Regional Growth in China: An Empirical Investigation using Multiple Imputation and Province-level Panel Data

Chen, Baizhu and Phillips, Kerk L.

Marshall School of Business, University of Southern California, Brigham Young University

June 2008

Regional Growth in China:
An Empirical Investigation using Multiple Imputation and Province-level Panel Data

Kerk L. Phillips*

Department of Economics

P.O. Box 22363

Brigham Young University

Provo, UT 84602-2363

phone: (801) 378-5928

fax: (801) 378-2844

email: kerk_phillips@byu.edu

Baizhu Chen

Department of Finance and Business Economics

Marshall School of Business

University of Southern California

Los Angeles, CA 90089-1421

phone: (213) 740-7558

fax: (213) 740-6650

email: baizhu@marshall.usc.edu

June 2008

Keywords: growth, provinces, empirical, panel-data
JEL codes: H1, O0

* The financial support of the David M. Kennedy Center for International Studies and of the College of Family Home and Social Science at Brigham Young University is gratefully acknowledged. We thank Gene Chang, Gan Li, and David Li for helpful comments.

Abstract

This paper examines the contributions of various factors to China's economic growth. The methodology is discussed in papers by Levine and Renelt (1992) and Sala-i-Martin (1997). Using multiple imputation techniques on a panel data from 1978 to 1999 for 30 provinces, autonomous regions, and independently administered cities, we find that provinces with more innovation capital and more bank-deposit-to-GDP ratios tend to experience higher economic growth. Migration of people into a province, the number of higher education teachers, railroad density & local government revenue as a percent of total government spending are all negatively related to subsequent growth rates.

1. Introduction

One of the most important events in recent history is the emergence of China as an economic power. Growth rates since the reforms of the 1970s have been very high. The exact causes of this rapid growth are still being debated even as standards of living in China continue to rise. Growth has not, however, been uniform across regions and provinces in China. In fact, regional disparities in levels of per capita income are much higher in China than in many other countries. For example, the Gini coefficient for Chinese provinces[1] in 1998 was .25. In contrast, a similarly calculated Gini coefficient for states in the U.S. for 2000 is .08, and is .05 across European Union members for 2001. Not only are levels of GDP different, but growth rates of GDP are also uneven across provinces, with the most rapid growth in the costal provinces of Jiangsu, Zejiang, Fujian and Guangzhou. If we can come to understand the causes for regional disparities in growth rates and levels of per capita income, we will improve our understanding of the overall process of growth in China. A better understanding of China's growth will not only contribute to our understanding of long-run economic growth in general, but will also benefit policy makers in China and other countries.

Concurrent with the impressive growth in China, there has been a rising interest in describing and understanding its nature. Chow (1993) was the first to attempt a quantitative decomposition of the factors of the Chinese growth. He concluded that capital formation, and not technological progress, played the principle role in China's economic growth from 1952 to 1980. Borensztein and Ostry (1996) also performed growth accounting on Chinese data, and found that productivity has been the driving force for China's economic growth since the economic reforms of 1978. Chen and Fleisher (1996) compared the total factor productivity of China's coastal and noncoastal provinces. They found that domestic investment in higher education and foreign direct

[1] Using publicly available population and GDP data and treating per capita values within provinces as identical for all residents.

investment helped to explain the gap in productivity and long-term growth between these regions. In a cross-province study based on post-reform data, Chen and Feng (2000) found that factors contributing to the difference of growth rates included human capital, the degree of openness, the share of state-owned enterprises, and the fertility rate. Bao et. al. (2002), on the other hand, found that geographic location is the dominant explanation for the divergent provincial economic growth rates in China. These studies all suggest that there are many variables that could be important driving factors behind China's economic growth.

One criticism leveled at the early empirical growth literature was its *ad hoc* nature of including explanatory variables. Numerous empirical studies since the seminal work of Barro (1991), and Barro and Lee (1993) have identified a substantial number of variables that can explain long run economic growth. Often, though, many of these variables lose significance when certain other variables are included in the regression model. Upon applying Leamer's (1985) extreme-bounds test, Levine and Renelt (1992) indeed conclude that very few or no variables are able to robustly explain the long run growth rate. This result was a serious challenge to the existing empirical growth literature. Sala-i-Martin (1997), on the other hand, argues that the methodology employed by Levine and Renelt (1992) is too restrictive, particularly when there are large numbers of potential regressors. He therefore redesigned the test to look at the entire distribution of a model's coefficient estimators to calculate confidence levels. Chinese growth literature is similarly subject to criticisms of the ad hoc nature of its regression functions. To confidently identify the factors that drive economic growth in China, it is necessary to apply a methodology similar to that of Sala-i-Martin (1997) to the Chinese data. This is the objective of this paper.

Section 2 describes the dataset. Section 3 illustrates the methodology for carrying out the test. We provide the test results in Section 4. Section 5 concludes.

2. Data Set

In the past two decades, there has been a blossoming of research on economic growth. Much of this work has been empirical in nature, and the bulk of that has used data from cross-country regressions. Advances in statistical theory and increases in available computing power have made it possible to move away from cross-sectional studies which use long-run (30-year averages) growth across a sample of several dozen countries. Instead, focus has begun to shift to panel regressions that utilize data from several countries observed at several points in time.[2]

Consistent with this emphasis, our dataset consists of various data taken from Chinese statistical publications that are compiled at the provincial level every year. Our sample runs from 1978 to 1999 and includes 30 provinces, autonomous regions, and independently administered cities. The city of Chongqing was made independent from Sichuan province in 1996. We aggregate these two regions for 1996-99, making it consistent with earlier observations.

Most of our data are all ultimately traceable to the National Bureau of Statistics, though they have come to us through a number of different channels. The CD-ROM Fifty Years of Chinese Statistical Data was a primary source. As were the English/Chinese language China Statistical Yearbook in various printed and CD-ROM editions. Hsueh et. al. (1993)[3] was an excellent source for filling in missing data from these first two sources. Finally, we obtained the data on bank deposits used in Li & Liu (2001) and filled in some remaining missing data from the printed versions of provincial statistical yearbooks at the Nanjing University library which were used in Phillips & Shen (2005).

---

[2] Some recent studies, for example, Paap et. al. (2005), Kim (2005), and Apergis (2005), attempt to use panel time series analyses to provide better information on economic growth.
[3] An extremely helpful source of provincial data up to 1989.

We gathered data on as many series as we could find that could arguably be important for economic growth and development. There are, of course, literally thousands of kinds of data that fit such a broad criterion. However, the need for consistently reported data from all or most provinces for the bulk of the sample period turns out to be a great winnower of data. After adjustments[4], we end up with the 49 series reported in Table 1.

These variables can be broadly classified into eleven categories:

(i) Convergence: value of real GDP per capita from the previous year.

(ii) Investment: real investment per capita, real fixed investment per capita, and percentage of fixed capital investment classified as "construction of fixed capital."

(iii) State-owned Enterprises: SOE staff & workers as percentage of total employment, SEO industrial output as percentage of total industrial output, value of SOE construction as percentage of total construction value, real SOE industrial output, growth of real SOE industrial output, etc.

(iv) Demographics: annual growth of population, and percent of population that is male, and percent of population classified as agricultural.

(v) Openness: net exports as percentage of GDP, real value of foreign capital actually utilized per capita, real value of foreign loans, and real value of foreign direct investment.

(vi) Financial: national bank deposits as percentage of GDP, and enterprise bank deposits as percentage of GDP.

(vii) Education: Primary school students enrolled as percentage of population, secondary school students enrolled as percentage of population, secondary school teachers as percentage of population,

[4] Such as converting to per capita levels or growth rates, etc.

(viii) Health: Health institutions per capita, hospital beds per capita, etc.

(ix) Infrastructure: railroad per square kilometer of area, highway per square kilometer of area, and telephones per capita.

(x) Government: Government consumption as percentage of GDP, local government revenue as percentage of government consumption, etc.

(xi) Migration: Implied population rate migrating into a province per 1000[5].

We are able to gather reasonably complete data for the variables listed in Table 1. We have double-checked this data for accuracy, and, in cases where there are obvious, but uncorrectable, errors, have omitted the flawed observations. With 21 years and 30 provinces, we have potentially 630 observations, though we often have less than that in practice due to missing data in the original sources.[6]

3. Methodology

We have a large number of potential regressors, many of which are likely unrelated to provincial growth. In an attempt to reduce the number of potential regressors we first run a series of OLS regressions with the annual growth of GDP per capita as the dependent variable. We include each of our 49 potential regressors as independent varianbles in separate regressions. We also include the level of GDP per capita from the previous period. This is to control for convergence in the same way that Levin & Renelt and Sala-i-Martin include the initial level of GDP per capita. To control for panel data fixed effects, we include a set of province and year dummy variables in every regression. Since it is not reasonable to assume province or time-period error terms are uncorrelated with our regressors, random effects estimation techniques are inappropriate. We include

---

[5] Imputed from year-to-year changes in population along with birth and death rates.
[6] The data set is available upon request.

only lagged values of regressors to ensure that our results are not driven by reverse causality.[7]

We have a serious problem with missing observations in our dataset. In order to compare the results of thousands of regression combinations, it is essential that we use the same sample set for all of them. However, if we restrict ourselves to observations where data is available for all regressors, we lose almost two-thirds of our data (257 available observations out of 630). This method ignores useful information from non-missing regressors. We therefore also estimate (3.1) using multiple imputation techniques developed by Rubin (1987) and Schafer (1997). We use imputation by chained equations[8]. We use the full set of potential regressors, including province and year dummies in the chained equations and do 10 imputations of missing values.

Multiple imputation involves generating imputations of missing values. A single imputation is inappropriate as it does not reflect the uncertainty behind the imputed missing values. Multiple imputation solves this problem by imputing the data several times. A model of the data is estimated using the complete sample and then missing observations are generated based on this model using the non-missing data to conditionally predict the missing values for a given observation. This is done using Monte Carlo methods where the missing observations are generated using random draws of residuals from the complete sample estimation.

Table 2 reports the results of these OLS regressions. We find 11 of our potential regressors are significant using either the complete dataset or using multiple imputation. Five of these are significant using both datasets.

We also regressed the annual growth of GDP per capita on the full set of all 49 potential regressors, the level of GDP per capita from the previous period and a set of

[7] This does not, however, rule out they are driven by the dependent variable and the regressor both being driven by some unobserved factor with different lags.
[8] We STATA's multiple imputation package, specifically the "ice" command.

province and year dummies. These results are reported in table 3. Here 23 of our 49 potential regressors are significant using the complete dataset or multiple imputation and four are significant in both.

We cull our list of potential regressors by dropping those that were not significant by any of the regressions run. This leaves us with the set of 26 regressors marked with asterisks in table 1.

Levine & Renelt (1992) showed that very few things can be said to robustly explain growth. Using a cross-section of 119 countries, they include the initial level of GDP per capita, the average annual growth rate of the population, the initial secondary-school enrollment rate, and the investment share in GDP in all regressions and run a series of regressions for various conceptual subsets of the remaining regressors. Their definition of "robust" relies on Leamer's (1985) extreme bounds tests. By this criterion, a variable is considered robust only if it is statistically significant in all regressions[9]. They find initial GDP per capita, investment as a percentage of GDP, and secondary-school enrollment rates are the only robustly significant variables in their dataset. Other variables can be shown to be sometimes significant and other times insignificant, depending on exactly which set of explanatory factors are used.

Sala-i-Martin (1997) argues that the extreme bounds test is unreasonably restrictive, especially in cases where there are large numbers of potential regressors and all possible permutations are analyzed. For example, if there are 10 potential regressors all combinations of four are examined, any given coefficient must be significant in all 84 of the permuations in which it is included to be considered robustly significant. However,

[9] In an extreme-bounds test, one calculates the lower extreme bound as the lowest value of $\beta_{yj} - 2\sigma_{yj},$ and the upper extreme bound as the largest value of $\beta_{yj} + 2\sigma_{yj}$, for all the possible regressions. The extreme bounds test for variable *y* says that if the lower extreme bound is negative and the upper extreme bound is positive, then variable *y* is not robust. This implies that as long as there is one regression for which the sign of the coefficient changes or is not significant, then the variable is not robustly influential.

with 20 potential regressors it must be significant in 969 permutaions. With 100 potential regressors the number is 156,849. The latter criterion is obviously more restrictive that the first. Sala-i-Martin argues for using rejections rates and t-tests based on averages of coefficients and standard errors across regression permuations. He suggests weighting averages by the value of the regression's log-likelihood, so that statistics from better fitting regressions carry more weight than those from poorly fitting ones. When these criteria are used, many of variables can be said to have robust effects on growth. A large portion of these variables are national in nature, however. That is, they each have a roughly equal impact on all regions within a country. Examples of such are: variability of inflation rates, degree of property right enforcement, financial market efficiency, etc. One of the challenges this paper faces is that of determining which factors might explain differences *within* a country.

The form for the regression is given below.

$$g_{it} = F_{it}\beta_F + x_{it-1}\beta_x + \varepsilon_{it} \tag{3.1}$$

where g_{it} is the growth of GDP per capita, F_{it} is a vector of province & time-period dummy variables and the lagged value of real GDP per capita[10], x_{it} is a 1x4 vector of regressors, and β_F & β_x are the corresponding coefficient vectors.

4. Results of Estimation

We estimate equation (3.1) using OLS for all possible permutations of four regressors in our culled list of 26. We use the complete dataset (with no missing observations) and estimate again using multiple implementation[11]. This gives 2300

[10] Growth literature robustly found a convergence effect - that an economy with lower income tends to grow faster, other things constant. The lagged value of real GDP per capita is included in each regression to control for convergence effects.

[11] We do not report the results from multiple imputation because none of the regressors were found to be robustly significant. Since multiple imputation uses the correlation between regressors to impute missing values, the regressors are multicolinear by construction. The multicolinearity is so strong that none of the regressors is robustly significant.

regressions for each independent variable[12]. To examine significance and robustness, we adopt three different approaches following Sala-i-Martin.

First, we test the significance of β_y for each regression. We do this with a simple t-test[13]. We then tally the number of times we reject the null hypothesis that β_y is zero. These results are shown in Table 4. We show the percentage of rejections for two-tailed tests at 90%, 95% and 99% confidence levels. Table 4 is one step away from the restrictive extreme-bounds test in that it incorporates more information about the distribution of the estimate of β_y.

Our second & third tests are based on simple and log-likelihood weighted averages of the coefficient estimates and standard errors. t-statistics are calculated using these averages. The average estimates, t-statistics and significance from two-tailed tests are reported in Tables 5 and 6.

Based on tables 4-6, we identify six regressors as being robustly significant in explaining provincial growth. These are:

- INNINVP % of fixed capital investment classified as "innovation capital"
- DEPGDPP national bank deposits as % of GDP
- HETPC higher education teachers as % of population
- RRDDEN km of railroad per sq km of area
- LGRGOVP local gov't revenue as % of gov't consumption
- MRATE implied migration rate into a province

We now turn our attention to the effects of each of these regressors. We regress the annual growth of GDP per capita on these six regressors, the level of GDP per capita from the previous period and a set of province and year dummies. As before we use the complete sample and multiple imputation. The results are shown in table 7. We find that

[12] And a total of 14,950 regressions.
[13] In this case and in all cases below we use the heteroskedasticity-consistent estimates of the standard errors.

INNINVP, DEPGDPP and MRATE are significant in both cases. RRDDEN and LGRGOVP are significant in the complete sample. HETPC is significant only when multiple imputation is used.

The percent of investment in innovative capital has a positive effect on growth. The coefficient indicates that a one percentage point increase in innovation capital's share of total investment will raise a province's subsequent growth rate by almost two-tenths of a percent. Obviously, the effect this actually has on growth depends on how much innovation capital investment changes. To get an idea of the importance of innovation capital we calculate a standardized effect by multiplying the coefficient by the observed standard deviation in INNINVP. We report this in table 7 as the "standardized effect." The effect of a one standard deviation increase in INNINVP is an increase in a province's growth rate by .9 to 1.3 percentage points. We find it interesting that the innovation portion of investment is so highly correlated with growth while the overall level of investment is not. This may be due to classifying investment occurring via joint ventures with international corporations as innovation. The result could be driven by the resultant transfer of technology.

National bank deposits as a percent of GDP are also positively correlated with growth. The standardized effect is an increase in a province's growth of between 2.1 and 3.1 percentage points with a one standard deviation increase in DEPGDPP. These results are consistent a more sophisticated financial intermediation system leading to a more efficient allocation of factors of production..

The implied migration rate is negatively associated with growth, but the standardized effect is relatively small; less than one percentage point. Since the standard errors are very high, the small effect comes from very small coefficients, indicating that while the correlation is high, migration rates do not have large impacts on provincial growth rates. A negative impact can best be explained by the fact the increased migration increases the number of people and therefore decreases output per person. The small size indicates that increased migration actually raises total GDP, however.

Railroad density is negatively correlated with growth. The standardized effect is a decrease of almost two percentage points. We find this surprising.

Local government revenue as a percent of government spending in a province is also negatively correlated with growth having a standardized effect of a drop of 1.1 percentage points. This result was documented in earlier work by Phillips & Shen (2005) and is likely driven by the negative effects of taxes on investment.

Finally, the number of higher-education teachers as a percent of the population is also associated with lower rates of growth. The standardized effects are very large, though they are not significantly negative for the complete sample. Using multiple imputation we get a standardized effect of a drop of 3.4 percentage points. Since higher education teachers are a very small portion of the population (fewer than 6 per 1000 people on average), the result cannot be explained by the notion that more teachers means fewer production workers and therefore less output. We are puzzled by this result as well.

5. Conclusions

This paper has examined the differences of growth across cities and provinces in the People's Republic of China for the period 1978 – 1997. Because of sparse data and missing observations that to not match up across our various variables, we have used multiple imputation techniques developed by Rubin (1987) and Schafer (1997) to estimate the effects of a variety of potential regional variables on the growth rate of regions. We find that provinces with more innovation capital and more bank-deposit-to-GDP ratios tend to experience higher economic growth. Migration of people into a province, the number of higher education teachers, railroad density & local government revenue as a percent of total government spending are all negatively related to subsequent growth rates.

Table 1
Adjusted Data used in Regressions, 30 provinces, 1978 – 1997

Variable	Description	Units	Category
GRGDPPC	annual growth of real GDP per capita	percent	dependent
RGDPPC	real GDP per capita	10,000 1995 RMB	all regressions
RINVPC	real investment per capita	1995 RMB	i
RFINVPC	real fixed investment per capita	1995 RMB	i
CAPINVP	% of fixed capital investment classified as "constuction of fixed capital"	percent	i
INNINVP*	% of fixed capital investment classified as "innovation capital"	percent	i
SOEEMPP	SOE staff & workers as % of total employment	percent	ii
SOEPRODP*	SOE industrial output as % of total industrial output	percent	ii
SOECONP	Value of SOE construction as % of total construction value	percent	ii
SOERETP	SOE retail sales as % of total retail sales	percent	ii
SOEINVP	SOE investment as % of total fixed capital investment	percent	ii
RSOEIO*	real SOE industrial output	10,000 1995 RMB	ii
GRSOEIO	growth of real SOE industrial output	percent	ii
GPOP	annual growth of population	percent	iii
MPOPP	percent of population that is male	percent	iii
AGPOPP	percent of population classified as 'agricultural'	percent	iii
MRATE*	implied population migrating into a province	per 1000	iii
EXIM	sum of exports & imports as percent of GDP	percent	iv
RFCAUPC	real value of 'foreign capital actually utilized" per capita	RMB	iv
RFLONPC*	real value of foreign loans	RMB	iv
RFDIPC	real value of foreign direct investment	RMB	iv
DEPGDPP*	national bank deposits as % of GDP	percent	v
EDEPGDPP*	enterprise bank deposits as % of GDP	percent	v
PSEPC*	primary school students enrolled as % of population	percent	vi
SSEPC*	secondary school students enrolled as % of population	percent	vi
RSEPC	regular secondary school students enrolled as % of population	percent	vi
HEEPC	higher education students enrolled as % of population	percent	vi
PSTPC*	primary school teachers as % of population	percent	vi
SSTPC	secondary school teachers as % of population	percent	vi
RSTPC	regular secondary school teachers as % of population	percent	vi
HETPC*	higher education teachers as % of population	percent	vi
HINSPC*	health institutions per capita	number	vi
HOSPPC*	hospitals per capita	number	vi
HIBEDPC	health institution beds per capita	number	vi
HOBEDPC*	hospital beds per capita	number	vi
MEDPC*	medical technians per capita	number	vi
DOCPC*	doctors per capita	number	vi
RRDDEN*	km of railroad per sq km of area	number	vii
HWYDEN*	km of highway per sq km of area	number	vii
TELPC*	telephones per capita	number	vii
GOVTGDPP	govt consumption as % of GDP	percent	viii
IISGDPP*	change in inventories as % of GDP	percent	viii
LGRGOVP*	local gov't revenue as % of gov't consumption	percent	viii
LGEGOVP*	local gov't expenses as % of gov't consumption	percent	viii
LTAXLGRP*	local gov't taxes revenue as % of local gov't revenue	percent	viii
LGCCLGEP*	local gov't capital consumption as % of local gov't expenditures	percent	viii
LGINLGEP	local gov't innovation investment as % of local gov't expenditures	percent	viii
LGAGLGEP	local gov't agricultural supports as % of local gov't expenditures	percent	viii
LGOTLGEP*	local gov't other expenses as % of local gov't expenditures	percent	viii
LGADLGEP	local gov't administrative expenses as % of local gov't expenditures	percent	viii

* robustly significant by at least one criterion

Table 2
Results of Single-Regressor Estimations

Variable	Category	Complete Data coefficient	Complete Data s.e.	Complete Data p-value	Multiple Imputation coefficient	Multiple Imputation s.e.	Multiple Imputation p-value
RGDPPC	i	1.340E-07	2.350E-07	0.569	6.790E-08	1.910E-07	0.723
RINVPC	ii	1.430E-07	3.210E-07	0.657	2.080E-07	2.450E-07	0.396
RFINVPC	ii	1.800E-07	3.260E-07	0.582	2.270E-07	2.580E-07	0.380
CAPINVP	ii	-0.0730	0.0731	0.319	-0.0360	0.0323	0.266
INNINVP	ii	0.3886	0.0967	0.000	0.1221	0.0753	0.105
SOEEMPP	iii	-0.1640	0.1338	0.222	-0.1154	0.1080	0.286
SOEPRODP	iii	-0.0916	0.0507	0.072	-0.0310	0.0305	0.310
SOECONP	iii	0.0275	0.0381	0.471	0.0020	0.0210	0.925
SOERETP	iii	0.1389	0.0924	0.134	0.0166	0.0350	0.634
SOEINVP	iii	0.0582	0.0639	0.363	-0.0165	0.0292	0.574
RSOEIO	iii	1.230E-05	1.540E-05	0.424	8.210E-07	1.040E-05	0.937
GRSOEIO	iii	0.0400	0.0323	0.217	0.0288	0.0283	0.311
GPOP	iv	-0.0183	0.5836	0.975	0.0088	0.1277	0.945
MPOPP	iv	0.1680	1.0280	0.870	-0.4242	0.6126	0.489
AGPOPP	iv	-0.0530	0.1370	0.699	0.0329	0.0961	0.732
EXIM	v	0.0606	0.0445	0.174	0.0300	0.0216	0.166
RFCAUPC	v	9.570E-06	6.400E-06	0.136	2.310E-06	7.630E-06	0.762
RFLONPC	v	3.200E-05	1.760E-05	0.071	3.750E-06	1.890E-05	0.843
RFDIPC	v	1.050E-05	8.920E-06	0.243	3.450E-06	9.910E-06	0.728
DEPGDPP	vi	0.0775	0.0264	0.004	0.0601	0.0226	0.008
EDEPGDPP	vi	0.0971	0.0363	0.008	0.0766	0.0324	0.018
PSEPC	vii	0.5798	0.3138	0.066	0.3502	0.2140	0.102
SSEPC	vii	0.3536	0.5272	0.503	0.3587	0.2991	0.231
RSEPC	vii	-0.0070	0.6476	0.991	0.0617	0.3003	0.837
HEEPC	vii	-2.2678	9.0102	0.802	3.5154	3.4662	0.311
PSTPC	vii	-0.0003	0.0013	0.838	0.0008	0.0009	0.367
SSTPC	vii	-0.0017	0.0015	0.256	-0.0008	0.0008	0.317
RSTPC	vii	-0.0016	0.0018	0.369	-0.0001	0.0008	0.931
HETPC	vii	-0.0132	0.0037	0.000	-0.0049	0.0030	0.110
HINSPC	viii	-0.0080	0.0071	0.262	0.0114	0.0068	0.093
HOSPPC	viii	0.0164	0.0226	0.468	0.0001	0.0005	0.854
HIBEDPC	viii	-20.3492	23.4401	0.386	8.2781	9.8564	0.401
HOBEDPC	viii	1.0519	17.8980	0.953	0.0167	0.1004	0.868
MEDPC	viii	7.3310	18.3944	0.691	10.6497	11.6869	0.363
DOCPC	viii	0.1576	27.7123	0.995	9.5304	15.6353	0.542
RRDDEN	ix	-1.1381	0.8414	0.178	-0.0762	0.1806	0.673
HWYDEN	ix	-0.0034	0.1008	0.973	-0.0050	0.0539	0.926
TELPC	ix	1.080E-05	6.090E-06	0.078	1.200E-05	5.330E-06	0.024
GOVTGDPP	x	0.2873	0.2253	0.204	0.0178	0.0980	0.856
IISGDPP	x	-0.0636	0.1195	0.595	-0.0138	0.0810	0.865
LGRGOVP	x	-0.0130	0.0038	0.001	-0.0043	0.0022	0.051
LGEGOVP	x	-0.0010	0.0126	0.938	0.0083	0.0075	0.265
LTAXLGRP	x	0.0244	0.0281	0.386	0.0017	0.0035	0.631
LGCCLGEP	x	0.0349	0.1460	0.811	-0.0112	0.0497	0.822
LGINLGEP	x	-0.0611	0.1164	0.600	0.0209	0.0797	0.793
LGAGLGEP	x	-0.0257	0.2019	0.899	0.0502	0.0694	0.470
LGOTLGEP	x	0.1010	0.1422	0.478	0.0680	0.0716	0.343
LGADLGEP	x	-0.0344	0.1596	0.829	-0.0024	0.0824	0.977
MRATE	xi	-0.0010	0.0002	0.000	-0.0012	0.0003	0.000

Table 3
Results of 49-Regressors Estimation

Variable	Category	Complete Data			Multiple Imputation		
		coefficient	s.e.	p-value	coefficient	s.e.	p-value
RGDPPC	i	-4.270E-06	1.460E-06	0.004	-2.190E-06	1.110E-06	0.049
RINVPC	ii	2.470E-06	2.280E-06	0.279	3.190E-06	2.220E-06	0.152
RFINVPC	ii	-4.160E-06	2.790E-06	0.138	-2.950E-06	2.300E-06	0.200
CAPINVP	ii	-0.0701	0.0846	0.409	-0.0014	0.0726	0.984
INNINVP	ii	0.4171	0.1163	0.000	0.1157	0.0897	0.198
SOEEMPP	iii	-0.4852	0.4369	0.268	-0.2997	0.2028	0.140
SOEPRODP	iii	-0.1646	0.0645	0.012	-0.0055	0.0517	0.915
SOECONP	iii	0.0661	0.0424	0.121	0.0044	0.0291	0.881
SOERETP	iii	0.0109	0.1203	0.928	0.0337	0.0452	0.455
SOEINVP	iii	-0.0218	0.0946	0.818	-0.0355	0.0686	0.605
RSOEIO	iii	7.910E-05	3.120E-05	0.012	1.690E-05	1.890E-05	0.372
GRSOEIO	iii	0.0200	0.0342	0.560	0.0182	0.0338	0.589
GPOP	iv	-0.1169	0.5565	0.834	0.0758	0.1380	0.583
MPOPP	iv	0.8494	1.4501	0.559	-0.5256	1.1381	0.644
AGPOPP	iv	0.2271	0.2567	0.377	0.2435	0.1570	0.122
EXIM	v	0.0726	0.0443	0.103	0.0533	0.0340	0.118
RFCAUPC	v	-9.050E-05	1.120E-04	0.420	-3.350E-05	6.730E-05	0.619
RFLONPC	v	8.780E-05	1.250E-04	0.483	4.600E-06	5.140E-05	0.929
RFDIPC	v	9.760E-05	1.180E-04	0.409	2.730E-05	7.690E-05	0.723
DEPGDPP	vi	0.1048	0.0661	0.114	0.0747	0.0401	0.063
EDEPGDPP	vi	-0.0956	0.0987	0.334	0.0090	0.0791	0.910
PSEPC	vii	0.4405	0.6266	0.483	-0.0147	0.3159	0.963
SSEPC	vii	-4.4663	2.6840	0.098	1.2530	1.2126	0.302
RSEPC	vii	3.3808	2.9726	0.257	-1.1497	1.1475	0.317
HEEPC	vii	1.4911	16.2200	0.927	-1.0726	9.3925	0.909
PSTPC	vii	-0.0039	0.0022	0.073	-6.473E-04	1.264E-03	0.609
SSTPC	vii	0.0088	0.0079	0.266	3.748E-04	2.378E-03	0.875
RSTPC	vii	0.0056	0.0077	0.468	0.0021	0.0018	0.261
HETPC	vii	-0.0268	0.0105	0.012	-0.0139	0.0062	0.027
HINSPC	viii	-0.0180	0.0107	0.094	0.0046	0.0068	0.502
HOSPPC	viii	0.0588	0.0401	0.145	0.0485	0.0238	0.042
HIBEDPC	viii	-23.2476	35.2772	0.511	7.2695	18.7043	0.698
HOBEDPC	viii	-15.1161	35.8905	0.674	-17.3040	8.4837	0.042
MEDPC	viii	137.2150	43.1493	0.002	33.1390	23.2788	0.155
DOCPC	viii	-106.4696	57.3950	0.065	7.6181	30.9918	0.806
RRDDEN	ix	-2.2929	1.1194	0.042	-0.2102	0.3290	0.523
HWYDEN	ix	-0.3654	0.1817	0.046	-0.0920	0.1727	0.594
TELPC	ix	1.499E-04	3.640E-05	0.000	3.340E-05	3.150E-05	0.289
GOVTGDPP	x	0.2255	0.2571	0.382	0.2706	0.2224	0.224
IISGDPP	x	-0.2150	0.1269	0.092	-0.1218	0.0965	0.207
LGRGOVP	x	-0.0260	0.0122	0.035	-0.0149	0.0049	0.002
LGEGOVP	x	0.0238	0.0243	0.329	0.0328	0.0135	0.016
LTAXLGRP	x	0.0977	0.0491	0.048	-0.0071	0.0068	0.295
LGCCLGEP	x	-0.3190	0.1528	0.038	0.0222	0.0808	0.784
LGINLGEP	x	-0.1097	0.1282	0.393	-0.0084	0.0926	0.928
LGAGLGEP	x	0.1251	0.2253	0.579	-0.0382	0.0865	0.659
LGOTLGEP	x	-0.2029	0.2026	0.318	0.1812	0.0941	0.055
LGADLGEP	x	-0.1795	0.2173	0.410	0.0234	0.1456	0.872
MRATE	xi	-0.0013	0.0003	0.000	-0.0012	0.0003	0.001

Table 4
Percentage of OLS Regressions where Variable is Significantly Different from Zero

Variable	90%	95%	99%
RGDPPC	24.13%	19.00%	5.39%
INNINVP	100.00%	100.00%	100.00%
SOEPRODP	0.00%	0.00%	0.00%
RSOEIO	6.35%	0.96%	0.00%
RFLONPC	3.91%	0.96%	0.00%
DEPGDPP	80.04%	65.91%	28.52%
EDEPGDPP	55.87%	35.13%	14.74%
PSEPC	0.00%	0.00%	0.00%
SSEPC	0.30%	0.04%	0.00%
PSTPC	1.39%	0.17%	0.00%
HETPC	99.91%	98.43%	87.91%
HINSPC	6.13%	2.30%	0.00%
HOSPPC	0.00%	0.00%	0.00%
HOBEDPC	0.74%	0.22%	0.00%
MEDPC	3.35%	1.13%	0.00%
DOCPC	1.30%	0.30%	0.00%
RRDDEN	43.17%	29.96%	12.70%
HWYDEN	3.70%	2.17%	0.39%
TELPC	42.52%	30.09%	10.35%
IISGDPP	0.00%	0.00%	0.00%
LGRGOVP	99.91%	99.48%	95.09%
LGEGOVP	2.30%	0.13%	0.00%
LTAXLGRP	0.04%	0.00%	0.00%
LGCCLGEP	0.00%	0.00%	0.00%
LGOTLGEP	47.17%	12.39%	1.00%
MRATE	100.00%	100.00%	99.52%

Note: numbers in the shaded boxes refer to more than 90% of times rejecting the null hypothesis of insignificance of the variable in the growth regressions.

Table 5
Simple Averages

Variable	coefficient	s.e.	p-value
RGDPPC	-3.2200E-07	3.3500E-07	0.438
INNINVP	0.4203	0.0925	0.000
SOEPRODP	0.0220	0.0605	0.721
RSOEIO	1.0400E-05	1.5300E-05	0.508
RFLONPC	-4.1800E-05	4.8100E-05	0.416
DEPGDPP	0.0697	0.0320	0.025
EDEPGDPP	0.0588	0.0428	0.129
PSEPC	0.0656	0.3524	0.844
SSEPC	0.0660	0.6172	0.930
PSTPC	-1.4775E-03	0.0015	0.329
HETPC	-0.0199	0.0059	0.001
HINSPC	0.0125	0.0103	0.229
HOSPPC	0.0027	0.0233	0.906
HOBEDPC	0.7248	20.3260	0.966
MEDPC	6.9710	21.8105	0.779
DOCPC	0.0782	31.0388	0.989
RRDDEN	-1.3913	0.8184	0.085
HWYDEN	-0.0131	0.1293	0.894
TELPC	1.1100E-05	1.0800E-05	0.275
IISGDPP	-0.0297	0.1041	0.768
LGRGOVP	-0.0167	0.0045	0.000
LGEGOVP	-0.0095	0.0117	0.424
LTAXLGRP	0.0182	0.0264	0.485
LGCCLGEP	-0.0452	0.1695	0.791
LGOTLGEP	0.1736	0.1139	0.129
MRATE	-0.0013	0.0003	0.000

Table 6
Log-likelihood Weighted Averages

Variable	coefficient	s.e.	p-value
RGDPPC	-3.2400E-07	3.3500E-07	0.435
INNINVP	0.4203	0.0925	0.000
SOEPRODP	0.0219	0.0605	0.722
RSOEIO	1.0400E-05	1.5300E-05	0.508
RFLONPC	-4.1900E-05	4.8100E-05	0.415
DEPGDPP	0.0698	0.0320	0.025
EDEPGDPP	0.0589	0.0428	0.128
PSEPC	0.0656	0.3523	0.844
SSEPC	0.0659	0.6170	0.930
PSTPC	-0.0015	0.0015	0.329
HETPC	-0.0199	0.0059	0.001
HINSPC	0.0125	0.0103	0.229
HOSPPC	0.0027	0.0233	0.905
HOBEDPC	0.7063	20.3196	0.967
MEDPC	6.9747	21.8060	0.779
DOCPC	0.0877	31.0315	0.989
RRDDEN	-1.3921	0.8181	0.084
HWYDEN	-0.0134	0.1293	0.892
TELPC	1.1200E-05	1.0800E-05	0.275
IISGDPP	-0.0299	0.1041	0.766
LGRGOVP	-0.0167	0.0045	0.000
LGEGOVP	-0.0095	0.0117	0.425
LTAXLGRP	0.0182	0.0264	0.487
LGCCLGEP	-0.0449	0.1694	0.792
LGOTLGEP	0.1734	0.1139	0.129
MRATE	-0.0013	3.4730E-04	0.000

Table 7
Coefficients and Standardized Effects

Complete Sample (536 observations)

Variable	coefficient	s.e.	p-value	average	st. dev.	st. effect
INNINVP	0.1938	0.0714	0.007	0.1869	0.0705	0.0137
DEPGDPP	0.0951	0.0241	0.000	0.6365	0.3252	0.0309
HETPC	-0.0044	0.0029	0.135	4.5880	5.9811	-0.0264
RRDDEN	-1.5321	0.5779	0.008	0.0119	0.0124	-0.0190
LGRGOVP	-0.0068	0.0022	0.002	1.3756	1.5967	-0.0109
MRATE	-0.0011	0.0003	0.000	1.2700	6.4905	-0.0070

Multiple Imputation (602 observations)

Variable	coefficient	s.e.	p-value	average	st. dev.	st. effect
INNINVP	0.1296	0.0744	0.082	0.1869	0.0705	0.0091
DEPGDPP	0.0637	0.0234	0.007	0.6365	0.3252	0.0207
HETPC	-0.0058	0.0026	0.026	4.5880	5.9811	-0.0344
RRDDEN	-0.1227	0.2329	0.599	0.0119	0.0124	-0.0015
LGRGOVP	-0.0035	0.0024	0.147	1.3756	1.5967	-0.0055
MRATE	-0.0012	0.0003	0.000	1.2700	6.4905	-0.0080

References

Apergis, Nicholas, “Inflation Uncertainty and Growth: Evidence from Panel Data”, Australian Economic Papers, June 2005, v. 44, iss. 2, pp. 186-97

Bao, Shuming, Gene Hsin Chang, Jeffrey D. Sachs, and Wing Thye Woo, “Geographic Factors and China’s Regional development Under Market Reforms, 1978-98”, China Economic Review, October 17, 2002

Borensztein, Eduardo and Jonathan D. Ostry, “Accounting for China’s Growth Performance,” American Economic Review, 86(2):224-228.

Cashin, Paul & Ratna Sahay, 1996, “Internal Migration, Center-State Grants and Economic Growth in the States of India”, International Monetary Fund Staff Papers, Vol. 43 No. 1 March.

Chen, Baizhu & Feng, Yi, “Determinants of Economic Growth in China: Private Enterprises, Education, and Openness – A Cross-Province Study”, China Economic Review, Vol. 11, No. 1, 2000.

Coulombe, Serge, 2000, “New Evidence of Convergence across Canadian Provinces: The Role of Urbanization”, Regional Studies, Vol. 34 No. 8 November, pp. 713-25.

Fleisher, Belton and Jian Chen, “The Coast-Noncoast Income Gap, Productivity, and Regional Economic Policy in China”, Journal of Comparative Economics, 25, 220-236, 1997.

Hsueh, Tien-tun, et al., China’s Provincial Statistics: 1949-1989, Westview Press, 1993.

Kim, Ji, “Convergence Hypothesis of Regional Income in Korea”, Applied Economics Letters, June 2005, v. 12, iss. 7, pp. 431-35

Knight, M., Norman L., and Villanueva, D., 1993, ‘Testing the Neoclassical Theory of Economic Growth: A Panel Data Approach," IMF Staff Papers, Vol. 40 No. 3.

Johnson, Paul A., 2000, “A Nonparametric Analysis of Income Convergence across the US States”, Economics Letters, vol. 69 No. 2 November pp. 219-23.

Lamo, Ana, 2000, “On Convergence Empirics: Some Evidence for Spanish Regions”, Investigaciones Economicas, Vol. 24 No.3 September pp. 681-707.

Leamer, E. E., 1985, ‘Sensitivity Analyses Would Help,’ American Economic Review, Vol. 57, No. 3 .

Levine, R. and Renelt, D., 1992, 'A Sensitivity Analysis of Cross-Country Growth Regressions,' American Economic Review, Vol. 82, No. 4. pp. 942 – 963.

Li, Kui-Wai and Tung Liu, 2001. "Impact of liberalization of financial resources in China's economic growth: evidence from provinces," Journal of Asian Economics, Vol. 12, No. 2, pp. 245-262.

National Bureau of Statistics, People's Republic of China, China Statistical Yearbook, var. eds. and CD-ROMs, China Statistics Press

Paap, Richard; Franses, Philip Hans; van Dijk, Dick, "Does Africa Grow Slower Than Asia, Latin America and the Middle East? Evidence from a New Data-Based Classification Method", Journal of Development Economics, August 2005, v. 77, iss. 2, pp. 553-70

Pekkala, Sari, 1999, "Regional Convergence across the Finnish Provinces and Subregions, 1960-94", Finnish Economic Papers, Vol 12 No. 1 Spring, pp. 28-40.

Phillips, Kerk L. and Shen Kunrong, 2005, "What Effect does the Size of the State-Owned Sector Have on Regional Growth in China?", Journal of Asian Economics, vol. 15 no. 6 pp. 1079-1102.

Quah, Danny T., (1996), "Empirics for Economic Growth and Convergence", European Economic Review, Vol. 40, No. 6, pp. 1353-75.

Rubin, D.B. (1987) *Multiple Imputation for Nonresponse in Surveys*. J. Wiley & Sons, New York.

Sala-i-Martin, X, 1997, 'I Just Ran Two Million Regressions,' American Economic Review, Vol. 87, No. 2.

Schafer, J.L. (1997) *Analysis of Incomplete Multivariate Data*. Chapman & Hall, London.