

# Why has China Grown So Fast? The Role of Physical and Human Capital Formation

Sai Ding and John Knight

Department of Economics  
University of Oxford  
Manor Road Building  
Oxford OX1 3UQ  
[sai.ding@economics.ox.ac.uk](mailto:sai.ding@economics.ox.ac.uk)  
[john.knight@economics.ox.ac.uk](mailto:john.knight@economics.ox.ac.uk)

**Abstract** The paper estimates cross-province growth regressions for China over the period of economic reform. It first addresses the problem of model uncertainty by adopting two approaches to model selection, Bayesian Model Averaging and the automated General-to-Specific approach, to consider a wide range of candidate predictors of growth in China. The first-stage model selection results identify a role for conditional convergence, physical and human capital formation, population growth, degree of openness, and institutional change. Starting from this baseline equation, the growth impact of physical and human capital is examined in some detail using panel data system GMM. For instance, 'investment in innovation', foreign direct investment, and private investment are found to be particularly important. Secondary school enrolment contributes to growth, and higher education enrolment even more so. Finally, some illustrative counterfactual predictions are conducted to answer the question: why has China, as a whole, and indeed all its provinces, grown so fast?

**Keywords** Economic growth; Convergence; Physical capital; Human capital; China

**JEL Classification** O40; O53

**Acknowledgement** We are grateful to Jonathan Temple, Adrian Wood and participants in the ESRC Development conference in Sussex, and also Belton Fleisher, Xiaobo Zhang and participants in the Microeconomic Drivers of Growth of China Conference in Oxford, for their constructive comments on an earlier version of this paper. The financial support of the Leverhulme Trust is gratefully acknowledged.

*“For the period 1960-1980 we observe, for example, India 1.4 percent a year,..., South Korea 7.0 percent,...An Indian will on average be twice as well off as his grandfather; a Korean 50 times...The consequences for human welfare involved in questions like these are simply staggering: once one starts to think about them, it is difficult to think of anything else.”* (Lucas, 2002: 20-1).

## 1. Introduction

Since economic reform commenced in 1978, the Chinese economy has experienced remarkable economic growth. The growth rate of GDP per capita has averaged 8.6 percent per annum over the thirty-year period 1978-2007. Nor is there any sign of deceleration in growth: over the years 2000-07, the equivalent figure was 9.2 percent, and China accounted for about 35 percent of the growth in world GDP at PPP prices<sup>1</sup>. For a major country – China accounts for more than one-fifth of world population – such rapid progress is unprecedented. It is all the more remarkable in the light of China’s poverty – over 300 million people have been lifted out of one-dollar-a-day poverty since 1978<sup>2</sup> – and of its difficult transition from being a centrally planned, closed economy at the start of reform towards becoming a market economy.

In this paper we explore the reasons for China’s growth success using a cross-province dataset spanning three decades. Our purpose is to explain why China as a whole, and indeed all its provinces, has grown so fast. Our expectation is that the analysis of provincial time series data will reveal more information about the various determinants of economic growth than would an aggregate time series analysis. The use of provincial data expands the sample size substantially.

Economists are better able to analyse the direct than the indirect determinants of growth, and yet these conventional variables may simply represent associations that are themselves to be explained by causal processes. There are three possible empirical approaches: growth accounting, structural growth modelling, and informal growth regression. In contrast to the former two, the third approach permits the introduction of some explanatory variables that represent the underlying as well as the proximate causes of growth. Unlike the growth accounting method, it does not involve the task of measuring the capital stock and

---

<sup>1</sup> Based on new statistical calculations of PPP exchange rates published in December 2007 by the International Comparison Program (ICP), the World Bank and IMF recently revised downward their estimates for China’s PPP-based GDP by around 40 percent. Despite this revision, China remains the main driver of global growth. For example, it contributed nearly 27 percent of world GDP growth in 2007 using the new PPP figure.

<sup>2</sup> The figure is calculated from Ravallion and Chen (2007).

thus it avoids making several assumptions about unknown parameters such as factor shares of income and the depreciation rate of capital. Two further arguments make us less inclined to use the growth accounting approach. Firstly, when total factor productivity (TFP) growth is measured as a residual, i.e. as the growth rate in GDP that cannot be accounted for by the growth of the observable inputs, it should not be equated with technological change as many researchers have done. Rather it is 'a measure of ignorance' (Abramovitz, 1986), covering many factors like structural change, improvement in allocative efficiency, economies of scale, and other omitted variables and measurement errors. Secondly, although growth accounting provides a convenient way to allow for the breakdown of observed growth of GDP into components associated with changes in factor inputs and in production technologies, we are not convinced that technological change and investment are separable in reality, i.e. changing technology requires investment, and investment inevitably involves technological change. This is consistent with the view of Scott (1989) that technological change and investment are part and parcel of the same thing and that separation is meaningless. Hence, informal growth regression is the methodology that we adopt.

A feature of our study is to use recently developed approaches to model selection in order to construct empirical models based on robust predictors. It is widely believed that growth theories are not explicit enough about variables that should be included in the empirical growth models. The issue of model uncertainty has attracted much research attention in the context of cross-country growth regressions. However, to the best of our knowledge, it has been largely ignored in cross-province growth studies of China, i.e. the existing literature has not explicitly or systematically considered the issue of model selection before any investigation of particular causes of China's growth. We first use two leading model selection approaches, Bayesian Model Averaging and the automated General-to-Specific approach, to examine the association between the growth rate of real GDP per capita and a large range of potential explanatory variables. These include the initial level of income, fixed capital formation, human capital formation, population growth, the degree of openness, institutional change, sectoral change, financial development, infrastructure and regional advantage. The variables flagged as being important by these procedures are then used in formulating our baseline model, which is estimated using panel data system GMM to control for problems of omitted variables, endogeneity and measurement error of regressors. In the second stage, we also examine the robustness of our selected model and the contribution of

the main variables. In this paper, our focus is on the growth impact of various types of physical and human capital investment.

In Section 2 we provide the background to Chinese economic growth, as an aid to interpretation. Section 3 is a literature survey which offers guidance on the choice of variables in our general model. Section 4 explains the empirical methodology and describes the dataset. Section 5 reports the results of our basic equations and their interpretation. The contribution of various physical and human capital investment variables is then examined in detail using system GMM and counterfactual predictions. Section 6 briefly analyzes the factors that permit the rapid capital accumulation in China. Section 7 summarises and concludes.

## **2. Background to China's growth**

The growth of the Chinese economy since the start of its economic reform has been a process of 'crossing the river by groping for the stepping stones', as described by Deng Xiaoping: no stereotype reform package was adopted in advance. One reform begat the need, or the opportunity, for another, and the process became cumulative. The reforms were incremental but hardly slow: huge changes have occurred in less than three decades, as China has moved from central planning towards a market economy. It is relevant that China's had been a labour surplus economy *par excellence*: labour was underemployed in the farms and in the urban state enterprises: government preferred unemployment to be disguised and shared rather than open and threatening (Knight and Song, 2005, chs. 2, 6, and 8). New sectors could thus be expanded without loss of output elsewhere.

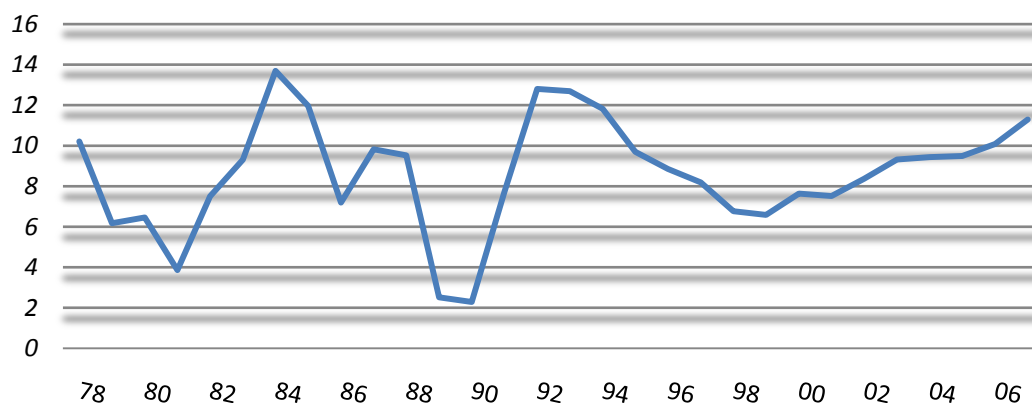
The first stage of economic reform (1978-84) concentrated on the rural areas. The communes were disbanded and individual incentives were restored. Farming households (then 82 percent of the population) were given use-rights to collectively-owned land under long term leases, and the right to sell their marginal produce on the open market. Rural non-farm enterprises were permitted, and they stepped in to produce the light manufactures that the urban state-owned enterprises (SOEs) generally failed to supply. Rural credit constraints encouraged household saving. Rural production rose rapidly as farms became more efficient, as surplus labour was used more productively in rural industry, and as rural entrepreneurship, saving and investment responded to the new opportunities.

The second stage of economic reform (1985-92) was an incremental process of reforming the urban economy, in particular the SOEs, which were gradually given greater

managerial autonomy. The principal-agent problem inherent in state ownership limited the efficiency of SOEs but competition from other market participants – initially village and township enterprises and later domestic and foreign privately owned enterprises as well as from imports – grew steadily.

The third stage of economic reform (1993- ) was ignited by Deng Xiaoping's 'Southern Tour' to mobilise support for more radical reforms. The private sector – for the first time acknowledged and accepted – was invigorated. Moreover, administrative and regulatory reform of rural-urban migration, the banking system, the tax system, foreign trade, and foreign investment lifted various binding constraints on economic growth. For instance, when the delayed effects of the 'one-child family policy' slowed down the growth of the urban-born labour force from the mid-1990s onwards, the relaxation of restrictions on temporary rural-urban migration permitted continued rapid growth of the urban economy.

**Figure 1. China's Annual Growth Rate of GDP Per Capita (%)**



Data source: World Bank Development Indicators (April 2008).

Figure 1, reflecting China's rapid growth of GDP per capita since 1978, shows a cyclical pattern of growth, more marked in the first and second stages of reform than in the third stage. Two peaks are evident, in 1984-5 and 1992-3, respectively reflecting the outcome of agricultural reforms and the green light given to capitalism. The growth rate troughed in 1989-90 owing to a surge of inflation, social unrest and international ostracism. A further examination of provincial growth trends shows that the growth rates of all provinces dropped dramatically in the late 1980s, indicating the general detrimental influence of such adverse shocks on economic growth.

In summary, the reforms created market institutions and incentives that had been lacking in the socialist planned economy. They improved both static allocative efficiency and dynamic factor accumulation. Growth was also facilitated by the absorption of the abundant resource, labour, into the expanding, more productive activities. There was drastic movement towards the economy's production frontier and dramatic movement of the frontier. It is plausible that together they were responsible for China's remarkably high rate of growth. This is the general hypothesis that we wish to explore.

### **3. The growth literature on China**

The starting point in this research is our paper utilising cross-country data to estimate an augmented Solow model of economic growth (Ding and Knight, 2008a). Our estimated growth equations were used to predict China's performance. By inserting China's values of the explanatory variables in the equation, we found that the growth rate could indeed be well explained. Five factors – conditional convergence from a low income level, high physical capital formation, high level of human capital, rapid structural change away from agriculture, and slow population growth – made the main contributions to the difference between China's growth and that of other developing regions. By providing these pointers, this cross-country analysis sets the scene for the current cross-province analysis.

There is a large literature on cross-province growth regressions for China<sup>3</sup>. Two empirical approaches have been used: some version of the neoclassical growth model, often in the form of the augmented Solow model as developed by Mankiw, Romer and Weil (1992) (MRW), or informal growth regressions (for instance, Barro, 1991; Barro and Sala-i-Martin, 2004), that contain among others the explanatory variables in which the researcher is most interested. Different periods are analysed, although most are confined to the period of economic reform, from 1978 onwards. The methods of analysis vary in sophistication, from cross-section OLS to panel data GMM analysis. Research focus covers a broad range of factors relating to variation in growth among Chinese provinces, such as convergence or divergence, physical and human capital investment, openness, economic reform, geographical location, infrastructure, financial development, labour market development, spatial dependence and preferential policies. An underlying problem in all the research is the need to find causal relationships as opposed to mere associations.

---

<sup>3</sup> The literature survey does not include growth accounting or time-series aggregate growth regressions for China.

Both Chen and Fleisher (1996) and Li *et al.* (1998) discovered that China's economic growth over the reform period generated significant forces for convergence of both provincial levels and growth rates of GDP per capita. The convergence was conditional on the variables controlling for the steady state in the augmented Solow model. The levels and growth rates of GDP per capita were higher in the provinces with lower population growth, more investment in physical and human capital and greater openness to foreign countries. Raiser (1998) and Zhang (2001) investigated both  $\beta$  and  $\sigma$  convergence across Chinese provinces and found that the rate of convergence declined after 1985, which decline they attributed to the shift from rural to urban reforms, the reduction of inter-provincial fiscal transfers, and different opportunities for international trade and foreign direct investment.

Yao (2006) examined the effect of exports and FDI on provincial growth rates over the period 1978-2000. By adopting Pedroni's panel cointegration test and Arellano and Bond's dynamic panel data estimating technique to control for the problems of non-stationarity and endogeneity, he found that both exports and FDI made a positive and significant contribution. He established the existence of simultaneous relationship between FDI and GDP growth and between exports and GDP growth, and concluded that the interaction among these three variables formed a virtuous circle of openness and growth in China.

Chen and Feng (2000) adopted a Barro-type framework to model the determinants of cross-province variation in economic growth over the period 1978-89. Their intention was to investigate the commonalities among China's and other countries' growth patterns as well as China's unique growth characteristics. They found that university education, industrialization and international trade raise growth, whereas high fertility and high inflation reduce it. A China-specific institutional factor, the presence of SOEs, was found to inhibit growth.

Bao *et al.* (2002) investigated the effect of geography on regional economic growth in China under market reforms and found that geographic factors such as coastline length, distance factor and population density close to the coastline are important in explaining growth disparities across provinces. Brun *et al.* (2002) focused on the spillover effects of economic growth from China's coastal to non-coastal regions and found that spillover effects were not sufficient to reduce income disparities among the provinces.

Démurger (2001) assessed the relationship between infrastructure development and economic growth in China. The fixed-effect and two-stage least-squares results showed that differences in geographical location, transport infrastructure and telecommunication facilities

accounted for a significant part of the observed differences in growth performance across provinces.

Hao (2006) examined the impact of the development of financial intermediation on growth over the period 1985-99 and found that it contributed to China's economic growth through two channels: the substitution of loans for state budget appropriation and the mobilization of household savings. Loan expansion itself did not contribute to growth because the distribution of loans, not being based on commercial criteria, was inefficient. Guariglia and Poncet (2006) showed that financial distortions represented an impediment to economic growth in China over the period 1989-2003, i.e. traditional indicators of financial development and China-specific indicators of state intervention are generally negatively associated with growth and its sources, whereas indicators of market-driven financing are positively correlated with them. The adverse effects of financial distortions on growth have gradually declined over time, possibly owing to the progressive reform of the banking sector especially after China's entry to the WTO in 2001. They argued that FDI is an important factor in alleviating the costs of financial distortions, so suggesting a partial explanation for why China can grow fast despite a malfunctioning financial system.

Cai *et al.* (2002) placed emphasis on the impact of lagged labour market reform on increasing regional disparity in China during the period 1978-98. They used the comparative productivity of agricultural labour as a proxy for labour market distortion and found that this measure impedes regional growth rates after controlling for a set of variables determining the steady state.

Jones *et al.* (2003) tested the augmented Solow model using city-level data for the decade 1989-99. They showed that policies giving preferential treatment to cities by promoting openness, such as special economic zone status or open coastal status, accounted for a large part of the differences in growth rates across cities. These policies affected growth directly by creating an environment more conducive to production, and indirectly by encouraging FDI to flow to these cities. Higher rates of FDI and lower rates of population growth were shown to be related to faster growth of per capita income. Surprisingly, they found no evidence for a positive effect of domestic investment on city growth, their explanation being that domestic investment in China is not primarily profit-driven and is thus inefficiently allocated.



These studies often use an assortment of economic theories to motivate a variety of variables that are included in the cross-province or cross-city growth regressions, and then test the robustness of their conclusions to the addition of an *ad hoc* selection of further controls. Although each study presents intuitively appealing results, none has directly posed the general question: can the variations among provinces highlighted by cross-province growth regressions explain why the economy as a whole has grown so fast? Moreover, no systematic consideration has been given to uncertainty about the regression specification, with the implication that conventional methods for inference can be misleading. We therefore attempt to fill these two gaps in the growth literature on China by using some recently developed methods of model selection. The baseline model will then be used to examine the deeper causes of rapid economic growth.

## **4. Empirical methodology and data**

### *4.1 Empirical methods*

There is no single explicit theoretical framework to guide empirical work on economic growth. The neoclassical model (Solow, 1956) predicts that the long-run economic growth rate is determined by the rate of exogenous technological progress, and that adjustment to stable steady-state growth is achieved by endogenous changes in factor accumulation. It is silent on the determinants of technological progress. Endogenous growth theory (for instance, Lucas, 1988; Romer, 1990) concentrates on technological progress and emphasizes the role of learning by doing, knowledge spillover, research and development, and education in driving economic growth. Because the theories are not mutually exclusive, the problem of model uncertainty concerning which variables should be included to capture the underlying data generating process presents a central difficulty for empirical growth analysis. This issue has gained increasing attention in the cross-country growth literature following the seminal work of Barro (1991), which identified a wide range of variables that are partially correlated with GDP per capita growth. A number of econometric and statistical methodologies have been developed and applied to handle model uncertainty, among which the Extreme Bounds Analysis, Bayesian Model Averaging, and General-to-Specific approach are most influential.

The issue of fragility of econometric inference with respect to modelling choices was first addressed by Leamer (1983, 1985), who proposed an Extreme Bounds Analysis (EBA) to test for the sensitivity of estimated results with respect to changes in the prior distribution of parameters. The extreme bounds for the coefficient of a particular variable are defined as

the lowest estimate of its value minus two times its standard error and the highest estimate of its value plus two times its standard error when different combinations of additional regressors enter the regression. A variable is regarded to be robust if its extreme bounds lie strictly to one side or the other of zero, i.e. the coefficient of interest displays small variation to the presence or absence of other regressors.

Levine and Renelt (1992) applied this methodology to cross-country growth regressions and investigated the robustness of a large number of variables that were found in the literature to be correlated with growth. In order to reduce the number of regressions required to compute the extreme bounds, they imposed several restrictions on the conditioning information set in modelling, for instance, four variables are *always* included in every regression based on past empirical studies and economic theory<sup>4</sup>; up to only three other control variables can be selected each time from a large pool of variables potentially important for growth. Using this variant version of EBA, they found that very few variables are robustly correlated with growth, i.e. almost all results are fragile to changes in the list of conditioning variables. The only exception is a positive and robust correlation between growth and the share of investment in GDP.

Sala-i-Martin (1997) challenged their pessimistic finding by adopting a less restrictive sensitivity test for the explanatory variables in the growth regressions. He claimed that the version of Levine and Renelt (1992) of EBA is too strong for any variable to pass; for instance, a single irrelevant outlier of the distribution of the coefficient estimates may make the variable non-robust. Therefore, rather than simply classifying variables as robust or fragile, Sala-i-Martin (1997) attempted to assign some level of confidence to each of the variables by computing the entire cumulative distribution function of the estimates of the variable of interest<sup>5</sup>. In this case, a variable is regarded to be robust if 95 percent of the distribution of the coefficient estimates lies to one or the other side of zero. In addition, Sala-i-Martin (1997) made a persuasive case that several of the variables chosen by Levine and Renelt (1992) are almost certainly endogenous. Instead he assembled a dataset that was less susceptible to that problem. Not surprisingly, a substantial number of variables turn out to be strongly related to economic growth using this method.

---

<sup>4</sup> They are the investment share of GDP, the initial level of real GDP per capita in 1960, the initial secondary school enrolment rate and the average rate of population growth.

<sup>5</sup> He constructed the weighted averages of all the estimates of the variable of interest and its corresponding standard deviations, using weights proportional to the likelihoods of each of the models.

One potential shortcoming of the approach used by Sala-i-Martin (1997) is that, owing to the lack of statistical theory, the statistical properties of the computed weighted averages are not well understood. To solve this problem, Fernández *et al.* (2001) and Sala-i-Martin *et al.* (2004) adopted an alternative technique, Bayesian Model Averaging (BMA), to re-examine the Sala-i-Martin (1997) dataset, in which model uncertainty is addressed using a formal statistical approach. The basic idea of BMA is that the posterior distribution of any parameter of interest is a weighted average of the posterior distributions of that parameter under each of the models with weights given by the posterior model probabilities, following strictly the rules of probability theory (see, for example, Leamer, 1978). Unlike the approaches which restrict the set of regressors to contain certain fixed variables and then add a small number of other variables, BMA allows for any subset of the variables to appear in the model. Fernández *et al.* (2001) set out a full BMA framework, which requires the specification of the prior distribution of all of the relevant parameters conditional on each possible model. Based on theoretical results and extensive simulations, they used the so-called 'improper uninformative priors' for the parameters, which are designed to be relatively uninformative so that, given informative data, the final results place relatively little weight on subjective prior knowledge. However, Sala-i-Martin *et al.* (2004) argued that acquisition of prior parameter information is difficult and sometimes infeasible when the number of possible regressors is large. Instead, they proposed a method of Bayesian Averaging of Classical Estimates (BACE), which assumes diffuse priors for the parameters, and for which only one prior parameter specification, the expected model size, is required. Both versions of BMA broadly support the more optimistic conclusion of Sala-i-Martin (1997) that a good number of economic variables have a robust partial correlation with long-run growth.

Another strand of research on model uncertainty is the General-to-Specific (GETS) search methodology emphasized by Hendry and Krolzig (2004). The basic idea is to specify a general unrestricted model (GUM), which is assumed to characterize the essential data generating process, and then to 'test down' to a parsimonious encompassing and congruent representation based on the theory of reduction. They claimed that using the GETS approach permits them to replicate most of the findings of Fernández *et al.* (2001) in just a few minutes. The huge efficiency gain makes the automatic procedures of model selection extremely attractive. Hoover and Perez (2004) compared the performance of the GETS approach with EBA adopted by Levine and Renelt (1992) and Sala-i-Martin (1997). Their Monte Carlo simulation results indicate that GETS algorithm not only usually finds the truth

but also discriminates between true and false variables extremely well. By contrast, EBA in the form advocated by Levine and Renelt (1992) is too stringent and rejects the truth too frequently, while that advocated by Sala-i-Martin (1997) is not able to discriminate and accepts the false too frequently along with the true. According to Hoover and Perez (2004), Sala-i-Martin (1997) was right to criticize Levine and Renelt (1992) for rejecting too many potential determinants of growth as non-robust, but his approach selected many variables that probably do not truly determine differences in growth rates among countries. They therefore concluded that GETS method is superior to EBA in getting at the truth.

In this paper we adopt both BMA and GETS approaches to consider the association between GDP per capita growth rates and a wide range of potential explanatory variables based on the cross-sectional data. The purpose of the first-stage model selection is to provide guidance on the choice of variables to include in the subsequent panel data analysis. BMA and GETS procedures have comparative advantages in dealing with model uncertainty and allow us to consider a wide range of candidate predictors in a rigorous way. However, neither of them is without limits nor immune from criticism. For example, one key disadvantage of BMA is the difficulty of interpretation, i.e. parameters are assumed to have the same interpretation regardless of the model they appear in; in addition, it does not lead to a simple model, making the interpretation of results harder (Chatfield, 1995). Criticisms of GETS modelling are commonly concerned with the problems of controlling the overall size of tests in a sequential testing process and of interpreting the final results from a classical viewpoint (Owen, 2003). Hence, the joint application of BMA and GETS model selection procedures in this paper is to combine the strengths of both methods and to circumvent the limitations of each to some extent (see Appendix 1 for brief discussion of the two methods).

Owing to the inclusion of potentially endogenous variables, no causal interpretation is attached to the results at this stage. When a subset of variables are identified as receiving the greatest support from the underlying data according to the model selection results, a further panel data analysis is conducted to investigate the deeper determinants of provincial GDP per capita growth in China. Although cross-sectional regression has the advantage of focusing on the long-run trends of economic growth, panel data methods can control for omitted variables that are persistent over time, and can alleviate measurement error and endogeneity biases by use of lags of the regressors as instruments (Temple, 1999).

We use a system GMM estimator, developed by Arellano and Bover (1995) and Blundell and Bond (1998), which combines the standard set of equations in first-differences

with suitably lagged levels as instruments, with an additional set of equations in levels with suitably lagged first-differences as instruments. By adding the original equation in levels to the system, they found dramatic improvement in efficiency and significant reduction in finite sample bias through exploiting these additional moment conditions. Bond *et al.* (2001) also claimed that the potential for obtaining consistent parameter estimates even in the presence of measurement error and endogenous right-hand-side variables is a considerable strength of the GMM approach in the context of empirical growth research. Finally, the robustness of our selected models and the contribution of main variables are carefully examined. Given the space limitations, in this paper we concentrate only on the role of physical and human capital formation in driving China's economic growth.

#### 4.2 The dataset

The original sample consists of a panel of 30 provinces with annual data for the period 1978-2006<sup>6</sup>. The data come mainly from *China Compendium of Statistics 1949-2004* compiled by National Bureau of Statistics of China. The data of 2005 and 2006 are obtained from the latest issues of *China Statistical Yearbook*. The reliability of Chinese official macroeconomic data is often under dispute. One important issue is the problem of data inconsistency over the sample period. For example, GDP figures for the years 2005 and 2006 were recompiled on the basis of China's 2004 Economic Census, while corresponding provincial data for earlier years remain unrevised. Another problem is data non-comparability across provinces. Take population as an example: the household registration population figure is provided for some provinces, whereas for others only permanent population data are available. In addition, the substantial 'floating population' of temporary migrants is not fully accounted for by the population data. These discrepancies can result in measurement error problems and may call into question the reliability of our estimation results. Therefore, on the one hand, we use a number of 'cleaning rules' (see Appendix 2) to get rid of potential outliers for each variable and, on the other hand, we employ the panel data System GMM estimator to deal with potential mismeasurement.

Our first-stage model selection analysis is based on cross-sectional data, in which observations are averaged over the entire sample period. For the subsequent panel-data study, we opt for the non-overlapping five-year time interval, which is widely used in the cross-

---

<sup>6</sup> China is administratively decomposed into 31 provinces, minority autonomous regions, and municipalities. Since Chongqing becomes a municipal city since 1997, we combine Chongqing with Sichuan for the period 1997-2006, so making it consistent with earlier observations.

country growth literature (for instance, Islam, 1995; Bond *et al.*, 2001; Ding and Knight, 2008a). On the one hand, by comparison with the yearly data, the five-year average setup alleviates the influence of temporary factors associated with business cycles. On the other hand, we are able to maintain more time series variation than would be possible with a longer-period interval.

**Table 1. Descriptive Statistics of Provincial GDP Per Capita Growth Rates**

	Full-sample period	Sub-sample periods		
	1978-2006	1978-1987	1988-1997	1998-2006
All provinces (30 provinces)	0.076 (0.058)	0.073 (0.056)	0.054 (0.065)	0.102 (0.037)
Coastal provinces (11 provinces)	0.082 (0.058)	0.076 (0.053)	0.065 (0.071)	0.105 (0.032)
Interior provinces (19 provinces)	0.072 (0.058)	0.071 (0.058)	0.047 (0.060)	0.100 (0.039)
Highest growth province	0.102 (0.061)	0.112 (0.057)	0.108 (0.075)	0.119 (0.042)
Lowest growth province	0.055 (0.060)	0.020 (0.025)	0.012 (0.089)	0.079 (0.036)

Note: Mean values and standard deviations (in parentheses) are provided; coastal provinces consist of Liaoning, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan, plus Beijing; and interior provinces include Anhui, Gansu, Guangxi, Guizhou, Heilongjiang, Henan, Hubei, Hunan, Inner Mongolia, Jiangxi, Jilin, Ningxia, Qinghai, Shaanxi, Shanxi, Sichuan, Tibet, Xinjiang and Yunnan; for the full-sample period, the highest growth province was Zhejiang, and the lowest growth province was Gansu; for the three sub-sample periods, Zhejiang, Fujian, Shandong were the highest growth provinces respectively, and Shanghai, Tibet, Yunnan were the corresponding lowest growth provinces.

All the variables are calculated in 1990 constant prices and price indices are province-specific<sup>7</sup>. The dependent variable is the growth rate of real GDP per capita. Table 1 shows descriptive statistics of provincial growth rates of real GDP per capita. The annual average per capita growth rates of all 30 provinces over the entire reform period was 7.6 percent, with an average value of 8.2 percent for the coastal provinces and 7.2 percent for interior provinces. China's economic reform generated across-the-nation rapid growth, i.e. both the coastal and inner regions grew fast by international standards. However, that a growth disparity did exist is indicated by the five-percent average growth difference between the highest growth province (Zhejiang) and lowest one (Gansu) over the full sample period. Table 1 also reveals interesting time patterns in China's growth. Rapid growth occurred in the first decade, slowed down and became more volatile in the second decade, and accelerated

<sup>7</sup> The deflator is the provincial consumer price index. The provincial price data of Tibet are missing for the period 1978-89; we use the national aggregate price index to substitute.

but stabilized in the third decade. In the period 1998-2006, the growth disparity across provinces became smaller and even the lowest-growing province (Yunnan) managed an average rate of 7.9 percent.

The explanatory variables can be broadly classified into ten categories: initial level of income, physical capital formation, human capital formation, population growth rate, the degree of openness, pace of economic reform or institutional change, sectoral change or degree of industrialization, infrastructure, financial development, and geographic location (see Appendix 2 for detailed definitions).

## 5. Empirical results

### 5.1 Baseline equation

The validity of a selected model depends primarily on the adequacy of the general unrestricted model as an approximation to the data generation process (Doornik and Hendry, 2007). A poorly specified general model stands little chance of leading to a good 'final' specific model. We consider ten different groups of explanatory variables, and rely on theory of economic growth (although sufficiently loose) and previous empirical findings to guide the specification of the general model. One important issue is that variables within each category are highly correlated, which may result in the problem of multicollinearity and thus inflate the coefficient standard errors if all variables are simultaneously included in one general regression. The strategy we adopt is to select one or two representative variables from each range (based on existing empirical literature and correlation results) to form the basic general model, and then to test for the robustness of the model selection results using other variables left in each group. Throughout this section, when we refer to growth we shall, unless indicated otherwise, mean annual average growth of real GDP per capita.

We start from a general model that includes 13 explanatory variables and searches for statistically acceptable reductions of this model. The included variables are the logarithm of initial level of income ( $\ln y_{i,t-1}$ ), ratio of fixed capital formation over GDP ( $fcf/GDP$ ), secondary school enrolment rate ( $stu_{SEC}/pop$ ), ratio of students enrolled in higher education to students enrolled in regular secondary education ( $stu_{HIGH}/stu_{REG\_SEC}$ ), population natural growth rate ( $pop\_ngr$ ), ratio of exports to GDP ( $export/GDP$ ), SOEs' share of total industrial output ( $ind_{SOE}/ind_{TOTAL}$ ), change in non-agricultural share of employment ( $mgrowth$ ), degree of industrialization ( $deofin$ ), railway density ( $railway\_area$ ), ratio of business

**Table 2. Correlation Matrix of the General Model**

	$g_{i,t}$	$\ln y_{i,t-1}$	$\frac{fcf}{GDP}$	$\frac{stu_{SEC}}{pop}$	$\frac{stu_{HIGH}}{stu_{REG\_SEC}}$	$pop\_ngr$	$\frac{export}{GDP}$	$\frac{ind_{SOE}}{ind_{TOTAL}}$	$mgrowth$	$deofin$	$\frac{railway}{area}$	$\frac{post \& tele}{GDP}$	$\frac{loan}{GDP}$	$dum\_coastal$
$g_{i,t}$	1.0000													
$\ln y_{i,t-1}$	0.3081	1.0000												
$\frac{fcf}{GDP}$	0.4081	0.5010	1.0000											
$\frac{stu_{SEC}}{pop}$	0.4147	0.5134	0.3645	1.0000										
$\frac{stu_{HIGH}}{stu_{REG\_SEC}}$	0.2943	0.7520	0.4465	0.3341	1.0000									
$pop\_ngr$	-0.4350	-0.7942	-0.3482	-0.5435	-0.6691	1.0000								
$\frac{export}{GDP}$	0.1992	0.6319	0.1539	0.2085	0.4495	-0.3600	1.0000							
$\frac{ind_{SOE}}{ind_{TOTAL}}$	-0.5287	-0.6915	-0.3504	-0.3978	-0.3940	0.6884	-0.4881	1.0000						
$mgrowth$	0.2666	-0.0246	0.0038	-0.0366	-0.0811	0.0285	0.1282	-0.0527	1.0000					
$deofin$	0.1077	0.6512	0.1398	0.3677	0.5552	-0.5904	0.5296	-0.3604	0.0440	1.0000				
$\frac{railway}{area}$	0.0091	0.4758	0.1461	0.2296	0.6114	-0.4482	0.3539	-0.1709	-0.1408	0.5666	1.0000			
$\frac{post \& tele}{GDP}$	0.5871	0.6663	0.6803	0.5524	0.5613	-0.5986	0.3393	-0.6854	-0.0055	0.2234	0.1231	1.0000		
$\frac{loan}{GDP}$	0.1783	0.3889	0.5121	0.1388	0.4140	-0.2781	0.1874	-0.1703	-0.0074	0.2545	0.1743	0.4563	1.0000	
$dum\_coastal$	0.0976	0.4791	-0.0323	0.0792	0.3360	-0.3351	0.5801	-0.3379	0.1477	0.4465	0.4204	0.0405	0.0098	1.0000



volume of post and telecommunications to GDP (*post & tele/GDP*), and a coastal dummy (*dum\_coastal*).

The correlation matrix of these variables is presented in Table 2. The positive correlation between growth and initial level of income suggests that there exists absolute (unconditional) divergence, rather than convergence, across the provinces over the entire sample period. Formal regression analysis using five-year average panel data system GMM shows similar results (Table 5): regressing the growth rate of GDP per capita on the logarithm of initial income term alone yields a positive and significant coefficient. In the growth literature, it is commonly believed that absolute convergence is more likely to apply across regions within countries than across countries because the former tend to have similar institutional arrangements, legal systems and cultures, giving rise to the possibility of their convergence to the same steady state. For instance, Barro and Sala-i-Martin (2004) found evidence that absolute convergence is the norm for the US states since 1880, the prefectures of Japan since 1930, and the regions of eight European countries since 1950. By contrast, some researchers have argued that heterogeneity in technology, preferences, and institutions may prevent unconditional convergence from occurring across regions (for example, Beenstock and Felsenstein, 2008). Our finding of unconditional regional divergence in China supports the latter view. Among other variables, only population growth and SOEs' share of total industrial output show a strong negative correlation with growth, which is consistent with our predictions.

**Table 3. Bayesian Model Averaging (BMA) Results**

Regressor	Posterior Probability of Inclusion	Posterior Mean	Posterior Standard Deviation
<i>Intercept</i>	100.0	0.207	0.042
<i>lny<sub>i,t-1</sub></i>	100.0	-0.019	0.007
<i>ind<sub>SOE</sub>/ind<sub>TOTAL</sub></i>	96.9	-0.053	0.022
<i>stu<sub>SEC</sub>/pop</i>	93.4	0.352	0.175
<i>dum_coastal</i>	62.2	0.006	0.006
<i>export/GDP</i>	29.7	0.011	0.022
<i>fcf/GDP</i>	29.3	0.015	0.029
<i>pop_ngr</i>	27.8	-0.432	0.904
<i>stu<sub>HIGH</sub>/stu<sub>REG_SEC</sub></i>	24.8	0.017	0.039
<i>loan/GDP</i>	10.8	-0.001	0.006
<i>post &amp; tele/GDP</i>	7.8	-0.006	0.041
<i>deofin</i>	7.3	0.001	0.006
<i>mgrowth</i>	5.4	0.009	0.123
<i>railway_area</i>	4.3	-0.002	0.022

Notes: Estimation is based on cross-sectional data; Dependent variable: growth rate of real provincial GDP per capita.

We first use BMA to isolate variables that have a high posterior probability of inclusion. In Table 3, we present a summary of the BMA results, where the posterior probability that the variable is included in the model, the posterior mean, and the posterior standard deviation for each variable are reported. We are aware of the difficulty of interpreting parameters in economic terms when the conditioning variables differ across models, so our emphasis here lies on the posterior probability of inclusion for each variable, i.e. the sum of posterior model probabilities for all models in which each variable appears. The results indicate a possibly important role for the initial level of income, SOEs' share of total industrial output, secondary school enrolment rate, coastal dummy, exports, fixed capital formation, and population growth. Each of these variables has a posterior probability of inclusion above 25 percent.

**Table 4. General-to-Specific (GETS) Model Selection Results**

Regressor	Coefficient	Standard Error	t-value	t-probability	Part. $R^2$
<i>Constant</i>	0.248	0.038	6.515	0.000	0.649
<i>lny<sub>i,t-1</sub></i>	-0.025	0.006	-4.307	0.000	0.447
<i>fcf/GDP</i>	0.059	0.027	2.234	0.036	0.178
<i>stu<sub>SEC</sub>/pop</i>	0.309	0.147	2.104	0.047	0.161
<i>pop_ngr</i>	-1.854	0.891	-2.082	0.049	0.159
<i>export/GDP</i>	0.041	0.021	1.934	0.066	0.139
<i>ind<sub>SOE</sub>/ind<sub>TOTAL</sub></i>	-0.056	0.018	-3.134	0.005	0.299
Sigma	0.007	RSS	0.001	$R^2$	0.767
F(6,23)	12.64 [0.000]	LogLik	108.627	T	30
AIC	-9.613	SC	-9.286	HQ	-9.508
Normality test			Chi <sup>2</sup> (2) = 1.872 [0.393]		
Testing for heteroscedasticity			F(12,10) = 0.558 [0.832]		

Notes: This is the OLS estimation of final specific model based on cross-sectional data; Dependent variable: growth rate of real provincial GDP per capita.

We then conduct an automatic model selection exercise using the GETS methodology. Starting from the same general model and searching for statistically acceptable reductions, *Autometrics* arrives at a final model with a set of explanatory variables broadly similar to those highlighted by the BMA analysis. The OLS estimation of the final specific model is reported in Table 4. We find that the initial income level, population growth and SOEs' share of industrial output are negatively correlated with GDP per capita growth, whereas fixed capital investment, secondary school enrolment rates and exports are positively correlated. The major difference between the results of the two methods lies in the role of the regional dummy variable in explaining cross-province growth rates, i.e. BMA analysis flags the

coastal dummy as potentially important (with a posterior inclusion probability of 62 percent), but GETS drops that variable during reductions. Other variables such as sectoral change, infrastructure and financial development are flagged as unimportant predictors of economic growth by both model selection methods. However, this outcome may simply reflect the highly endogenous nature of these variables, which cannot be accounted for at the model-selection stage. We will re-examine the role of these variables in determining output growth in the panel data context in a different paper (Ding and Knight, 2008b).

The existence of a robust partial correlation does not imply that the variables of interest cause growth (Levine and Renelt, 1992). Based on the model selection results delivered by BMA and GETS, we therefore estimate the baseline model using panel data system GMM, in which the endogeneity of regressors can be controlled for. Note that all estimated standard errors are corrected for heteroscedasticity and that time dummies are included. We treat the population natural growth rate as an exogenous variable, the initial level of income as a predetermined variable, and all other variables including physical and human capital accumulation, exports and SOEs' share of industrial output as potentially endogenous variables. Since the  $p$  values of over-identifying tests may be inflated when the number of moment conditions is large (Bowsher, 2002), we restrict the number of instruments used for each first differenced equation by including a subset of instruments for each predetermined or endogenous variable. Several studies have found that the two-step standard errors tend to be biased downwards in finite samples (Arellano and Bond, 1991; Blundell and Bond, 1998). By applying a correction to the two-step covariance matrix derived by Windmeijer (2005), we find very similar results obtained from the one-step and two-step GMM estimators. To conserve space we report only the heteroscedasticity-robust one-step system GMM results.

Interestingly, our panel data system GMM results support the model selected by the GETS procedure, i.e. the coastal dummy appears insignificant, and there is not much effect on other parameters with or without this regional variable. Controlling for other explanatory variables, the initial level of income is found to have a negative effect on subsequent provincial growth rates, providing evidence of conditional convergence over the reform period. The estimated coefficient implies that a one percent lower initial level of GDP per capita raises the subsequent growth rate of GDP per capita by 0.06 percentage points. Conditional convergence is an implication of the neoclassical growth model, deriving from the assumption of diminishing returns to capital accumulation. The controls imply that the provinces have different steady states, and that convergence will lead them to their respective

steady state levels of income per capita. Despite the challenge posed by endogenous growth theory, the neoclassical paradigm of convergence is widely supported by empirical evidence in both the cross-country growth literature (for example, MRW, 1992; Islam, 1995; Bond *et al.*, 2001; Ding and Knight, 2008a) and the cross-province growth study on China (for example, Chen and Fleisher, 1996; Chen and Feng, 2000; Cai *et al.*, 2002). Table 5 also shows estimates of the effects of initial income per capita in the absence of controls for other variables: the coefficient is significantly positive, indicating absolute divergence.

**Table 5. System GMM Estimation Results of the Baseline Model**

Regressor	Without Coastal Dummy				With Coastal Dummy	
	Coefficient	Robust S.E.	Coefficient	Robust S.E.	Coefficient	Robust S.E.
<i>Constant</i>	0.404**	0.078			0.403**	0.078
$\ln y_{i,t-1}$	-0.055**	0.008	0.036**	0.004	-0.056**	0.009
<i>fcf/GDP</i>	0.151**	0.036			0.157**	0.048
<i>stu<sub>SEC</sub>/pop</i>	1.329**	0.301			1.334**	0.298
<i>pop_ngr</i>	-4.349**	1.282			-4.217**	1.329
<i>export/GDP</i>	0.083**	0.019			0.079**	0.019
<i>ind<sub>SOE</sub>/ind<sub>TOTAL</sub></i>	-0.036**	0.013			-0.037**	0.012
<i>dum_coastal</i>					0.003	0.012
<i>AR(2) p value</i>	0.944		0.811		0.955	
<i>Hansen p value</i>	0.835		0.362		0.904	
<i>Dif Sargan p value</i>	0.316		0.966		0.276	
<i>Num of Obs</i>	148		150		148	

Notes: 5-year interval panel data is used for estimation; robust standard error refers to heteroskedasticity-consistent standard error;  $\ln y_{i,t-1}$  is treated as pre-determined, *pop\_ngr* is treated as exogenous, and all other variables are treated as endogenous; \*\* and \* indicate that the coefficient is significantly different from zero at the 5 and 10 percent significance level respectively.

Our findings of absolute divergence and conditional convergence reveal an interesting growth pattern in China: poor provinces did not grow faster than rich ones, but they tended to converge in a relative sense towards their own steady states. One possible explanation for this pattern is that relatively poor provinces have lower stocks of physical and human capital, so that the marginal product of capital is higher for them. Another explanation might lie in central government's regional development policies. During the period 1978-1993, fiscal decentralization reform gave provincial governments more discretionary power in tax administration and revenue collection. The 'fiscal contracting system' reduced central government's share of revenue and curtailed fiscal transfers away from rich and towards poor provinces (Raiser, 1998; Knight and Li, 1999). In 1994, the 'tax assignment system' reform strengthened central government's fiscal capacity, which enabled it to increase fiscal

redistribution towards poor provinces and to promote economic development in poor regions such as the western provinces and minority areas. This might help to explain why absolute divergence has been weaker in recent years (Table 1).

Fixed capital formation is an important determinant of China's growth, i.e. a one percentage point rise in the ratio of fixed capital formation to GDP in a province raises its growth rate of GDP per capita by 0.2 percentage points. Human capital investment appears to be even more important, i.e. a one percentage point increase in secondary school enrolment rates is associated with a higher growth rate of GDP per capita by 1.3 percentage points. Since both physical and human capital accumulation are the focus of this paper, detailed discussion will follow in the next two subsections.

Natural increase in population has a negative consequence for growth, i.e. reducing the rate of population growth by one percent is associated with an increase in GDP per capita growth of 0.5 percent<sup>8</sup>. Rapid population growth rate can be referred to as an opportunity cost of economic growth, i.e. faster growth of the labour force means more capital has to be used to equip the growing labour force, and hence there is less scope for capital deepening, with resultant slower growth of capital per worker and thus output per worker. From the neoclassical point of view and within the standard Solow model, slower population growth implies a higher equilibrium level of the output per worker and capital stock per worker. This means that a province is further from its equilibrium and the forces of convergence are therefore stronger, so raising the output per worker faster. China has been keen to curb its population growth mainly through the family planning policy, implemented since the late 1970s. Despite the controversy over the humanity of the 'one-child family policy', such tightened demographic policy has been efficient in slowing down population growth and reducing the strain on resources in China, which has a positive impact on its growth of GDP per capita.

Exports are conducive to provincial growth, i.e. a one percentage point increase in the ratio of exports to GDP leads to an increase in GDP per capita growth of 0.08 percentage points. According to the report of Commission on Growth and Development (2008), a flourishing export sector is an important ingredient of high and sustained growth, especially in the early stages. In endogenous growth theory, international trade, especially exports, is viewed as an important source of human capital augmentation, technological change and

---

<sup>8</sup> We calculate the elasticity of  $y$  with respect to  $x_j$ , equivalent to  $\partial \log(y) / \partial \log(x_j)$ .

knowledge spillover across countries (Grossman and Helpman, 1995). China's open-door policy, adopted after 1978, created an excellent opportunity to exploit its comparative advantage in the labour-intensive manufacturing industry, making exports a driver of China's growth.

The SOEs' share of industrial output has a significant and negative impact on output growth, i.e. a decrease of one percentage point in the variable raises GDP per capita growth rate by 0.04 percentage points. This variable is a proxy for the pace of economic reform or institutional change. In the mid-1980s, SOEs were given successively greater autonomy in production and a greater share of the profits they generated through a variety of profit remittance contracts and management responsibility systems (Riedel *et al.*, 2007). However, owing to the principal-agent problem inherent to state ownership, the effect of the industrial reform in improving the efficiency and profitability of SOEs remained limited. By contrast, non-state-owned enterprises such as collectively-owned rural township and village enterprises in the 1980s and domestic and foreign privately-owned industrial enterprises in the 1990s grew rapidly in response to market opportunities and better incentive structures. Therefore, the declining share of SOEs in industrial output is conducive to the growth of GDP per capita.

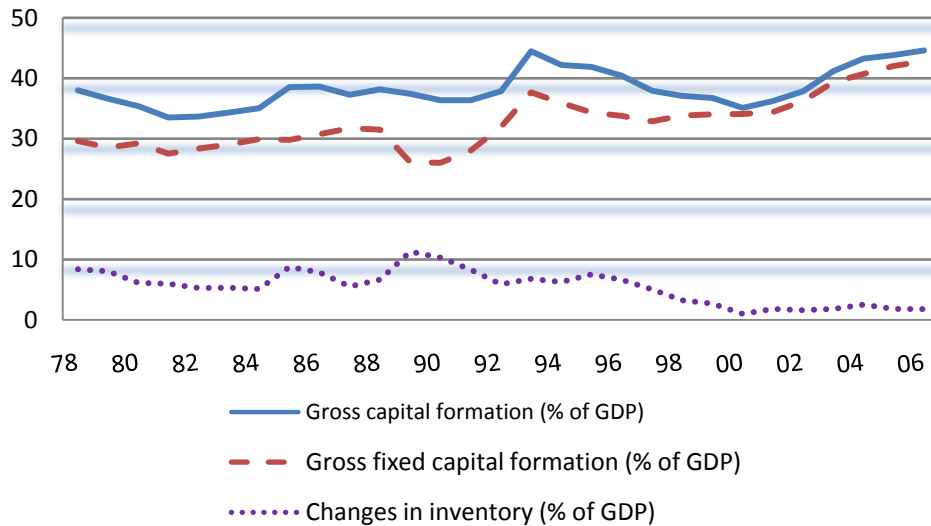
Our system GMM estimation shows that there is no evidence of second order serial correlation in the first-differenced residuals and neither the Hansen test nor the Difference Sargan test rejects the validity of instruments, all of which results suggest the consistency of the estimators being used. In brief, our panel data results favour the model selected by GETS procedure and highlights the role of conditional convergence, physical and human capital formation, population growth, degree of openness, and institutional change in determining economic growth across Chinese provinces.

## *5.2 Physical capital accumulation*

It is widely believed that China's exceptional growth performance over the past three decades is most fundamentally a reflection of the high investment rates that have characterised the economy. As Figure 2 illustrates, real gross capital formation over the entire reform period averaged a fairly steady 38.3 percent of real GDP, which is very high by international standards. The rate of gross fixed capital formation has increased significantly in recent years, rising from an average of 29.3 percent between 1978 and 1993 to an average of 36.6 percent thereafter. Inventory accumulation amounted to, on average, 5.5 percent of

GDP. It peaked at the end of 1980s, reflecting the severe economic recession, and declined gradually thereafter thanks to reforms away from the planned economy. Hence, it is not implausible to hypothesize that China's growth success is mainly investment-driven and that a major part of the answer to the question 'why does China grow so fast?' is simply 'because it invests so much' (Naughton, 2006; Riedel *et al.*, 2007).

**Figure 2. Gross Capital Formation and Its Composition**



Data source: World Bank Development Indicators (April 2008).

In the cross-country growth literature, there is a substantial empirical evidence that capital accumulation has a positive and significant effect on growth, for instance Barro (1991), Levine and Renelt (1992), Sala-i-Martin (1997), and Bond *et al.* (2007). However, Blomström *et al.* (1996) argued that fixed investment does not cause economic growth: they found that growth induces subsequent capital formation more than capital formation induces subsequent growth. In an influential review of the recent empirical literature, Easterly and Levine (2001) claimed that 'the data do not provide strong support for the contention that factor accumulation ignites faster growth in output per worker'. Given this controversy, the issue of causality is crucial in examining the role of capital formation in the Chinese context. Controlling for the set of variables selected in the baseline model, we focus on the impact of various types of physical capital investment on GDP per capita growth. Because the association between fixed investment and growth does not prove causality, all measures of physical capital formation are treated as endogenous variables in our system GMM estimation, i.e. levels of physical capital investment lagged by 10-year and 15-year periods are used as instruments in the first-differences equations, and first-differenced physical

investment variables lagged by 5-year period are used as additional instruments for the levels equations.

In Table 6, we test for the role of gross capital formation ( $gcf/GDP$ ) and its two components, fixed capital formation ( $fcf/GDP$ ) and inventory accumulation ( $inven/GDP$ ). In Model 1, fixed capital formation has a positive and significant effect on growth, i.e. a one

**Table 6. Robustness Tests for Physical Capital Investment**  
(By national account classification)

Regressor	Model 1	Model 2	Model 3	Model 4	Model 5
<i>Constant</i>	0.403** (0.078)	0.301** (0.085)	0.349** (0.092)	0.179** (0.083)	0.349** (0.067)
$\ln y_{i,t-1}$	-0.056** (0.009)	-0.032** (0.009)	-0.046** (0.011)	-0.044** (0.008)	-0.041** (0.008)
$fcf/GDP$	0.157** (0.048)				0.076* (0.042)
$inven/GDP$		-0.513** (0.072)			-0.433** (0.053)
$gcf/GDP$			0.038 (0.048)		
$fcf/gcf$				0.251** (0.027)	
$stu_{SEC}/pop$	1.334** (0.298)	1.330** (0.180)	1.744** (0.330)	0.859** (0.239)	1.133** (0.192)
$pop\_ngr$	-4.217** (1.329)	-2.843** (1.406)	-4.476** (1.481)	-3.244** (1.325)	-2.783** (1.258)
$export/GDP$	0.079** (0.019)	0.047* (0.028)	0.092** (0.029)	0.048** (0.024)	0.045** (0.023)
$ind_{SOE}/ind_{TOTAL}$	-0.037** (0.012)	-0.004 (0.023)	-0.029* (0.017)	-0.022 (0.018)	-0.021 (0.015)
$dum\_coastal$	0.003 (0.012)	0.001 (0.011)	-0.019 (0.015)	0.012 (0.009)	0.013 (0.011)
<i>AR(2) p value</i>	0.955	0.970	0.487	0.522	0.955
<i>Hansen p value</i>	0.904	0.869	0.899	0.828	0.993
<i>Dif Sargan p value</i>	0.276	0.224	0.203	0.155	0.552
<i>Num of Obs</i>	148	148	148	148	148

Notes: 5-year interval panel data is used for estimation; heteroskedasticity-consistent standard error is in parentheses;  $\ln y_{i,t-1}$  is treated as pre-determined,  $pop\_ngr$  is treated as exogenous, and all other variables are treated as endogenous; \*\* and \* indicate that the coefficient is significantly different from zero at the 5 and 10 percent significance level respectively.

percentage point rise in the fixed investment ratio results in a 0.2 percentage point increase in GDP per capita growth. Thus, fixed capital formation is indeed a source of economic growth. By contrast, inventory accumulation has a large negative impact (Model 2), i.e. an increase in inventory investment of one percentage point reduces GDP per capita growth by 0.5



percentage points. While some inventory accumulation, buildup of stockpiles and stores, is a necessary consequence of economic growth, most of inventory accumulation in China over the 30-year period can be viewed as a command-economy-type inefficiency, attributable to the government's policy of requiring the loss-making SOEs to maintain or expand employment and production even when they were unable to sell all their products. Lastly, we find that gross capital formation is insignificant in determining output growth (Model 3) owing to the offsetting effects of its two components. In Model 4, following Cai *et al.* (2002), we define an investment efficiency term as the proportion of fixed capital formation in gross capital formation. As expected, the coefficient of this term is positive and highly significant, reflecting the fact that growth is accelerated when more investment is directed towards productivity-enhancing fixed assets. The positive effect of fixed capital formation and negative effect of inventory accumulation on growth remain robust when both variables enter the regression jointly (Model 5).

In Table 7, total investment in fixed assets ( $finv_{Total}/GDP$ ) is decomposed into investment in capital construction<sup>9</sup> ( $finv_{CC}/GDP$ ), investment in innovation<sup>10</sup> ( $finv_{INNO}/GDP$ ), and investment in other fixed assets<sup>11</sup> ( $finv_{OTHER}/GDP$ ). We find that the former three have positive and significant impacts on growth (Models 6-8), whereas the last appears insignificant (Model 9). These results highlight the role of investment spending on capital construction and technological innovation in promoting growth and imply that fixed investment in other areas such as the real estate sector and natural resource exploitation is not growth-enhancing, possibly because the full effects are longer term. Moreover, the growth impact of investment in innovation is much greater than that of total investment in fixed assets and that of investment in capital construction, i.e. a one percentage point rise in the innovation investment ratio is associated with 0.3 percentage point higher growth rate of GDP per capita. Although, according to Scott (1989, 1993), almost every investment is bound up with technological change, our 'investment in innovation' variable is particularly likely to identify productivity-enhancing innovation. Our finding suggests that the contribution which investment can make to technological progress is a powerful driver of growth. China began economic reform from a position far below the technological frontier. In a transition

---

<sup>9</sup> Capital construction investment refers to the new construction projects or extension projects of enterprises and other institutions with the purpose of expanding production capacity or improving project efficiency.

<sup>10</sup> Innovation investment consists of the renewal of fixed assets and technological innovation of the original facilities by enterprises and other institutions.

<sup>11</sup> Investment in other fixed assets includes investment in real estate development, natural resource (oil, coal, ore, etc) maintenance and exploitation, and other construction and purchases of fixed assets not listed in capital construction investment and innovation investment.

economy, much investment is necessary to maintain the value of the firm in response to a violent process of obsolescence created by new products, new production processes, and huge changes in relative prices (Riedel *et al.*, 2007). Table 7 provides the evidence that investment and investment-driven improvement in technology are important for China's growth.

**Table 7. Robustness Tests for Physical Capital Investment**  
(By usage classification)

Regressor	Model 6	Model 7	Model 8	Model 9
<i>Constant</i>	0.369** (0.086)	0.427** (0.085)	0.384** (0.088)	0.349** (0.086)
$\ln y_{i,t-1}$	-0.052** (0.011)	-0.053** (0.009)	-0.045** (0.011)	-0.039** (0.008)
$f_{inv_{Total}}/GDP$	0.169** (0.036)			
$f_{inv_{CC}}/GDP$		0.177** (0.031)		
$f_{inv_{INNO}}/GDP$			0.316** (0.099)	
$f_{inv_{OTHER}}/GDP$				-0.029 (0.079)
$stu_{SEC}/pop$	1.244** (0.281)	1.335** (0.253)	0.925** (0.330)	1.458** (0.314)
$pop\_ngr$	-3.480** (1.158)	-4.578** (1.287)	-3.192** (1.214)	-5.023** (1.552)
$export/GDP$	0.056** (0.016)	0.061** (0.016)	0.068** (0.022)	0.075** (0.027)
$ind_{SOE}/ind_{TOTAL}$	-0.039** (0.016)	-0.054** (0.015)	-0.063** (0.016)	-0.032** (0.022)
$dum\_coastal$	0.011 (0.012)	0.009 (0.009)	-0.001 (0.011)	-0.014 (0.011)
<i>AR(2) p value</i>	0.917	0.641	0.414	0.480
<i>Hansen p value</i>	0.929	0.900	0.882	0.884
<i>Dif Sargan p value</i>	0.202	0.164	0.194	0.168
<i>Num of Obs</i>	149	149	149	149

Notes: 5-year interval panel data is used for estimation; heteroskedasticity-consistent standard error is in parentheses;  $\ln y_{i,t-1}$  is treated as pre-determined,  $pop\_ngr$  is treated as exogenous, and all other variables are treated as endogenous; \*\* and \* indicate that the coefficient is significantly different from zero at the 5 and 10 percent significance level respectively.

In Table 8, we first distinguish the roles of domestic investment ( $f_{inv_{DOM}}/GDP$ ) and foreign direct investment ( $f_{di}/GDP$ ) in driving growth. Ideally, we would choose measures of the *stock* of FDI. But in the absence of such data at the provincial level, the *flows* of FDI are adopted. Theoretically speaking, FDI can contribute to growth by bringing not only capital but also a familiarity with foreign production techniques, overseas markets, and

international supply chains to the host country. Empirical evidence of knowledge or productivity spillovers from FDI is rather mixed. For instance, Javorcik (2008) utilised survey data to examine various channels through which the entry of multinationals affects domestic enterprises in the same industry or in upstream or downstream sectors in the Czech Republic and Latvia. She found both positive and negative effects of foreign entry on domestic producers, depending on host country conditions and the type of FDI inflows. In the Chinese literature, the aggregate growth impact of FDI is commonly found to be positive. For instance, Wei (1995) investigated the growth impact of China's open door policy using city-level datasets and found that higher rates of FDI were related to faster growth of per capita income in the late 1980s. Similar findings were discovered by Jones *et al.* (2003) who tested the Solow model augmented by FDI using city-level data for the decade 1989-99. Using quarterly national data from the years 1981-97, Liu *et al.* (2002) showed that there is a long-run bi-directional causal relationship among growth, imports, exports, and FDI in a time-series cointegration framework. From a cross-province panel data analysis for the period 1978-2000, Yao (2006) found that both exports and FDI have a strong and positive effect on economic growth. Our hypothesis is therefore that a unit of FDI is more important for growth than a unit of domestic investment on account of technological transfers and knowledge spillovers.

Domestic investment has a positive influence on growth, i.e. an increase in domestic investment of one percentage point raises GDP per capita growth by 0.2 percentage points (Model 10). Surprisingly, we find no evidence of a significant and positive impact of FDI (Model 11). One possible explanation of our finding might lie in variable specification. The share of FDI inflow to total GDP ( $fdi/GDP$ ) is a useful indicator to measure the importance of FDI in a host country's national economy and to reflect a country's openness towards inward FDI. FDI grew steadily through the 1980s but never exceeded one percent of GDP (Figure 3), and it was highly concentrated in four Special Economic Zones in Guangdong and Fujian provinces. Deng Xiaoping's 'Southern Tour' in the spring of 1992 unleashed a surge of inward FDI in China, which reached its peak of 6.2 percent of GDP in 1994. Since 1995, the share of FDI in GDP has drifted downwards because China's GDP growth has been more rapid than the growth of FDI inflow. This may partly explain the insignificant and negative coefficient of the FDI term in Model 11.

**Table 8. Robustness Tests for Physical Capital Investment**  
(Domestic vs foreign investment)

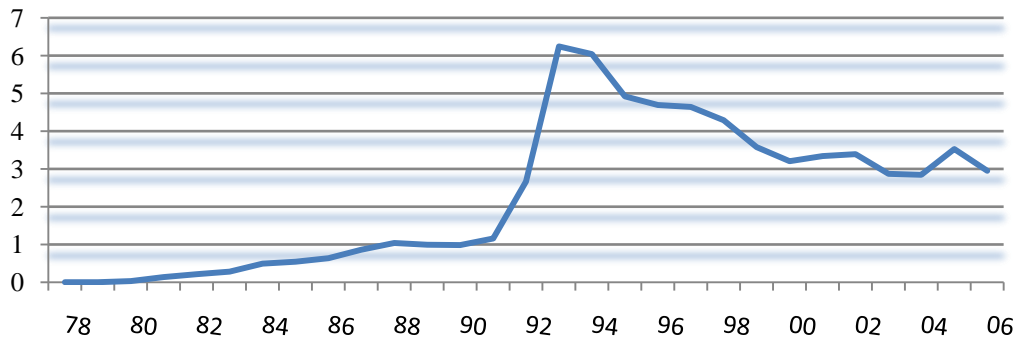
Regressor	Model 10	Model 11	Model 12	Model 13
<i>Constant</i>	0.426** (0.099)	0.254** (0.086)	0.270** (0.084)	0.349** (0.092)
<i>lny<sub>i,t-1</sub></i>	-0.059** (0.011)	-0.032** (0.009)	-0.033** (0.010)	-0.039** (0.011)
<i>finv<sub>DOM</sub>/GDP</i>	0.221** (0.032)			
<i>fdi/GDP</i>		-0.079 (0.099)		
<i>fdi</i>			0.006** (0.003)	0.006** (0.002)
<i>finv<sub>DOM</sub></i>				0.001** (0.000)
<i>stu<sub>SEC</sub>/pop</i>	1.063** (0.284)	1.775** (0.250)	1.394** (0.393)	0.898** (0.374)
<i>pop_ngr</i>	-3.672** (1.289)	-3.246** (1.257)	-3.210** (1.277)	-4.084** (1.471)
<i>export/GDP</i>	0.059** (0.019)	0.071** (0.024)	0.069** (0.025)	0.044* (0.025)
<i>ind<sub>SOE</sub>/ind<sub>TOTAL</sub></i>	-0.053** (0.019)	-0.029* (0.017)	-0.023 (0.018)	-0.015 (0.015)
<i>dum_coastal</i>	0.028** (0.014)	-0.014 (0.013)	-0.010 (0.013)	-0.004 (0.010)
<i>AR(2) p value</i>	0.863	0.215	0.659	0.839
<i>Hansen p value</i>	0.869	0.859	0.893	0.991
<i>Dif Sargan p value</i>	0.219	0.138	0.235	0.444
<i>Num of Obs</i>	149	148	148	148

Notes: 5-year interval panel data is used for estimation; heteroskedasticity-consistent standard error is in parentheses; *lny<sub>i,t-1</sub>* is treated as pre-determined, *pop\_ngr* is treated as exogenous, and all other variables are treated as endogenous; \*\* and \* indicate that the coefficient is significantly different from zero at the 5 and 10 percent significance level respectively.

To address this problem, we use instead the volume of inward FDI as a proxy. The trend of FDI is reflected in Figure 4: the net inflow was under US\$10 billion before 1992, but reached US\$80 billion in 2006. In Model 12, we find a positive effect of FDI on provincial growth, i.e. for a one-billion-RMB rise in FDI, the growth rate of GDP per capita increases by 0.006 percentage points. In order to compare the growth impacts of domestic investment and FDI, it is necessary to convert the two variables into the same units. In Model 13, we include both the absolute volume of FDI (*fdi*) and domestic investment (*finv<sub>DOM</sub>*) and find that a unit of FDI is about 6 times more important for growth than a unit of domestic investment, i.e. a one-billion-RMB rise in domestic investment is associated with an increase in growth rate of GDP per capita by 0.001 percentage points. This is consistent with our

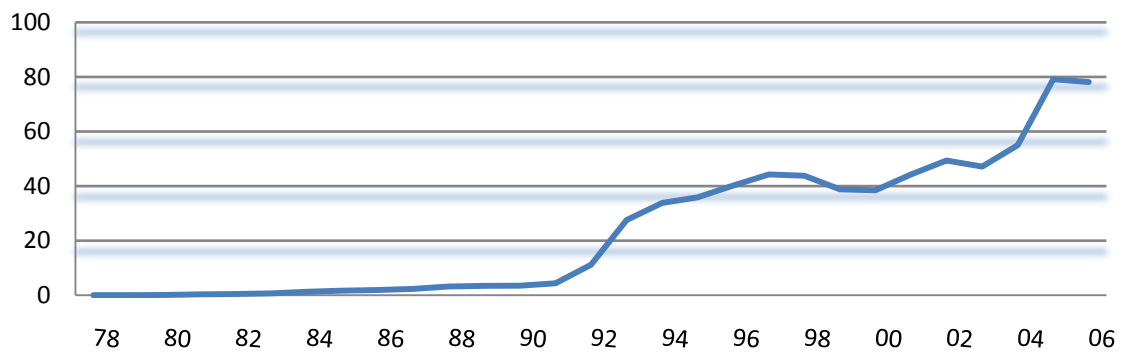
prediction that an investment has a greater impact on growth if it comes from abroad, given the role of FDI in export promotion, technology transfer, productivity improvement, and the inducement of structural change.

**Figure 3. Foreign Direct Investment Net Inflow (% of GDP)**



Data source: World Bank Development Indicators (April 2008).

**Figure 4. Foreign Direct Investment Net Inflow (Unit: Billion US\$)**



Data source: World Bank Development Indicators (April 2008).

We now classify investment in fixed assets according to ownership: investment spending by SOEs ( $finv_{SOE}/finv_{Total}$ ), collectively-owned enterprises ( $finv_{COL}/finv_{Total}$ ), private enterprises ( $finv_{PRIV}/finv_{Total}$ ) and domestic private enterprises ( $finv_{DOM\_PRIV}/finv_{Total}$ ) in Table 9. Since these variables may contain similar information as our proxy for the extent of economic reform ( $ind_{SOE}/ind_{TOTAL}$ , the SOEs' share of industrial output), we drop that term in these regressions to avoid the problem of multicollinearity. We find that the investment made by SOEs has a significantly negative effect on growth (Model 14), i.e. reducing the share of SOEs in total fixed investment by one percentage point is associated with an increase in GDP per capita growth of 0.08 percentage points. This result confirms the widespread perception that the efficiency of investment in the SOEs is far below that in the

**Table 9. Robustness Tests for Physical Capital Investment**  
(By ownership classification)

Regressor	Model 14	Model 15	Model 16	Model 17	Model 18	Model 19	Model 20	Model 21
<i>Constant</i>	0.329** (0.095)	0.317** (0.072)	0.232* (0.131)	0.236** (0.119)	0.416** (0.096)	0.376** (0.089)	0.264** (0.127)	0.271** (0.112)
<i>lny<sub>i,t-1</sub></i>	-0.032** (0.009)	-0.039** (0.008)	-0.028** (0.013)	-0.029** (0.012)	-0.047** (0.011)	-0.054** (0.009)	-0.042** (0.013)	-0.042** (0.012)
<i>fcf/GDP</i>					0.142** (0.036)	0.148** (0.058)	0.165** (0.047)	0.181** (0.049)
<i>finv<sub>SOE</sub>/finv<sub>Total</sub></i>	-0.083** (0.028)				-0.107** (0.026)			
<i>finv<sub>COL</sub>/finv<sub>Total</sub></i>		-0.021 (0.078)				-0.008 (0.062)		
<i>finv<sub>PRIV</sub>/finv<sub>Total</sub></i>			0.126* (0.071)				0.188* (0.111)	
<i>finv<sub>DOM_PRIV</sub>/finv<sub>Total</sub></i>				0.141** (0.070)				0.172** (0.077)
<i>stu<sub>SEC</sub>/pop</i>	1.294** (0.298)	1.817** (0.273)	1.529** (0.369)	1.565** (0.332)	0.970** (0.308)	1.536** (0.301)	1.325** (0.312)	1.246** (0.308)
<i>pop_ngr</i>	-4.179** (1.679)	-5.245** (1.256)	-5.749** (1.972)	-6.298** (2.254)	-3.786* (1.953)	-5.179** (1.592)	-5.386** (1.516)	-5.614** (1.549)
<i>export/GDP</i>	0.055** (0.027)	0.095** (0.029)	0.077** (0.032)	0.108** (0.035)	0.062** (0.031)	0.094** (0.023)	0.099** (0.026)	0.111** (0.028)
<i>dum_coastal</i>	-0.015 (0.011)	-0.022 (0.015)	-0.013 (0.014)	-0.025 (0.016)	0.001 (0.012)	0.002 (0.011)	-0.006 (0.012)	-0.001 (0.012)
<i>AR(2) p value</i>	0.636	0.501	0.317	0.171	0.941	0.991	0.607	0.483
<i>Hansen p value</i>	0.888	0.866	0.863	0.927	0.910	0.904	0.902	0.896
<i>Dif Sargan p value</i>	0.193	0.172	0.159	0.196	0.260	0.262	0.223	0.276
<i>Num of Obs</i>	149	149	149	149	148	148	148	148

Notes: 5-year interval panel data is used for estimation; heteroskedasticity-consistent standard error is in parentheses; *lny<sub>i,t-1</sub>* is treated as pre-determined, *pop\_ngr* is treated as exogenous, and all other variables are treated as endogenous; \*\* and \* indicate that the coefficient is significantly different from zero at the 5 and 10 percent significance level respectively.

non-state sector, so that the growth rate of GDP per capita should fall as the share of investment by SOEs increases (for instance, Brandt and Zhu, 2000). Consequently, the recent decline in SOEs' share of fixed investment is a positive development. The coefficient on investment by collective firms appears insignificant (Model 15). The collective economy consists of both township and village enterprises (TVEs) and urban collective firms. The former are generally said to have been dynamic, especially in the 1980s, whereas the latter are run by local governments and still suffer from the disincentives associated with soft budget constraints and principal-agent problems. Hence, the impact of collective firms on growth is ambiguous.

Private-sector investment affects growth positively, i.e. a one percentage point increase in this variable is associated with an increased growth rate of 0.13 percentage points (Model 16). Similar results hold for domestic private firms (Model 17). The expansion of private sector investment has a favourable impact, given the evidence that the average return on investment in the private sector is higher than that in the SOEs (Riedel et al., 2007, pp. 40-42). Our empirical evidence thus supports the view that private sector is the driving force in the Chinese economy. The results remain robust when we further control for the total investment in fixed assets in the growth regressions (Models 18-21). It is therefore a positive development that the centre of gravity of the economy has been shifting from the state to the private sector. However, a major imbalance in the allocation of resources between the public and private sectors remains. For example, bank loans constitute a major share of investment financing only for the relatively inefficient and unprofitable SOEs, while private firms are discriminated against by the formal financial system and rely predominantly on their 'own funds' to finance investment (Allen *et al.*, 2005; Guariglia *et al.*, 2008). Therefore, if China is to sustain its rapid economic growth, it is important for government to address the misallocation of financial resources so as to improve overall investment efficiency.

### 5.3 Human capital accumulation

Human capital accumulation can be treated analogously to physical capital accumulation, and can be incorporated accordingly into growth models and their empirical testing (for example, MRW, 1992). Whereas in this paper the assumed relationship is between changes in human capital and changes in output, it is also possible that the stock of human capital itself contributes to economic growth through the generation, absorption and dissemination of knowledge. Human capital is assumed to play such a role in some

endogenous models. For instance, according to Romer (1990), human capital is an input into the research and development activity which generates technological progress. However, research based on cross-country data has produced surprisingly mixed results on the effect of education on economic growth. For example, MRW (1992) found a significantly positive effect of secondary school enrolment rates on growth, whereas other researchers (Benhabib and Spiegel, 1994; Pritchett, 1999) claimed that increases in measured educational attainment (changes in the average years of schooling) are not related to output growth, especially in developing countries. We now examine the impact of human capital investment on economic growth in China.

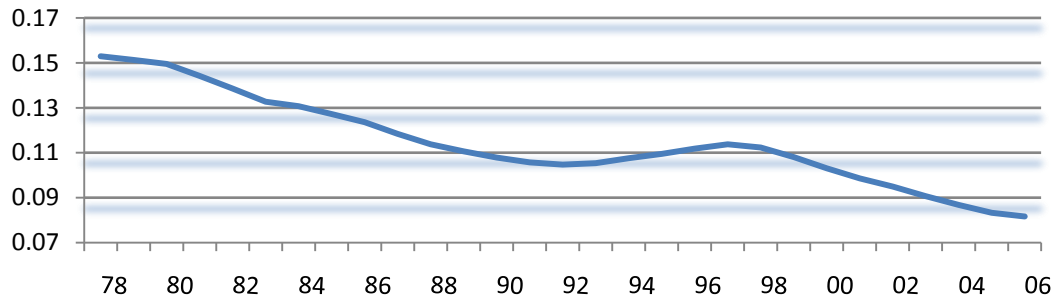
It is difficult to find a variable that adequately represents human capital. In reality, investment in human capital can take many forms, including formal and informal education, on-the-job training, health improvements and learning-by-doing. In most empirical studies human capital is normally proxied by average years of schooling, and increments to human capital either by changes in average years of schooling or by educational enrolment rates. Thus, the quality of education and other types of human capital investment are largely ignored. Given the data availability, we adopt various educational enrolment rates to measure certain aspects of human capital in China. Although enrolment is normally measured as a proportion of the relevant age group, enrolment as a proportion of the total population is a better guide to the increase in human capital. School enrolment rates may conflate human capital *stock* and *accumulation* effects and can be a poor proxy for either (see, for instance, Gemmell, 1996 and Temple, 1999). However, this is the information that is available to us annually at the province level.

Figure 5a illustrates a declining trend of primary school enrolment as a proportion of total population. Since primary education is universal in China, the decline simply reflects the falling numbers of children as a result of the one-child-family policy, introduced in the late 1970s. Figure 5b shows that the ratio of secondary school enrolment to total population was declining in the 1980s but trended upwards thereafter as participation soared. Figure 5c reflects the fact that the ratio of enrolment in higher education to total population was very low until 1997 but then shot up as a result of policy reversal despite the fact that the number of 18-year olds was now falling. To deal with the endogeneity problem of these variables, levels of human capital investment variables lagged by 10-year and 15-year periods are used as instruments in the first-differences equations, and first-differenced human capital

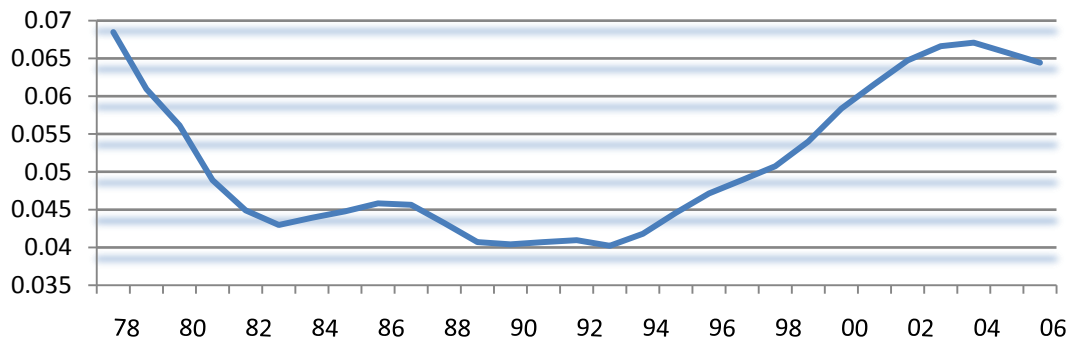


investment variables lagged by 5-year period are used as additional instruments for the levels equations in the system GMM estimation.

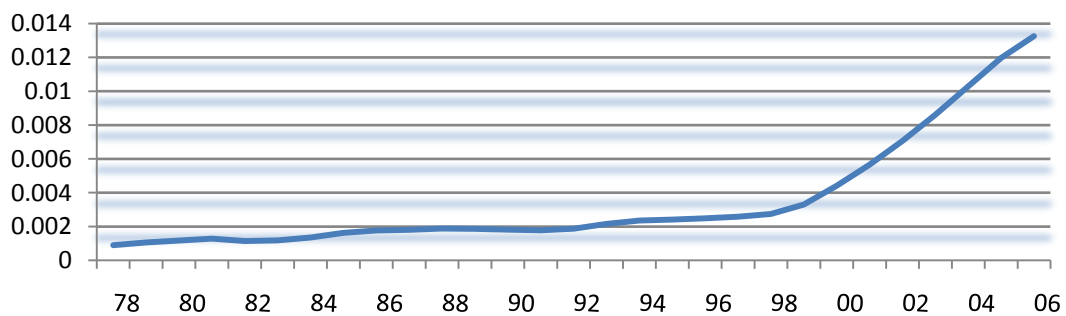
**Figure 5a. Ratio of Primary School Enrolment to Total Population**



**Figure 5b. Ratio of Secondary School Enrolment to Total Population**



**Figure 5c. Ratio of Higher Education Enrolment to Total Population**



Data source: National Bureau of Statistics of China.

In Table 10, our human capital measurements consist of primary school enrolment rates ( $stu_{PRIM}/pop$ ), secondary school enrolment rates ( $stu_{SEC}/pop$ ), regular secondary school enrolment rates ( $stu_{REG\_SEC}/pop$ ), number of students enrolled in higher education over population ( $stu_{HIGH}/pop$ ), university and college enrolment as a percentage of population ( $stu_{UNI \& COL}/pop$ ), the number of students enrolled in higher education over the number of

**Table 10. Robustness Tests for Human Capital Investment**

Regressor	Model 22	Model 23	Model 24	Model 25	Model 26	Model 27	Model 28	Model 29	Model 30	Model 31	Model 32	Model 33
<i>Constant</i>	0.623** (0.102)	0.403** (0.078)	0.393** (0.085)	0.558** (0.071)	0.479** (0.072)	0.565** (0.083)	0.418** (0.072)	0.454** (0.075)	0.448** (0.076)	0.449** (0.071)	0.546** (0.089)	0.508** (0.074)
<i>lny<sub>i,t-1</sub></i>	-0.069** (0.012)	-0.056** (0.008)	-0.054** (0.009)	-0.067** (0.009)	-0.058** (0.009)	-0.065** (0.010)	-0.060** (0.008)	-0.062** (0.009)	-0.061** (0.009)	-0.062** (0.009)	-0.062** (0.010)	-0.069** (0.009)
<i>fcf/GDP</i>	0.201** (0.044)	0.157** (0.049)	0.168** (0.044)	0.169** (0.049)	0.201** (0.051)	0.177** (0.047)	0.139** (0.047)	0.129** (0.041)	0.127** (0.038)	0.135** (0.036)	0.160** (0.050)	0.164** (0.035)
<i>stu<sub>PRIM</sub>/pop</i>	-0.243 (0.248)							0.060 (0.278)	0.009 (0.271)	-0.025 (0.280)	0.080 (0.299)	-0.303 (0.194)
<i>stu<sub>SEC</sub>/pop</i>		1.334** (0.298)						0.956** (0.360)		1.156** (0.349)		
<i>stu<sub>REG_SEC</sub>/pop</i>			1.322** (0.296)						1.118** (0.359)			
<i>stu<sub>HIGH</sub>/pop</i>				3.862** (0.696)				3.393** (0.974)	3.528** (0.918)			
<i>stu<sub>UNI &amp; COL</sub>/pop</i>					3.077** (0.558)					2.771** (0.863)		
<i>stu<sub>HIGH</sub>/stu<sub>REG_SEC</sub></i>						0.099** (0.029)					0.099** (0.045)	
<i>stu<sub>SEC &amp; HIGH</sub>/pop</i>							1.526** (0.273)					1.391** (0.290)
<i>pop_ngr</i>	-6.624** (2.169)	-4.217** (1.329)	-4.337** (1.278)	-2.032 (1.727)	-1.293 (1.744)	-4.597** (2.085)	-2.846** (1.450)	-1.708 (1.226)	-1.438 (1.189)	-1.565 (1.134)	-5.118** (1.808)	-2.433* (1.319)
<i>export/GDP</i>	0.086** (0.024)	0.079** (0.019)	0.076** (0.020)	0.033 (0.027)	0.021 (0.025)	0.063** (0.023)	0.067** (0.017)	0.047** (0.017)	0.045** (0.018)	0.037** (0.017)	0.068** (0.023)	0.068** (0.013)
<i>ind<sub>SOE</sub>/ind<sub>TOTAL</sub></i>	-0.063 (0.018)	-0.037** (0.012)	-0.040** (0.012)	-0.107** (0.017)	-0.106** (0.019)	-0.093** (0.014)	-0.047** (0.011)	-0.073** (0.016)	-0.075** (0.017)	-0.073** (0.018)	-0.087** (0.015)	-0.052** (0.013)
<i>dum_coastal</i>	0.008 (0.011)	0.003 (0.012)	0.012 (0.011)	0.012 (0.015)	0.012 (0.014)	0.004 (0.011)	0.005 (0.012)	0.009 (0.009)	0.014* (0.008)	0.011 (0.008)	0.001 (0.011)	0.012 (0.008)
<i>AR(2) p value</i>	0.622	0.955	0.918	0.597	0.679	0.960	0.606	0.437	0.367	0.401	0.906	0.549
<i>Hansen p value</i>	0.889	0.904	0.892	0.938	0.864	0.920	0.915	0.999	0.999	0.999	0.995	0.994
<i>Dif Sargan p value</i>	0.166	0.276	0.297	0.147	0.231	0.113	0.211	0.798	0.794	0.796	0.471	0.551
<i>Num of Obs</i>	149	148	149	149	147	149	148	148	149	146	149	148

students enrolled in regular secondary education ( $stu_{HIGH}/stu_{REG\_SEC}$ ), and students enrolled in both secondary and higher education over population ( $stu_{SEC \& HIGH}/pop$ )<sup>12</sup>.

In line with Chen and Feng (2000) in their cross-province study, we find that the coefficient of primary school enrolment rates is insignificant in the growth regression (Model 22). This is to be expected because primary education is mandatory and the variation of primary enrolment across provinces in China is very low. According to official data, China had achieved virtually universal enrolment in primary education prior to economic reform. In line with both the cross-country evidence (for instance, Barro, 1991; MRW, 1992) and the cross-province evidence (for instance, Chen and Fleisher, 1996; Démurger, 2001), we find that secondary/regular secondary school enrolment rates have positive and significant impacts on output growth, i.e. a one percentage point increase in secondary/regular secondary school enrolment rates is associated with a higher growth rate of GDP per capita by 1.3 percentage points (Models 23 and 24). Higher education enrolment rates/university and college enrolment rates have bigger positive effects, i.e. a one percentage point rise in these variables leads to higher GDP per capita growth by 3-4 percentage points, holding other conditions constant (Models 25 and 26). Our coefficients on higher education is consistent with that of Chen and Feng (2000) based on the cross-sectional OLS, but, since it controls for endogeneity, is more likely to reflect causality. Following Yao (2006), we use the ratio of number of students enrolled in higher education over the those enrolled in regular secondary education to reflect the propensity for secondary school students to be enrolled in higher education (Model 27). It too is found to have a positive and significant impact on growth, i.e. an increase of one percentage point in this ratio raises growth by 0.1 percentage points. Lastly, we examine the impact of students enrolled in both secondary and higher education over population (Model 28), and find that a one percentage point rise in this enrolment rate is associated with a 1.5 percentage points rise in the growth rate of GDP per capita. Our results remain robust when all, and subsets of, human capital variables enter the regression simultaneously (Models 29-33). The important contribution to growth made by higher education might be explained by the remarkable relative neglect of higher education, and consequent scarcity of tertiary graduates, throughout the first two decades of economic reform. This result is consistent with Chi (2008), who used educational attainment to measure

---

<sup>12</sup> Secondary education includes junior and senior secondary schools, specialized secondary schools, vocational secondary schools, and technical training schools. Regular secondary schools include merely junior and senior secondary schools. Higher education includes universities and colleges as well as self-taught programmes for undergraduates.

human capital and found evidence that tertiary education has a positive and bigger impact on both GDP growth and fixed investment than primary and secondary education. He therefore argued that China's production function exhibited some degree of capital-skill complementarity.

To test the robustness of our human capital results, census information on the percentage of population aged 6 and above with primary, secondary, or tertiary educational attainment are adopted. These data are proxies for the stock of human capital but they are only available for the census years 1982, 1990, 1995 and 2000. We interpolate the census data to derive the observations in the years required by the analysis. For these reasons, inaccuracies and measurement errors are to be expected when these alternative human capital variables are deployed. Nevertheless, we find that the share of population with junior secondary education has a positive and significant effect on provincial growth, and that when changes in share are included in the growth equation, an increase in the relative stock of people with both junior secondary and tertiary education raises growth significantly.

#### *5.4 Illustrative counterfactual predictions*

We return to the underlying question: can cross-province growth regressions help us to understand why China as a whole has grown so fast? We attempt to answer the question by means of counterfactual predictions in Tables 11 and 12. The methodology is to predict growth rates by changing mean values of key variables based on model estimation. Because these simulations contain the questionable assumption that a change in one variable would not alter the other variables in the equation, they can merely illustrate the rough orders of magnitude of the impact on the growth rate. However, insofar as a fall in one variable (say, human capital formation) induces fall in another variable (say, physical capital formation), the simulations would understate the impact on growth.

The average value of fixed investment in relation to GDP over the full sample period was 34.3 percent. If instead it had been 10 percentage points lower (24.3 percent), the coefficient of the baseline model in Table 5 implies that China's growth of GDP per capita would have fallen by 1.5 percentage points, from 8.0 to 6.5 percent, holding other variables constant. Similarly, secondary school enrolments averaged 5.8 percent of total population. If it had been 2 percentage points lower, controlling for other variables, the growth rate of GDP per capita would have declined to 5.4 percent. Had both the physical and human capital

variables been reduced in this way, China's per capita growth rate would have fallen to 3.9 percent, holding other controls constant.

**Table 11. Counterfactual predictions of growth rate of GDP per capita (Baseline Model)**

Predicted growth rates of GDP per capita	<i>stu<sub>SEC</sub>/pop</i>			
	Mean (5.79)	Reduce by 1 unit (4.79)	Reduce by 2 units (3.79)	Reduce to the mean of LDCs (2.26)
<i>fcf/GDP</i>				
Mean (34.27)	8.05	6.72	5.39	3.36
Reduce by 1 unit (33.27)	7.89	6.57	5.24	3.21
Reduce by 5 units (29.27)	7.29	5.97	4.64	2.61
Reduce by 10 units (24.27)	6.54	5.22	3.89	1.86
Reduce to the mean of LDCs (17.76)	5.56	4.23	2.91	0.88

Note: Units in this table are percentage points; LDCs refer to least developed countries.

China was a low-income country at the start of economic reform<sup>13</sup>. The mean values of the fixed capital formation and secondary school enrolment ratios of all least developed countries (United Nations 2008 classification) over the entire period are 17.8 percent and 2.3 percent respectively<sup>14</sup>. Plugging these values into the baseline model, we find that China's growth rate of GDP per capita would have been only 0.9 percent per annum, i.e. roughly the same as that for the least developed countries (1.0 percent). On these assumptions, China's status as a growth outlier would have been lost.

Consider the effects of changes in the composition of physical investment in Table 12. The ratio of innovation investment averaged 6.9 percent of total GDP in China over the whole sample period. Our estimated coefficient shows that a reduction in this variable to half of its mean would have resulted in a 1.1 percentage points lower growth rate of GDP per capita. Combining the effect of reducing the secondary school enrolment ratio to half, the growth rate would have fallen to 4.3 percent holding other conditions the same. The mean value of provincial FDI over the period 1978-2006 was 4.3 billion RMB. Had the figure been held down at half of the mean, the growth rate would have been 1.3 percentage points lower, at 6.8 percent per annum. The joint effects of reducing both FDI and the secondary school enrolment rate by half would have been a growth rate of 2.8 percent, *ceteris paribus*. Private fixed investment averaged 13.2 percent of total GDP in China over the full period. If it had

<sup>13</sup> According to the *World Development Report* (1981) published by World Bank, China's GNP per capita was \$260 in 1979, ranking at the 22nd poorest country.

<sup>14</sup> The ratio of fixed capital formation to GDP and total population data for least developed countries come from *World Development Indicators* (April 2008 edition); secondary school enrolment figures come from UNESCO.

remained at its 1978 level (5.1 percent), growth would have been 1.2 percentage points lower, at 6.9 percent per annum. Combining the effect of reducing secondary school enrolment rate to half, China's growth rate would have been down to 3.4 percent.

**Table 12. Counterfactual predictions of growth rate of GDP per capita (Other models)**

Predicted growth rates of GDP per capita (Unit: <i>pps</i> )		<i>stu<sub>SEC</sub>/pop</i>			
		Mean (5.79 <i>pps</i> )	Reduce by 1 <i>pp</i>	Reduce by 2 <i>pps</i>	Reduce to half of the mean
Model 8 (Table 7)					
<i>finv<sub>INNO</sub>/GDP</i>	Mean (6.85 <i>pps</i> )	8.02	7.09	6.17	5.35
	Reduce by 1 <i>pp</i>	7.71	6.78	5.86	5.03
	Reduce by 2 <i>pps</i>	7.39	6.47	5.54	4.71
	Reduce by 3 <i>pps</i>	7.07	6.15	5.22	4.39
	Reduce to half of the mean	6.94	6.02	5.09	4.26
Model 12 (Table 8)					
<i>fdi</i>	Mean (4.26 billion RMB)	8.05	6.63	5.24	3.99
	Reduce by 1 billion RMB	7.46	6.07	4.67	3.43
	Reduce by 2 billion RMB	6.87	5.48	4.08	2.84
	Reduce to half of the mean	6.79	5.40	4.00	2.76
	Reduce by 3 billion RMB	6.28	4.89	3.49	2.25
Model 17 (Table 9)					
<i>finv<sub>DOM_PRIV</sub>/</i> <i>finv<sub>Total</sub></i>	Mean (13.24 <i>pps</i> )	8.05	7.45	5.88	4.48
	Reduce by 1 <i>pp</i>	7.91	7.33	5.77	4.37
	Reduce by 3 <i>pps</i>	7.63	7.05	5.49	4.09
	Reduce to half of the mean	7.12	6.54	4.97	3.58
	Reduce to 1978 Mean (5.06 <i>pps</i> )	6.89	6.32	4.75	3.36
Model 29 (Table 10)					
<i>stu<sub>HIGH</sub>/pop</i>	Mean (0.44 <i>pps</i> )	8.08	7.12	6.16	5.60
	Reduce by 0.1 <i>pps</i>	7.68	6.72	5.77	5.21
	Reduce by 0.2 <i>pps</i>	7.34	6.38	5.43	4.87
	Reduce to half of the mean	7.27	6.31	5.36	4.79
	Reduce by 0.3 <i>pps</i>	7.00	6.04	5.09	4.53

Note: *pp(s)* refers to percentage point(s).

How important was post-primary education to growth? Secondary and higher education enrolments averaged 5.8 and 0.4 percent of population respectively over the full period. Had both the secondary and tertiary ratios been half of their mean values instead, the growth rate would have fallen by 3.3 percentage points, to 4.8 percent per annum.

The conclusion to be drawn from these simple counterfactual exercises is that both the quantity and composition of physical and human capital formation are potentially important to China's rate of economic growth. A reduction of these inputs to levels commonly found in

the countries that, unlike China, remained least developed, could have reduced the growth rate to that of those countries.

## **6. How was rapid capital accumulation possible?**

In this section we pose the questions that flow logically from our results: how was China's rapid physical, and also human, capital accumulation possible? As each of these questions deserves a separate study on its own, we merely provide an outline sketch of the possible components of an answer.

With the physical capital stock being well below its equilibrium level, there were powerful profit incentives to invest in many sectors of the economy. However, China's rapid capital accumulation would not have been possible without high domestic saving. What made such a high national saving rate possible? Households have become a principal source of saving since the start of economic reform. Household saving increased from 5 percent of income in 1978 to a peak of 34 percent in 1994, and remained above 24 percent in 2000 (Modigliani and Cao, 2004). Over the reform period, China's GDP per capita rose nearly tenfold, from \$165 in 1978 to \$1598 in 2006<sup>15</sup>. With higher income, households chose to save a higher proportion of their income. Riedel *et al.* (2007) claimed that a virtuous circle has operated in China, whereby higher income leads to more saving, hence permitting more investment and faster growth, which in turn leads to higher income. By contrast, Modigliani and Cao (2004) attributed China's high household saving rate to the accelerating economic growth and demographic change, rather than income per capita. Within the framework of the Life Cycle Hypothesis, they argued that household wealth would bear a constant ratio to income, so that faster growth would involve a higher saving rate. They also argued that the birth-control policy undermined the traditional role of children as old-age support and, in the absence of a well-developed social security system, this encouraged households to provide for retirement through saving. Another contribution might come from the increasing scope for household business and housing investment. Given credit constraints, households responded to the new opportunities by saving for investment (Naughton, 2006).

High enterprise and government saving also contributes to China's high saving rate. On the one hand, the imperfect capital market makes firms, especially private firms, rely mainly on their own funds (i.e. retained earnings) to finance investment. This provides them with a

---

<sup>15</sup> Data come from *World Development Indicators* (April 2008 Edition); GDP per capita is gross domestic product divided by midyear population and is in constant 2000 US dollar.

strong incentive to save. On the other hand, the profitability of firms has increased significantly since enterprise reform began in earnest in the mid-1990s. Moreover, as the government has not sought dividends from SOEs, their rising profits are either reinvested or sit in their savings accounts. Government saving has also been high since 1978 as a result of a policy favouring government-financed investment over government consumption (Kuijs, 2005). The Chinese government was willing and able to take a long run view because it expected to remain in power for many years, it was not subject to democratic pressures for 'jam today', and the rapid growth of household incomes provided a shield against social discontent.

Educational enrolment rates and their growth over time, so important for economic growth, have in turn to be explained: both demand and supply factors played a role. In 1988, early in the process of urban reform, the wage premium of upper secondary education over primary education for urban residents was very low, at 4 percent, but with labour market reform it rose to 26 percent in 1995 and to 33 percent in 2002. The urban wage premium of higher education over upper secondary education rose from only 5 percent in 1988 to 17 percent in 1995 and to 42 percent in 2002 (Knight and Song, 2005, table 3.2; Knight and Song, 2008, table 2). Although the returns to education remained low in farming, they were higher in the non-farm activities that were opening up to rural workers, and education also improved their access to these higher-income activities (Knight *et al.*, 2008, tables 15-17, 21). As opportunities for local non-farm employment and rural-urban migrants grew, education became an increasingly important means of raising the incomes of rural workers (Knight and Song, 2005, table 8.8).

These labour market reforms and structural changes raised the private demand for education. This demand was in any case strong on account of the respect and status commonly accorded to education, which had been embedded by Chinese history. Thus, for instance, when a sharp change in higher education policy took place in the late 1990s, the remarkable increase in the supply of college places was fully met by the pent-up demand, with enrolment rising over four-fold between 1998 and 2005<sup>16</sup>.

## 7. Conclusion

In this paper, we attempt to answer a broad question: why has China grown so fast? Despite the diversity in growth among provinces, the economies of all provinces grew rapidly

---

<sup>16</sup> National Bureau of Statistics, *China Statistical Yearbook 2006*, table 21-6.



by international standards over the period of economic reform. Model uncertainty is a fundamental problem for empirical research in economic growth because economic theories are not precise enough to pinpoint the exact determinants of growth. We adopted two recently developed approaches to model selection, BMA and GETS, to consider a wide range of candidate predictors of economic growth in China. The first-stage model selection results identified a role for conditional convergence, physical and human capital formation, population growth, degree of openness, and institutional change in determining output growth across China's provinces. Using the basic equation, we proceeded to examine the growth impact of physical and human capital investment in some detail using panel data system GMM.

Among results of the baseline model (Table 5), three major findings form the basis of our story: there is conditional convergence among provinces, and both physical and human capital accumulation promote economic growth. The first result is consistent with the neoclassical model of economic growth, i.e. the hypothesis that each province is converging towards its equilibrium steady state. Initial income per capita is used in our regression to control for the transitional dynamics induced by factor accumulation. It might, however, have other explanations, e.g. that convergence reflects the effects of fiscal transfer from the central government to poor provinces and minority areas.

The last two results are inconsistent with the steady state version of neoclassical growth theory, with its prediction that only technological progress governs the equilibrium growth rate. They are consistent, however, with movement towards equilibrium growth. Such transitional movement is indeed to be expected given the extreme disequilibrium of the Chinese economy at the start of economic reform and the slow pace of conditional convergence suggested by the coefficient on initial income per capita. An alternative interpretation of the positive effects of investment, drawing on endogenous growth theory, is that it has generated not only capital accumulation but also technological progress.

Our more detailed investigation of the effects of physical and human capital accumulation was intended to throw further light on the mechanisms at work. It was fixed investment that made the contribution to growth: inventory investment, reaching its peak in 1990, was a drain on the economy (Table 6). The greatest contribution was made by investment identified as 'investment in innovation' as opposed to 'investment in capital construction', and 'investment in other fixed assets', such as real estate, made no contribution (Table 7). This result suggests that physical investment makes the greatest contribution to

growth when it is most closely bound up with technological progress. Consistent with this argument, we expected that a unit of foreign direct investment would have a greater effect on growth than a unit of domestic investment. Our evidence supports this hypothesis after we solve a specification problem (Table 8). Breaking down physical investment by ownership, we found the contribution of investment by SOEs to be negative, that of collective enterprises to be negligible, and only that of private enterprises to be positive (Table 9). Thus, the reform process that unleashed a private sector was important for growth, and the distorted financial system which continued to favour the state-owned enterprises held back growth.

Whereas the primary school enrolment rate has no effect on economic growth, both the secondary school and the higher education enrolment rates had a positive effect, the latter more than the former (Table 10). Indeed the coefficient on higher education enrolment in relation to population implied that a rise of one percentage point would raise the growth rate of GDP per capita by 3.8 percentage points. This sensitivity might be explained by the neglect of higher education until the late 1990s: in 1997 higher education enrolment was still only 5 percent of the relevant age group. However, despite our use of GMM estimation and instrumenting of the human capital variables using lags, some caution is required in interpreting the estimated effect as a causal one.

To address our question: can cross-province growth regressions explain why China as a whole has grown so fast, various counterfactual exercises were conducted (Tables 11 and 12). We found that a significant reduction of these inputs could have reduced China's growth rate dramatically, indicating the important role of physical and human capital formation in determining China's remarkable rate of economic growth. The factors which made rapid physical and human capital accumulation possible were discussed briefly. Incentives for saving have been strong for households, enterprises and governments over the reform period, and labour market reform, by increasing the wage premia on education, has produced rapid growth in the demand for education, to which government has responded by increasing the supply.

In this paper we have used and extended the baseline growth equation to examine the contribution that factor accumulation has made to China's economic growth. These are the proximate determinants of growth. Underlying them, however, are the other influences on growth that enter our baseline equation. These other determinants are explored in a companion paper (Ding and Knight, 2008b).

## References

- Abramovitz, M. (1986). Catching Up, Forging Ahead, and Falling Behind. *Journal of Economic History* , 46 (2), 385-406.
- Allen, F., Qian, J. and M. Qian. (2005). Law, Finance, and Economic Growth in China. *Journal of Financial Economics* , 77, 57-116.
- Arellano, M. and O. Bover. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics* , 68, 29-52.
- Arellano, M. and S. Bond. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and An Application to Employment Equations. *Review of Economic Studies* , 58, 277-97.
- Bao, S. M., Chang, G. H., Sachs, J. D. and W. T. Woo. (2002). Geographic Factors and China's Regional Development under Market Reforms, 1978-1998. *China Economic Review* , 13, 89-111.
- Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics* , 106 (2), 407-43.
- Barro, R. J. and X. Sala-i-Martin. (2004). *Economic Growth* (second ed.). Cambridge, MA: The MIT Press.
- Beenstock, M. and D. Felsenstein. (2008). Regional Heterogeneity, Conditional Convergence and Regional Inequality. *Regional Studies*, 42(4), 475-88.
- Benhabib, J. and M. M. Spiegel. (1994). The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data. *Journal of Monetary Economics* , 34 (2), 143-73.
- Blomström, M., Lipsey, R. and M. Zejan. (1996). Is Fixed Investment the Key to Economic Growth? *The Quarterly Journal of Economics* , 111 (1), 269-76.
- Blundell, R. and S. Bond. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics* , 87 (1), 115-43.
- Bond, S., Hoeffler, A. and J. Temple. (2001). GMM Estimation of Empirical Growth Models. *CEPR discussion paper no. 3048* .
- Bond, S., Leblebicioglu, A. and F. Schiantarelli. (2007). Capital Accumulation and Growth: A New Look at the Empirical Evidence. *Working Paper, University of Oxford*.
- Bowsher, C. G. (2002). On Testing Overidentifying Restrictions in Dynamic Panel Data Models. *Economics Letters* , 77, 211-20.
- Brandt, L. and X. D. Zhu. (2000). Redistribution in a Decentralized Economy: Growth and Inflation in China under Reform. *Journal of Political Economy* , 108 (2), 422-51.

- Brun, J. F., Combes, J. L. and M. F. Renard. (2002). Are There Spillover Effects Between Coastal and Noncoastal Regions in China? *China Economic Review* , 13, 161-69.
- Cai, F., Wang, D. W. and Y. Du. (2002). Regional Disparity and Economic Growth in China: The Impact of Labour Market Distortions. *China Economic Review* , 13, 197-212.
- Chatfield, C. (1995). Model Uncertainty, Data Mining, and Statistical Inference. *Journal of the Royal Statistical Society* , 158, 419-44.
- Chen B. Z. and Y. Feng. (2000). Determinants of Economic Growth in China: Private Enterprise, Education and Openness. *China Economic Review* , 11, 1-15.
- Chen, J. and B. M. Fleisher. (1996). Regional Income Inequality and Economic Growth in China. *Journal of Comparative Economics* , 22, 141-64.
- Chi, W. (2008). The Role of Human Capital in China's Economic Development: Review and New Evidence. *China Economic Review*, 19, 421-36.
- Commission on Growth and Development. (2008). *The Growth Report: Strategies for Sustained Growth and Inclusive Development (Conference edition)*.
- Démurger, S. (2001). Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China. *Journal of Comparative Economics* , 29, 95-117.
- Ding, S. and J. Knight. (2008a). Can the Augmented Solow Model Explain China's Economic Growth? A Cross-Country Panel Data Analysis. *Department of Economics Discussion Paper No. 380, University of Oxford*.
- Ding, S. and J. Knight. (2008b). Why has China Grown so Fast? The Role of Structural Change.
- Doornik, J. A. and D. F. Hendry. (2007). *Pcgive 12 Voume I: Empirical Econometric Modelling*. London: Timberlake Consultants Press.
- Easterly, W. and R. Levine. (2001). What Have We Learned from a Decade of Empirical Research on Growth? It's not Factor Accumulation: Stylized Facts and Growth Models. *World Bank Economic Review* , 15, 177-219.
- Fernández, C., Ley, E. and M. F. J. Steel. (2001). Model Uncertainty in Cross-Country Growth Regressions. *Journal of Applied Econometrics* , 16 (5), 563-76.
- Gemmell, N. (1996). Evaluating The Impacts of Human Capital Stocks and Accumulation on Economic Growth: Some New Evidence. *Oxford Bulletin of Economics and Statistics* , 58 (1), 9-28.
- Grossman, G. and E. Helpman. (1995). *Innovatoin and Growth in the Global Economy*. Cambridge, MA: MIT Press.

- Guariglia, A. and S. Poncet. (2006). Could Financial Distortions be No Impediment to Economic Growth After All? Evidence from China. *The University of Nottingham China and the World Economy Research Paper Series No. 36* .
- Guariglia, A., Liu, X. and L. Song. (2008). Is the Growth of Chinese Firms Constrained by Internal Finance? *Mimeograph, University of Nottingham* .
- Hao, C. (2006). Development of Financial Intermediation and Economic Growth: The Chinese Experience. *China Economic Review* , 17, 347-62.
- Hendry, D. F. and H. M. Krolzig. (2004). We Ran One Regression. *Oxford Bulletin of Economics and Statistics* , 66 (5), 799-810.
- Hoover, K. D. and S. J. Perez. (2004). Truth and Robustness in Cross-Country Growth Regressions. *Oxford Bulletin of Economics and Statistics* , 66 (5), 765-98.
- Huang, Y. F. (2005). What Determines Financial Development. *Department of Economics Discussion Paper No. 05/580, University of Bristol* .
- Islam, N. (1995). Growth Empirics: A Panel Data Approach. *The Quarterly Journal of Economics* , 110 (4), 1127-70.
- Javorcik, B. (2008). Can Survey Evidence Shed Light on Spillovers from Foreign Direct Investment? *The World Bank Research Observer* , 23 (2), 139-159.
- Jones, D. C., Li, C. and A. L. Owen. (2003). Growth and Regional Inequality in China during the Reform Era. *China Economic Review* , 14, 186-200.
- Knight, J. and S. Li. (1999). Fiscal Decentralization: Incentives, Redistribution and Reform in China. *Oxford Development Studies* , 27 (1), 5-32.
- Knight, J. and L. Song. (2005). *Towards a Labour Market in China*. Oxford: Oxford University Press.
- Knight, J., Li, S. and Q. Deng. (2008). Education and the Poverty Trap in Rural China. *University of Oxford Centre for the Study of African Economics Working Paper Series No. 2008-02*.
- Knight, J. and L. Song. (2008). China's Emerging Urban Wage Structure, 1995-2002. In B. Gustafsson, S. Li, & T. Sicular, *Inequality and Public Policy in China* (pp. 221-242). Cambridge: Cambridge University Press.
- Kuijs, L. (2005). Investment and Saving in China. *World Bank Policy Research Working Paper No. 3633*.
- Leamer, E. E. (1978). *Specification Searches*. New York: John Wiley & Sons.
- Leamer, E. E. (1983). Let's Take the Con Out of Econometrics. *American Economic Review* , 73 (1), 31-43.

- Leamer, E. E. (1985). Sensitivity Analysis Would Help. *American Economic Review* , 75 (3), 308-13.
- Levine, R. and D. Renelt. (1992). A Sensitivity Analysis of Cross-Country Growth Regressions. *American Economic Review* , 82 (4), 942-63.
- Li, H., Liu, Z. N. and I. Rebelo. (1998). Testing the Neoclassical Theory of Economic Growth: Evidence from Chinese Provinces. *Economics of Planning* , 31, 117-32.
- Lucas, R. E. (1988). On the Mechanics of Economic Development. *Journal of Monetary Economics* , 22, 3-42.
- Lucas, R. E. (2002). *Lectures on Economic Growth*. MA: Harvard University Press.
- Malik, A. and J. R. W. Temple. (2005). The Geography of Output Volatility. *CSAE Working Paper No. 2005-07, University of Oxford* .
- Mankiw, N. G., Romer, D. and D. N. Weil. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics* , 107 (2), 407-37.
- Modigliani, F. and S. L. Cao. (2004). The Chinese Saving Puzzle and the Life-Cycle Hypothesis. *Journal of Economic Literature*, 42(1), 145-170.
- Naughton, B. (2006). *The Chinese Economy: Transitions and Growth*. Cambridge, MA: The MIT Press.
- Nickell, S. (1981). Biases in Dynamic Models with Fixed Effects. *Econometrica* , 49 (6), 1417-1426.
- Owen, P. D. (2003). General-to-Specific Modelling Using Pcgets. *Journal of Economic Surveys* , 17 (4), 609-27.
- Pritchett, L. (1999). Where Has All the Education Gone? *The World Bank Economic Review* , 15 (3), 367-91.
- Raftery, A. E. (1995). Bayesian Model Selection in Social Research. In P. V. Marsden, *Sociological Methodology*. Cambridge: Blackwells.
- Raiser, M. (1998). Subsidising Inequality: Economic Reforms, Fiscal Transfers and Convergence Across Chinese Provinces. *The Journal of Development Studies* , 34 (3), 1-26.
- Ravallion, M. and S. H. Chen. (2007). China's (Uneven) Progress Against Poverty. *Journal of Development Economics* , 82, 1-42.
- Riedel, J., Jin, J. and J. Gao. (2007). *How China Grows: Investment, Finance and Reform*. Princeton: Princeton University Press.
- Romer, P. M. (1990). Are Nonconvexities Important for Understanding Growth? *American Economic Review*, 80(2), 97-103.

- Sala-i-Martin, X. X. (1997). I Just Ran Two Million Regressions. *American Economic Review* , 87 (2), 178-83.
- Sala-i-Martin, X., Doppelhofer, G. and R. Miller. (2004). Determinants of Long-Run Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach. *American Economic Review* , 94 (4), 813-35.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics* , 6, 461-64.
- Scott, M. F. (1989). *A New View of Economic Growth*. Oxford: Clarendon Press.
- Scott, M. F. (1993). Explaining Economic Growth. *American Economic Review, Papers and Proceedings of the Hundred and Fifth Annual Meeting of the American Economic Association* , 83 (2), 421-25 .
- Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics* , 70, 65-94.
- Temple, J. (1999). The New Growth Evidence. *Journal of Economic Literature* , 37 (1), 112-56.
- Wei, S. J. (1995). Open Door Policy and China's Rapid Growth: Evidence from City-Level Data. In T. Ito, & A. Krueger, *Growth Theories in Light of the East Asian Experience*. University of Chicago Press.
- Windmeijer, F. (2005). A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators. *Journal of Econometrics* , 126, 25-51.
- Yao, S. (2006). On Economic Growth, FDI and Exports in China. *Applied Economics* , 38, 339-51.
- Zhang, W. (2001). Rethinking Regional Disparity in China. *Economics of Planning* , 34, 113-38.

## Appendix 1 : Model Selection Procedures

### Bayesian model averaging (BMA)

The following brief discussion of the theory behind BMA draws heavily on Raftery (1995), Sala-i-Martin *et al.* (2004), Malik and Temple (2005) and Huang (2005).

A natural way to think about model uncertainty is to admit that we do not know which model is 'true' and instead, attach probabilities to different possible models. BMA treats parameters and models as random variables and summarizes the uncertainty about the model in terms of a probability distribution over the space of all possible models.

Suppose we want to make inference about an unknown quantity of interest (such as a parameter),  $\Delta$ , given data  $D$ . There are a large number of possible statistical models,  $M_1, \dots, M_K$  for the data space. If we consider only linear regression models but are unsure about which  $p$  possible regressors to include, there could be as many as  $2^p$  models considered. Bayes' rule and basic probability theory suggest that the posterior distribution of the parameters is the weighted average of all the possible conditional posterior densities with the weights given by the posterior probabilities of each of the possible models. Then the posterior distribution of  $\Delta$  given data  $D$  is

$$P(\Delta | D) = \sum_{k=1}^K P(\Delta | D, M_k) P(M_k | D) \quad , \quad (\text{A.1})$$

where  $P(\Delta | D, M_k)$  is the posterior distribution of  $\Delta$  given the model  $M_k$ , and  $P(M_k | D)$  is the posterior model probability. Thus the BMA posterior distribution of  $\Delta$  is a weighted average of the posterior distributions of  $\Delta$  under each of the models, weighted by their posterior model probabilities.

Based on Bayes' theorem, the posterior model probability is given by

$$P(M_k | D) = \frac{P(D | M_k) P(M_k)}{\sum_{i=1}^K P(D | M_i) P(M_i)} \quad , \quad (\text{A.2})$$

where  $P(M_k)$  is the prior probability of model  $M_k$ , and  $P(D | M_k)$  is the integrated likelihood of model  $M_k$ , obtained by integrating over the unknown parameters

$$P(D | M_k) = \int P(D | \theta_k, M_k) P(\theta_k | M_k) d\theta_k \quad , \quad (\text{A.3})$$

where  $\theta_k$  is the parameter of model  $M_k$ ,  $P(D | \theta_k, M_k)$  is the likelihood of  $\theta_k$  under model  $M_k$ , and  $P(\theta_k | M_k)$  is the prior distribution over the parameter space associated with model  $M_k$ . The integrated likelihood  $P(D | M_k)$  is a high dimensional integral that can be hard to calculate analytically, and therefore some simplification and approximations are required. Raftery (1995) proposes that a convenient solution is to approximate twice the log Bayes factor using the Bayesian Information Criterion (*BIC*) due to Schwarz (1978). One important advantage of the *BIC* approximation is that it avoids the need for an explicit specification for the prior distributions  $P(\theta_k | M_k)$ . To represent no prior preference for any model, each model presumed equally likely before examining the data, i.e. all possible



models have equal prior probabilities or  $P(M_i) = 1/K$ . Then the posterior model probability can be calculated as

$$P(M_k | D) \approx \frac{\exp(-0.5BIC_k)}{\sum_{i=1}^K \exp(-0.5BIC_i)} \quad . \quad (A.4)$$

Then we are ready to implement a systematic form of inference for different parameters of interest, which is superior to the *ad hoc* strategies often used in cross-province growth study on China. One potential difficulty in implementing BMA is the sheer range of possible models. To deal with this problem, Occam's Window technique and Markov Chain Monte Carlo techniques can be adopted. The former focuses on a subset defined by Occam's Window technique and treats all the worst-fitting models outside the subset as having zero posterior probability. Embodying the principle of parsimony, this method considerably reduces the number of possible models, and in the meantime encompasses the inherent model uncertainty present. The latter has the advantage of simultaneously selecting variables and identifying outliers, but requires a larger sample size relative to the regressor set. Given our small sample size (N=30), we use the package *bicreg* for S-Plus or R written by Adrian Raftery, where the computational procedure for Occam's Window technique is implemented to exclude the relatively unlikely models.

### **General-to-specific approach (GETS)**

The following brief discussion of general-to-specific methodology draws heavily on Owen (2003), Hendry and Krolzig (2004), Hoover and Perez (2004), and Doornik and Hendry (2007).

The general-to-specific model selection is also referred to as the LSE approach to econometric modelling. It begins with the idea that the truth can be characterized by a sufficiently rich regression (the general regression), i.e. if every possible variable is included in the regression, then the general regression must contain all the information about the true determinants. However, the model may not be in a perspicacious form, therefore the information content can be sharpened by a more parsimonious regression (the specific regression). The specific regression is a valid restriction of the general model if it is statistically well specified and it encompasses every other parsimonious regression.

The specification of the general unrestricted model (GUM) from which reductions commence is crucial to the performance of GETS approach, i.e. the specific model will not be able to improve on a bad GUM. Economic theory and previous empirical findings can play a central role in providing 'prior simplification'. Once a GUM is specified, insignificant variables are eliminated to reduce complexity, and diagnostic checks (normality test, heteroscedasticity test, *F* test for parameter constancy and Reset test for function form) on the validity of these reductions ensures congruence of the final model. In order to keep all promising variables in the final model, we set the target size as huge (level of significance: 0.1).

The computing software we use to implement GETS modelling is *Autometrics* (part of *Pcgive 12* in *OxMetrics 5*, which was recently released in late 2007). It is an upgraded version of

*Pcgets*, taking many features of the earlier implementations, but also differing in several important aspects. For example, *Autometrics* relied much less on presearch as the simulation experiments show almost the same operating characteristics with and without presearch; *Autometrics* does not implement the multiple-path search (which is an unstructured way of searching the model space), instead, it considers the whole search space from the outset using a tree search, discarding parts in a systematic way; while using roughly the same battery of diagnostic tests, *Autometrics* postpones the testing until a candidate terminal model has been found, and if necessary, backtracking is used to find a valid model, making the implementation faster and resulting in more parsimonious models; and a block-search algorithm is used by *Autometrics* to handle the case of more variables than observations. In brief, simulation results show that *Autometrics* is similar with *Pcgets* in terms of power, but had better size performance in some cases.

## Appendix 2

### Adjusted Data List used in the Regressions, 30 Provinces, 1978-2006

Variable	Definition	Units
<i>Dependent variables</i>		
$g_{i,t}$	Growth rate of real provincial GDP per capita	percent
<i>Independent variables</i>		
1. Initial income variable		
$\ln y_{i,t-1}$	Logarithm of beginning-period real GDP per capita	1990 RMB
2. Physical capital formation		
(1) By national account classification		
gcf_gdp	Gross capital formation to GDP	percent
fcf_gdp	Fixed capital formation to GDP	percent
inven_gdp	Inventory investment to GDP ( $\text{inven\_gdp} = \text{gcf\_gdp} - \text{fcf\_gdp}$ )	percent
fcf_gcf	Fixed capital formation to gross capital formation	percent
(2) By usage classification		
finv <sub>TOTAL</sub> _gdp	Total investment in fixed assets to GDP	percent
finv <sub>CC</sub> _gdp	Fixed investment in capital construction to GDP	percent
finv <sub>INNO</sub> _gdp	Fixed investment in innovation to GDP	percent
finv <sub>OTHER</sub> _gdp	Fixed investment in other usage to GDP ( $\text{finv}_{\text{OTHER\_gdp}} = \text{finv}_{\text{TOTAL\_gdp}} - \text{finv}_{\text{CC\_gdp}} - \text{finv}_{\text{INNO\_gdp}}$ )	percent
(3) Domestic vs foreign investment		
finv <sub>DOM</sub> _gdp	Ratio of domestic fixed investment to GDP	percent
fdi_gdp	Ratio of foreign direct investment to GDP (FDI converted to RMB using official exchange rate from IFS, IMF)	percent
fdi	Volume of foreign direct investment (FDI converted to RMB using official exchange rate from IFS, IMF)	billion RMB
finv <sub>DOM</sub>	Volume of domestic investment	billion RMB

---

#### (4) By ownership classification

$finv_{SOE\_finv_{TOTAL}}$	Investment in fixed assets by state-owned units / Total investment in fixed assets	percent
$finv_{COL\_finv_{TOTAL}}$	Investment in fixed assets by collectively-owned units / Total investment in fixed assets	percent
$finv_{PRIV\_finv_{TOTAL}}$	Investment in fixed assets by private units / Total investment in fixed assets	percent
$finv_{DOM\_PRIV\_finv_{TOTAL}}$	Investment in fixed assets by domestic private units / Total investment in fixed assets	percent

#### 3. Human capital formation

$stu_{PRIM\_pop}$	Students Enrolled in Primary Education / Year-end total population	percent
$stu_{SEC\_pop}$	Students Enrolled in Secondary Education / Year-end total population	percent
$stu_{REG\_SEC\_pop}$	Students Enrolled in Regular Secondary Education / Year-end total population	percent
$stu_{HIGH\_pop}$	Students Enrolled in Higher Education / Year-end total population	percent
$stu_{UNI\&COL\_pop}$	Students Enrolled in Universities and Colleges / Year-end total population	percent
$stu_{HIGH\_stu_{REG\_SEC}}$	Students Enrolled in Higher Education / Students Enrolled in Regular Secondary Education	percent
$stu_{SEC\&HIGH\_pop}$	Students Enrolled in both secondary and higher Education / Year-end total population	percent

#### 4. Population growth rate

$pop\_ngr\_nbs$	Population natural growth rate = Birth rate - death rate	percent
$pop\_gr$	Annual population growth rate = Log difference of total population	percent

#### 5. Degree of openness

##### (1) Trade volumes

$trade\_gdp$	Ratio of exports and imports to GDP (Exports and imports converted to RMB using official exchange rate from IFS, IMF)	percent
$export\_gdp$	Ratio of exports to GDP (Exports converted to RMB using official exchange rate from IFS, IMF)	percent
$import\_gdp$	Ratio of imports to GDP (Imports converted to RMB using official exchange rate from IFS, IMF)	percent

##### (2) Changes of trade volumes

$trade\_gr$	Growth rate of trade volumes (Exports and imports converted to RMB using official exchange rate from IFS, IMF)	percent
$export\_gr$	Growth rate of exports (Exports converted to RMB using official exchange rate from IFS, IMF)	percent
$import\_gr$	Growth rate of imports (Imports converted to RMB using official exchange rate from IFS, IMF)	percent

##### (3) Foreign direct investment

$fdi\_gdp$	Ratio of foreign direct investment to GDP (FDI converted to RMB using official exchange rate from IFS, IMF)	percent
$fdi$	Volume of foreign direct investment (FDI converted to RMB using official exchange rate from IFS, IMF)	billion RMB

#### 6. Institutional change

---

---

<b>(1) Of investment</b>		
$finv_{SOE\_}finv_{TOTAL}$	Investment in fixed assets by state-owned units / Total investment in fixed assets	percent
$finv_{COL\_}finv_{TOTAL}$	Investment in fixed assets by collectively-owned units / Total investment in fixed assets	percent
$finv_{PRIV\_}finv_{TOTAL}$	Investment in fixed assets by private units / Total investment in fixed assets	percent
<b>(2) Of industrial output</b>		
$ind_{SOE\_}ind_{TOTAL}$	Output value of state-owned enterprises / Gross industrial output value	percent
$ind_{COL\_}ind_{TOTAL}$	Output value of collective enterprises / Gross industrial output value	percent
$ind_{PRIV\_}ind_{TOTAL}$	Output value of private enterprises / Gross industrial output value	percent
<b>(3) Of employment</b>		
$wok_{SOE\_}wok_{TOTAL}$	State-owned enterprise workers / Total staff and workers	percent
$wok_{COL\_}wok_{TOTAL}$	Collective enterprise workers / Total staff and workers	percent
$wok_{PRIV\_}wok_{TOTAL}$	Private enterprise workers / Total staff and workers	percent
<b>7. Sectoral change</b>		
<b>(1) Temple and Wößmann (2006)'s specification</b>		
$s$	Agricultural share of GDP (Primary sector GDP / Total GDP)	percent
$a$	Agricultural share of employment (Primary sector employment / Total number of employed persons)	percent
$m$	Non-agricultural share of employment ( $m=1-a$ )	percent
$p$	Migration propensity ( $p=-\Delta a/a$ )	
$MGROWTH$	Linear sectoral change term: Change of non-agricultural share of employment ( $\Delta m$ )	percent
$DISEQ$	Non-linear sectoral change term: Change of non-agricultural share of employment adjusted by migration propensity ( $p/(1-p) * \Delta m$ )	percent
$MGROWTH2$	Linear sectoral change term: Change of non-agricultural share of employment * Average labour productivity in agricultural sector ( $\Delta m * s/a$ )	percent
$DISEQ2$	Non-linear sectoral change term: Change of non-agricultural share of employment adjusted by migration propensity * Average labour productivity in agricultural sector ( $p/(1-p) * \Delta m * s/a$ )	percent
<b>(2) Poirson (2001)'s specification</b>		
$MGROWTH * RLP$	Change in employment share in non-agricultural sector weighted by relative labour productivity (RLP = ratio of average labour productivity in non-agriculture to that in agriculture)	percent
<b>(3) Degree of industrialization</b>		
$deofin$	Degree of industrialization (Gross industrial output value / (Gross industrial output value + Gross agricultural output value))	percent
$gr\_deofin$	Growth rate of degree of industrialization ( $\Delta deofin$ )	percent
<b>8. Infrastructure</b>		
$railway\_area$	Mileages of railways per square kilometre (Total railway length /	percent

---

---

highway_area	Area) Mileages of highways per square kilometre (Total highways length / Area)	percent
post&tele_gdp	Business volume of post and telecommunication / GDP	percent
<b>9. Financial development</b>		
loan_gdp	Total bank loan outstanding / GDP	percent
saving_gdp	Savings deposit in urban and rural areas / GDP	percent
<b>10. Geographic location</b>		
dumcoastal	A dummy variable which is equal to one for coastal provinces (Liaoning, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan, plus Beijing), and zero otherwise.	0 or 1

---

Note: All the variables are calculated in 1990 constant prices and price indices are province-specific.

### **Pre-test data cleaning rules**

- Treat any observation of annual growth rate of GDP per capita / per worker above - / + 25% as outlier;
- Treat any observation of annual population growth rate above - / + 8% as outlier;
- Treat any observation of annual employment growth rate above - / + 8% as outlier.