–40, 1976.

Fujita, Masahisa and Daping Hu, "Regional Disparity in China 1985–1994: The Effects of Globalization and Economic Liberalization," *Annals of Regional Science*, 35(1), 3–37, 2001.
Griliches Z., "Notes on the Role of Education in Production Functions and Growth Accounting," in Lee Hansen (ed.), *Education and Income*, NBER Studies in Income and Wealth, Vol. 35, New York, 1970.
Holz, C.A., "Fast, Clear and Accurate: How Reliable are Chinese Output and Economic Growth Statistics?" *China Quarterly*, 173, 122–63, 2003.
———, "Deconstructing China's GDP Statistics," *China Economic Review*, 15, 164–202, 2004.
Im, Kyung So, M. H. Pesaran, and Yongcheol Shin, "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics*, 115(1), 53–74, 2003.
Jian, Tianlun, Jeffrey D. Sachs, and Andrew M. Warner, "Trends in Regional Inequality in China," *China Economic Review*, 7(1), 1–21, 1996.
Jorgenson, D. W. and B. M. Fraumeni, "Investment in Education and U.S. Economic Growth," *Scandinavian Journal of Economics*, Supplement, 94, S51–70, 1992.
Keidel, Albert, "China's GDP Expenditure Accounts," *China Economic Review*, 12(4), 355–67, 2001.
Lardy, Nicholas R., "Replies," *The China Quarterly*, 66, 340–54, 1976.
———, *Economic Growth and Distribution in China*, Cambridge University Press, Cambridge, 1978.
Lin, Yifu and Fang Cai, *Zhongguo jingji: gaige yu fazhan [Chinese Economy: Reform and Development*], Zhongguo caizheng jingji chubanshe, Beijing, 2003.
Lin, Yifu, Fang Cai, and Zhou Li, "Zhongguo jingji zhuanxingqi de diqu chayi fenxi" ["Analysis of China's Regional Inequality in the Period of Transition"], *Jingji yanjiu* [Economic Research], 6, 3–10, 1999.
Long, Gen Ying, "China's Changing Regional Disparities during the Reform Period," *Economic Geography*, 75(1), 59–70, 1999.
Lyons, Thomas P., "Interprovincial Disparities in China: Output and Consumption," *Economic Development and Cultural Change*, 39, 471–506, 1991.
National Bureau of Statistics, Department of Comprehensive Statistics, *Xin zhongguo wushinian tongji ziliao huibian [Comprehensive Statistical Data and Materials on 50 Years of New China]*, Zhongguo tongji chubanshe, Beijing, 1999.
National Bureau of Statistics, Department of Fixed Assets Investment Statistics, *Zhongguo guding zichan touzi shudian 1950–2000 [China's Fixed Asset Investment Statistics 1950–2000]*, Zhongguo tongji chubanshe, Beijing, 2002.
National Bureau of Statistics, Department of National Accounts, *Zhonggou guonei shengchan zongzhi hesuan lishi ziliu 1952–1995 [The Gross National Product of China]*, Dongbei daxue chubanshe, Dalian, 1997.
National Bureau of Statistics, Department of National Economic Balance, *Guomin shouru tongji ziliao huibian 1949–1985 [Collection of National Income Statistics]*, Zhonggou tongji chubanshe, Beijing, 1987.
National Bureau of Statistics, Urban Socio-Economic Survey Team, *Quanguo guding zichan jiazhi chonggu xishu biaozhun mulu 1984–2000 [A Catalogue of Coefficients for the National Revaluation of Fixed Assets 1984–2000]*, Zhongguo tongji chubanshe, Beijing, 2001.
Naughton, Barry, "The Third Front: Defence Industrialization in the Chinese Interior," *China Quarterly*, 115, 351–86, 1988.
———, "Provincial Economic Growth in China: Causes and Consequences of Regional Differentiation," in Mary-Françoise Renard (ed.), *China and Its Regions: Economic Growth and Reform in Chinese Provinces*, E. Elgar, Cheltenham, 2002.
———, "How Much Can Regional Integration Do to Unify China's Market?" in Nicholas C. Hope, Dennis Tao Yang, and MuYang Li (eds), *How Far Across the River? Chinese Policy Reform at the Millennium*, Stanford University Press, Stanford, 2003.
Nehru, Vikram and Ashok Dhareshwar, "A New Database on Physical Capital Stock: Sources, Methodology and Results," *Revista de Analisis Economico*, 8, 37–59, 1993.
Poncet, Sandra, "Measuring Chinese Domestic and International Integration," *China Economic Review*, 14(1), 1–21, 2003.
Raiser, Martin, "Subsidising Inequality: Economic Reforms, Fiscal Transfers and Convergence Across Chinese Provinces," *The Journal of Development Studies*, 34(3), 1–26, 1998.
Qian, Xiaolei and Russell Smyth, "Growth Accounting for the Chinese Provinces 1990–2000: Incorporating Human Capital Accumulation," *Journal of Chinese Economic and Business Studies*, 4(1), 21–37, 2006.
Rawski, Thomas G., "What is Happening to China's GDP Statistics," *China Economic Review*, 12(4), 347–54, 2001.
Sen, Amartya, *On Economic Inequality*, Clarendon Press, Oxford, 1997.

Shorrocks, Anthony, "Inequality Decomposition by Population Subgroups," *Econometrica*, 52(6), 1369–85, 1984.

Temple, Jonathan, "The New Growth Evidence," *Journal of Economic Literature*, 37(1), 112–56, 1999.

Tsui, Kai Yuen, "China's Regional Inequality, 1952–85," *Journal of Comparative Economics*, 15(1), 1–21, 1991.

———, "Economic Reform and Interprovincial Inequalities in China," *Journal of Development Economics*, 50(2), 353–68, 1996.

Wang, Shaoquang and Angang Hu, *The Political Economy of Uneven Development. The Case of China*, M.E. Sharpe, Armonk, 1999.

Wang, Yan and Yudong Yao, "China's Economic Growth 1952–1999: Incorporating Human Capital Accumulation," *China Economic Review*, 14(1), 32–52, 2003.

Wooldridge, Jeffrey M., *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, 2002.

Wu, Harry X., "China's GDP Level and Growth Performance: Alternative Estimates and the Implications," *Review of Income and Wealth*, 46(4), 475–99, 2000.

Young, Alwyn, "The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China," *Quarterly Journal of Economics*, 115(4), 1091–135, 2000a.

———, "Gold into Base Metals: Productivity Growth in the People's Republic of China during the Reform Period," NBER Working Paper No. 7856, 2000b.