University of Glasgow

Ding, S. and Knight, J. (2011) *Why has China grown so fast? The role of physical and human capital formation.* Oxford Bulletin of Economics and Statistics, 73 (2). pp. 141-174. ISSN 0305-9049

Why has China Grown So Fast? The Role of Physical and Human Capital Formation[*]

SAI DING [a],[*] and JOHN KNIGHT [b]

[a] *Department of Economics, University of Glasgow, Adam Smith Building, Glasgow, UK, G12 8RT . (Email: sai.ding@glasgow.ac.uk)*
[b] *China Growth Centre, St Edmund Hall, University of Oxford, Oxford, UK, OX1 4AR. (Email: john.knight@economics.ox.ac.uk)*
[*]*Corresponding author. Telephone: 0044 141 3305066; Fax: 0044 141 3304940.*

Abstract

Cross-province growth regressions for China are estimated for the reform period. Two research questions are asked. Can the regressions help us to understand why China as a whole has grown so fast? What types of investment matter for China's growth? We address the problem of model uncertainty by adopting two approaches to model selection to consider a wide range of candidate predictors of growth. Starting from the baseline equation, the growth impact of physical and human capital is examined using panel data techniques. Both forms of capital promote economic growth. 'Investment in innovation' and private investment are found to be particularly important. Secondary school enrolment contributes to growth, and higher education enrolment even more so.

JEL Classification   O40; O53
Keywords

Economic growth; Physical capital; Human capital; Model Uncertainty; China

'For the period 1960-1980 we observe, for example, India 1.4% a year,…, South Korea 7.0%,…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% 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%, and China accounted for about 35% of the growth in world GDP at PPP prices[1]. 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[2] – and of its difficult transition from being a centrally planned, closed economy at the start of reform towards becoming a market economy.

This paper is a natural follow-up to our study utilising cross-country data to explore the reasons for China's growth success compared with other countries in the world (Ding and Knight, 2009). In this paper we rely on a cross-province dataset spanning three decades to explain why China as a whole, and indeed all its provinces, has grown so fast. Our justification for this approach is at least twofold.

Firstly, from the economic perspective, all of China's provinces have grown rapidly. Even though each province would correspond to a country in most other parts of the world, they have certain characteristics in common. Can we learn why China as a whole has grown so fast by comparing the performance of its provinces? They share several features. All provinces are part of the 'development state' in which, since 1978, rapid economic growth has received primacy at all levels of government. However, some provinces reformed and marketised earlier and further than others. They are all subject to central government policies with regard to foreign trade, family planning, macroeconomic management, financial policies, etc. Nevertheless, there are province differences in openness to foreign trade, natural

---

[1] 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%. Despite this revision, China remains the main driver of global growth. For example, it contributed nearly 27% of world GDP growth in 2007 using the new PPP figure.

[2] The figure is calculated from Ravallion and Chen (2007).

increase in population, level of economic activity, and the investment-output ratio. Even if China's high growth is mainly due to the high average rate of investment in physical and human capital, it is still informative to use province differences in investment to investigate, for instance, what the effect on China's growth rate would be if it had the much lower investment rate typical of other poor countries.

Nevertheless, a qualification is in order. It is a common practice in the growth literature on China to extrapolate results from cross-province regressions to the economy as a whole (see, for instance, Guariglia and Poncet, 2006; Hao, 2006; and Yao, 2006). However, a national policy to increase a growth determinant might be more synchronised across provinces than is the case in our data set. Even in an economy as large as China's and in one with a spatially segmented capital market (Allen *et al.*, 2005; Guariglia *et al.*, 2009), there is greater mobility of goods and factors across province borders than across the national border. An economy-wide increase in a growth determinant could therefore have effects different from those of the same proportionate increase confined to one province alone. For instance, additional physical investment at national level might be more subject to diminishing returns, or might raise aggregate demand sufficiently to induce growth-retarding policy responses. Caution is therefore required in applying estimates of the growth impact of a variable at the province level to the national level.

Secondly, from the statistical perspective, the reliability of China's macroeconomic data poses a significant challenge to empirical researchers in this field. We have adopted various measures to mitigate potential mismeasurement problems (see Section 4 for detail). In terms of data selection, we favour provincial over national data for the following reasons. First, in 2006, China's National Bureau of Statistics (NBS) undertook a benchmark revision of national income and product accounts statistics based on the 2004 economic census. This revision validates the pre-economic census provincial aggregate output values and invalidates the corresponding national figures (Holz, 2008). Second, 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 our 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.

Each has its strengths and weaknesses; each deserves to be explored. 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 economic growth. That is well suited to our purposes and therefore used in our research.

A feature of our study is to use recently developed approaches to model selection in order to construct empirical models based on robust predictors. There are many growth theories and few grounds on which to choose among them. 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 and model averaging 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 address the problems of omitted variables, endogeneity and measurement error of regressors. In the second stage, the robustness of the selected models and the contribution of main variables are examined in detail. In this paper, our focus is on the growth impact of various types of physical and human capital investment. And the key question we ask is: which types matter? The contribution of other underlying variables such as the degree of openness, institutional change and structural change is analyzed in a companion paper (Ding and Knight, 2008b). Finally, some counterfactual predictions are conducted to answer the underlying question: can the cross-province growth regressions help us to understand why China as a whole has grown so fast?

In Section 2 we provide the background to Chinese economic growth, as an aid to interpretation. Section 3 explains and justifies our empirical methodology. Having discussed the dataset, section 4 reports the empirical results: the baseline equation, the physical capital variables and the human capital variables. Section 5 briefly analyzes the factors that have permitted the rapid capital accumulation in China. Section 6 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 was relevant that China 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% 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, 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 the effect on investor confidence of the social unrest and ensuing international ostracism, and the policy response to a previous surge in inflation. 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. Methodology

We draw on the relevant literatures to explain and justify our methodology. This concerns the choice of informal growth regressions, the method of dealing with model uncertainty and the panel data estimation approach.

Why informal growth regressions?

The starting point in this research is our cross-country analysis of the extent to which the growth difference between China and other countries can be explained by the neoclassical growth model (Ding and Knight, 2009). We found that the Solow model augmented by both human capital and structural change provides a fairly good account of China's remarkable growth performance. Moreover, 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 China's

relative growth success. 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. 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. The research 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 (see, for example, Chen and Fleisher, 1996; Li *et al.*, 1998; Raiser, 1998; Chen and Feng, 2000; Démurger, 2001; Zhang, 2001; Bao *et al.*, 2002; Brun *et al.*, 2002; Cai *et al.*, 2002; Jones *et al.*, 2003; Guariglia and Poncet, 2006; Hao, 2006; Yao, 2006; and Fleisher *et al*., 2009). An underlying problem in all the research is the difficulty in establishing causal relationships as opposed to mere associations.

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.

Another strand of growth research on China adopts the growth accounting approach to break down the observed growth of GDP into components associated with changes in factor inputs and in production technologies (see, for instance, Borenzstein and Ostry, 1996; Hu and Khan, 1997; Woo, 1998; Wang and Yao, 2001; Young, 2003; and Brandt *et al*., 2008). Average annual total factor productivity (TFP) growth in China for the reform period is found to range from a high of 3.9% to a low of 1.5% in these studies. This disconcertingly wide variation is partly the result of the different assumptions made. The growth accounting

approach involves measuring the capital stock and making assumptions about unknown parameters such as output elasticities and capital depreciation rates. Two further arguments make us disinclined to use growth accounting. First, when the 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 any misspecification of the production function. This is particularly true for China. According to Borenzstein and Ostry (1996), technological progress in China has been substantially lower than TFP growth, with the difference representing structural change and unmeasured input growth. Second, technological change and investment may not be 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. For instance, Ding and Knight (2009) found that investment is a major carrier of structural change in China: structural transformation requires investment in new, normally high-productivity activities. Employment growth in the high-productivity industrial and service sectors is determined by the rate of investment in those sectors, and the new job opportunities are largely taken by migrant workers from the low-productivity agricultural sector.

Dealing with model uncertainty

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 following the seminal work of Levine and Renelt (1992) which applied an Extreme Bounds Analysis (EBA) 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. This work is further extended by Sala-i-Martin (1997) and Temple (2000). Other econometric and statistical methods have been developed and applied to handle model uncertainty as well, among which Bayesian Model Averaging (Raftery, 1995; Fernández *et al.*, 2001; Sala-i-Martin *et al.*, 2004) and General-to-Specific approach (Hendry and Krolzig, 2004; Hoover and Perez, 2004) are among the most influential[3]. In this paper we adopt Bayesian Model Averaging (BMA) and General-to-Specific approach (GETS) to consider the association between GDP per capita growth rates and a wide range of potential explanatory variables. 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.

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. Thus 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. By treating parameters and models as random variables, the uncertainty about the model is summarized in terms of a probability distribution over the space of all possible models. The idea of the GETS procedure 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. The specific regression is a valid restriction of the general model if it is statistically well specified and also encompasses every other parsimonious regression. One attractive feature of the automatic procedure of model selection is argued to be the huge efficiency gain.

Each of the two procedures has comparative advantages and disadvantages in dealing with model uncertainty. 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

---

[3] Detailed discussion of various model selection methods is available in the working paper version of this paper (Ding and Knight, 2008a).

designed to combine the strengths of both methods and to circumvent the limitations of each to some extent (see Appendix B for brief discussion of the two methods).

Since neither method can handle the problem of endogenous regressors during the model selection process, no causal interpretation can be attached to the results at this stage. We therefore adopt a two-stage testing approach to solve this problem. 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).

The panel data estimation approach

It is a challenge to estimate a short dynamic panel with fixed effects and multiple endogenous regressors, especially when the number of cross sections is relatively small. Several econometric problems require attention. For instance, the correlation between the lagged dependent variable and the time-invariant region-specific effects renders the OLS estimator biased and inconsistent (Hsiao, 1986). In the cross-country or cross-province growth regressions, the OLS estimate of the coefficient of initial income term is likely to be biased upward (Bond *et al*., 2001; Hoeffler, 2002). Nickell (1981) showed that the within-groups estimator will be biased for fixed T and large N. Although the bias diminishes with T, for the typical growth regression with small T, the within-groups estimate of the coefficient of the initial income term is likely to be seriously biased downwards. This problem also applies to the instrumental variable methods based on fixed effects.

The system GMM estimator for dynamic panels has become popular in the empirical growth literature so as to overcome the Nickell (1981) bias and to address the problems of endogeneity and mismeasurement. It 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 and exploiting these additional moment conditions, Arellano and Bover (1995) and Blundell and Bond (1998) found a dramatic improvement in efficiency and a significant reduction in finite sample bias compared with first-differenced GMM. However,

caution is needed in at least two aspects in applying system GMM to our study. First, the instrument proliferation problem is likely to be severe when the cross-section dimension is small. According to Bowsher (2002) and Roodman (2009), as T rises, the instrument count can easily grow large relative to the sample size, making some asymptotic results about the estimators and related specification tests misleading. Second, the problem of cross-sectional error dependence can lead to serious problems in the estimation of short dynamic panels. Sarafidis and Robertson (2009) demonstrate that under cross-sectional error dependence, the GMM estimator is inconsistent as $N \longrightarrow \infty$ for fixed T, which holds for any lag length of the instruments used.

To address the instrument proliferation problem, we adopt two approaches to restrict the number of instruments used in our system GMM estimation. The first is to collapse the instrument sets, i.e. the GMM estimator is based on one instrument per variable instead of one instrument for each variable at each period. The second approach is to use only certain lags instead of all possible lag lengths for instruments in each first-differenced equation. For example, for potentially endogenous variables, levels of that variable lagged by 10-year, 15-year and 20-year periods are used as instruments in the first-differenced equations, and first-differenced variables lagged by a 5-year period are used as additional instruments for the levels equations. Following the suggestion of Roodman (2009), we report the number of instruments generated for our regressions together with the Hansen and Difference Sargan statistics. Regarding the second problem, we include time-specific effects in our regressions to capture common variations in the dependent variable and to reduce the asymptotic bias of the estimator in the presence of cross-sectional error dependence. Besides, all the standard errors are robust to heteroskedasticity and clustering on province.

4. Empirical results

The dataset

The original sample consists of a panel of 30 provinces with annual data for the period 1978-2007[4]. The data come mainly from *China Compendium of Statistics 1949-2004* compiled by the National Bureau of Statistics of China. The data for 2005-07 are obtained from the latest issues of *China Statistical Yearbook*. The reliability of Chinese official macroeconomic data

---

[4] 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 1997-2007, so making it consistent with earlier observations.

is often under dispute. One important issue is the problem of data inconsistency over the sample period. For example, GDP figures for the years 2005-07 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 A) 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 non-overlapping five-year time intervals, which have been widely used in the cross-country growth literature (for instance, Islam, 1995; Bond *et al.*, 2001; Ding and Knight, 2009). 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.

All the variables are calculated in 1990 constant prices and price indices are province-specific[5]. 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 rate of all 30 provinces over the entire reform period was 7.7%, with an average value of 8.1% for the coastal provinces and 7.5% for interior provinces. China's economic reform generated across-the-nation rapid growth, i.e. both the coastal and inner

---

[5] The deflator is the provincial consumer price index (CPI). Assessing China's GDP growth rates relies on reliable GDP deflator data. However, it is widely believed that China's implicit GDP deflator based on the Material Product System approach has understated inflation in China, therefore exaggerating the real GDP growth (Wu, 1997; Maddison, 1998; Woo, 1998; Rawski, 2001). Using a different approach, Maddison (1998) predicts that the average annual real GDP growth rate for China is 2.4 percentage points below the official one. His GDP figure of China has been accepted internationally by the Penn World Tables and the World Bank. However, that figure is an aggregate one for China as a whole rather than province-specific. According to Wu (1997) and Holz (2006), a simple and possibly relatively acceptable approach to deriving China's real GDP growth rate is to use China's official CPI as a single deflator. They both show that the use of price indices instead of the official implicit deflators gives a figure of China's real growth rate similar to Maddison's. For this reason, we deflate the nominal GDP and other variables for each province by province-specific CPI. There are no provincial price data for Tibet for the period 1978-89; we use the national aggregate price index to substitute.

regions grew fast by international standards. However, that a growth disparity did exist is indicated by the 4% 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 in the second decade, and accelerated in the third decade. In the period 1998-2007, the growth disparity across provinces became smaller and even the lowest-growing province (Yunnan) managed an average rate of 8.2%.

The explanatory variables can be broadly classified into ten categories: initial level of income, physical capital formation, human capital formation, population growth rate, degree of openness, pace of economic reform or institutional change, sectoral change or degree of industrialization, infrastructure, financial development, and geographic location. A geographical distinction is made between 'coastal' and other provinces. This classification follows that of the literature, the underlying rationale being that the provinces deemed to be coastal have advantages in the form of lower-cost access to markets (see Appendix A for detailed definitions).

Baseline equation

The prospects for selecting a good model depend 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 growth theory (although sufficiently loosely) 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 cause problems 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 average annual 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 ($lny_{i,t-1}$), ratio of fixed capital formation over GDP (fcf/GDP), secondary school enrolment as a proportion of the total population (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 ($export/GDP$), the SOE share of 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 volume of post and telecommunications to GDP ($post\&tele/GDP$), and a coastal dummy ($dum\_coastal$).

We first use BMA to isolate variables that have a high posterior probability of inclusion. In Table 2, 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. Given that the prior probability of a variable being in the true model is set at 0.5, its robustness may be assessed in terms of how the data update this prior. We therefore refer to a specific variable as being important if the posterior probability of inclusion is greater than 0.5. The results indicate a possibly important role for the initial level of income, the SOE share of total industrial output, secondary school enrolment, fixed capital formation, and population growth.

We then conduct an automatic model selection exercise using the GETS methodology. Starting from the same general model and searching for statistically acceptable reductions, the software package *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 3. We find that growth in GDP per capita is negatively associated with the initial income level, population growth and the SOE share of industrial output, whereas fixed capital investment and secondary school enrolment are positively correlated. The major difference between the results of the two methods lies in the role of exports in explaining cross-province growth rates, i.e. despite the statistical insignificance, exports as a proportion of GDP is retained by GETS in the final specific model, but BMA analysis flags the export ratio as potentially unimportant (with a posterior inclusion probability of 28%). Other variables such as sectoral change, infrastructure and financial development are identified 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 re-examine the role of these variables in determining output growth in the panel data context in a different paper (Ding and Knight, 2008b).

Based on the model selection results delivered by BMA and GETS, we now estimate the baseline model using various panel data techniques in Table 4. Consistent with the prediction of Bond *et al*. (2001) and Hoeffler (2002), we find that our system GMM estimator yields a consistent estimate of the coefficient on the initial level of income which lies in between the upper bound provided by the OLS estimator and the lower bound given by the within-groups estimator. The instrumental variable method (IV-2SLS) yields an estimate of the initial level of income even lower than that of within-groups, indicating the potential bias of this kind of fixed-effect estimator in short dynamic panels[6]. Thus, the panel data system GMM with a restricted instrument set is our preferred estimation method.

Interestingly, the GMM results support the model selected by the GETS procedure, i.e. the ratio of exports to GDP appears positive and significant. 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 percentage point lower initial level of GDP per capita raises the subsequent growth rate of GDP per capita by 0.05 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 conditional 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, 2009) and cross-province growth studies of China (for example, Chen and Fleisher, 1996; Chen and Feng, 2000; Cai *et al*., 2002).

One possible explanation for conditional convergence 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 for conditional convergence might lie in the central

---

[6] Because the Nickell (1981) bias would diminish as *T* becomes larger, as a robustness check we also implement both the within-groups and IV methods using 3-year time intervals. However, no significant improvement of the results is found, perhaps because the 3-year averages are too short to eliminate business cycle effects.

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 the 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 the central government's fiscal capacity, which enabled it to promote economic development in poor regions such as the western provinces and minority areas. From about 1998 onwards there was a fiscal redistribution towards poor provinces (Wong and Bird, 2008: 456). This might help to explain the convergence between lowest and highest growth provinces in recent years (Table 1).

Fixed capital formation is an important determinant of China's growth: 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 0.1 percentage points. Human capital investment appears to be even more important, i.e. a one percentage point increase in the secondary school enrolment ratio is associated with a higher growth rate of GDP per capita by 0.3 percentage points. Since both physical and human capital accumulation are the focus of this paper, detailed discussion will follow in the next two subsections.

The increase in population has a negative consequence for growth: reducing the rate of population growth by 0.1 percentage points is associated with an increase in GDP per capita growth of 0.4 percentage points. A rapid population growth rate involves a cost, 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. Within the standard Solow model, slower population growth implies a higher equilibrium level of output per worker and capital per worker. This means that if two provinces have the same initial income level, but one has a lower population growth rate, it will grow more quickly than the other. 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: 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 the 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 knowledge spillover across countries (Grossman and Helpman, 1993). China's open-door policy, adopted after 1978, created an excellent opportunity to exploit its comparative advantage in labour-intensive manufacturing industry, making exports a driver of China's growth.

The SOE share of industrial output has a significant and negative impact on output growth: a decrease of one percentage point in the variable raises the 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. In brief, our panel data results favour the model selected by the 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.

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% of real GDP (gcf/GDP), which is very high by international standards. The rate of gross fixed capital formation (fcf/GDP) has increased significantly in

recent years, rising from an average of 29.3% between 1978 and 1993 to an average of 36.6% thereafter. Inventory accumulation amounted to, on average, 5.5% of GDP ($inven/GDP$). It peaked at the end of 1980s, reflecting the severe economic recession, and declined gradually thereafter thanks to the process of marketization. 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).

In the cross-country growth literature, there is 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*. (forthcoming). 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. We are particularly interested in the question of what sort of physical capital formation matters for China's 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. To save space, we report only the coefficients of interest along with relevant specification tests.

We first decompose total investment in fixed assets ($finv_{Total}/GDP$) into investment in capital construction ($finv_{CC}/GDP$), investment in innovation ($finv_{INNO}/GDP$), and investment in other fixed assets[7] ($finv_{OTHER}/GDP$) in Panel 1 of Table 5. We find that the former three have positive and significant impacts on growth, whereas the last appears insignificant. 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

---

[7] 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. Innovation investment consists of the renewal of fixed assets and technological innovation of the original facilities by enterprises and other institutions. 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.

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: 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 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). Our results provide evidence that investment and investment-driven improvements in technology are important for China's growth.

We then classify investment in fixed assets according to ownership: investment spending by SOEs ($finv_{SOE}/finv_{Total}$), collectively-owned enterprises ($finv_{COL}/finv_{Total}$) and private enterprises ($finv_{PRIV}/finv_{Total}$) in Panel 2 of Table 5. Since these variables may contain information similar to our proxy for the extent of economic reform (ind_{SOE}/ind_{TOTAL}, the SOE share of industrial output), we drop that term in these regressions. We find that the share of investment made by SOEs has a significantly negative effect on growth: reducing the SOE share of fixed investment by one percentage point is associated with an increase in GDP per capita growth of 0.13 percentage points. This result is consistent with the widespread perception that the efficiency of investment in the SOEs is far below that in the non-state sector, so that the growth rate of GDP per capita should fall as the SOE share of investment increases (for instance, Brandt and Zhu, 2000). Consequently, the recent decline in the SOE share of fixed investment is a positive development. The coefficient on the share of investment by collective firms appears insignificant. The collective economy consists of both township and village enterprises (TVEs) and urban collectives 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.

The share of private-sector investment affects growth positively: a one percentage point increase in this variable is associated with an increased growth rate of 0.20 percentage points. 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 the private sector is the driving force in the Chinese economy. It is therefore a positive development that the centre of gravity of the economy has been shifting from the state to the private sector.

A caveat is in order: in the late 1990s some SOEs were corporatized, and would have been reclassified as private, although the state, being the major shareholder, generally retained control. Insofar as the most profitable or promising SOEs were selected for listing, the results are likely to exaggerate the incentive effects of different ownership status. 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). The estimates imply that financial sector reform would raise the growth rate further.

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 on output growth of secondary school enrolment as a proportion of working-age population, whereas other researchers (Benhabib and Spiegel, 1994; Pritchett, 1999) claimed that output growth did not seem to be strongly related to increases in measured educational attainment (changes

in the average years of schooling), 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 use enrolment at different educational levels 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 and its effect on the economic growth of a province. School enrolment may conflate human capital *stock* and *accumulation* effects and can be a poor proxy for either (see, for instance, Gemmell, 1996 and Temple, 1999). Nevertheless, we make use of what information is available to us annually at the province level.

In Table 6, our human capital measurements consist of primary school enrolment (stu_{PRIM}/pop), secondary school enrolment (stu_{SEC}/pop), regular secondary school enrolment (stu_{REG_SEC}/pop), higher education enrolment (stu_{HIGH}/pop), university and college enrolment ($stu_{UNI\,\&\,COL}/pop$), and enrolment in both secondary and higher education ($stu_{SEC\,\&\,HIGH}/pop$), each expressed as a proportion of the total population[8]. To deal with the possible endogeneity of these variables, levels of human capital investment variables lagged by 10-year, 15-year and 20-year periods are used as instruments in the first-differences equations, and first-differenced human capital investment variables lagged by a 5-year period are used as additional instruments for the levels equations in the system GMM estimation.

In line with Chen and Feng (2000) in their cross-province study, we find that the coefficient of the primary enrolment variable is insignificant in the growth regression. This is to be expected because primary education is mandatory and the negative coefficient may reflect the falling number of children as a result of the one-child-family policy, introduced in

---

[8] 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.

the late 1970s[9]. Consistent 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 the secondary and regular secondary school enrolment variables have positive and significant impacts on output growth.

The higher education enrolment and university and college enrolment proportions have bigger positive effects. For instance, a one percentage point rise in the ratio of enrolment in higher education to population leads to higher GDP per capita growth by 3.6 percentage points, holding other conditions constant. 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. Higher education enrolment remained below 0.3% of total population until 1998, and then shot up to 1.4% in 2006 as a result of a sharp change in higher education policy. Our finding is consistent with Chi (2008), who used educational attainment to measure 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. Lastly, we examine the impact of students enrolled in both secondary and higher education over population, and find that a one percentage point rise in this variable is associated with a 2.0 percentage point rise in the growth rate of GDP per capita.

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 the shares are included in the growth equation, an increase in the relative stock of people with both junior secondary and tertiary education raises growth significantly. These results are consistent with our findings based on school enrolment data.

Illustrative counterfactual predictions

---

[9] The ratio of primary school enrolment to total population fell from 15% in 1978 to 8% in 2006.

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 7 and 8. 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 an adverse change in one variable (say, human capital formation) induces an adverse change 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%. If instead it had been 10 percentage points lower (24.3%), the system GMM coefficient of the baseline model in Table 4 implies that China's growth of GDP per capita would have fallen by 0.9 percentage points, from 8.0 to 7.1%, holding other variables constant. Similarly, secondary school enrolment averaged 5.8% 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 6.0%. Had both the physical and human capital variables been reduced in this way, China's per capita growth rate would have fallen to 5.1%, holding other controls constant.

China was a low-income country at the start of economic reform[10]. 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% and 2.3% respectively[11]. Plugging these values into the baseline model, we find that China's growth rate of GDP per capita would have been only 2.9% per annum, to be compared with 1.0% for the least developed countries. On these assumptions, China's status as a growth outlier would have been much weaker.

Consider the effects of changes in the composition of physical investment in Table 8. The ratio of innovation investment averaged 6.9% 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 one percentage point lower growth rate of GDP per capita.

---

[10] According to the *World Development Report* (1981) published by the World Bank, China's GNP per capita was $260 in 1979, ranking as the 22nd poorest country.
[11] The ratio of fixed capital formation to GDP and the total population data for least developed countries come from the *World Development Indicators* (April 2008 edition); secondary school enrolment figures come from UNESCO.

Combining the effect of reducing the secondary school enrolment ratio to half, the growth rate would have fallen to 3.4% holding other conditions the same. Private fixed investment averaged 16.7% of total GDP in China over the full period. If it had remained at its 1978 level (5.1%), growth would have been 2 percentage points lower, at 6.1% per annum. Combining the effect of reducing secondary school enrolment to half, China's growth rate would have been down to 1.2%.

How important was post-primary education to growth? Secondary and higher education enrolments averaged 5.8% and 0.4% 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 5.8 percentage points, to 2.2% 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.

5.   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% of income in 1978 to a peak of 34% in 1994, and remained above 24% 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 2007[12]. 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

---

[12] Data come from the *World Development Indicators* (April 2008 edition); GDP per capita is gross domestic product divided by midyear population and is in constant 2000 US dollars.

growth, which in turn leads to higher income. By contrast, Modigliani and Cao (2004) attributed China's high household saving rate to 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 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 and its growth over time, so important for economic growth, has 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%, but with labour market reform it rose to 26% in 1995 and to 33% in 2002. The urban wage premium of higher education over upper secondary education rose from only 5% in 1988 to 17% in 1995 and to 42% 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*., 2010, tables 2, 5, 6). 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[13].

6. Conclusion

In this paper, we have attempted 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 by international standards over the period of economic reform. To address the problem of model uncertainty, 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, 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. They are consistent with the implications of the transitional dynamics of neoclassical growth theory. Such transitional movement is indeed to be expected given the likely disequilibrium of the Chinese economy at the start of economic reform. Our evidence of conditional convergence implies that each province is converging towards its equilibrium steady state. It might, however, have other explanations, e.g. that convergence reflects the effects of fiscal transfers from the central government to poor provinces and minority areas. The growth impacts of physical and human capital accumulation are in line with the conditional convergence argument. 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.

---

[13] NBS, *China Statistical Yearbook 2006*, table 21-6.

Our more detailed investigation of the effects of physical and human capital accumulation was intended to throw further light on the mechanisms at work. Among the types of fixed investment, 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. This result suggests that physical investment makes the greatest contribution to growth when it is most closely bound up with technological progress. Breaking down physical investment by ownership, we found that an increased share of SOEs decreases the contribution of investment to growth, an increased share of collective enterprises has a negligible effect, and an increased share of private enterprises raises the contribution. 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 primary school enrolment has no effect on economic growth, both secondary school and higher education enrolment had a positive effect, the latter more than the former. 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 2.8 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% of the relevant age group. Our use of GMM estimation and instrumenting of the human capital variables using lags provided the best means of estimating the causal effect of human capital on growth.

To address our title question - why has China grown so fast? - it was necessary to assume that the growth impact of a variable estimated on the basis of its variation among provinces would be a guide to its impact in the economy as a whole. Various counterfactual exercises were conducted on that basis. We found that a significant reduction in capital 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* , Vol. 46, pp. 385-406.

Allen, F., Qian, J. and Qian, M. (2005). 'Law, finance, and economic growth in China', *Journal of Financial Economics* , Vol. 77, pp. 57-116.

Arellano, M. and Bover, O. (1995). 'Another look at the instrumental variable estimation of error-components models', *Journal of Econometrics* , Vol. 68, pp. 29-52.

Arellano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an Application to employment equations', *Review of Economic Studies* , Vol. 58, pp. 277-297.

Bao, S. M., Chang, G. H., Sachs, J. D. and Woo, W. T. (2002). 'Geographic factors and China's regional development under market reforms, 1978-1998', *China Economic Review* , Vol. 13, pp. 89-111.

Barro, R. J. (1991). 'Economic growth in a cross section of countries', *Quarterly Journal of Economics* , Vol. 106 , pp. 407-43.

Barro, R. J. and Sala-i-Martin, X. X. (2004). *Economic Growth* (second ed.). Cambridge, MA: The MIT Press.

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

Blomström, M., Lipsey, R. and Zejan, M. (1996). 'Is fixed investment the key to economic growth?' *The Quarterly Journal of Economics* , Vol. 111 , pp. 269-276.

Blundell, R. and Bond, S. (1998). 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics* , Vol. 87, pp. 115-143.

Bond, S., Hoeffler, A. and Temple, J. (2001). 'GMM estimation of empirical growth models', *CEPR Discussion Paper No. 3048* .

Bond, S., Leblebicioglu, A. and Schiantarelli, F. (forthcoming). 'Capital accumulation and growth: a new look at the empirical evidence', *Journal of Applied Econometrics.*

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

Bowsher, C. G. (2002). 'On testing overidentifying restrictions in dynamic panel data models', *Economics Letters* , Vol. 77, pp. 211-220.

Brandt, L. and Zhu, X. D. (2000). 'Redistribution in a decentralized economy: growth and inflation in China under Reform', *Journal of Political Economy* , Vol. 108 , pp. 422-451.

Brandt, L., Hsieh, C. T. and Zhu, X. D. (2008). 'Growth and structural transformation in China', In Brandt, L. and Rawski, T., China's Great Economic Transformation (pp. 683-728). New York: Cambridge University Press.

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

Cai, F., Wang, D. W. and Du, Y. (2002). 'Regional disparity and economic growth in China: the impact of labour market distortions', *China Economic Review* , Vol. 13, pp. 197-212.

Chatfield, C. (1995). 'Model uncertainty, data mining, and statistical inference', *Journal of the Royal Statistical Society* , Vol. 158, pp. 419-444.

Chen B. Z. and Feng, Y. (2000). 'Determinants of economic growth in China: private enterprise, education and openness', *China Economic Review* , Vol. 11, pp. 1-15.

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

Chi, W. (2008). 'The role of human capital in China's economic development: review and new evidence', *China Economic Review*, Vol. 19, pp. 421-436.

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* , Vol. 29, pp. 95-117.

Ding, S. and Knight, J. (2008a). 'Why has China grown so fast? The role of physical and human capital formation', *Department of Economics Discussion Paper No. 414,* University of Oxford.

Ding, S. and Knight, J. (2008b). 'Why has China grown so fast? The role of structural change', *Department of Economics Discussion Paper No. 415,* University of Oxford.

Ding, S. and Knight, J. (2009). 'Can the augmented Solow model explain China's economic growth? A cross-country panel data analysis', *Journal of Comparative Economics*, Vol. 37 (3), pp. 432-452.

Doornik, J. A. and Hendry, D. F. (2007). *Pcgive 12 Voume I: Empirical Econometric Modelling.* London: Timberlake Consultants Press.

Dowrick, S. and Gemmell, N. (1991). 'Industrialization, catching up and economic Growth: a comparative study across the world's capitalist economies', *Economic Journal* , Vol. 101 (405), pp. 263-275.

Easterly, W. and Levine, R. (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* , Vol. 15, pp. 177-219.

Fernández, C., Ley, E. and Steel, M. F. J. (2001). 'Model uncertainty in cross-country growth regressions', *Journal of Applied Econometrics* , Vol. 16, pp. 563-576.

Fleisher, B., Li, H. and Qiang, M. (2009). 'Human capital, economic growth, and regional inequality in China', *Journal of Development Economics , in press*.

Gemmell, N. (1996). 'Evaluating the impacts of human capital stocks and accumulation on economic growth: some new evidence', *Oxford Bulletin of Economics and Statistics*, Vol. 58, pp. 9-28.

Grossman, G. and Helpman, E. (1993). *Innovation and Growth in the Global Economy.* Cambridge, MA: MIT Press.

Guariglia, A. and Poncet, S. (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 Song, L. (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* , Vol. 17, pp. 347-362.

Hendry, D. F. and Krolzig, H. M. (2004). 'We ran one regression', *Oxford Bulletin of Economics and Statistics*, Vol. 66 , pp. 799-810.

Hoeffler, A. (2002). 'The augmented Solow model and the African growth debate', *Oxford Bulletin of Economics and Statistics*, Vol. 64 (2), pp. 135-158.

Holz, C. A. (2006). 'China's reform period economic growth: how reliable are Angus Maddison's estimates?' *Review of Income and Wealth*, Vol. 52 (1), pp. 85-119.

Holz, C. A. (2008). 'China's 2004 economic census and 2006 benchmark revision of GDP statistics: more questions than answers?' *China Quarterly*, doi: 10.1017/ S030574100 800009X.

Hoover, K. D. and Perez, S. J. (2004). 'Truth and robustness in cross-country growth regressions', *Oxford Bulletin of Economics and Statistics*, Vol. 66, pp. 765-798.

Hsiao, C. (1986). *Analysis of Panel Data*. Cambridge: Cambridge University Press.

Hu, Z., and Khan, M. (1997). 'Why is China growing so fast?' *International Monetary Fund Staff Papers*, Vol. 44 (1), pp. 103-131.

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*, Vol. 110, pp. 1127-1170.

Javorcik, B. (2008). 'Can survey evidence shed light on spillovers from foreign direct investment?' *The World Bank Research Observer*, Vol. 23, pp. 139-159.

Jones, D. C., Li, C. and Owen, A. L. (2003). 'Growth and regional inequality in China during the reform era', *China Economic Review*, Vol. 14, pp. 186-200.

Knight, J. and Li, S. (1999). 'Fiscal decentralization: incentives, redistribution and reform in China', *Oxford Development Studies*, Vol. 27, pp. 5-32.

Knight, J. and Song, L. (2005). *Towards a Labour Market in China*. Oxford: Oxford University Press.

Knight, J., Li, S. and Deng, Q. (2010). 'Education and the poverty trap in rural China: closing the trap', *Oxford Development Studies*, Vol. 38, No. 1, March, pp. 1-24.

Knight, J. and Song, L. (2008). 'China's emerging urban wage structure, 1995-2002', In Gustafsson, B., Li, S. and Sicular, T. (eds), *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*.

Levine, R. and Renelt, D. (1992). 'A sensitivity analysis of cross-country growth regressions', *American Economic Review*, Vol. 82, pp. 942-963.

Li, H., Liu, Z. N. and Rebelo, I. (1998). 'Testing the neoclassical theory of economic growth: evidence from Chinese Provinces', *Economics of Planning*, Vol. 31, pp. 117-132.

Liu, X. H., Burridge, P. and Sinclair, P. J. N. (2002). 'Relationships between economic growth, foreign direct investment and trade: evidence from China', *Applied Economics*, Vol. 34, pp.1433-1440.

Lucas, R. E. (1988). 'On the mechanics of economic development', *Journal of Monetary Economics*, Vol. 22, pp. 3-42.

Lucas, R. E. (2002). *Lectures on Economic Growth*. MA: Harvard University Press.

Maddison, A. (1998). *Chinese Economic Performance in the Long Run*. Paris: Development Centre of the Organisation for Economic Co-operation and Development.

Malik, A. and Temple, J. (2009). 'The geography of output volatility', *Journal of Development Economics*, Vol. 90 (2), pp. 163-178.

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

Modigliani, F. and Cao, S. L. (2004). 'The Chinese saving puzzle and the life-cycle hypothesis', *Journal of Economic Literature*, Vol. 42, pp. 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*, Vol. 49(6), pp. 1417-1426.

Owen, P. D. (2003). 'General-to-specific modelling using Pcgets', *Journal of Economic Surveys*, Vol. 17, pp. 609-627.

Poirson, H. (2001). 'The Impact of Intersectoral Labour Reallocation on Economic Growth', *Journal of African Economies*, Vol. 10 (1), pp. 37-63.

Pritchett, L. (1999). 'Where has all the education gone?' *The World Bank Economic Review*, Vol. 15, pp. 367-391.

Raftery, A. E. (1995). 'Bayesian model selection in social research', In Marsden, P. V., *Sociological Methodology.* Cambridge: Blackwell.

Raiser, M. (1998). 'Subsidising inequality: economic reforms, fiscal transfers and convergence across Chinese provinces', *Journal of Development Studies*, Vol. 34, pp. 1-26.

Ravallion, M. and Chen, S. H. (2007). 'China's (uneven) progress against poverty', *Journal of Development Economics*, Vol. 82, pp. 1-42.

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

Riedel, J., Jin, J. and Gao, J. (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*, Vol. 80, pp. 97-103.

Roodman, D. (2009). 'Practitioners' corner: A note on the theme of too many instruments', *Oxford Bulletin of Economics and Statistics*, Vol. 71 (1), pp. 135-158.

Sala-i-Martin, X. X. (1997). 'I just ran two million regressions', *American Economic Review*, Vol. 87, pp. 178-183.

Sala-i-Martin, X. X., Doppelhofer, G. and Miller, R. (2004). 'Determinants of long-run growth: A Bayesian averaging of classical estimates (BACE) approach', *American Economic Review*, Vol. 94 (4), pp. 813-835.

Sarafidis, V. and Robertson, D. (2009). 'On the impact of error cross-sectional dependence in short dynamic panel estimation', *Econometrics Journal*, Vol. 12, pp. 62-81.

Schwarz, G. (1978). 'Estimating the dimension of a model', *Annals of Statistics*, Vol. 6, pp. 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*, Vol. 83, pp. 421-425.

Solow, R. M. (1956). 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, Vol. 70, pp. 65-94.

Temple, J. (1999). 'The new growth evidence', *Journal of Economic Literature*, Vol. 37, pp. 112-156.

Temple, J. (2000). 'Growth regressions and what the textbooks don't tell you', *Bulletin of Economic Research*, Vol. 52(3), pp. 181-205.

Temple, J. and Wöβmann, L. (2006). 'Dualism and cross-country growth regressions', *Journal of Economic Growth*, Vol. 11, pp. 187-228.

Wang, Y. and Yao, Y. (2001). 'Sources of China's economic growth, 1952-99: Incorporating human capital accumulation', *World Bank Policy Research Working Paper No. 2650.*

Wei, S. J. (1995). 'Open door policy and China's rapid growth: evidence from city-level data', In Ito, T. and Krueger, A., *Growth Theories in Light of the East Asian Experience.* University of Chicago Press.

Wong, C. P. and Bird, R. M. (2008). 'China's fiscal system: a work in progress', In Brandt, L. and Rawski, T. (eds), *China's Great Economic Transformation* (pp. 429-466). New York: Cambridge University Press.

Woo, W. T. (1998). Chinese Economic Growth: Sources and Prospects. In Fouquin, M. and Lemoine, F. (eds), *The Chinese Economy.* London: Economica.

Wu, H. (1997). 'Measuring China's GDP', *Briefing Paper Series No. 8*, Department of Foreign Affairs and Trade, Austrilia.

Yao, S. (2006). 'On economic growth, FDI and exports in China', *Applied Economics* , Vol. 38, pp. 339-351.

Young, A. (2003). 'Gold into base metals: productivity growth in the People's Republic of China during the reform period', *Journal of Political Economy* , Vol. 111 (1), pp. 1220-1261.

Zhang, W. (2001). 'Rethinking regional disparity in China', *Economics of Planning* , Vol. 34, pp. 113-138.

Appendix A

Detailed variable definitions and data cleaning rules (30 Provinces, 1978-2007)

Variable	Definition	Units
Dependent variables		
$g_{i,t}$	Growth rate of real provincial GDP per capita	%
Independent variables		
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	%
fcf_gdp	Fixed capital formation to GDP	%
inven_gdp	Inventory investment to GDP (inven_gdp = gcf_gdp-fcf_gdp)	%
(2) By usage classification		
$finv_{TOTAL}$_gdp	Total investment in fixed assets to GDP	%
$finv_{CC}$_gdp	Fixed investment in capital construction to GDP	%
$finv_{INNO}$_gdp	Fixed investment in innovation to GDP	%
$finv_{OTHER}$_gdp	Fixed investment in other usage to GDP ($finv_{OTHER}$_gdp = $finv_{TOTAL}$_gdp-$finv_{CC}$_gdp- $finv_{INNO}$_gdp)	%
(3) By ownership classification		
$finv_{SOE}$_$finv_{TOTAL}$	Investment in fixed assets by state-owned units / Total investment in fixed assets	%
$finv_{COL}$_$finv_{TOTAL}$	Investment in fixed assets by collectively-owned units / Total investment in fixed assets	%
$finv_{PRIV}$_$finv_{TOTAL}$	Investment in fixed assets by private units / Total investment in fixed assets	%
3. Human capital formation		
stu_{PRIM}_pop	Students Enrolled in Primary Education / Year-end total population	%
stu_{SEC}_pop	Students Enrolled in Secondary Education / Year-end total population	%
stu_{REG_SEC}_pop	Students Enrolled in Regular Secondary Education / Year-end total population	%
stu_{HIGH}_pop	Students Enrolled in Higher Education / Year-end total population	%
$stu_{UNI\&COL}$_pop	Students Enrolled in Universities and Colleges / Year-end total population	%
$stu_{SEC\&HIGH}$_pop	Students Enrolled in both secondary and higher Education / Year-end total population	%
4. Population growth rate		
pop_ngr	Population natural growth rate = Birth rate - death rate	%
pop_gr	Annual population growth rate = Log difference of total population	%
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)	
`export_gdp`	Ratio of exports to GDP (Exports converted to RMB using official exchange rate from IFS, IMF)	%
`import_gdp`	Ratio of imports to GDP (Imports converted to RMB using official exchange rate from IFS, IMF)	%

(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)	%
`export_gr`	Growth rate of exports (Exports converted to RMB using official exchange rate from IFS, IMF)	%
`import_gr`	Growth rate of imports (Imports converted to RMB using official exchange rate from IFS, IMF)	%

6. Institutional change

(1) Of investment

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

(2) Of industrial output

$ind_{SOE}\_ind_{TOTAL}$	Output value of state-owned enterprises / Gross industrial output value	%
$ind_{COL}\_ind_{TOTAL}$	Output value of collective enterprises / Gross industrial output value	%
$ind_{PRIV}\_ind_{TOTAL}$	Output value of private enterprises / Gross industrial output value	%

(3) Of employment

$wok_{SOE}\_wok_{TOTAL}$	State-owned enterprise workers / Total staff and workers	%
$Wok_{COL}\_wok_{TOTAL}$	Collective enterprise workers / Total staff and workers	%
$Wok_{PRIV}\_wok_{TOTAL}$	Private enterprise workers / Total staff and workers	%

7. Sectoral change

(1) Temple and Wöβmann (2006)'s specification

`s`	Agricultural share of GDP (Primary sector GDP / Total GDP)	%
`a`	Agricultural share of employment (Primary sector employment / Total number of employed persons)	%
`m`	Non-agricultural share of employment (`m=1-a`)	%
`p`	Migration propensity (`p=-(da/dt)/a`)	
`MGROWTH`	Linear sectoral change term: Change of non-agricultural share of employment (`dm/dt`)	%
`DISEQ`	Non-linear sectoral change term: Change of non-agricultural share of employment adjusted by migration propensity (`p/(1-p)*(dm/dt)`)	%
`MGROWTH2`	Linear sectoral change term: Change of non-agricultural share of employment * Average labour productivity in agricultural sector (`(dm/dt)*s/a`)	%
`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)*(dm/dt)*s/a`)	%

(2) Dowrick and Gemmell (1991) or Poirson (2001)'s specification

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)	%

(3) Degree of industrialization

deofin	Degree of industrialization (Gross industrial output value / (Gross industrial output value + Gross agricultural output value))	%
gr_deofin	Growth rate of degree of industrialization	%

8. Infrastructure

railway_area	Mileages of railways per square kilometre (Total railway length / Area)	%
highway_area	Mileages of highways per square kilometre (Total highways length / Area)	%
post&tele_gdp	Business volume of post and telecommunication / GDP	%

9. Financial development

loan_gdp	Total bank loan outstanding / GDP	%
saving_gdp	Savings deposit in urban and rural areas / GDP	%

10. Geographic location

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

o Because of the gap between pre- and post-census GDP data, we treat any observation of annual growth rate of provincial GDP per capita above - / + 25% as an outlier. As a result, 33 observations are removed before the first-stage cross-sectional analysis and before calculating the five-year averages for subsequent panel estimation.

o The population data at the province level are problematic as some provinces report household registration population whereas others report permanent population only. We therefore treat any observation of annual population growth rate above - / + 8% as an outlier. For population natural growth rate (pop_ngr), one observation is removed; and for population growth rate (pop_gr), 38 observations are removed.

o As for the employment data, data before 1998 are the figures of all staff and workers for each province whereas after 1998 the figures include on-post staffs and workers only. We therefore extrapolate to make the series consistent before and after 1998.

o Regarding the gross industrial output value, before 1998 the data are for all independent account enterprises, whereas after 1998 only enterprises above designated size (annual sales income of over 5 million RMB) are covered. We therefore extrapolate to make the series consistent before and after 1998.

o The business volume of post and communication are reported at 1990 constant prices during 1978-99 and at 2000 constant prices thereafter. We therefore use our own price deflator to transform all values to 1990 constant prices.

Appendix B: 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) and Malik and Temple (2009).

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), Δ , given data D. There are a large number of possible statistical models, M_1 , ..., 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 Δ given data D is

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

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

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

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

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

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

where θ_k is the parameter vector of model M_k , $P\,(D\mid\theta_k,\,M_k)$ is the likelihood of θ_k under model M_k , and $P\,(\theta_k\mid M_k)$ is the prior distribution over the parameter space associated with model M_k . The integrated likelihood $P\,(D\mid M_k)$ 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 $P\,(\theta_k\mid M_k)$. To represent no prior preference for any model, each model can be 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

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

(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 studies of 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 informative, and 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 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 to *Pcgets* in terms of power, but had better size performance in some cases.

TABLE 1

Descriptive statistics of provincial GDP per capita growth rates

	Full-sample period	*Sub-sample periods*		
	1978-2007	*1978-1987*	*1988-1997*	*1998-2007*
All provinces (30 provinces)	0.077	0.072	0.054	0.106
	(0.037)	(0.027)	(0.031)	(0.032)
Coastal provinces (11 provinces)	0.081	0.078	0.061	0.119
	(0.033)	(0.027)	(0.030)	(0.028)
Interior provinces (19 provinces)	0.075	0.073	0.055	0.109
	(0.038)	(0.026)	(0.032)	(0.034)
Highest growth province	0.103	0.112	0.108	0.131
	(0.024)	(0.001)	(0.046)	(0.055)
Lowest growth province	0.061	0.019	0.011	0.082
	(0.049)	(0.007)	(0.007)	(0.040)

Notes: 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, Shaanxi were the highest growth provinces respectively, and Shanghai, Tibet, Yunnan were the corresponding lowest growth provinces.

TABLE 2

Bayesian Model Averaging (BMA) results

Regressor	*Posterior probability of inclusion*	*Posterior mean*	*Posterior standard deviation*
$Intercept$	100.0	0.223	0.036
$lny_{i,t-1}$	100.0	-0.021	0.005
ind_{SOE}/ind_{TOTAL}	100.0	-0.064	0.013
stu_{SEC}/pop	100.0	0.483	0.135
fcf/GDP	69.2	0.035	0.031
pop_ngr	59.9	-0.859	0.917
stu_{HIGH}/stu_{REG_SEC}	36.3	0.024	0.041
$export/GDP$	27.6	0.007	0.015
$railway\_area$	20.2	-0.020	0.056
$loan/GDP$	8.5	-0.001	0.003
$dum\_coastal$	8.4	0.001	0.002
$mgrowth$	7.7	-0.005	0.136
$post\,\&\,tele/GDP$	6.9	-0.002	0.025
$deofin$	5.6	-0.001	0.004

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

TABLE 3

General-to-Specific (GETS) model selection results

Regressor	Coefficient	Standard error	t-value	t-probability	Part.R^2
$Constant$	0.249	0.029	8.35	0.000	0.752
$lny_{i,t-1}$	-0.025	0.004	-5.45	0.000	0.564
fcf/GDP	0.059	0.021	2.86	0.009	0.262
stu_{SEC}/pop	0.418	0.122	3.44	0.002	0.339
pop_ngr	-1.823	0.701	-2.60	0.016	0.227
$export/GDP$	0.025	0.018	1.43	0.167	0.081
ind_{SOE}/ind_{TOTAL}	-0.055	0.012	-4.31	0.000	0.446
Sigma	0.006	RSS	0.001	R^2	0.845
F(6,23)	20.97 [0.000]	LogLik	115.804	AIC	-7.197
Normality test			Chi^2(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, T=30; Dependent variable: growth rate of real provincial GDP per capita; RSS: residual sum of squares; F(6,23): joint significance test; LogLik: log-likelihood; and AIC: Akaike's information criterion.

TABLE 4

Panel data estimation of the selected baseline model

Regressors	OLS	Within Groups	IV (2SLS)	SYS-GMM
$lny_{i,t-1}$	-0.035** (0.004)	-0.061** (0.012)	-0.074** (0.018)	-0.046** (0.007)
fcf/GDP	0.074** (0.016)	0.067** (0.026)	0.227** (0.100)	0.093** (0.026)
stu_{SEC}/pop	0.463** (0.139)	0.035 (0.253)	0.898* (0.500)	1.008** (0.284)
pop_ngr	-3.095** (0.597)	-3.843** (1.162)	-3.475* (2.078)	-4.057** (1.036)
$export/GDP$	0.027** (0.011)	0.005 (0.016)	0.075 (0.061)	0.044** (0.018)
ind_{SOE}/ind_{TOTAL}	-0.052** (0.011)	-0.053** (0.025)	0.017 (0.084)	-0.054** (0.021)
R^2	0.746	0.784		
$AR(2)$ p $value$				0.866
$Hansen$ p $value$			0.352	0.665
Dif $Sargan$ p $value$				0.249
$No.of$ $observations$	150	150	149	150

Notes: 5-year interval panel data is used for estimation and all time dummies are included but not reported to save space; standard errors are in parentheses, which are heteroskedasticity-consistent and clustering on province; in both the IV and system GMM estimation, $lny_{i,t-1}$ 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 or 10% significance level respectively.

TABLE 5

The growth impact of physical capital formation

Panel 1. Investment in capital construction, in innovation, and in other fixed assets

	$finv_{Total}/GDP$	$finv_{CC}/GDP$	$finv_{INNO}/GDP$	$finv_{OTHER}/GDP$
Coefficients	0.184 ** (0.042)	0.174** (0.047)	0.277** (0.082)	0.077 (0.111)
$AR(2)$ *p value*	0.636	0.778	0.810	0.911
Hansen p value	0.108	0.100	0.261	0.167
Dif Sargan p value	0.607	0.862	0.367	0.922
No. of instruments	26	26	26	26
No. of observations	150	150	150	150

Panel 2. Total investment in fixed assets: by ownership classification

	$finv_{SOE}/finv_{Total}$	$finv_{COL}/finv_{Total}$	$finv_{PRIV}/finv_{Total}$
Coefficients	-0.129** (0.026)	-0.077 (0.111)	0.197** (0.096)
$AR(2)$ *p value*	0.610	0.919	0.953
Hansen p value	0.167	0.172	0.193
Dif Sargan p value	0.393	0.586	0.791
No. of instruments	26	26	26
No. of observations	149	149	149

Notes: panel data system GMM results are reported; standard errors are in parentheses, which are heteroskedasticity-consistent and clustering on province; the control variables are those selected by the model selection procedures, i.e. $lny_{i,t-1}$, stu_{SEC}/pop, pop_ngr, $export/GDP$ and ind_{SOE}/ind_{TOTAL}, in which $lny_{i,t-1}$ is treated as pre-determined, pop_ngr is treated as exogenous, and all other variables are treated as endogenous; time dummies are included; ** and * indicate that the coefficient is significantly different from zero at the 5 or 10% significance level respectively.

TABLE 6

The growth impact of human capital formation

Panel 1. Primary and secondary enrolments

	stu_{PRIM}/pop	stu_{SEC}/pop	stu_{REG_SEC}/pop
Coefficients	-0.611** (0.254)	1.538** (0.326)	2.171** (0.337)
$AR(2)$ *p value*	0.186	0.523	0.913
Hansen p value	0.138	0.179	0.203
Dif Sargan p value	0.895	0.706	0.905
No. of instruments	26	26	26
No. of observations	149	149	149

Panel 2. Higher education enrolments

	stu_{HIGH}/pop	$stu_{UNI\,\&\,COL}/pop$	$stu_{SEC\,\&\,HIGH}/pop$
Coefficients	3.568** (0.579)	2.201** (0.617)	1.953** (0.285)
$AR(2)$ *p value*	0.286	0.155	0.803
Hansen p value	0.232	0.196	0.225
Dif Sargan p value	0.514	0.653	0.801
No. of instruments	26	26	26
No. of observations	149	147	149

Notes: panel data system GMM results are reported; standard errors are in parentheses, which are heteroskedasticity-consistent and clustering on province; the control variables are those selected by the model selection procedures, i.e. $lny_{i,t-1}$, fcf/GDP, pop_ngr, $export/GDP$ and ind_{SOE}/ind_{TOTAL}, in which $lny_{i,t-1}$ is treated as pre-determined, pop_ngr is treated as exogenous, and all other variables are treated as endogenous; time dummies are included; ** and * indicate that the coefficient is significantly different from zero at the 5 or 10% significance level respectively.

TABLE 7

*Counterfactual predictions of growth rate of GDP per capita (baseline model)*

Predicted growth rates of GDP per capita (Unit: pps)	stu_{SEC}/pop			
	Mean (5.79 pp)	Reduce by 1 pp	Reduce by 2 pps	Reduce to the mean of LDCs (2.26 pps)
fcf/GDP				
Mean (34.27 pps)	8.03	7.02	6.02	4.48
Reduce by 1 pp	7.94	6.93	5.92	4.39
Reduce by 5 pps	7.57	6.56	5.55	4.01
Reduce by 10 pps	7.10	6.10	5.09	3.55
Reduce to the mean of LDCs (17.76 pps)	6.50	5.49	4.48	2.95

Notes: *pp(s)* refers to percentage point(s); LDCs refer to the least developed countries.

TABLE 8

Counterfactual predictions of growth rate of GDP per capita (other models)

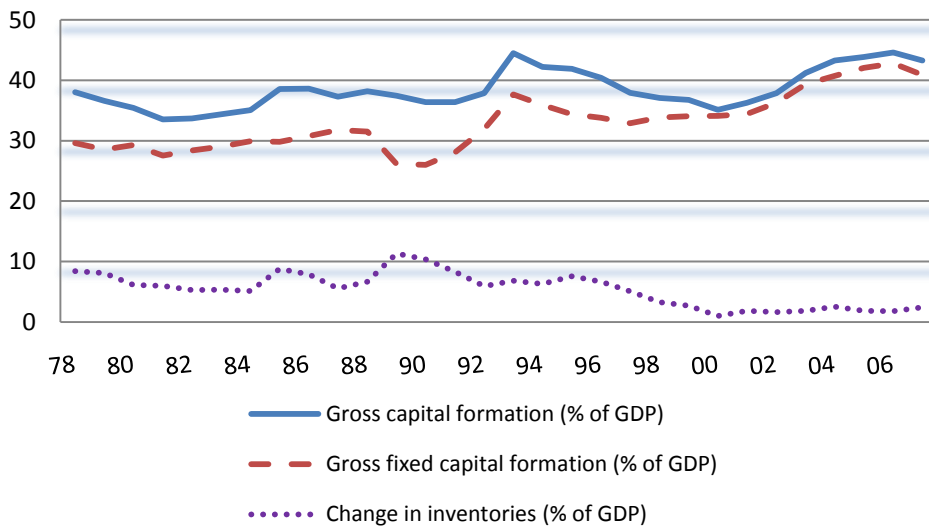| Predicted growth rates of GDP per capita (Unit: pps)		stu_{SEC}/pop			
		Mean (5.79 pps)	Reduce by 1 pp	Reduce by 2 pps	Reduce to half of the mean
Innovation investment (Panel 1, Table 5)					
$finv_{INNO}/GDP$	Mean (6.85 pps)	8.03	6.76	5.50	4.36
	Reduce by 1 pp	7.76	6.49	5.22	4.09
	Reduce by 2 pps	7.48	6.21	4.94	3.81
	Reduce by 3 pps	7.20	5.93	4.66	3.53
	Reduce to half of the mean	7.08	5.81	4.55	3.41
Private investment (Panel 2, Table 5)					
$finv_{PRIV}/$	Mean (16.7 pps)	8.06	5.30	4.16	3.15
$finv_{Total}$	Reduce by 1 pp	7.85	5.13	4.00	2.98
	Reduce by 3 pps	7.46	4.74	3.60	2.58
	Reduce to half of the mean	6.55	3.83	2.70	1.68
	Reduce to 1978 Mean (5.06 pps)	6.05	3.34	2.20	1.18
*Higher education enrolments (Table 6)**					
stu_{HIGH}/pop	Mean (0.44 pps)	8.06	6.09	4.12	2.96
	Reduce by 0.1 pps	7.74	5.76	3.79	2.63
	Reduce by 0.2 pps	7.41	5.42	3.45	2.30
	Reduce to half of the mean	7.34	5.35	3.38	2.22
	Reduce by 0.3 pps	7.08	5.09	3.12	1.96

Notes: *pp(s)* refers to percentage point(s); * the regression that simultaneously includes both secondary and higher education enrolments is estimated but not reported in Table 6 to save space, but the estimated coefficients of the two variables are similar to the case when they enter the regression separately.

Figure 1. China's Annual Growth Rate of GDP Per Capita (%)
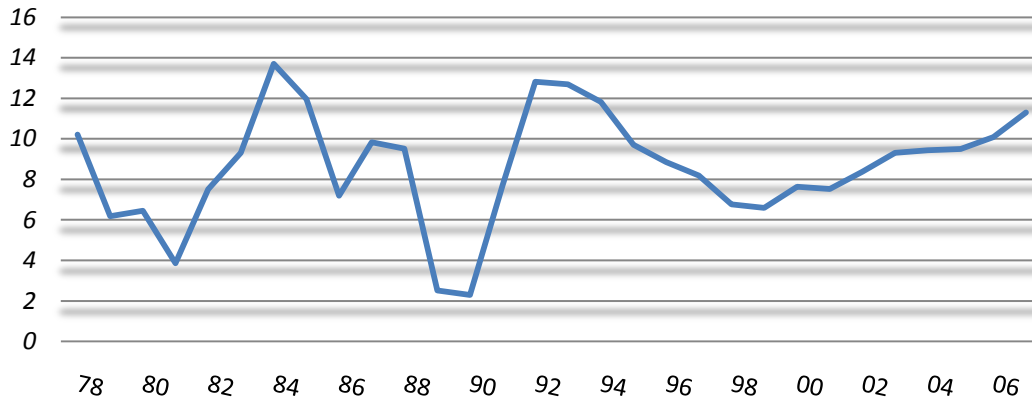
Data source: World Bank, *World Development Indicators* (June 2009 edition).

Figure 2. Gross Capital Formation and Its Composition

— Gross capital formation (% of GDP)

- - - Gross fixed capital formation (% of GDP)

····· Change in inventories (% of GDP)

Data source: World Bank, *World Development Indicators* (June 2009 edition).