1999

The effects of public infrastructure development on regional economic development in China

Chin Kok Kam
The University of Montana

Let us know how access to this document benefits you.
Follow this and additional works at: https://scholarworks.umt.edu/etd

Recommended Citation
https://scholarworks.umt.edu/etd/8764

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.
Permission is granted by the author to reproduce this material in its entirety, provided that this material is used for scholarly purposes and is properly cited in published works and reports.

** Please check "Yes" or "No" and provide signature **

Yes, I grant permission  
No, I do not grant permission  

Author's Signature  

Date  

Any copying for commercial purposes or financial gain may be undertaken only with the author's explicit consent.
The Effects of Public Infrastructure Development on Regional Economic Development in China

by

Chin Kok Kam

B.A., University of Montana, 1997

Presented in partial fulfillment of the requirements for the degree of Masters of Arts

University of Montana

1999

Approved by

Chairman, Board of Examiners

Dean, Graduate School

Date
Deng Xiao Peng’s Reform in China in 1978 addressed structural problems in the Chinese economy. It aims to 1) create private ownership, 2) allow market forces to influence allocation of resources and determine prices, and 3) use material incentives to promote labor productivity and efficiency. The 1978 Reform’s success led to rapid economic expansion in China. The real economic growth rate since 1979 averaged 9% a year. However, this impressive economic growth has led to tremendous strain and unfulfilled demand in vital economic support systems such as transport and energy. Existing bottlenecks in the Chinese economy due to insufficient highways, railroads and power supply need to be promptly dealt with to maintain continuous economic growth.

This paper is a step in assessing this problem of public infrastructure bottlenecks in the Chinese economy. It attempts to empirically estimate the effects of different types of public infrastructures’ growth on regional economic development. The three types of public infrastructures that will be examined are highways, railroads and power supply. A pooled-data sample, extracted from the Statistical Yearbook of China, will be used for the analysis. The sample consists of regional data for 30 provinces over the period of 1993 through 1997.

The statistical model used to measure the effects of public infrastructure growth on regional economic growth follows the model established by Lu (1995). The model is derived from the Harrod-Domar production function. Regression using the Ordinary Least Squares method will be applied to estimate the relationship between economic growth and input growth.
To my wife, Mei Leng, who was always there in my times of struggle
and
my parents who never gave up on me.
TABLE OF CONTENTS

ABSTRACT ........................................................................................................................................ ii
ACKNOWLEDGMENTS .................................................................................................................... iii
TABLE OF CONTENTS .................................................................................................................... iv
LIST OF TABLES .................................................................................................................................. vi
LIST OF FIGURES ........................................................................................................................... vii
CHAPTER 1 INTRODUCTION ........................................................................................................... 1
   A Brief History of China’s Great Leap Forward ................................................................. 1
   Economic Factors Contributing to the Failure of the Great Leap Forward ...................... 3
CHAPTER 2 PUBLIC INFRASTRUCTURE IN CHINA ................................................................... 5
   Effects of public infrastructure improvements on short-run average cost of a dispersed production system ................................................................. 5
   Public infrastructure condition in China today ................................................................. 9
CHAPTER 3 LITERATURE REVIEW ............................................................................................. 11
   The Role of Public Infrastructure in the Production Process ........................................... 11
   The Production function approach ..................................................................................... 12
   Review on concerns and criticisms of contemporary public infrastructure research papers .......................................................................................................................... 14
      - Wide range of estimates ................................................................................................. 15
      - Spurious Correlation and the first difference approach ............................................ 17
      - Direction of Causation ................................................................................................. 18
   The Cost-Function Approach ......................................................................................... 19
   Literature Review: Cross Culture ..................................................................................... 21
   Literature Review: China ................................................................................................. 21
CHAPTER 4 THEORY AND METHODOLOGY ........................................................................... 24
   The General Framework ................................................................................................. 24
   Expected Outcome of Coefficients ($B$) ........................................................................... 26
      - Labor Coefficients ..................................................................................................... 26
      - Capital Coefficients ................................................................................................... 27
- Public Infrastructure Coefficients .............................................. 27
The Modified Production Function .............................................. 29

CHAPTER 5 DATA .............................................................................. 35
Regional Real Output growth (Y) .......................................... 35
Labor Growth Rate (L) ................................................................. 39
Capital Growth Rate (INV/Y) ....................................................... 41
Power Generation, Gas and Water Production and
Supply Infrastructure Growth Rate (PGW/Y) ................... 46
Growth in Length of Highways (HW) ........................................ 48
Growth in Length of Railways (RW) .......................................... 50
Regional Dummies ......................................................................... 52
Year Dummies ............................................................................. 52
Highway Density Dummy ............................................................. 52

CHAPTER 6 ESTIMATION .................................................................. 53
Using the Ordinary Least Squares Method ......................... 53
Heteroskedasticity ....................................................................... 55
RESULTS .......................................................................................... 56
INTERPRETATION .......................................................................... 58
Insignificant Production Input variables .......................... 60
- Estimation problems ................................................................. 61
  Multicollinearity ......................................................................... 61
  Simultaneity ................................................................................. 63
  Mis-Specification ......................................................................... 64
  Errors of Measurement ............................................................ 65
  Data Problems (Outliers) .......................................................... 65

CHAPTER 7 CONCLUSION ................................................................. 67
Recommendation for Further Research ................................. 68
APPENDIX A ................................................................................... 70
APPENDIX B ................................................................................... 73
BIBLIOGRAPHY ................................................................................. 77
LIST OF FIGURES

Figure | Page
-------|------
FIGURE 1: Short-run Average Cost Curves of concentrated and dispersed Production system | 7
FIGURE 2: Shift in the Short-run Average Cost Curves (dispersed Production system) after improvements in public infrastructure | 8
FIGURE 3: Annual Real GDP Growth Rates by province (1995-97) | 38
FIGURE 4: Annual Labor growth rate by province (1995-97) | 40
FIGURE 5: Annual INV/Y Growth Rates by province (1995-97) | 44
FIGURE 6: Annual PWG Growth Rates by province (1995-97) | 46
FIGURE 7: Annual Highway Growth Rates by province (1995-97) | 48
FIGURE 8: Annual Railway Growth Rates by province (1995-97) | 50
FIGURE 9: Highway Density | 52

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
CHAPTER 1
INTRODUCTION

A Brief History of China’s Great Leap Forward

"After the establishment of the People’s Republic in 1949, the Chinese Communist Party chose to rely heavily on Soviet models of economic structure and development strategy." (Harding, p.14) Among the characteristics of the Soviet model are high investment rate through mandatory savings, and advancement of heavy industry as a top priority, even at the expense of all other industries. Thus agriculture, which in 1952 provided more than 55 percent the country’s national output and provided employment for more than 85 percent of China’s labor force, yet received only about 7 percent of total gross investment capital during the First Five Year plan in 1953 to 1957. (Harding, p.16) “Within only a few year’s time, many Chinese leaders had become uneasy with what they regarded as serious shortcomings in the Soviet model.” (Harding, p.16) They believed that the enforcement of the Soviet model of urban development had "enlarged the gap between city and countryside, brought economic activity under excessively tight central control, created an unresponsive bureaucracy interested only in perpetuating its own power and privilege and produced tensions and contradictions between the Party and the ordinary citizen.” (Harding, p.16)

The Great Leap Forward in the late 1950s was an effort to reform the economic structure and development model introduced by the Soviets. The Great Leap Forward attempted to industrialize China through a dispersed production system. Unlike other
development models typically used, such as the 'Two Sector Economic Development with Unlimited Supply of Labor Model' developed by W. Arthur Lewis, the Great Leap Forward was designed to bring the industrial production to the villages where the majority of the population resided. Instead of moving workers and investment capital to large cities, a large number of small factories were set up in rural areas creating small production units all over the country. "The Great Leap Forward, introduced by Mao and other party leaders, toward the end of the 1950's was a product of a theoretical vision rather than a rationally thought-out operational plan." (Richmond, p.47)

Mao's plan to develop the rural areas with small scale production factories was very different from the typical economic development model used at that time. The Lewis model suggested that capital be concentrated in the urban areas for higher production efficiency. Labor and financial surpluses from the rural areas should be moved to urban areas. (Gillis, Perkins, Roemer and Snodgrass, p.54) The Great Leap Forward on the other hand was an attempt to industrialize China while maintaining a peasant society. Capital was dispersed throughout the country to create numerous small scale production units. Labor surpluses from the countryside were not relocated to urban areas since capital for production was made available in the rural areas.

"[Mao's Great Leap Forward] vision encompassed political, social, as well as economical aspects of ideology, and it led to severe economic crisis." (Richmond, p.47) The failure of the Great Leap Forward led to the death of 20 to 25 million (Harding, p.12) people who died from a horrifying famine due to a dramatic decline in agricultural output and the failure of the experiment with rural industrial production at small scale.
In part, this problem was due to inadequate infrastructure in the form of roads, electricity and water supply to support the production and movement of goods into and out of the rural areas. Lack of infrastructure is still a serious problem facing the rapidly growing Chinese economy in the 1990s. This paper is a step in assessing this issue for China.

**Economic Factors Contributing to the Failure of the Great Leap Forward**

Hindsight tells us that the failure of the Great Leap Forward can be attributed to a number of factors. The focus of this paper will be on the economic factors that contributed to the failure of the Great Leap Forward although we know that there were many other non-economic factors involved.

Among the economic factors that contributed to the failure of the Great Leap Forward were 1) a distribution system that was not developed to accommodate the movement of output from the factory to the consumer, 2) a dispersed production system that was inefficient and therefore was producing at a much higher short-run average cost compared to a concentrated production system, and 3) a production system that lacked market incentives.¹ Only the first two economic problems are relevant to the focus of this paper, which is public infrastructure.

A possible solution for both the first and second economic problem mentioned above is improvements in the quality and quantity of public infrastructure. First, and most importantly, improvements in the quality and quantity of roads and railroads would

---

¹ See Harding (1987), and Richmond (1969)
accelerate the movement of final goods and raw materials. Second, improvement in roads and railroads would lower the average cost of production by lowering the cost of transportation of raw materials to the factories.

Unlike the Lewis urban economic development system, Mao’s idea of many small production units in rural areas requires an extensive transportation network to facilitate the movement of raw materials and final goods. The quality and quantity of public infrastructure thus determines the fate of a dispersed production system. Excellent public infrastructure (quantity and quality) is essential in sustaining economic growth in a dispersed production system like China’s. Improvements in public infrastructure can lower the production cost of a dispersed system making it more competitive. The following section explains in more detail why a dispersed system has higher production cost and how improvements in public infrastructure can reduce it.
CHAPTER 2
PUBLIC INFRASTRUCTURE IN CHINA

Effects of public infrastructure improvements on short-run average cost of a dispersed production system

The effects of public infrastructure improvements on short-run average cost of a dispersed production system can be explained with the help of FIGURE 1 and FIGURE 2. FIGURE 1 illustrates the inefficiency of a dispersed system in the 1950s. Assume two different production systems, both have the same level of public capital: one is a dispersed system consisting of many small units of production in a large geographical area, and the other is a concentrated production system consisting of only a few large-scale production unit located within a small area. Suppose also that both systems are capable of producing at the same level, \( Q_A \). As shown in the graph, the small units of production prevent the dispersed system as a whole from attaining economies of scale. Although both production systems have the same output capacity, the dispersed system is producing at a much higher average cost.

Concentrated production systems are able to produce at a lower average cost because of agglomeration economies. "Agglomeration economies are cost reductions that occur because economic activities are carried on at one place." (Blair. p. 95) In particular, the type of agglomeration economies that lowers the cost of production of a urban production system is called urbanization economies. Urbanization economies are "cost savings that accrue to a wide variety of firms when the volume of activity in an
entire urban area increases.” (Blair, p.100) Urbanization economies are mainly a result of economies of scale in public infrastructure and division of labor.

The increase in the activity and size of an urban area results in lower per-unit infrastructure cost. This savings may be passed down to producers in the form of lower taxes per unit of output. Also, cost of production are lower because the division of labor is more extensive in an urban economy. For example, in a small town, it is not economically feasible to build a port because of limited demand. Activities that are not available in a small town will have to be purchased elsewhere. This extra cost of importing will cause the firm to be less competitive. (Blair, p.101)

FIGURE 2 illustrates the effects of an improvement in public infrastructure to the average cost curves (holding all else constant) in the dispersed system that existed in China in the 1950s. The graph shows that the improvements in public infrastructure have lowered the average cost of production in China. This implies that a dispersed production system needs public infrastructure that is superior to that of a concentrated production system just to stay competitive.
FIGURE 1: Short-run Average Cost Curves of concentrated and dispersed production system.

Short-run Average Cost Curves of the small individual production units of a dispersed production system (China: 1950's)

Short-run Average Cost curve of a centralized production system (Large scale)

Long-run Average Cost
FIGURE 2: Shift in the Short-run Average Cost Curves (dispersed production system) after improvements in public infrastructure.

Comments: FIGURE 2 illustrates the improvements in public infrastructure for the dispersed production system. In theory, improvements in public infrastructure like highways and railways will reduce the cost of production by reducing the cost of inputs and transportation cost of final goods. The reduction in production cost will shift both the Short-Run Average Cost and the Long-run Average Cost curve of a dispersed production system (like the one in the 1950's) down.
The Condition of Public infrastructure in China today

I believe that insufficient public infrastructure contributed significantly to the economic problems in the Great Leap Forward era in the 1950s. However, the question remains: Is the level of public infrastructure still insufficient and causing economic problems today?

Deng Xiao Ping’s Reform in China in 1978 addressed structural problems in the economy. Its goals were to 1) create private ownership, 2) allow market forces to influence allocation of resources and determine prices, and 3) use material incentives to promote labor productivity and efficiency. The 1978 Reform’s success led to rapid economic expansion in China. The real economic growth rate since 1979 averaged 9% a year. However, this impressive economic growth has led to tremendous strain and unfulfilled demand in vital economic support systems such as transport and energy. Existing bottlenecks in the Chinese economy today due to insufficient highways, railroads and power need to be promptly dealt with to maintain continuous economic growth.

Recent studies suggest that the level of public infrastructure is far below a satisfactory level despite forty years of effort by the Chinese government to improve on the quality and quantity of public infrastructure. The World Development Report 1994 published by the World Bank notes that:

“The coverage of China’s intercity transport networks is one of the thinnest in the world: the total route length per capita or per unit of arable length - for highways or railways - is similar to, or lower than, that in Brazil, India, and Russia. This has resulted mainly from chronic underinvestment in
China’s transport infrastructure. China’s transport investment amounted to only 1.3 percent of GNP annually during 1981-90, a period of rapid growth in transport demand. Since the onset of China’s open door policy in 1979, economic growth averaging 9 percent a year has resulted in an unprecedented expansion in intercity traffic —with growth averaging 8 percent a year for freight and 12 percent a year for passengers. This traffic growth has imposed tremendous strains on the transport infrastructure, as manifested by the growth of bottlenecks in the railway network, the severe rationing of transport capacity on railway lines, and the poor quality of service experienced by shippers and passengers.” (World Bank, p.18)

The remainder of this research will be an empirical analysis of the effects of public infrastructure improvements on regional economic growth in China.
CHAPTER 3

LITERATURE REVIEW

The Role of Public Infrastructure in the Production Process

Since the work of Aschauer (1989), there has been greater acceptance of the importance of public infrastructure for economic growth. Other works like Mera (1973), Eberts (1990) and Munnell (1990) on the effects of public infrastructure on the economy have shown that public infrastructure is an important and significant variable in explaining output growth.

"Everyone agrees that public capital investment can expand the productive capacity of an area, both by increasing resources and by enhancing the productivity of existing resources. A well constructed highway allows a truck driver to avoid circuitous back roads and to transport goods to market in less time. The reduction in required time means that the producer pays the driver lower wages and the truck experiences less wear and tear. Hence, public investment in highways enables private companies to produce at lower total cost. The condition of the highway, of course, is just as important as its existence. Similar stories can be told for mass transit, water and sewer systems, and other components of public capital."

(Munnell, 1992, p.191)
However, public infrastructure is different from the other input variables because the entrepreneur has no control over the level of its provision, unlike other inputs where the entrepreneur chooses the levels that would maximize profit. Other than the fact that the government controls the level of public infrastructure instead of the entrepreneur, the role of public infrastructure in the production process is very similar to the roles of the other production inputs. (Meade, p.56)

Like the other inputs, public infrastructure is a key factor in the production process. The absence of public infrastructure will cause output to cease. Without highways and railroads, final goods cannot be distributed. If output cannot be sold, a profit maximizing producer will completely lose the incentive to produce.

Public infrastructure can be best compared to the capital input. Both capital and public infrastructure are used in the production process but do not become part of the final good. Both are physical assets where usage in the production process causes depreciation. Both require maintenance and servicing. For all the reasons given above, public infrastructure can be considered as an input in the production process.

**The Production function approach**

Since the work of Aschauer (1989), many studies have estimated the effects of public infrastructure improvements on economic growth. The method most commonly used in these studies is the estimation of a production function with public infrastructure as an additional input. Many of these studies have found that the public infrastructure
variable is statistically significant in explaining variation in output and has positive effects on economic growth.

Aschauer (1989) and Munnell (1990) both estimated the national level output elasticity of non-military public infrastructure in the United States using time-series data. The model used in Aschauer is the Cobb-Douglas production function. The dependent variable is the log of output minus log non-residential private capital ($\ln Y - \ln K$), which is equal to the ratio of output to non-residential capital ($Y/K$) in linear terms. The independent variables include the log of labor minus non-residential private capital ratio ($\ln N - \ln K = \text{the labor to non-residential private capital ratio (N/K) in linear terms}$), and the log of non-military public capital minus non-residential private capital ($\ln G - \ln K = \text{non-military public capital to non-residential private capital ratio (G/K) in linear terms}$).

Aschauer's estimated equation: $\ln Y_t - \ln K_t = a_t + e_N*(\ln N - \ln K) + e_G * (\ln G - \ln K)$

where: $a_t$ is a measure of productivity or Hicks-neutral technical change, coefficient $e_N$ measures the output elasticity of labor, and coefficient $e_G$ measures the output elasticity of non-military public capital.

Similarly, Munnell also estimated the Cobb Douglas production function using log levels. Instead of subtracting the log of non-residential private capital from each side of the production function, Munnell subtracted the log of labor on the left hand side ($\ln Y - \ln L$) to obtain labor productivity. In linear terms, the left-hand side is the output to labor ratio ($Y/L$). The empirical model used by Munnell to measure the effects of public capital on labor productivity is as follows.

---

1 See Aschauer (1989).
2 See Munnell (1990)
Estimated equation: \( \ln Q_t - \ln L_t = \ln MFP + a(\ln K - \ln L) + c(\ln G - \ln L) \)

where, \( \ln MFP \) is the % growth in Multifactor Productivity, \( a \) measures the output elasticity of private capital, and \( c \) measures the output elasticity of public capital.

The output elasticity of public capital estimated by both studies are strikingly close to one another. Aschauer’s measure of output elasticity of public capital is 0.39. This means that an increase of one percent in public capital will lead to a 0.39 percent increase in output. Munnell found that the output elasticity of public capital is 0.34. In both studies, the independent variable representing public capital is positive and statistically significant.

Other similar studies at the state level in the United States show that the output elasticity of public infrastructure ranges from 0.15 to 0.20.\(^4\) (Munnell, 1992, p.194) Similar studies conducted using city level data found that the output elasticity of public capital to be in the range of 0.03 to 0.08.\(^5\) (Munnell, 1992, p.194)

**Criticisms of contemporary public infrastructure research papers**

“Numerous authors have included infrastructure as an additional argument of the production function, declaring that public infrastructure can be taken as an input in the production process that contributes independently to output. In such regression, infrastructure variables are generally found to be significant, though controversy exists over the methods of estimating the expanded function....” (Lu, p.1)

---

\(^4\) Examples are Munnell, 1990b; Martin, 1987; and Eisner, 1991.

\(^5\) Examples are Duffy-Deno and Eberts, 1989; Eberts, 1990.
According to Munnell (1992) there are three main controversial issues concerning the estimation methods of the production function. “First, they contend that common trends in the output and public infrastructure data have led to a spurious correlation. Second, they argue that the wide range of estimates (in the Unites States) emerging from the various studies renders the coefficients suspect. Finally, they suggest that the causation runs not from public capital to output, but rather in the other direction.” (p.192)

Wide range of estimates

In the World Development Report 1994 published by the World Bank, the reliability of the research done on this topic in the United States was also questioned:

“While there is still no consensus on the magnitude or the exact nature of the impact of infrastructure on growth, many studies on the topic have concluded that the role of infrastructure on growth is substantial, significant…. Although the indications to date are suggestive, there is still a need to explain why the results vary so much from study to study. Until this problem is resolved, results are neither specific nor solid enough to serve as a basis for designing policies for infrastructure investment.”

(p. 15)

Table 1 shows the various studies done in estimating the output elasticities of public capital in the United States.
Table 1. Output elasticity of public capital (various studies)

<table>
<thead>
<tr>
<th>Author</th>
<th>Level of Aggregation</th>
<th>Output elasticity of public capital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aschauer (1989)</td>
<td>National</td>
<td>0.39</td>
</tr>
<tr>
<td>Holtz-Eakin (1988)</td>
<td>National</td>
<td>0.39</td>
</tr>
<tr>
<td>Munnell (1990a)</td>
<td>National</td>
<td>0.34</td>
</tr>
<tr>
<td>Costa, Ellson, and Martin (1987)</td>
<td>States</td>
<td>0.20</td>
</tr>
<tr>
<td>Munnell (1990b)</td>
<td>States</td>
<td>0.15</td>
</tr>
<tr>
<td>Duffy-Deno and Eberts (1989)</td>
<td>Metropolitan Areas</td>
<td>0.08</td>
</tr>
<tr>
<td>Eberts (1990)</td>
<td>Metropolitan Areas</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Note: Quoted in Munnell (1992), p. 194.

According to Munnell,

"[T]he critics are seriously misreading the evidence. In almost all cases the impact of public capital on private sector output and productivity has been positive and statistically significant. This finding is amazing, given that much public capital spending is designed to alleviate environment problems or enhance the quality of life, and therefore contributes little to national output as conventionally measured." (p. 193)
In addition, Munnell argues that the estimates tend to be very similar at each level of government being studied and that:

"variation between estimates occurs as the unit of observation moves from the nation to states to cities. [Referring to Table 1] As the geographic focus narrows, the estimated impact of public capital becomes smaller. The most obvious explanation is that, because of leakages, one cannot capture all of the payoffs to an infrastructure investment by looking at a small geographic area." (p. 193)

**Spurious Correlation and the First Difference approach**

Spurious correlation is a problem where a common factor causes growth in both the dependent and independent variable. The two variables will have high correlation but in actuality, changes in one variable do not affect the other. (Gujarati. p.387) Spurious correlation causes time-series estimations to be biased.

Criticism of spurious correlation has led some researchers to use the first-difference form to estimate the production function. They argue that the first difference form should be used to estimate the production equation to avoid spurious correlation. "Specifically, they contend that the data are not stationary but tend to drift over time, and that it is necessary to remove this trend to eliminate spurious correlation and determine the true relationship between the two variables." (Munnell.1992. p.193)

Among the researchers that supported the first-differencing method to estimate the production function are Aaron (1990), Hulten and Schwab (1991), Jorgenson (1991), and
Tatom (1991). In a comment on Aschauer's work. Aaron (1991) has discounted the reliability of Aschauer's time-series analysis. Aaron argues that, "Time series are dominated by trend and produce marvelous fits that tend to distract one from their meager power to explain much of the relevant variance. The econometric devices used for avoiding these problems are many and varied...." (Aaron. p.53) Aschauer's original paper used none of those devices.

While the critics are saying the results generated from time-series analysis are too good to be true, the first-differencing method has produced "results showing that public capital's effect is quite small, sometimes negative, and generally not statistically significant." (Munnell. p.193) According to Munnell, the first-difference model irrationally assumes that the growth in capital stock, private or public, in a given year affects the growth in output in that same year. "In fact, equations estimated in this form often yield implausible coefficients for labor and private capital as well as for public capital (Evans and Karras. 1991; Hulten and Schwab. 1991; Tatom. 1991)." (Munnell. 1992. p.193)

**Direction of Causation**

The third major criticism involves the question of the direction of causation. Does higher public infrastructure spending stimulate higher output, or does higher output causes more public infrastructure spending?

"A number of studies have found that causation runs in both directions. Yet more sophisticated estimates that address these issues either have concluded the positive results
were not much affected by the different econometric methods or have found no noticeable impact of infrastructure on growth.” (World Bank, p.15)

Eberts and Fogarty (1987) examined the question of causality by examining public and private data from 1904 to 1978 for 40 metropolitan areas. Their study showed causation running in both directions. “Their analysis indicated that public investment led private investment in cities that experienced most of their growth before the 1950s, while the reverse was true for southern cities that grew faster since 1950.” (Munnell, 1992, p.195)

The Cost-Function Approach

In response to the criticisms of use of aggregate time-series and the direction of causation, a new generation of methods for estimating the effects of public capital on the economy emerged. Instead of using production functions, researchers argued that the cost function is a better method to “disentangle the effects of infrastructure, scale economies, and fixed effects on cost and the cost-output relationship.” (Munnell, 1992, p. 195)

Dalenberg and Eberts (1992) used the cost function approach and found that higher levels of public capital stocks were associated with lower production costs. The model used for estimation is as follows:

\[ CT = f(Q, PL, PK, PE, G) + PG \cdot G + PA \cdot A + TX \]

where: CT = total cost,
Q = manufacturing output,
PL = price of manufacturing labor,
PK = service price of manufacturing labor.
PE = price of manufacturing energy,
PG = service price of public capital.
PA = service price of land
A = land
G = public infrastructure.
TX = federal, state and local taxes.

Using data from 31 SMSA’s (Standard Metropolitan Statistical Area) over the period 1976 through 1978, they found that the total production cost elasticity of public capital is -0.2. This is interpreted as a one percent increase in public capital will decrease total production cost by 0.2%.

Morrison and Schwartz (1992) also attempted to measure the effects of public capital on the economy using the cost function approach. Data for this research was gathered from the 48 states in the U.S. for years 1970 through 1987. This analysis is based on “…cost-side productivity growth measures, which are designed to capture the reduction in input use (and thus cost) used to produce a given output level when technical change occurs.” (Morrison and Schwartz, 1992, p.2) The result of this study also found that the increase in the stock of infrastructure increases the efficiency of production through the reduction of production cost.

Nadiri and Mamuneas (1992) also adopted the cost-function analysis to examine the effects of public financed infrastructure and R&D capital on the cost structure and productivity performance of twelve manufacturing industries in the United States. Their research found that the two types of public capitals are significant in reducing the
production cost in every manufacturing industry in their model. In all of the manufacturing industries, they found that increases in public capital (both types) caused the cost function to shift down, thus generating greater productivity.

**Literature Review: Cross Culture**

Several studies for other countries have been conducted on the effect of public capital on economic performance. They include Berndt and Hansson (Sweden), 1992; Shah (Mexico), 1992; Ramirez (Mexico), 1992; Seitz and Licht (West Germany), 1993; Mera (Japan), 1972; and Lu (China), 1995.

Mera (1973) used the Cobb-Douglas production function to estimate effects of public capital on output for Japan. Using regional data and log levels (on both side of the equation) to estimate the production function, he found the output elasticity of public capital in Japan to be 0.2.

Shah in 1992 used a cost function model quite similar to the approach used by Dalenberg and Eberts (1992). Shah’s translog cost-function study is based on cross-sectional data for twenty six three-digit Mexican industries for the period 1970 to 1987. The result showed that a one percent increase in the stock of public capital will reduce the cost of production by 0.05 percent.

A study done on Sweden by Berndt and Hansson in 1992 found that increase in public capital tends to decrease the production cost by requiring less labor input for the same amount of output. The econometric model used was a labor input requirement
function, derived from a cost function. The study shows that an increase in public infrastructure will decrease the required amount of labor input.

**Literature Review: China**

Studies on the effects of public infrastructure improvements on economic growth in China are still not extensive. There are only two papers published recently on this subject: Man (1998) and Lu (1995). Man attempts to estimate the effects of transportation investment on regional economic growth. The empirical model used by Man is as follows:

\[
\text{GDP}_j = \text{TRINV}_j + \text{FINV}_j + \text{EMPLMPC}_j + \text{EMPLAPC}_j + \text{UNEMPLR}_j + \text{LPOP}_j
\]

Where, subscript \( j \) is the notation for province, GDP is the log of GDP in 1992, TRINV is the log of transportation investment, FINV is the log of foreign investment, EMPLMPC is the ratio of manufacturing to total employment, EMPLAPC is the ratio of farming, forestry, husbandry and fishery to total employment. UNEMPLR is the unemployment rate, and LPOP is the log of regional population.

The outcome of her regression shows that the output elasticity of transportation investment is 0.25. This implies that an increase in transportation investment of 1% leads to 0.25 % increase in regional GDP. However, the model used by Man to estimate output elasticity of transportation seems to be questionable.

Lu (1995) estimated the output elasticity of public infrastructure by measuring the production function of the thirty provinces of China. The dependent variable used in his study is the average growth rate of provincial GDP, from 1990 to 1994. The independent
variables used are measures of average growth rates of labor, capital, exports, and public infrastructure (differentiated by transportation and telecommunications infrastructure).

Lu's estimated equation: GDP = \( B_0 + B_1 \text{Labor} + B_2 \text{Capital} + B_3 \text{Telecommunication} + B_4 \text{Highway} + B_5 \text{Exports} \)

- where the dot denotes average growth rates from 1990-1994.

The result of Lu's study revealed that the output elasticity of transportation to be 0.43. This implies that an increase in the level of transportation infrastructure of 1% leads to 0.43% increase in regional GDP. Output elasticity of telecommunications of 0.21 implies that a 1% increasing in telecommunications infrastructure leads to 0.21% increase in regional GDP. These high output elasticities for transport and telecommunications indicate the seriousness of the public infrastructure bottleneck. In sum, Lu's study found that for the case of China, a small improvement in public infrastructure will cause a much larger increase in real GDP compared to other production inputs, such as capital. His study of China also found that the output elasticity of labor and capital to be 0.32 and 0.1 respectively. (Lu, p.12)
The General Framework

The simple Harrod-Domar production function model (Gillis, Perkins, Roemer and Snodgrass: 1992) will be used as the basic framework of this macro analysis of the effects of public infrastructure improvement on regional economic growth. A production function describes the available production technology and how factors of production determine the level of output produced. (Mankiw, p.45) The production function can be mathematically expressed as:

\[ Y = f(K, L, R, A) \]

where

- \( Y = \) output or national product
- \( K = \) stock of capital
- \( L = \) size of the labor force
- \( R = \) stock of arable land and natural resources
- \( A = \) technology.

The mathematical expression above shows that the level of output is a function of the levels of inputs employed. Under normal circumstances, ceteris paribus, when the level of an input is increased, we expect to gain additional output. Output is therefore a positive function of inputs.
Due to the lack of data for the Land variable, it will be left out of the model. Taking into account public infrastructure as an input variable that affects the level of output, the production function that will be estimated is thus:

\[ Y = f(\text{Labor}, \text{Capital}, \text{Public Infrastructure}) \]

Assuming the logarithmic form for the production technology, the following expressions may be derived to get us to the modified production function:

\[
\begin{align*}
\ln Y_t &= B_1 + B_2 \ln \text{Labor}_t + B_3 \ln \text{CAP}_t + B_4 \ln \Pi_t \quad \ldots \ldots \ldots \ldots (1) \\
\ln Y_{t-1} &= B_1 + B_{2,t-1} \ln \text{Labor}_t + B_3 \ln \text{CAP}_{t-1} + B_4 \ln \Pi_{t-1} \quad \ldots \ldots \ldots \ldots (2)
\end{align*}
\]

(1) - (2) equals:

\[
\ln \left( \frac{Y_t}{Y_{t-1}} \right) = B_2 \ln \left( \frac{\text{Labor}_t}{\text{Labor}_{t-1}} \right) + B_3 \ln \left( \frac{\text{CAP}_t}{\text{CAP}_{t-1}} \right) + B_4 \ln \left( \frac{\Pi_t}{\Pi_{t-1}} \right)
\]

The log of the quotients approximates the percentage changes of the variables in the model. Thus, we can rewrite the above logarithmic function in linear form as:

\[
\% \Delta Y = B_0 + B_1 \% \Delta \text{Labor} + B_2 \% \Delta \text{Capital} + B_3 \% \Delta \text{Public Infrastructure}
\]

Where the percentage change in output is a positive function of the percentage change in labor employed, capital available, and public infrastructure capital available. The coefficients represented by beta’s \((B)\) is interpreted as the percentage point change in output growth due to a percentage point increase in the growth of inputs. The Constant \((B_1)\) has been added for estimation purposes.

---

7 See Aschauer (1989) pg. 20.
**Expected Outcome of Coefficients (B)**

According to the simple Harod-Domar production function, output is a positive function of the inputs. **Labor, capital and public infrastructure capital** coefficients should therefore be positively related to output. Numerous empirical studies such as Aschauer (1989), Munnell (1990) and Lu (1995) support this conclusion using the Harod-Domar production function.

*Labor Coefficient*

The current estimate of China's population is 1.23 billion people with 696 million people employed. Due to China's Communist ideology, the labor force is employed even when laborers are not producing at their full potential. For this reason, the unemployment rate in China has always been very low. (Hughes, p. 71) It is plausible that China's problem of redundant labor fits into the definition of 'disguised unemployment,' where the redundant labors would have been unemployed if the decision to hire was based on market efficiency rather than political motives. (Gillis, Perkins, Roemer and Snodgrass, p. 54)

If disguised unemployment exists in China due to its communist ideology, the effect of an increase in the quantity of labor on economic growth is expected to be very low, zero (insignificant) or even possibly negative.

---

8 Situation where there is underutilization of labor and reduction in the quantity of labor has no effect on total production. (Gillis, Perkins, Roemer and Snodgrass 54)
**Capital Coefficient**

The capital variable in the production function refers to the stock of plants and equipment used in the production of output. It does not refer to the funds used in financing investments. In a market economy, the capital used for production is usually called private capital. However, in a planned economy such as China, only a small percentage of the total capital stock available for production is privately owned. Most production capital is owned by the state. Therefore, to avoid confusion or misunderstanding, it is important to note that the capital variable used in the study of China is the capital stock of production regardless of ownership. The coefficient for the capital variable is expected to be positive\(^9\) in a market economy. In a planned economy the stock may include capital that is not being used efficiently and thus not generating positive coefficient. The inability to separate out old state owned from new private owned capital may cause this variable’s coefficient to be insignificant—unfit to represent the effects of capital changes on economic growth.

**Public Infrastructure Coefficient**

Public infrastructure are for example roads, hospitals and schools. For the purpose of estimating the production function, only economic infrastructure—the types of public infrastructure directly relating to the production process—will be used in the estimation process. For example, investments in school buildings and hospitals are not

---

\(^9\) Examples of studies that found the capital coefficient to be positive and significant are Aschauer (1989), Munnell (1990), and Shah (1992).
included. The process of measuring economic public infrastructure will be explained in detail in the Data section of this paper.

As mentioned earlier, the lack of public infrastructure in the 1950s was one of the contributing factors to the failure of the Great Leap Forward. Forty years have passed since this experience and much has changed about the quantity and quality of railroads, highways and ports. Despite much improvement in public infrastructure, there still seems to be a massive need for more due to the rapid economic growth in China since the early 1990's.

"The picture of infrastructure under strain is reflected in the numerous reports citing under-developed infrastructure as a major bottleneck of the Chinese economy. For instance, in recent years only about 60 percent of demand for railway shipments has been met; the gap between electricity demand and supply was as high as 25 percent in 1994, resulting in one-quarter of the nation's total production capacity being idle; roads and transportation has remained under-developed, with the average rate of speed for motor vehicles in large cities dropping to 15km per hour; and the number of telephones per 100 persons is still extremely low (only 4.99 in 1995)." (Lu, p.3)

Since there is high excess demand for public infrastructure in China's economy, the expected coefficient for the public infrastructure variable is positive and relatively high.
The 'Modified' Production Function

The model that will be used in the paper will build on the modified production function model established by Lu (1995). Output and input variables will be based on growth rates measures similar to Lu's growth model, but for a different time period. Lu's estimation is based on data for years 1990 to 1994, while this present study uses data from 1993 to 1998. In addition, a variable has been included in this study to measure the effects of power, gas and water production infrastructure on regional economic growth. Also Lu's transport infrastructure variable has been differentiated into two types: highways and railways. The variable for exports and telecommunications will not be included in this study for the following reasons. The export variable was not included in this study because the argument for including this variable given by Lu is not convincing. Lu's reason for the inclusion of the Exports variable was that Feder (1982) had included Exports as an independent variable in his estimation of the production function and found that it was statistically significant. I argue that the export variable in Lu's model should not be used because there is no theoretical reason for exports to be included in the estimation of the production function. The telecommunications variable was excluded from my model because data for this variable is only available for 1994 and 1995.

The Ordinary Least Squares regression will be used to analyze the effects of the production inputs growth on output growth. The dependent variable will be the growth in output. Growth in output (Y) will be measured by the real change in GDP from a given year to the next. The real GDP growth rate was used instead of nominal GDP growth
rates because it provides measurement of output growth that is not affected by the changes in the price level.

The variable for labor growth \( L \) will be measured by the percentage change in aggregate labor employed in a province from a given year to the next. The use of aggregate labor growth rate only measures changes in the quantity of labor. Since the price of labor is not taken into account, this model assumes away the changes in the quality of labor or human capital. "Human capital is the economist's term for the knowledge, skills that workers acquire through education, training and experience." (Mankiw, p.105) The quality of labor plays a substantial role in determining the level of output in long-term productivity growth in rapidly growing economies where education investment are documented. However, this is a short-term model where human capital of the labor pool is assumed to be constant.

Although data on the national stock of production capital in China is available, the number of years (observations) is too limited to provide a data set equivalent to that used in this study. Therefore, the capital growth variable will be represented by the total capital investment to GDP \( (INV/Y) \) ratio. This proxy for percentage change in capital was also used by Aschauer (1989) and Lu (1995). The rationale for the proxy is as follows:

We know that: \[ \text{Investment} = \text{change in capital} \]

or \[ INV = dK \]

(1)
where $INV$ is the gross investment in capital and $dK$ is the change in capital stock.\(^{10}\)

Dividing both sides of equation (1) by income, we get the expression:

$$INV/Y = dK/Y \quad \quad (2)$$

Expression (2) can be written as:

$$INV/Y = (dK/K) \times K/Y \quad \quad (3)$$

where $K$ is the capital stock. Assuming the capital to output ($K/Y$) ratio is constant, the investment ratio ($INV/Y$) becomes a good proxy for capital growth.\(^{11}\) The $INV/Y$ variable is a flow concept since it is measured over a consistent period of time.

Public infrastructure can be divided into two groups, social and economic infrastructure. Examples of social infrastructure are schools and hospitals. Examples of economic infrastructure are highways, railroads and power plants. The type of public infrastructures that the production function refers to are only those that play a direct role in the production process as an input. Therefore, social infrastructure is not part of the production function equation. This is not to say that social infrastructure is not important in the production process. For example, schools provide the economy with educated and skillful labors, which are the crucial ingredients in the production of output. Social infrastructure itself is not a direct input and its effects are long-term.

\(^{10}\) Gross investment has been used in place of net investment because this is the data availability. Depreciation of capital is assumed to be a constant part of the capital stock. By using the first difference (or growth rate, as we will discuss later in this section) model, the shortcomings of using gross instead of net investment is reduced. This is because as we reduce the time between each measurement of capital, the assumption that depreciation is constant strengthens.

\(^{11}\) See Aschauer (1989) and Lu (1995)
Economic public infrastructure will be classified into two types for the purpose of this study. This will allow a detailed analysis of the bottleneck effect associated with economic public infrastructure. The first type consists of capital built to generate electric power, gas and water production and supply. Adopting an approach equivalent to that used in the measurement of the stock of capital, the investment ratio \( \frac{PGW}{Y} \) will be used as a proxy for growth in capital stock of power, gas and water production and supply infrastructure.

The second type of economic public infrastructure involves public infrastructure for transportation. This paper will attempt to analyze two types of transportation public infrastructure: highways and railways. The improvements of highways \( (HW) \) will be measured by the annual growth rate in the length of highways measured in kilometers by province. The improvements of railways \( (RW) \) will be measured by the annual growth rate in the length of operating railways by kilometers by province.

Regional Dummies will be used to take into account the effects of differences in the qualitative characteristics between regions that remain constant over time and that cannot be easily quantified. Examples of these characteristics are political environment, ancestry, and religion.\(^\text{12}\)

Year Dummies have a similar role in the analysis to the Regional Dummies utilized. Year dummies take into account the effects of qualitative characteristics or events that happened in a particular year. Some examples of these events are natural disaster, political upheaval, election, and government economic policy.

**Highway Density Dummy** will be used to measure the effects of agglomeration economies on regional economic growth. We are assuming that the higher the highway density of a province, the higher the degree of urbanization and thus the higher the agglomeration economies effect will be. The Highway Density dummy is expected to be positive. This means that increases in agglomeration economies effects will increase regional economic growth. The method of calculating this variable will be discussed in greater detail in the Data section of this paper.

Given the assumptions of no changes in 1) production technology, 2) demand and supply conditions, and 3) other exogenous variables that may affect the growth rate of GDP, (for example, the pattern of economic interdependence between regions, quality of inputs, and currency exchange rate) the equation that will be estimated is then:

\[
y_{t2|t1} = B_0 + B_1 L_{t2|t1} + B_2 \frac{\text{INV}}{Y_{t1|t0}} + B_3 \frac{\text{PGW}}{Y_{t1|t0}} + B_4 \text{HW}_{t1|t0} + B_5 \text{RW}_{t1|t0} + \\
B_6 \text{RW}_{t1|t0} + B_7 \text{Regional Dummies} + B_8 \text{Year Dummies} + B_9 \text{Highway Density Dummies} + e
\]

where, \( Y \) is the percentage change in GDP from year 1 to year 2,

\( L \) is the percentage change in aggregate labor employed from year 1 to year 2,

\( \text{INV}/Y \) is the ratio of capital investment to GDP for year 1, proxy for percentage change in total gross investment in capital from year 0 to year 1,

\( \text{PGW}/Y \) is the ratio of capital investment in power, gas and water production, and supply infrastructure to GDP for year 1, proxy for growth in capital stock of power, gas and water production and supply infrastructure from year 0 to year 1,
HW is the percentage change in the length of highways from year 0 to year 1,

RW is the percentage change in the length of railways from year 0 to year 1,

e is the error term.

Variables INV/Y, PGW/Y, HW and RW are lagged one year. The model assumes that growth in these variables will affect the following year’s GDP growth.

(The abbreviated variables above will be used for the rest of this paper when referring to variables in the data.)
CHAPTER 5

Data

The sample data is constructed from provincial level data from the thirty provinces of China. This sample of pooled time series cross sectional data was obtained from the 30 provinces from 1993 to 1997. The number of observation was 90\(^3\). Table 2 provides the descriptive statistics to each variable in the sample. Appendix A Table 7 shows the complete data sample used in this study.

Table 2. Descriptive Statistics of variables

<table>
<thead>
<tr>
<th>NAME</th>
<th>N</th>
<th>MEAN</th>
<th>ST. DEV</th>
<th>VARIANCE</th>
<th>MINIMUM</th>
<th>MAXIMUM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>90</td>
<td>11.363</td>
<td>2.7145</td>
<td>7.3684</td>
<td>4.3000</td>
<td>18.100</td>
</tr>
<tr>
<td>L</td>
<td>90</td>
<td>1.1104</td>
<td>1.4427</td>
<td>2.0813</td>
<td>-5.6800</td>
<td>5.8200</td>
</tr>
<tr>
<td>K</td>
<td>90</td>
<td>23.366</td>
<td>7.9610</td>
<td>63.378</td>
<td>8.4900</td>
<td>52.900</td>
</tr>
<tr>
<td>PWG</td>
<td>90</td>
<td>3.6586</td>
<td>3.4500</td>
<td>11.902</td>
<td>1.1800</td>
<td>25.040</td>
</tr>
<tr>
<td>RAIL</td>
<td>90</td>
<td>2.3293</td>
<td>7.4680</td>
<td>55.771</td>
<td>-0.9200</td>
<td>55.560</td>
</tr>
<tr>
<td>HW</td>
<td>90</td>
<td>2.7622</td>
<td>3.4434</td>
<td>11.857</td>
<td>0.0000</td>
<td>22.980</td>
</tr>
<tr>
<td>Y95</td>
<td>90</td>
<td>0.33333</td>
<td>0.47405</td>
<td>0.22472</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Y96</td>
<td>90</td>
<td>0.33333</td>
<td>0.47405</td>
<td>0.22472</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Y97</td>
<td>90</td>
<td>0.33333</td>
<td>0.47405</td>
<td>0.22472</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>ROAD DEN.</td>
<td>90</td>
<td>0.20000</td>
<td>0.40224</td>
<td>0.16180</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Regional Real Output growth (Y)

The measure of change in real output (Y) used in this study is the annual percentage change in the value of real GDP by province. These numbers were obtained from GDP indices from the Statistical Yearbook of China (1994-1997).

\(^3\) The number of observation is equal to 90 because of the use of growth rates and a one year lag in the model.
The GDP indices are based on \textit{comparable prices}\textsuperscript{14}. This GDP indices show the GDP of a given year in terms of the previous year's (base year) prices. This allows economic performance comparisons between the current and base year. The GDP indices of any given year reported in the Statistical Yearbook of China are calculated using the preceding year as the base year.

The method of calculating the annual percentage change in real GDP growth ($Y$) from the GDP indices is as follows:

$$Y_{t_2t_1} = \frac{\text{GDP Index of year } t_2 - \text{GDP Index of year } t_1 (\text{base year})}{\text{GDP Index of year } t_1 (\text{base year})}$$

Since the GDP Index of year $t_1$ (base year) is always equal to 100,

$$Y_{t_2t_1} = \frac{\text{GDP Index of year } t_2 - 100}{100}$$

$$Y_{t_2t_1} = \text{GDP Index of year } t_2 - 100$$

For example calculation of the real GDP growth rate of Beijing province from 1996 to 1997 is as follows:

$$Y_{96-97} = \text{GDP Index of year 1997} - 100$$

$$= 109.6 - 100$$

$$= 9.6\%$$

\textit{(where the base year is 1996, which is equal to 100.)}

\textsuperscript{14} Comparable Prices are applied when comparing indicators over time to reflect accurately the changes in real terms. Two methods are used for calculating comparable prices: 1. output by constant price of certain year; 2. output in current prices divided by relevant price index." (State Statistical Bureau, 1994, pp.54)
The following table presents the basic statistics of the output (Y) variable.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>11.95</td>
<td>4.30</td>
<td>17.90</td>
</tr>
<tr>
<td>1996</td>
<td>11.64</td>
<td>4.80</td>
<td>18.10</td>
</tr>
<tr>
<td>1997</td>
<td>10.50</td>
<td>6.70</td>
<td>14.50</td>
</tr>
</tbody>
</table>

FIGURE 3. Annual Real GDP Growth Rates by province (1995-97)

North Region

Northeast Region

East Region

Central Region

South Region

Southwest Region

**Labor Growth rate (L)**

The data used for the percentage change in labor (L) variable is the percentage change in the number of people employed by region. People employed is defined as the total number of persons engaged in social labor which generates income, including:

1) *Total staffs and workers* (employees of the state ownership enterprises, collective ownership enterprises, joint ownership enterprises, shareholding enterprises, foreign ownership enterprises, and other ownership enterprises and their affiliated units)

2) *Employees in urban private enterprises*

3) *Urban individuals laborers*

4) *Rural laborers*

5) *Other social laborers*

(Statistical Yearbook of China, 1994, p.133)

The method of calculating the percentage change in the number of people employed by region is as follows:

\[
L_{t2} = \frac{\text{Num. people employed (Year 2)} - \text{Num. people employed (Year 1)}}{\text{Num. people employed (Year 2)}} \times 100
\]

The following table presents the basic statistics of the L variable.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1995</td>
<td>1.11</td>
<td>-5.68</td>
<td>3.44</td>
</tr>
<tr>
<td>1996</td>
<td>0.88</td>
<td>-1.28</td>
<td>3.52</td>
</tr>
<tr>
<td>1997</td>
<td>0.34</td>
<td>-1.61</td>
<td>5.82</td>
</tr>
</tbody>
</table>

FIGURE 4: Annual Labor growth rate by province (1995-97)

Data Problems with the Labor variable

There are some outliers in the data set that seems to require explanation. Some of these values may well be printing errors. These errors may cause the estimation to be biased. Correction was made to the one outlier that was clearly a typing mistake. All other outliers were left. The following is the description of the correction made to this variable before the regression.

The reported growth rate of labor in Sichuan province from 1996 to 1997 is -26%. What happened to the millions of labors in this province in one year? Since there was no war or famine in that province during that time, I could only conclude that the data was printed wrongly. In this case, the problematic observation is replaced with a value generated with a moving average. The value of labor growth rate for Sichuan in 1997 was thus replaced with the generated value of 0.91%.

Capital Growth rate (INV/Y)

As explained earlier, the capital investment to income (INV/Y) ratio will be used as a proxy for the gross capital growth variable. The reason for using this proxy is because the capital stock data needed to calculate capital growth rate is unavailable. The measure used to represent income (Y) is regional Gross Domestic Product (GDP). The data for both Y and INV are obtained from the Statistical Yearbook of China (1994-1997)\textsuperscript{15}.

\textsuperscript{15} Years 1994-1997 were used instead of 1995-1998 because of this variable is lag one year.
Capital Investment (INV) is defined as *total economic investment (gross) in fixed assets excluding public infrastructure investments, by region*. This data for INV is derived from *total (gross) investment in fixed assets by region* obtained from the Statistical Yearbook of China. Included in the *total gross investment in fixed assets by region* measure are some non-economic *investments in fixed assets* and also investments in fixed assets in public infrastructure. Both of these needs to be excluded from the INV measure to obtain a variable relevant to this study. Investments in fixed assets in public infrastructure were excluded in the measure of INV because public infrastructure will be branched out as three independent variables: HW, RW and PWG.

The non-economic items that are deducted are *capital construction investments* and *technological updates and transformation investments by state-owned units, and investments in fixed assets by urban collective-owned units* in:

1. Social services.
2. Health care, sports and social welfare.
3. Education, culture, art, radio, film and television.
4. Government agencies, party agencies and social organizations. (Statistical yearbook of China, 1994)

---

16 Some other non-economic investments may be still included in the measure of INV due to generalized classification of investment of other forms of ownership in the Statistical Yearbook of China. For example, investment in fixed assets by *joint ventures with foreigners units* are not differentiated by types of investments. However, there is no reason to believe that these 'non-government owned organizations' should have significant investment in the four non-economic items mentioned above.
Investments in fixed assets in public infrastructure that are excluded in the measure of INV are:

1) *capital construction investments of state-owned units in Electric power, Gas and Water production and supply.*

2) *technological updates and transformation investments by state-owned units in Electric power, Gas and Water production and supply.*

3) *Investment in fixed assets in Electric power, Gas and Water production and supply by urban collective owned units.*

4) *capital construction investments of state-owned units in Transportation, Storage, Postal and Telecommunications services.*

5) *technological updates and transformation investments by state-owned units in Transportation, Storage, Postal and Telecommunications services.*

6) *Investment in fixed assets in Transportation, Storage, Postal and Telecommunications services by urban collective owned units.* (Statistical yearbook of China, 1994)

---

The following table presents the basic statistics of the INV/Y variable.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>24.28</td>
<td>8.49</td>
<td>40.27</td>
</tr>
<tr>
<td>1995</td>
<td>23.52</td>
<td>9.18</td>
<td>50.47</td>
</tr>
<tr>
<td>1996</td>
<td>28.30</td>
<td>10.82</td>
<td>52.90</td>
</tr>
</tbody>
</table>


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
FIGURE 5. Annual INV/Y Growth Rates by province (1994-96)

North Region

Northeast Region

East Region

Central Region

South Region

Southwest

Power Generation, Gas and Water Production and Supply Infrastructure Growth

Rate (PGW/Y)

The Power Generation, Gas and Water Production and Supply Infrastructure Growth Rate (PWG/Y) variable is represented by an investment ratio generated using the similar method in generating the INV/Y variable. PGW is the investments in fixed assets in Power Generation, Gas and Water Production and Supply Regional Infrastructure. Y represents the income measured by GDP.

PWG is constructed by summing up all investments related to electric power, gas and water production and supply. PGW includes:

1. Capital construction investments in electric power, gas and water production and supply by a) State owned units, and b) Urban collective-owned units.

2. Technological updates and transformation investments electric power, gas and water production and supply by state-owned units. (Statistical yearbook of China, 1994)

The following table presents the basic statistics of the PWG/Y variable.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>3.39</td>
<td>1.18</td>
<td>12.67</td>
</tr>
<tr>
<td>1995</td>
<td>3.46</td>
<td>1.37</td>
<td>16.79</td>
</tr>
<tr>
<td>1996</td>
<td>4.13</td>
<td>1.50</td>
<td>25.04</td>
</tr>
</tbody>
</table>

FIGURE 6. Annual PWG Growth Rates by province (1994-96)

**Growth in Length of Highways (HW)**

The data on growth in the length of highways is represented by the annual growth rate of *total highways in kilometers (km)*. The method of calculation of Growth in Length of Highways (HW) is as follows:

\[
\text{HW}_{97-96(\text{Beijing})} = \frac{\text{Total highways (km) in 1997} - \text{Total highways (km) in 1996}}{\text{Total highways (km) in 1996}} \times 100
\]

\[
= \frac{(12306 - 12084)}{12084} \times 100
\]

\[= 1.83\%
\]

*The following table presents the basic statistics of the HW variable.*

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.39</td>
<td>-0.71</td>
<td>4.84</td>
</tr>
<tr>
<td>1995</td>
<td>2.58</td>
<td>-0.20</td>
<td>55.56</td>
</tr>
<tr>
<td>1996</td>
<td>4.01</td>
<td>-0.92</td>
<td>33.84</td>
</tr>
</tbody>
</table>

(Statistical yearbook of China: 1994 - 1997)
FIGURE 7. Annual Highway Growth Rates by province (1994-96)

Growth in Length of Railways (RW)

The data on growth in the length of total railways is represented by the annual growth rate of *total railways in operation in kilometers (km)*. The method of calculation of Growth in Length of total railways (RW) is as follows:

\[
\text{RW}_{97-96}^{(Beijing)} = \frac{\text{Total railways in operation (km) in 97} - \text{Total railways in operation (km) in 96}}{\text{Total railways in operation (km) in 1996}} \times 100
\]

\[
= \frac{(1069 - 1067) \times 100}{1067} = 0.18\%
\]

The following table presents the basic statistics of the RW variable.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>National Average</th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>2.51</td>
<td>0.00</td>
<td>22.98</td>
</tr>
<tr>
<td>1995</td>
<td>3.51</td>
<td>0.27</td>
<td>13.93</td>
</tr>
<tr>
<td>1996</td>
<td>2.27</td>
<td>0.00</td>
<td>6.74</td>
</tr>
</tbody>
</table>

FIGURE 8. Annual Railway Growth Rates by province

(Statistical yearbook of China: 1994 -1997)
**Regional Dummies**

The 30 provinces were divided into 6 geographical regions for the purpose of producing the Regional Dummies. The provinces are divided as follows:\(^{17}\):

**North** - Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia

**Northeast** - Liaoning, Jilin, Heilongjiang

**East** - Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong

**Central** - Henan, Hubei, Guangdong, Guangxi, Hainan

**South** - Sichuan, Guizhou, Yunnan, Tibet

**Southwest** - Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang

The purpose of the regional dummies is to capture the fixed effects between regions.

**Year Dummies**

The year dummy was used to capture the economic growth trend of the country. The programming method used to construct the Year dummy is presented in Appendix B.

**Highway Density Dummy**

Highway density was computed by dividing the total length of highway by the size of the province. The following provides the basic statistics of highway density for the 30

---

\(^{17}\) The 30 provinces are divided into 6 geographical regions according to the classification used in the Statistical Yearbook of China.
provinces from 1995 to 1997. The highway density dummy was constructed using the following criteria:

If density is greater than 0.48, than dummy =1

If density is less than 0.48, than dummy =0

Highway density of 0.48 was chosen as the divider so that the Highway Density dummy represents the upper 20% of the scale\(^\text{18}\).

FIGURE 9: Highway Density

(Statistical yearbook of China: 1994)

\(^{18}\) Road density: low = 0.02, high = 0.61.
CHAPTER 6

Estimation

Using the Ordinary Least Squares Method

The modified production function equation will be estimated using the Ordinary Least Squares (OLS) method. The OLS is a method of finding the best fit line for a given set of data by minimizing the sum of the squared residuals. (Gujarati, p. 52)

Given the dependent and independent variable relationships discussed above, the suitable regression model that represents the hypothesis presented is:

\[ Y = B_0 + B_1L + B_2 \text{INV/Y} + B_3 \text{PGW/Y} + B_4 \text{HW} + B_5 \text{RW} + \]
\[ B_6 \text{Regional Dummies} + B_7 \text{Year Dummies} + B_8 \text{Highway Density Dummy} + e \]

Where the dot above the variables denotes the growth rate from year 1 to year 2; the squiggle above variables denotes growth rate from year 0 to year 1; and \( e \) represents the residuals.
Table 3. OLS Estimation of the Production Function (Dependent Variable = GDP Growth Rate)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Coefficients</th>
<th>$t$-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0.17511</td>
<td>0.5594</td>
</tr>
<tr>
<td>K</td>
<td>-0.03435</td>
<td>-0.7219</td>
</tr>
<tr>
<td>PWG</td>
<td>0.04988</td>
<td>0.6515</td>
</tr>
<tr>
<td>RAIL</td>
<td>0.07877</td>
<td>0.991</td>
</tr>
<tr>
<td>HW</td>
<td>0.11507</td>
<td>1.353</td>
</tr>
<tr>
<td>NORTHEAS</td>
<td>-1.2039</td>
<td>-1.353</td>
</tr>
<tr>
<td>EAST</td>
<td>2.4366</td>
<td>3.414</td>
</tr>
<tr>
<td>CEN</td>
<td>1.4595</td>
<td>1.696</td>
</tr>
<tr>
<td>SOUTH</td>
<td>-1.6343</td>
<td>-1.665</td>
</tr>
<tr>
<td>SW</td>
<td>-1.2490</td>
<td>-1.420</td>
</tr>
<tr>
<td>¥96</td>
<td>-0.5828</td>
<td>-0.9896</td>
</tr>
<tr>
<td>¥97</td>
<td>-1.8554</td>
<td>-3.158</td>
</tr>
<tr>
<td>ROADEN</td>
<td>0.9981</td>
<td>0.5458</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>11.762</td>
<td>7.786</td>
</tr>
<tr>
<td>R-square</td>
<td>0.4363</td>
<td></td>
</tr>
<tr>
<td>R-square adjusted</td>
<td>0.3399</td>
<td></td>
</tr>
</tbody>
</table>

The $R^2$ of 0.4363 shows that the variation in the independent variables were able to explain 43.6% of the variation in the dependent variable ($Y$). After conducting the F-test,
the results showed that there is sufficient evidence to claim that the estimated coefficients are not all equal to zero.

**Heteroskedasticity**

"...heteroskedasticity does not destroy the unbiasness and consistency of the properties of the OLS estimators, but they are no longer efficient, not even asymptotically." (Gujarati, p.381) The value of the estimated coefficients itself are not affected by the heteroskedasticity problem. (Gujarati, p.365) Heteroskedasticity may arise due to outliers in the data. (Gujarati, p.358) If Heteroskedasticity exists in the data used in this study, the standard errors of the estimated coefficients will be biased. Biased standard errors will effect the coefficients' (B's) level of significance (t-ratio).

Since the problem of heteroskedasticity is more likely to affect cross-sectional than in time series data. (Gujarati, p.59) the test for Heteroskedasticity was carried out for the data used in this study. White's General Heteroskedasticity test was employed.

White's test was implemented by regressing the squared residuals from the original regression on the original X variables (regressors), their squared values, and the cross product(s) of the regressors. The actual implementation of this test can be seen in Appendix B, page 75.

The results of White's test for heteroskedasticity shows evidence¹⁹ of heteroskedasticity in this data. White's correction for Heteroskedasticity was used to correct the biased standard errors of the coefficients. White's correction will change the

---

¹⁹ Confidence level of 95%.
standard errors of the coefficients, and thus the t-ratio. (Gujarati, p.382) The value of the estimated coefficients themselves will remain the same after the correction.\textsuperscript{20}

**Results**

The results of the regression are presented in Table 4. The estimated coefficient for Railway and Highway showed significant t-ratios at the 10% error level. The modified production function model that was used did not offer a high explanatory power on GDP growth compared to other studies\textsuperscript{21} discussed in the Literature Review section of this paper. The R-square of 0.4363 means that variation in the independent variables are able to explain 43.6% of the variation in GDP growth.

\textsuperscript{21} See for example Aschauer (1989), Munnell (1990), and Lu (1995).
Table 4. OLS Estimation of the Production Function (Dependent Variable = GDP Growth Rate) Using Heteroskedasticity - Consistent Covariance Matrix

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Coefficients</th>
<th>t-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>0.17511</td>
<td>0.8559</td>
</tr>
<tr>
<td>K</td>
<td>-0.03435</td>
<td>-0.7716</td>
</tr>
<tr>
<td>PWG</td>
<td>0.04988</td>
<td>0.9276</td>
</tr>
<tr>
<td>RAIL</td>
<td>0.07877</td>
<td>2.226*</td>
</tr>
<tr>
<td>HW</td>
<td>0.11507</td>
<td>1.326*</td>
</tr>
<tr>
<td>NORTHEAS</td>
<td>-1.2039</td>
<td>-1.680*</td>
</tr>
<tr>
<td>EAST</td>
<td>2.4366</td>
<td>4.490*</td>
</tr>
<tr>
<td>CEN</td>
<td>1.4595</td>
<td>2.339*</td>
</tr>
<tr>
<td>SOUTH</td>
<td>-1.6343</td>
<td>-1.473*</td>
</tr>
<tr>
<td>SW</td>
<td>-1.2490</td>
<td>-2.033*</td>
</tr>
<tr>
<td>Y96</td>
<td>-0.5828</td>
<td>-0.9961</td>
</tr>
<tr>
<td>Y97</td>
<td>-1.8554</td>
<td>-3.366*</td>
</tr>
<tr>
<td>ROADEN</td>
<td>0.9981</td>
<td>0.8995</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>11.762</td>
<td>8.908*</td>
</tr>
<tr>
<td>R-square</td>
<td>0.4363</td>
<td></td>
</tr>
<tr>
<td>R-square adjusted</td>
<td>0.3399</td>
<td></td>
</tr>
<tr>
<td>F-test</td>
<td>Not all coefficients are equal to zero.\textsuperscript{22}</td>
<td></td>
</tr>
<tr>
<td>$F_{\text{crit}}$</td>
<td>1.85</td>
<td></td>
</tr>
<tr>
<td>$F_{\text{test}}$</td>
<td>4.54</td>
<td></td>
</tr>
<tr>
<td>Condition: If $F_{\text{test}} &gt; F_{\text{crit}}$, then reject the null hypothesis of all coefficients are equal to zero.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: * Statistical significance at the 10 percent error level

The following figure is a scatter plot of the residuals from the above regression.

\textsuperscript{22} The details on the F-test are reported in Appendix.
**Interpretation**

The labor coefficient ($L$) is interpreted as the percentage point change in output due to one percentage point increase in Labor. The coefficient of Labor cannot be meaningfully interpreted due to the insignificant $t$-ratio. The coefficient may be insignificant due to the redundant labor or disguised unemployment problem in China today.\(^{23}\)

The INV/Y coefficient is interpreted as the percentage point change in output due to one percentage point increase in capital. The INV/Y coefficient cannot be meaningfully interpreted due to the insignificant $t$-ratio. The inability to separate state and private capital data may have caused this variable to appear unrelated to output growth.\(^{24}\)

\(^{23}\) Disguised unemployment was discussed in chapter 4 of this paper (p. 26).

\(^{24}\) This problem was discussed in Chapter 4 of this paper (p.27).
The $\text{PGW}$ coefficient is interpreted as the percentage point change in output due to one percentage point increase in investment in capital for electric power, gas and water production and supply. $\text{PGW}$ coefficient also cannot be meaningfully interpreted due to the insignificant $t$-ratio.

The $\text{RW}$ coefficient is interpreted as the percentage point change in output due to one percentage point increase in railroads. $\text{RW}$ coefficient of 0.078 means a one percentage point increase in $\text{RW}$ capital investment will increase the GDP by 0.078 percentage points.

The $\text{HW}$ coefficient is interpreted as the percentage point change in output due to one percentage point increase in railroads. $\text{HW}$ coefficient of 0.115 means that a one percentage point increase in the length of highway will cause GDP to increase by 0.115 percentage point.

The negative and decreasing coefficients of the Year Dummies ($\text{Y96}$ and $\text{Y97}$) show that there is a downward trend in regional GDP growth in China.

The $\text{Roaden}$ coefficient is positive but its $t$-ratio is insignificant. This study cannot therefore prove the positive relationship between the degree of agglomeration economies and regional economic growth.

All Regional Dummies show statistical significance and are important in explaining regional GDP growth. This shows that there are significant differences between regions that affects growth. The Eastern region, consisting of Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi and Shangdong, tends to have higher regional GDP growth compared to the other regions. The Southern region, consisting of Sichuan,
Guizhou, Yunnan and Tibet, tends to have smaller regional GDP growth compared to the other regions. The base region used in this regression is the Southwestern region consisting of Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

The following table compares the results (estimated coefficient) obtained from Lu and this study.

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Present Study</th>
<th>Lu</th>
</tr>
</thead>
<tbody>
<tr>
<td>L</td>
<td>-0.012909</td>
<td>0.8893</td>
</tr>
<tr>
<td>K</td>
<td>-0.019719</td>
<td>0.1105</td>
</tr>
<tr>
<td>PWG</td>
<td>0.12526</td>
<td></td>
</tr>
<tr>
<td>RAIL</td>
<td>0.088746*</td>
<td>0.4263 * (Transportation)</td>
</tr>
<tr>
<td>HW</td>
<td>0.12439*</td>
<td></td>
</tr>
</tbody>
</table>

Note: * Statistical significance at the 90 percent confidence level

Note: The coefficients are interpreted as the percentage point change in GDP due to one percentage point increase in the independent variable

**Insignificant Production Input Variables**

Only two out of the five production input variables showed significant t-ratios. This raises the question of the reliability of the model and data used in this study. The
following sections discuss the possible estimation problems which caused the poor results.

**Estimation Problems**

Possible explanation of the poor results is that the estimation process is flawed due to the presence of the multicollinearity, simultaneity, mis-specification and data problems. These problems may cause the estimates to be biased and inefficient.

**Multicollinearity**

The term multicollinearity originally meant the existence of a perfect or exact linear relationship between some or all independent variables of a regression model. Today, the term is used in a broader sense to include the case where independent variables are correlated but not perfectly so. (Gujarati, p.320) In the event where multicollinearity exists, the estimators are still unbiased and efficient. However, the standard errors of the coefficients tends to be large and $t$-ratio tends to be insignificant. The test that will be used to test for Multicollinearity is the Auxiliary R-square. A ‘High Pair-wise correlation among regressors’\(^{25}\) rule of thumb will also be discussed.

The *Auxiliary R-square method*\(^{26}\) was used to test for multicollinearity. The Auxiliary R-squares are generated for each independent variable by regressing it on the other independent variables in the model. For example, given a simple model of:

$$Y = B_1 + B_2X_1 + B_3X_2 + B_4X_3 + e$$

\(^{25}\) See Gujarati (1995), pg. 335.

\(^{26}\) See Gujarati (1995), pg. 337.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Auxiliary R-square of independent variable $X_1$, $X_2$ and $X_3$ is generated by running the following regression:

\[
\text{OLS} \quad X_1 = A_1 + A_2 X_2 + A_3 X_3 + e_1
\]
\[
\text{OLS} \quad X_2 = G_1 + G_2 X_1 + G_3 X_3 + e_2
\]
\[
\text{OLS} \quad X_3 = H_1 + H_2 X_1 + H_3 X_2 + e_3
\]

Test criteria: Kline's rule of thumb: if any of the Auxiliary R-squares is higher than the original R-square (from the original model), then Multicollinearity is serious. (Gujarati, p.337)

Test results: The following is the test values generated using the Auxiliary R-square method.

<table>
<thead>
<tr>
<th>Independent Variable (XXX)</th>
<th>R-SQUARE OF (XXX) ON OTHER INDEPENDENT VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labor</td>
<td>0.1358</td>
</tr>
<tr>
<td>INV/Y</td>
<td>0.4474</td>
</tr>
<tr>
<td>PWG/Y</td>
<td>0.4188</td>
</tr>
<tr>
<td>RW</td>
<td>0.0856</td>
</tr>
<tr>
<td>HW</td>
<td>0.2658</td>
</tr>
<tr>
<td>NORTHEAS</td>
<td>0.2395</td>
</tr>
<tr>
<td>EAST</td>
<td>0.3940</td>
</tr>
<tr>
<td>CENTRAL</td>
<td>0.1452</td>
</tr>
<tr>
<td>SOUTH</td>
<td>0.4279</td>
</tr>
<tr>
<td>SW</td>
<td>0.2205</td>
</tr>
<tr>
<td>Y96</td>
<td>0.2789</td>
</tr>
<tr>
<td>Y97</td>
<td>0.2969</td>
</tr>
<tr>
<td>ROADEN3</td>
<td>0.3584</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>0.0000</td>
</tr>
</tbody>
</table>
**High Pair-wise correlation among regressors:** If the zero-order correlation coefficient between two independent variables is higher than 0.8, then multicollinearity is a serious problem. This simple test showed no evidence of multicollinearity in the data.

Conclusion: There are some evidence of *Multicollinearity* using the Auxiliary R-square method. Multicollinearity does not bias the estimated coefficients, and the standard errors are still correctly measured. However, in cases of high or near perfect multicollinearity, “the OLS estimators have large variances and covariances, making precise estimation difficult.” (Gujarati, p.327) Because of that, “the confidence level tends to be much wider, leading to the acceptance of ‘zero null hypothesis’ more readily.” (Gujarati, p.327) Also, the t-ratio of one or more coefficients tends to be insignificant.

**Simultaneity**

Simultaneity is a “situation in which the dependent variables are determined by the simultaneous interactions of several relationships” within the model. (Kennedy, 1992, p.54) In the model presented in this paper, simultaneity may be a problem because of the direction of causation between output and the inputs. For example, it may be true that growth in output induces the government to invest more into public infrastructure. It may also be true that output growth attracts labor migration into a province, instead of
labor growth causing output growth. Simultaneity can cause estimated coefficients to be biased.\footnote{As discussed in the Methodology section of this paper the Cost-function approach of estimation would avoid the simultaneity problem. Given the data that is available, we cannot pursue the Cost-function approach.}

In this analysis, precautions were taken to avoid the simultaneity problem. The independent variables (with the exception of Labor and dummy variables) are lagged one period. This means that we are analyzing the effects of public infrastructure growth on output growth of the next period. This will minimize the possibility of any feedback effect of output growth on public infrastructure growth.

\textit{Mis-Specification}

The actual production function model consists of two independent variables not included in this study. The reason for excluding these relevant variables is that the data is not available. The excluded variables are the quantity (stock) of land used directly in the production process, and the amount of raw materials used in the production of goods and services.

If the omitted variables are correlated with the included variables, the mis-specification will cause the estimated coefficients to be biased. If the omitted variables are not correlated with the included variables, the estimated coefficients will still be inefficient.
Errors of Measurement

Errors of measurement of the independent and dependent variables could also have caused the insignificant input coefficient. Even though the estimation would still give unbiased coefficients, they are inefficient. Errors of measurement in the dependent variable will not only cause estimates to be “bias[ed] but also inconsistent, that is, they remain bias[ed] even if the sample size n increases indefinitely.” (Gujarati, p.469)

In this paper, I suspect that the variables PWG/Y, INV/Y and Y are measured incorrectly. The use of the investment to output ratio seems to generate implausible high measures for the INV/Y and the PWG/Y variables. China’s high real GDP growth rates (Y) also been a topic of debate among scholars. The reliability of the method used to generate the GDP growth rate figures are still questionable.

Data Problems (Outliers)

The following exercise re-estimates the model established earlier in this paper after eliminating the outliers. Although each observation in the sample may contain valuable information for the determination of the relationship between the dependent and independent variables, they also cause the estimated coefficients to be inefficient and bias.

The following are the criteria use to remove the outliers.

Labor variable: remove if value of L is greater than 5%.

Rail variable: remove if value is greater that 15%.

---

Highway variable: remove if value is greater that 10%.

PWG variable: remove if value is greater that 10%.

Table 6 presents the result of the estimation. Removing the outliers did not seem to help in solving the relationship between the dependent and independent variables. The result of this regression is still not satisfactory. Most of the coefficients of the inputs turned out to be negative and insignificant, which is contradictory to the production function theory. However, the R-square did improve after the outliers were removed from the regression.

Table 6. OLS Estimation of the Production Function after removing outliers  
(Dependent Variable = GDP Growth Rate)

<table>
<thead>
<tr>
<th>Independent Variable</th>
<th>Coefficients</th>
<th>( t )-ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>( L )</td>
<td>0.143</td>
<td>0.6474</td>
</tr>
<tr>
<td>( K )</td>
<td>-0.024</td>
<td>-0.5159</td>
</tr>
<tr>
<td>( PWG )</td>
<td>-0.295</td>
<td>-1.587 *</td>
</tr>
<tr>
<td>( RAIL )</td>
<td>-0.001</td>
<td>-0.164 *</td>
</tr>
<tr>
<td>( HW )</td>
<td>0.091</td>
<td>0.6623</td>
</tr>
<tr>
<td>( NORTHEAS )</td>
<td>-1.627</td>
<td>-1.755 *</td>
</tr>
<tr>
<td>( EAST )</td>
<td>2.387</td>
<td>3.367 *</td>
</tr>
<tr>
<td>( CEN )</td>
<td>1.568</td>
<td>1.848 *</td>
</tr>
<tr>
<td>( SOUTH )</td>
<td>-1.410</td>
<td>-1.431 *</td>
</tr>
<tr>
<td>( SW )</td>
<td>-1.284</td>
<td>-1.554 *</td>
</tr>
<tr>
<td>( Y96 )</td>
<td>-0.128</td>
<td>-0.2257</td>
</tr>
<tr>
<td>( Y97 )</td>
<td>-1.210</td>
<td>-2.058 *</td>
</tr>
<tr>
<td>( ROADEN3 )</td>
<td>0.898</td>
<td>0.5203</td>
</tr>
<tr>
<td>CONSTANT</td>
<td>12.404</td>
<td>7.287 *</td>
</tr>
</tbody>
</table>

R-square 0.4706  
R-square adjusted 0.3631  
F-test Coefficients are not equal to zero.

Note: * Statistical significance at the 10 percent error level
CHAPTER 7

CONCLUSION

This paper examined the effects of public infrastructure investment on regional economic growth in China. Since the implementation of its dispersed, rural production system in the 1950s, China seems to be continuously lacking in public infrastructure to support its rapidly growing economy.

The World Bank (1994) noted that China's transportation network is among the thinnest in the world due to chronic underinvestment. (p. 18) Lu (1995) noted that there is an acute unfulfilled demand of railway shipments of 40%, and 25% in electricity in China in 1994. (p.3) This lack of public infrastructure in China has grown to be so severe that it is causing a bottleneck effect on the economy, restricting growth to a level much less than its potential.

The modified production function model (Lu, 1995) was used to estimate the effects of public infrastructure investment on regional economic growth in China. The dependent variable, GDP growth rate, was regressed against independent variables: Labor growth, Capital growth (proxy), Highway growth, Railway growth, PWG growth\(^{29}\) (proxy), Year dummies, Regional dummies, and Highway Density dummies. The Ordinary Least Squared (OLS) method was used to estimate the coefficients for the variables. Pooled data of 30 provinces from 1993 to 1997 were the sample.

\(^{29}\) PWG growth - Growth in power, water and gas infrastructure.
The result of the regression was less than satisfactory. The explanationary power (R-square) of the model employed was moderate (0.43). The estimated coefficient for Labor, Capital and PWG\textsuperscript{30} turned out to be insignificant. The estimated coefficients of the different public infrastructures were expected to be large because of the bottleneck effect in the public infrastructure sector. Instead, the estimated coefficients for the public infrastructure variables - Highway and Railway were small compared to Lu’s estimated public infrastructure coefficient. They are 0.11 and 0.07 respectively. However, the Regional Dummies and Year Dummies variables were all statistically significant in explaining regional growth rates in China.

The poor results indicated that there may be some estimation errors in the regression. Multicollinearity, simultaneity, mis-specification and data problems were some potential problems that may have caused the flawed results. Due to the insignificant t-ratios of most the production input variables, this present study are not able to verify the seriousness of the bottleneck in the public infrastructure sector in China.

**Recommendations for Further Study**

Perhaps the one most significant modification that can improve this present model is not to use the investment income ratio to proxy for capital growth rates. The investment income ratio proxy generated for the capital variable seemed to be implausibly large. The mean of the capital growth rate proxy is approximately 23% while the maximum is 52%. I strongly believe that if the measurement for capital stock is

\textsuperscript{30} Power, Water and Gas infrastructure.
available to calculate the actual growth rate, the results will be greatly improved. Unfortunately, capital stock measurements for China are not available. In fact, it may be impossible to estimate the capital stock for a planned economy since there is no market price for one to use in estimating the value of an infrastructure.

Another problem encountered in the regression is high growth rates measurements in the \textit{HIGHWAY} and \textit{RAILWAY} variables for less developed provinces. This is because these provinces usually have very small amount of infrastructure stock. A good example is the case of Tibet. The length of railways in Tibet in 1994 was 459 km. This figure increased to 714 km in 1995. This small increase of 255 km shows as 55% in terms of growth rates. These high growth rates in \textit{HIGHWAY} and \textit{RAILWAY} variables in less developed provinces may have biased the estimation. This problem can be solved by running different regressions for provinces with different level of developments. By dividing the 30 provinces to 3 categories of level of developments, the effects of the high infrastructure growth rates in less developed provinces can be isolated.
<table>
<thead>
<tr>
<th>Province</th>
<th>1995</th>
<th>1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.4</td>
<td>9.2</td>
</tr>
<tr>
<td>2</td>
<td>14.9</td>
<td>14.3</td>
</tr>
<tr>
<td>3</td>
<td>13.9</td>
<td>12.7</td>
</tr>
<tr>
<td>4</td>
<td>11.1</td>
<td>11.2</td>
</tr>
<tr>
<td>5</td>
<td>9.1</td>
<td>9.9</td>
</tr>
<tr>
<td>6</td>
<td>7.1</td>
<td>7.5</td>
</tr>
<tr>
<td>7</td>
<td>9.7</td>
<td>9.9</td>
</tr>
<tr>
<td>8</td>
<td>9.6</td>
<td>9.5</td>
</tr>
<tr>
<td>9</td>
<td>14.1</td>
<td>15.3</td>
</tr>
<tr>
<td>10</td>
<td>19.5</td>
<td>19.5</td>
</tr>
<tr>
<td>11</td>
<td>16.7</td>
<td>15.3</td>
</tr>
<tr>
<td>12</td>
<td>14.3</td>
<td>15.3</td>
</tr>
<tr>
<td>13</td>
<td>15.2</td>
<td>15.2</td>
</tr>
<tr>
<td>14</td>
<td>14.5</td>
<td>14.5</td>
</tr>
<tr>
<td>15</td>
<td>14.2</td>
<td>14.2</td>
</tr>
<tr>
<td>16</td>
<td>14.8</td>
<td>14.8</td>
</tr>
<tr>
<td>17</td>
<td>14.6</td>
<td>14.6</td>
</tr>
<tr>
<td>18</td>
<td>10.9</td>
<td>10.9</td>
</tr>
<tr>
<td>19</td>
<td>14.9</td>
<td>14.9</td>
</tr>
<tr>
<td>20</td>
<td>15.3</td>
<td>15.3</td>
</tr>
<tr>
<td>21</td>
<td>14.3</td>
<td>14.3</td>
</tr>
<tr>
<td>22</td>
<td>19.5</td>
<td>19.5</td>
</tr>
<tr>
<td>23</td>
<td>7.5</td>
<td>11.2</td>
</tr>
<tr>
<td>24</td>
<td>15.9</td>
<td>15.9</td>
</tr>
<tr>
<td>25</td>
<td>17.9</td>
<td>17.9</td>
</tr>
<tr>
<td>26</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>27</td>
<td>9.9</td>
<td>9.9</td>
</tr>
<tr>
<td>28</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>29</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>30</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>31</td>
<td>9.2</td>
<td>9.2</td>
</tr>
<tr>
<td>32</td>
<td>14.3</td>
<td>14.3</td>
</tr>
<tr>
<td>33</td>
<td>13.5</td>
<td>13.5</td>
</tr>
<tr>
<td>34</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>35</td>
<td>12.7</td>
<td>12.7</td>
</tr>
<tr>
<td>36</td>
<td>8.6</td>
<td>8.6</td>
</tr>
<tr>
<td>37</td>
<td>13.7</td>
<td>13.7</td>
</tr>
<tr>
<td>38</td>
<td>10.5</td>
<td>10.5</td>
</tr>
<tr>
<td>39</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>40</td>
<td>12.2</td>
<td>12.2</td>
</tr>
<tr>
<td>41</td>
<td>12.7</td>
<td>12.7</td>
</tr>
<tr>
<td>42</td>
<td>14.4</td>
<td>14.4</td>
</tr>
<tr>
<td>43</td>
<td>15.4</td>
<td>15.4</td>
</tr>
<tr>
<td>44</td>
<td>13.4</td>
<td>13.4</td>
</tr>
<tr>
<td>45</td>
<td>12.2</td>
<td>12.2</td>
</tr>
<tr>
<td>46</td>
<td>13.9</td>
<td>13.9</td>
</tr>
<tr>
<td>47</td>
<td>13.2</td>
<td>13.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Province</th>
<th>1995</th>
<th>1996</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yr</td>
<td>GDP</td>
<td>GDP</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>INV/Y*</td>
<td>PWG*</td>
<td>RAIL*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HW*</td>
<td>HIGHWAY DENSITY</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 7. Complete sample data**
<table>
<thead>
<tr>
<th></th>
<th>1996</th>
<th>1997</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>12.2</td>
<td>0.18</td>
</tr>
<tr>
<td>19</td>
<td>10.7</td>
<td>0.93</td>
</tr>
<tr>
<td>20</td>
<td>10.3</td>
<td>1.44</td>
</tr>
<tr>
<td>21</td>
<td>4.8</td>
<td>-0.09</td>
</tr>
<tr>
<td>22</td>
<td>10.1</td>
<td>-0.63</td>
</tr>
<tr>
<td>23</td>
<td>8.9</td>
<td>1.88</td>
</tr>
<tr>
<td>24</td>
<td>10.4</td>
<td>1.26</td>
</tr>
<tr>
<td>25</td>
<td>13.2</td>
<td>3.52</td>
</tr>
<tr>
<td>26</td>
<td>10.2</td>
<td>1.32</td>
</tr>
<tr>
<td>27</td>
<td>11.5</td>
<td>1.35</td>
</tr>
<tr>
<td>28</td>
<td>10.1</td>
<td>-0.63</td>
</tr>
<tr>
<td>29</td>
<td>18.1</td>
<td>2.67</td>
</tr>
<tr>
<td>30</td>
<td>6.4</td>
<td>1.42</td>
</tr>
<tr>
<td>1</td>
<td>9.6</td>
<td>-0.02</td>
</tr>
<tr>
<td>2</td>
<td>12.1</td>
<td>1.38</td>
</tr>
<tr>
<td>3</td>
<td>12.5</td>
<td>0.70</td>
</tr>
<tr>
<td>4</td>
<td>10.5</td>
<td>0.35</td>
</tr>
<tr>
<td>5</td>
<td>9.7</td>
<td>-0.72</td>
</tr>
<tr>
<td>6</td>
<td>8.9</td>
<td>1.60</td>
</tr>
<tr>
<td>7</td>
<td>9.2</td>
<td>-1.61</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>5.82</td>
</tr>
<tr>
<td>9</td>
<td>12.7</td>
<td>0.77</td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>-0.06</td>
</tr>
<tr>
<td>11</td>
<td>11.1</td>
<td>-0.06</td>
</tr>
<tr>
<td>12</td>
<td>12.7</td>
<td>2.33</td>
</tr>
<tr>
<td>13</td>
<td>14.5</td>
<td>1.25</td>
</tr>
<tr>
<td>14</td>
<td>11.5</td>
<td>0.64</td>
</tr>
<tr>
<td>15</td>
<td>11.2</td>
<td>1.23</td>
</tr>
<tr>
<td>16</td>
<td>10.4</td>
<td>3.89</td>
</tr>
<tr>
<td>17</td>
<td>13</td>
<td>0.61</td>
</tr>
<tr>
<td>18</td>
<td>10.8</td>
<td>1.22</td>
</tr>
<tr>
<td>19</td>
<td>10.6</td>
<td>2.54</td>
</tr>
<tr>
<td>20</td>
<td>9</td>
<td>1.47</td>
</tr>
<tr>
<td>21</td>
<td>6.7</td>
<td>-1.22</td>
</tr>
<tr>
<td>22</td>
<td>10.2</td>
<td>0.91</td>
</tr>
<tr>
<td>23</td>
<td>9</td>
<td>1.85</td>
</tr>
<tr>
<td>24</td>
<td>9.4</td>
<td>1.53</td>
</tr>
<tr>
<td>25</td>
<td>11.5</td>
<td>2.21</td>
</tr>
<tr>
<td>26</td>
<td>10</td>
<td>0.78</td>
</tr>
<tr>
<td>27</td>
<td>8.5</td>
<td>0.92</td>
</tr>
<tr>
<td>28</td>
<td>9</td>
<td>1.51</td>
</tr>
<tr>
<td>29</td>
<td>7.6</td>
<td>4.12</td>
</tr>
<tr>
<td>30</td>
<td>11</td>
<td>2.84</td>
</tr>
</tbody>
</table>

Note: * signifies growth rates.
**F-test**

The F-test is conducted to test if all the coefficients are equal to zero.

Critical value at 95% confidence level.

\[
F_{\text{crit}} = F_{0.05, \text{df1} = 13, \text{df2} = 76} = 1.85
\]

\[
F_{\text{test}} = \frac{\text{ESS}/(k-1)}{\text{RSS}/(n-k)} = 4.54
\]

Condition: If \( F_{\text{test}} > F_{\text{crit}} \), then reject \( H_0 \).

where \( K \) is the number of variable, \( n \) is the sample size,

\[
\text{ESS} = B_2 \sum y_i X_{2i} + B_3 \sum y_i X_{3i} + B_4 \sum y_i X_{4i} + \ldots + B_k \sum y_i X_{ki},
\]

\[
\text{RSS} = \sum e_i^2
\]

**Conclusion:** Reject \( H_0 \). There is sufficient evidence to prove that not all coefficients are equal to zero.
APPENDIX B

Sample SHAZAM Program

read (china3.txt) province year dy dl dk dpwg drail dhw density

* Generating Regional dummies
if (province.eq.1) north=1
if (province.eq.2) north=1
if (province.eq.3) north=1
if (province.eq.4) north=1
if (province.eq.5) north=1
if (province.eq.6) northeast=1
if (province.eq.7) northeast=1
if (province.eq.8) northeast=1
if (province.eq.9) east=1
if (province.eq.10) east=1
if (province.eq.11) east=1
if (province.eq.12) east=1
if (province.eq.13) east=1
if (province.eq.14) east=1
if (province.eq.15) east=1
if (province.eq.16) cen=1
if (province.eq.17) cen=1
if (province.eq.18) cen=1
if (province.eq.19) south=1
if (province.eq.20) south=1
if (province.eq.21) south=1
if (province.eq.22) sw=1
if (province.eq.23) sw=1
if (province.eq.24) sw=1
if (province.eq.25) northwest=1
if (province.eq.26) northwest=1
if (province.eq.27) northwest=1
if (province.eq.28) northwest=1
if (province.eq.29) northwest=1
if (province.eq.30) northwest=1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
* Generating Year dummies
if (year.eq.1995) y95=1
if (year.eq.1996) y96=1
if (year.eq.1997) y97=1

* Generating Highway Density dummy
if(density.gt.0.3) roaden3 = 1

* Estimating the ‘Modified’ Production function
-----------------------------------------------
ols dy dl dk dpwg drail dhw northeast east cen south sw y96 y97 &
roaden3/auxrsqr resid=e

* Testing for Heteroskedasticity: White’s Test

gen e2 = e*e
gen dl2=dl*dl
gen dk2=dk*dk
gen dpwg2=dpwg*dpwg
gen drail2=drail*drail
gen dhw2=dhw*dhw
gen dlk=dl*dk
gen dlp=dl*dpwg
gen dlhw=dl*dhw
gen dlrail=dl*drail
gen dkp=dk*dpwg
gen dkral=dk*drail
gen dkhw=dk*dhw
gen dprail=dpwg*drail
gen dphw=dpwg*dhw
gen drailhw=drail*dhw
ols e2 dl2 dk2 dpwg2 drail2 dhw2 dlk dlp dlhw drail dkp dkral dkhw &
dprail dphw drailhw dl dk dpwg drail dhw northeast east cen south sw &
y96 y97 roaden3

gen1 nr= $n*$r2
print nr

gen1 a=.05
gen1 df1=.$k-1$
distrib a/type=chi df=df1 inverse
* Estimation using Heteroskedasticity-consistent covariance matrix
(White’s Correction)

```
ols dy dl dk dpwg drail dhw northeast east cen south sw y96 y97 roaden3 &
/het resid=e1 anova
```

* F test
```
distrib a/type=f df1=13 df2=76 inverse
```

** Re-estimating the model without the outliers

```
* Removing outliers
skipif(dl.gt.5)
skipif(drail.gt.15)
skipif(dhw.gt.10)
skipif(dpwg.gt.10)
stat/all

ols dy dl dk dpwg drail dhw northeast east cen south sw y96 y97 &
roaden3/auxrsqr resid=e
```

* Testing for Heteroskedasticity: White's Test

```
gen e2 = e*e
gen dl2=dl*dl
gen dk2=dk*dk
gen dpwg2=dpwg*dpwg
gen drail2=drail*drail
gen dhw2=dhw*dhw
gen dlk=dl*dk
gen dlp=dl*dpwg
gen dlhw=dl*dhw
gen dlrail=dl*drail
gen dkp=dk*dpwg
gen dkrail=dk*drail
gen dkhw=dk*dhw
gen dprail=dpwg*drail
gen dphw=dpwg*dhw
```
gen d railhw = d rail * d hw
ols e2 dl2 dk2 dpwg2 d rail2 d hw2 dlk d lp d h hw d rail d kp d krail d khw &
d prail d phw d railhw dl dk dpwg d rail d hw northeast east cen south sw &
y96 y97 roaden3

gen1 nr = n * r2
print nr

gen1 a = .05
gen1 df1 = k - 1
distrib a / type = chi df = df1 inverse

stop
BIBLIOGRAPHY


