



Research Paper No. 2006/153

## China's Regional Inequality in Innovation Capability, 1995-2004

Peilei Fan<sup>1</sup> and Guanghua Wan<sup>2</sup>

December 2006

### Abstract

Relying on a recently developed decomposition framework, this paper explores spatial distribution of innovation capability in China. It is found that at the regional level, China's inequality in innovation capability increased from 1995 to 2004. At the provincial level, the inequality decreased from 1995 to 2000, but increased from 2000 to 2004. Location, industrialization and urbanization, human capital, and openness (foreign direct investment) are significant contributors to the inequality in innovation capability. Unbalanced development in high-tech parks exerts a growing explanatory power in driving innovation disparity, which implies that institutional factor plays a direct role.

Keywords: innovation, regional disparity, inequality, decomposition, Asia, China

JEL classification: R12

---

Copyright © UNU-WIDER 2006

<sup>1</sup>UNU-IAS, Yokohama; <sup>2</sup>UNU-WIDER, Helsinki.

This study has been prepared within the UNU-WIDER project on Inequality and Poverty in China.

UNU-WIDER acknowledges with thanks the financial contributions to its research programme by the governments of Denmark (Royal Ministry of Foreign Affairs), Finland (Ministry for Foreign Affairs), Norway (Royal Ministry of Foreign Affairs), Sweden (Swedish International Development Cooperation Agency—Sida) and the United Kingdom (Department for International Development).

ISSN 1810-2611

ISBN 92-9190-937-8

ISBN-13 978-92-9190-937-7

*The World Institute for Development Economics Research (WIDER) was established by the United Nations University (UNU) as its first research and training centre and started work in Helsinki, Finland in 1985. The Institute undertakes applied research and policy analysis on structural changes affecting the developing and transitional economies, provides a forum for the advocacy of policies leading to robust, equitable and environmentally sustainable growth, and promotes capacity strengthening and training in the field of economic and social policy making. Work is carried out by staff researchers and visiting scholars in Helsinki and through networks of collaborating scholars and institutions around the world.*

*www.wider.unu.edu*

*publications@wider.unu.edu*

UNU World Institute for Development Economics Research (UNU-WIDER)  
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Lorraine Telfer-Taivainen at UNU-WIDER

The views expressed in this publication are those of the author(s). Publication does not imply endorsement by the Institute or the United Nations University, nor by the programme/project sponsors, of any of the views expressed.

## 1 Introduction

It has been widely acknowledged that innovation has become an increasingly important determinant of economic growth (Schumpeter 1942; Barsberg 1987; Malecki 1987; Fargerberg 1994; Fargerberg et al. 1997; Malecki 1997). OECD (1988, 2001) documents the contributions of innovation to GDP growth for several of its member countries. For instance, Japan's miraculous development after the Second World War was fuelled by its technical progress. Especially in the 1960s, technological advances constituted 6.1 per cent of Japan's GDP growth, with the remaining 3.5 per cent attributable to labour and capital inputs. Similarly, a substantial proportion of growth in the USA can be attributed to technological progress, particularly in the late 1990s, when the USA led the world in information and communication technology (ICT).

Recognizing the importance of innovation, the Chinese government has been instrumental in directing the country towards a knowledge economy. Recent policy initiatives include the National High-tech Research and Development Plan (863 Plan), the National Basic Sciences Initiative (973 Plan), and the Torch programme that specifically aims at facilitating commercialization of scientific research outcomes. Furthermore, the government has set up 53 national high-tech parks to attract foreign high-tech firms and to encourage the development of domestic high-tech companies. These policy initiatives have undoubtedly promoted innovation activities in China. According to a recent study, technological progress contributed more than 40 per cent of the remarkable economic growth rate in China during the period 1981-2000 (Fan and Watanabe 2006).

Given the predominant role of innovation in economic growth in China and the fast rising regional inequality in China (Wan et al. 2006, 2007), two issues deserve special research attention. First, it is interesting to examine innovation inequality in China. This has been undertaken in terms of patent application (Sun 2000). Second, more importantly, it is crucial to analyze sources or contributing factors of innovation inequality. Although studies exist, which focus on determinants of innovation capabilities in China and elsewhere (Guerrero and Sero 1997), no previous attempt has been made to quantify contributions of various determinants to the inequality of innovation. The typical regression model in the current literature (e.g., Sun 2000; Guerrero and Sero 1997) can only be used to gauge the impacts of independent variables on the level of innovation, not the inequality of innovation. The latter requires the regression-based inequality decomposition of Wan (2002, 2004).

This paper represents a first attempt to measure and analyze factor contributions to China's regional inequality in innovation capability from 1995 to 2004. To be more precise, we intent to focus on the following research questions:

What was the status of China's regional inequality in innovation capability and how did it change from 1995 to 2004? How much did the relevant factors contribute to the level of innovation capability and how much did they contribute to the regional inequality in innovation capability?

The plan of the paper is as follows. In Section 2, we will review the limited literature on innovation inequality in China and provide preliminary data analysis. This is followed by measurement of innovation inequality in China in Section 3. Section 4 presents regression analysis as well as decomposition of innovation inequality. The summary and policy recommendations are provided in Section 5.

## **2 Literature review and preliminary data analysis**

Few studies explored innovation inequality in China with the exception of Sun (2000, 2003), who used a primary index, top-five index, top-ten index, and coefficient of variation to indicate spatial pattern of innovation. Sun (2000) found that patents in China were highly clustered in the east-coastal region and the inland provinces, although the degree of spatial concentration declined during 1985-95. When other indicators of innovation such as new products sales and R&D (research and development) spending were used, the spatial concentration is found to be on the rise in the 1990s (Sun 2003).

Sun (2003) classifies the provinces into two groups and applies the logistic regression to model the cluster membership resulting from the classification. Provinces in Cluster 1 spend more on in-house R&D, and those in Cluster 2 spend more on imported technologies. The four independent variables for the logistic regression are GDP per capita, size of S&T (science and technology) staff, ratio of international trade to GDP, and a coast-inland dummy variable. However, the model did not work well as none of the independent variables were significant.

Innovation capability can be measured by different indexes such as R&D inputs, patent counts, patent citations, and new product announcement. In this paper, we follow Audretsch and Feldman (2004: 2713-39) and use patent data as a proxy measure for innovation capability.<sup>1</sup> Hagedoorn and Cloudt (2003) find that statistical overlap between various innovativeness indicators is strong and any of these indicators, including patent, may be used to measure innovation capability. In fact, patent is generally accepted as one of the most appropriate measure for innovation capability (Mansfield 1986; Barsberg 1987; Griliches 1990).

---

<sup>1</sup> Patent data have some weaknesses, such as sectoral difference in patenting behaviour, difference in patenting between large and small firms. Further, not all the innovations are patented and the patent counts equalize the economic significance of different new technologies.

Table 1: Patents in China, 1995 and 2004

Province	1995				2004				R	
	Total	inventions	uti. models	designs	Total	inventions	uti. models	designs	1995	2004
Beijing	4025	328	3169	528	9005	3216	3956	1833	10.40	5.72
Tianjin	1034	63	785	186	2578	432	1587	559	3.55	2.36
Hebei	1580	56	1341	183	3407	357	2064	986	0.79	0.47
Liaoning	2745	131	2362	252	5749	911	3752	1086	2.17	1.26
Shanghai	1436	72	1025	339	10625	1687	4040	4898	3.28	5.74
Jiangsu	2413	72	1884	457	11330	1026	5474	4830	1.10	1.41
Zhejiang	2131	54	1455	622	15249	785	5492	8972	1.59	3.01
Fujian	933	17	439	477	4758	160	1776	2822	0.93	1.26
Shandong	2861	84	2222	555	9733	788	6028	2917	1.06	0.99
Guangdong	4611	56	1447	3108	31446	1941	9307	20198	2.17	3.66
Guangxi	665	20	457	188	1272	127	666	479	0.47	0.24
Hainan	108	4	44	60	278	36	93	149	0.48	0.32
Shanxi	569	47	480	42	1189	295	636	258	0.60	0.33
Inner Mongolia	415	8	293	114	831	108	437	286	0.59	0.32
Jilin	824	38	723	63	2145	451	1179	515	1.03	0.73
Heilongjiang	1403	44	1248	111	2809	326	1997	486	1.23	0.68
Anhui	574	18	469	87	1607	150	972	485	0.31	0.23
Jiangxi	509	19	402	88	1169	105	625	439	0.40	0.25
Henan	1145	34	1009	102	3318	306	2117	895	0.41	0.32
Hubei	1017	55	868	94	3280	744	1966	570	0.57	0.51
Hunan	1515	51	1318	146	3281	436	1801	1044	0.77	0.46
Sichuan	2019	79	1486	454	8031	730	3069	4232	0.58	0.63
Guizhou	274	12	207	55	737	179	364	194	0.25	0.18
Yunnan	569	35	346	188	1264	235	586	443	0.46	0.27
Shaanxi	1085	52	934	99	2007	459	1193	355	1.00	0.50
Gansu	257	7	215	35	514	127	322	65	0.34	0.18
Qinghai	65	2	61	2	70	21	30	19	0.44	0.12
Ningxia	111	4	98	9	293	46	119	128	0.70	0.47
Xinjiang	312	9	286	17	792	75	530	187	0.61	0.38
E. China	24542	957	16630	6955	105430	11466	44235	49729	1.60	1.82
C. China	7971	314	6810	847	19629	2921	11730	4978	0.60	0.40
W. China	4692	200	3633	859	13708	1872	6213	5623	0.55	0.43
Total	37205	1471	27073	8661	138767	16259	62178	60330	1.00	1.00

## Notes:

1. Because of the unavailable data for Tibet for early years, we exclude Tibet from our analysis.
2. The total patents granted do not include Hong Kong, Macau, and Taiwan. The figures of Sichuan include the figures of Chongqing City.
3. n.a. = not applicable or not available. Location quotient is calculated for total patents only.
4. The total population used in our calculation does not include the military population, nor the population of Hong Kong, Macau, and Taiwan.

Source: National Bureau of Statistics (various).

In China, the State Bureau of Intellectual Properties examines and certifies different types of patents. The invention patents refer to those that show ‘novelty’ and have been developed to the point where they can be utilized in industry. The utility model patents are creations or improvements relating to the form, construction, or fitting of an object, with a lower technical requirement than invention patents. The design patents refer to original designs relating to the shape, pattern, colour or a combination thereof of objects.<sup>2</sup>

Relying on *China Statistical Yearbook on Science and Technology* (1995-2005) and *China Statistical Yearbook* (1994-2005),<sup>3</sup> we compiled a set of data on innovation and relevant determinants at the regional (by region, we mean east, central and west China in this paper) or provincial level. Table 1 displays numbers of granted patents in 1995 and 2004, including numbers of invention, utility model, and design. The table reveals that eastern China dominated certified patents, especially Beijing, Liaoning, Shandong, and Guangdong. In sharp contrast, central and particularly west China fell behind. We also report regional per capita patents relative to the national average, denoted by  $R$ . Thus a region with  $R > 1$  performs better in creating patents than the national average, and vice versa.

The  $R$  values for a particular year indicate innovation inequality between regions in China. Over time, eastern China increased its  $R$  value from 1.60 in 1995 to 1.82 in 2004, which means that the gap between the eastern regions and others had expanded. Looking into individual regions, some provinces/cities (e.g. Beijing, Tianjin, Heilongjiang, and Shaanxi) underwent significant drop while others (such as Shanghai, Guangdong, and Zhejiang) experienced substantial gains in  $R$  values. In 2004, all provinces in eastern China, except Hebei, Shandong, Guangxi, and Hainan, possessed  $R$  values greater than one. On the contrary, provinces in central and western China all had  $R$  values smaller than one.

### **3 Measuring innovation inequality**

To formally measure innovation inequality, we use the following five indices: the Atkinson index ( $A$ ), the generalized entropy family  $GE(0)$  and  $GE(1)$ , the squared coefficient of variation ( $CV^2$ ), and the Gini coefficient ( $Gini$ ).

Let  $y$  denote the number of patents per 10,000 persons,  $N$  denote the total number of the provinces/regions, and  $\bar{y}$  denote the mean of  $y$ . The Atkinson index is defined as:

---

<sup>2</sup> China State Intellectual Property Office; see website at [www.sipo.gov.cn](http://www.sipo.gov.cn).

<sup>3</sup> National Bureau of Statistics (1994-2005a,b).

$$A = 1 - \frac{1}{\bar{y}} \left( \prod_{i=1}^N y_i \right)^{1/N} \quad (1)$$

$$GE(0) = \frac{1}{N} \sum_{i=1}^N \ln \frac{\bar{y}}{y_i} \quad (2)$$

$$GE(0) = \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\bar{y}} \ln \frac{\bar{y}}{y_i} \quad (3)$$

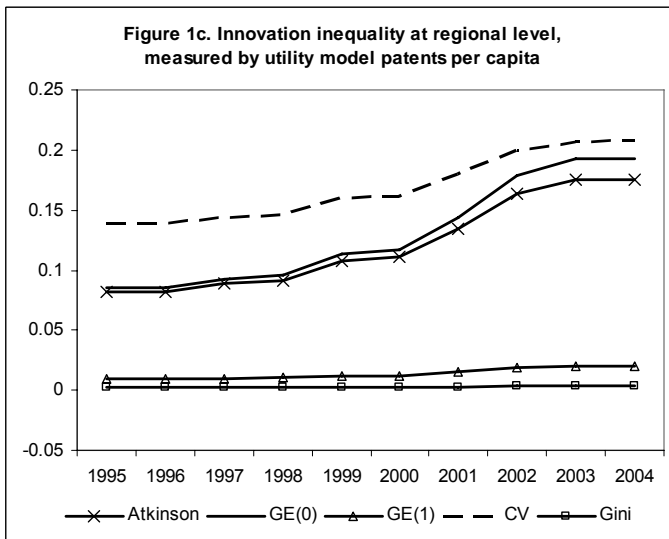
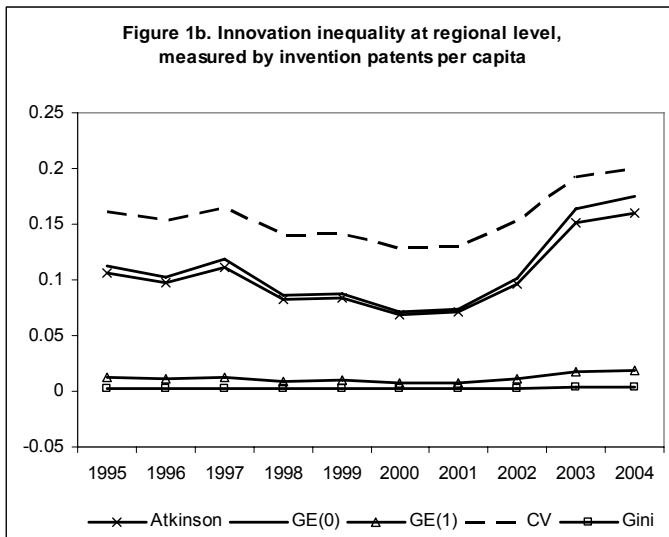
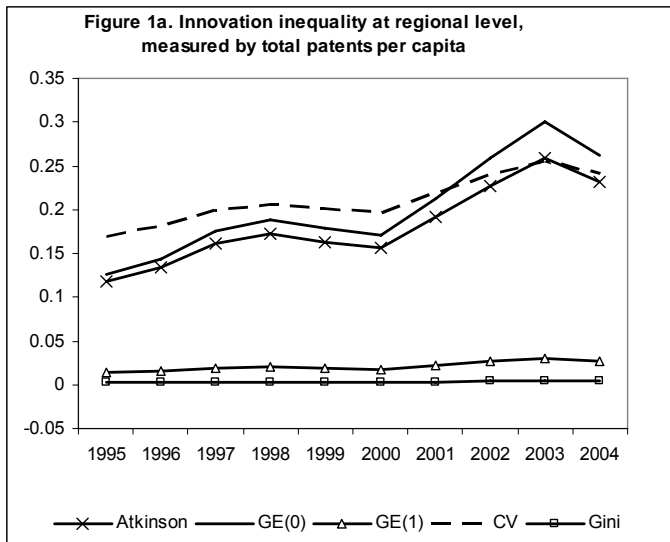
The squared coefficient of variation ( $CV^2$ ) can be expressed as:

$$CV^2 = \frac{1}{\bar{y}^2} \left[ \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 \right] \quad (4)$$

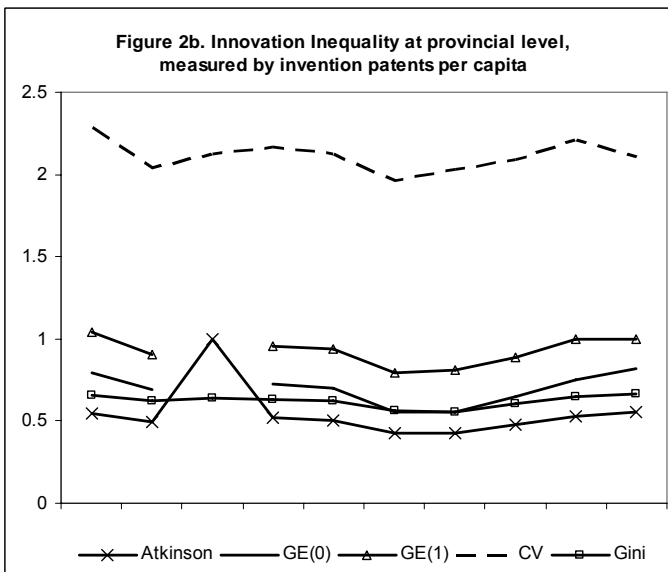
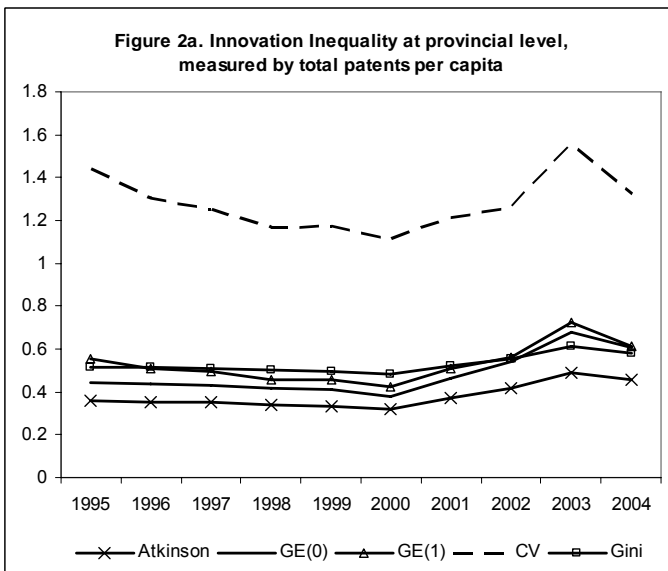
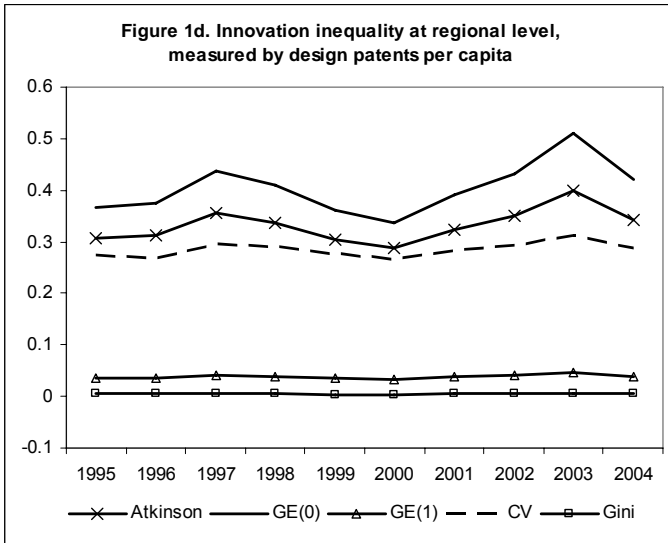
The Gini index is defined as:

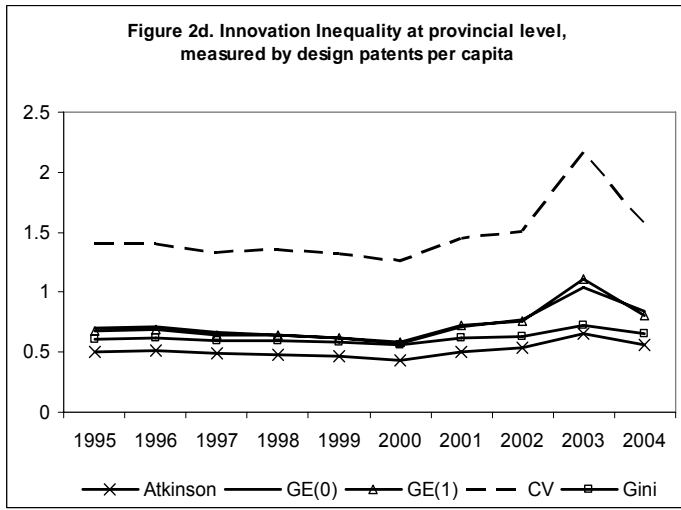
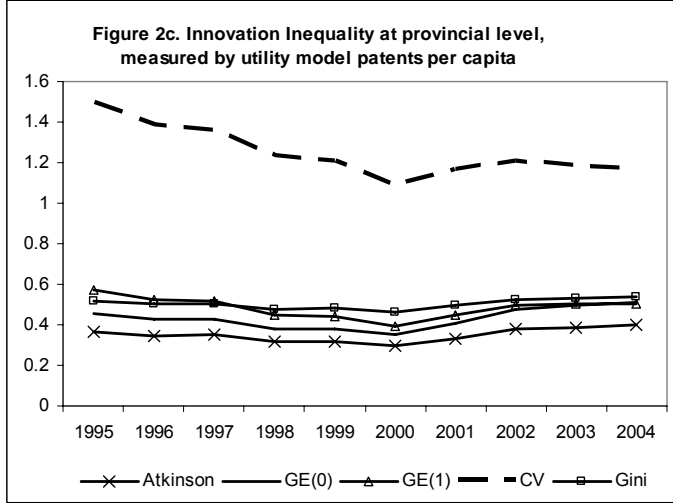
$$Gini = \frac{\sum_{i=1}^N \sum_{j=1}^N |y_i - y_j|}{2N\bar{y}^2} \quad (5)$$

Regional inequality in innovation in terms of total patent number and its components are displayed in Figures 1 and 2. All indexes show that the inequality increased from 1995 to 2004 (Figure 1) with a notable drop in 2000, although the trend differs somehow for different indicators of innovation. Interestingly, design patents are found to be distributed most unequally. On the other hand, Figure 2 conveys a mixed message. For invention patents, all indexes showed that inequality oscillated around the same level. For utility model and design patents, the inequality decreased from 1995 to 2000, but increased from 2000 to 2004. This pattern is shared by the total patent, which is not surprising as utility model and design patents constitute a great portion of the total patent. These results seem to be in line with Sun (2000, 2003), which found a decline in patent concentration from 1985 to 1995 and an increasing concentration afterwards. It should be noted that Sun used total patent number rather than patent per capita in calculating the inequality indexes.









#### 4 Decomposing innovation inequality

In order to implement regression-based inequality decomposition, it is necessary to model innovation as a knowledge production function. A conventional functional form is:

$$Y_i = \alpha RD_i^\beta HK_i^\gamma \varepsilon_i \quad (6)$$

where  $Y$  stands for the output of innovation activity,  $RD$  represents R&D inputs,  $HK$  represents human capital inputs. The subscript  $i$  represents the unit of observation, such as regions, industries, or enterprises (Audretsch and Feldman 2004). To incorporate the spatial dimension, Jaffe (1989) modified the above traditional model and used:

$$Y_{si} = \alpha IRD_i^{\beta 1} * UR_{si}^{\beta 2} * (UR_{si} * GC_{si}^{\beta 3}) * \varepsilon_{si} \quad (7)$$

where *IRD* is private corporate expenditure on R&D, *UR* is the research expenditures occurred at universities, and *GC* measures the geographic coincidence of university and corporate research. The subscripts *s* (*i*) represents a state (industry). Equation (7) was also employed by Acs et al. (1992) and Feldman (1994) in modelling innovation.

In studying China's patent distribution, Sun (2000) used the same model specification and found that the level of regional development, R&D, openness (import, export, and presence of foreign enterprises), and agglomeration (urbanization) are significant determinants of patent production. However, Sun's study and the alike do not permit quantification of the impacts of various input factors on the spatial inequality of innovations. This is because the regression model can only be used to explain the level of innovation capability, not its spatial variation. In what follows, we will first estimate a flexible regression model and then use the estimated model to conduct inequality decomposition.

#### **4.1 Regression analysis: innovation capability**

Using the published consumer price indexes (CPIs) and the spatial price index of Brandt and Holz (2006) for 2004, we deflated all observations in value terms. To deal with the delay between innovation inputs and output, we follow Guerrero and Sero (1997) by lagging independent variables by one year in our empirical model. This also helps to address the possible endogeneity problem.

We follow the 'general-to-specific' modelling strategy in this paper. In other words, we start with as many explanatory variables as possible subject to data availability. To explain the per capita number of patents (*Y*), education and R&D funding ought to be considered. Consequently, average years of schooling of labour force (*Edu*) and per capita public R&D funding (*RD*) are included as independent variables. Following Sun (2000), rate of urbanization (*Urban*) is used to capture possible agglomeration effects. Since most inventions occur in the non-farming sectors, it is necessary to control for structure of economic activities. For this purpose, share of agricultural GDP (*Stru*) is taken as an independent variable. We also included per capita GDP (*GDP*) and the trade/GDP ratio (*Opent*) and FDI/GDP ratio (*Openf*). *GDP* represents level of development and *Opent* and *Openf* may crowd out domestic innovative capability. Needless to say, it is necessary to incorporate location dummy variables. In this paper, we define *DB*, *DG*, *DS*, and *DT* for Beijing, Guangdong, Shanghai, and Tianjin, respectively. To allow for nonlinearity, we also include the squares of education and R&D in the model. Difference in the impacts of education and R&D on inventions across regional belts are considered by adding *D1\*Edu* and *D2\*Edu*, where *D1* is the dummy for east China and *D2* for central China. Finally, time trend (*Year*) is used to denote reform and other time-dependent forces underlying innovation. To facilitate robustness test, several variables are considered and they include total population (*POP*), consumer price index (*CPI*) and per capita value of hi-tech product (*HT*). Population size may bring about economies of scale or economies of specialization in

innovative activities. *CPI* signals macroeconomic environment and *HT* may reflect non-public R&D inputs.

Regarding functional forms, we consider four possibilities instead of, as in almost all earlier studies of innovation, sticking to a particular specification. The Lin-Lin specification involves no transformation to the dependent or independent variables. The Log-Log specification involves taking logarithms to both the dependent and independent variables, with the exception of dummy variables. As discussed earlier, this specification is commonly used in the literature. By the same token, we can have the Lin-Log and Log-Lin specifications.

Fitting the four specifications to the Chinese data produce results as tabulated in Table 2. Models with the same dependent variable can be easily compared by examining the log-likelihood values. In doing so, we eliminate the Lin-Log and Log-Log models. To choose between the Lin-Lin and Log-Lin models, the  $\chi^2$  test of Cox can be utilized. The test result indicates preference of the Log-Lin model. Based on the Log-Lin model, R&D is not a significant determinant of innovation in China. This is understandable for at least two reasons. First, public R&D accounts for a fairly small percentage of total R&D input. More importantly, R&D in China is not appropriately managed. Misallocation and corruption are well-known. To many academics, it would be surprising to find R&D to be significant. It is useful to note that this insignificance is robust to model specification (Table 2). Like R&D, the variables *POP*, *CPI*, and *D2\*Edu* are not significant in any of the equations.

Now, we drop the five variables which are insignificant (*RD*, *RD*<sup>2</sup>, *POP*, *CPI*, and *D2\*Edu*) and estimate the four models again, with results tabulated in Table 3. Repeating the earlier model selection procedure, we end up with the Log-Lin functional form again. As is evident from Table 3, the selected model is of good quality in terms of common sense, statistical properties and innovation production theory. The only insignificant variable is the *HT* variable. However, *HT* is significant in other models and in earlier estimation (see Table 3). More importantly, we believe that *HT* is a relevant and important determinant of innovation outputs. Therefore, we decide to keep this variable in the model.

Table 2: Summary of regression models, with initial 19 variables

VARIABLE NAME	Log-Lin			Lin-Lin			Log-Log			Lin-Log		
	Standardized Coefficient	T- RATIO	P- VALUE	Standardized Coefficient	T- RATIO	P- VALUE	Standardized Coefficient	T- RATIO	P- VALUE	Standardized Coefficient	T- RATIO	P- VALUE
RD	-0.02	-0.53	0.600	0.02	0.28	0.782	0.07	0.98	0.328	0.13	1.00	0.317
RD <sup>2</sup>	0.01	0.32	0.746	0.00	-0.06	0.952	-0.07	-0.93	0.353	-0.10	-0.80	0.424
EDU	0.72	2.45	0.014	-0.51	-1.10	0.270	0.30	0.67	0.505	-2.38	-2.72	0.006
EDU2	-0.90	-2.98	0.003	0.42	0.88	0.377	-0.34	-0.73	0.468	2.27	2.55	0.011
OPENF	-0.13	-4.34	0.000	-0.06	-1.08	0.279	-0.03	-0.85	0.396	-0.01	-0.09	0.930
OPENT	-0.02	-1.42	0.156	0.01	0.41	0.685	-0.06	-2.16	0.031	-0.06	-1.38	0.168
URBAN	0.15	4.01	0.000	0.05	0.60	0.547	0.15	5.01	0.000	0.05	0.56	0.578
STRU	-0.31	-9.85	0.000	0.22	1.94	0.053	-0.57	-11.73	0.000	-0.83	-5.01	0.000
GDP	0.06	2.65	0.008	-0.02	-0.44	0.657	0.05	2.17	0.030	0.00	-0.03	0.979
YEAR	0.24	8.31	0.000	0.22	3.49	0.000	0.09	4.19	0.000	0.10	1.97	0.049
DB	0.25	8.01	0.000	0.31	4.79	0.000	0.09	2.75	0.006	0.13	1.48	0.138
DG	0.20	6.73	0.000	0.24	2.38	0.017	0.18	3.74	0.000	0.13	1.21	0.226
DS	0.13	2.98	0.003	0.35	5.55	0.000	-0.13	-3.05	0.002	-0.17	-1.52	0.130
DT	0.07	3.26	0.001	0.10	1.63	0.102	-0.07	-3.00	0.003	-0.22	-3.07	0.002
HT	0.04	1.50	0.133	0.34	6.76	0.000	0.28	8.69	0.000	0.21	2.99	0.003
POP	-0.01	-0.53	0.594	-0.04	-1.57	0.117	-0.02	-1.03	0.304	-0.03	-1.01	0.312
CPI	0.00	0.00	0.997	-0.03	-0.61	0.540	0.02	0.67	0.503	0.03	0.57	0.570
D1*EDU	0.50	10.04	0.000	0.43	3.65	0.000	0.23	4.09	0.000	0.22	1.54	0.123
D2*EDU	0.05	1.27	0.205	0.03	0.24	0.809	0.00	-0.03	0.977	0.05	0.33	0.743
CONSTANT	0.00	-3.15	0.002	0.00	0.80	0.421	0.00	-0.61	0.539	0.00	3.46	0.001

Same est. $\rho$ for all cross-sections	0.87	0.67	0.84	0.72
BUSE [1973] $R^2$	0.91	0.79	0.89	0.63
BUSE raw moment $R^2$	0.92	0.87	0.92	0.77
Variance of the estimate- $\sigma^2$	0.03	0.14	0.03	0.17
Standard error of the estimate- $\sigma$	0.18	0.37	0.19	0.41
Sum of squared errors (SSE)	9.00	40.46	10.03	47.96
Mean of dependant variable	-0.85	0.76	-0.85	0.76
Log of the likelihood function	74.42	-151.97	61.74	-179.09

---

Source: Authors' calculations.

Table 3. Summary of regression models, with 14 variables

VARIABLE NAME	Log-Lin			Lin-Lin			Log-Log			Lin-Log		
	Standardized Coefficient	T-RATIO	P-VALUE	Standardized Coefficient	T-RATIO	P-VALUE	Standardized Coefficient	T-RATIO	P-VALUE	Standardized Coefficient	T-RATIO	P-VALUE
EDU	0.70	2.34	0.019	-0.25	-0.49	0.621	0.29	0.65	0.516	-2.43	-2.79	0.005
EDU <sup>2</sup>	-0.87	-2.84	0.005	0.15	0.30	0.766	-0.33	-0.71	0.481	2.35	2.65	0.008
OPENF	-0.13	-4.25	0.000	-0.07	-1.25	0.211	-0.03	-0.88	0.381	-0.01	-0.10	0.922
OPENT	-0.02	-2.01	0.044	0.00	0.07	0.944	-0.03	-2.07	0.038	-0.02	-0.69	0.491
URBAN	0.14	3.76	0.000	0.00	0.05	0.961	0.14	4.82	0.000	0.04	0.55	0.580
STRU	-0.31	-9.80	0.000	0.20	1.81	0.071	-0.57	-11.84	0.000	-0.80	-4.71	0.000
GDP	0.05	2.67	0.008	0.00	-0.04	0.968	0.06	2.64	0.008	0.01	0.21	0.836
YEAR	0.24	8.27	0.000	0.23	3.47	0.001	0.09	4.21	0.000	0.09	1.84	0.066
DB	0.25	7.87	0.000	0.35	5.17	0.000	0.08	2.65	0.008	0.20	2.98	0.003
DG	0.20	6.73	0.000	0.23	2.39	0.017	0.17	3.79	0.000	0.20	1.93	0.053
DS	0.14	3.01	0.003	0.37	5.50	0.000	-0.13	-3.02	0.003	-0.07	-0.76	0.449
DT	0.07	3.22	0.001	0.11	1.94	0.053	-0.07	-3.02	0.003	-0.13	-2.47	0.013
HT	0.04	1.48	0.139	0.34	6.61	0.000	0.28	8.83	0.000	0.20	3.03	0.002
D1*EDU	0.46	11.75	0.000	0.41	3.99	0.000	0.24	6.13	0.000	0.01	0.08	0.937
CONSTANT	0.00	-2.96	0.003	0.00	0.21	0.837	0.00	-0.61	0.544	0.00	3.47	0.001
Same est. p for all cross-sections		0.88			0.67			0.84			0.74	
BUSE [1973] R <sup>2</sup>		0.91			0.78			0.88			0.61	
BUSE raw moment R <sup>2</sup>		0.92			0.87			0.92			0.76	
Variance of the estimate- $\sigma^2$		0.03			0.14			0.03			0.17	
Standard error of the estimate- $\sigma$		0.18			0.37			0.19			0.41	
Sum of squared errors (SSE)		9.03			40.68			10.10			48.20	
Mean of dependant variable		-0.85			0.76			-0.85			0.76	
Log of the likelihood function		73.87			-152.13			60.78			-179.84	

Source: Authors' calculations.

Table 4: Robustness test

VARIABLE NAME	Add POP and CPI			Add POP			Add CPI		
	Standardized Coefficient	T-RATIO	P-VALUE	Standardized Coefficient	T-RATIO	P-VALUE	Standardized Coefficient	T-RATIO	P-VALUE
EDU	0.66	2.31	0.021	0.66	2.31	0.021	0.66	2.33	0.020
EDU2	-0.05	-2.81	0.005	-0.05	-2.81	0.005	-0.05	-2.83	0.005
OPENF	-0.03	-4.31	0.000	-0.03	-4.27	0.000	-0.03	-4.25	0.000
OPENT	0.00	-1.36	0.173	0.00	-2.08	0.038	0.00	-1.66	0.098
URBAN	0.01	3.63	0.000	0.01	3.76	0.000	0.01	3.75	0.000
STRU	-0.04	-9.75	0.000	-0.04	-9.74	0.000	-0.04	-9.80	0.000
GDP	0.12	2.59	0.010	0.11	2.57	0.010	0.11	2.50	0.013
YEAR	0.08	8.37	0.000	0.08	8.34	0.000	0.08	8.27	0.000
DB	1.31	7.96	0.000	1.31	7.91	0.000	1.31	7.85	0.000
DG	1.03	6.56	0.000	1.04	6.56	0.000	1.04	6.72	0.000
DS	0.73	3.14	0.002	0.71	3.09	0.002	0.71	3.00	0.003
DT	0.39	3.38	0.001	0.38	3.30	0.001	0.38	3.20	0.001
HT	0.00	1.48	0.138	0.00	1.48	0.138	0.00	1.48	0.139
D1*EDU	0.12	11.89	0.000	0.12	11.79	0.000	0.12	11.72	0.000
POP	0.00	-0.86	0.389	0.00	-0.76	0.447			
CPI	-0.05	-0.41	0.685				0.01	0.08	0.940
CONSTANT	-3.07	-2.87	0.004	-3.12	-2.93	0.003	-3.17	-2.96	0.003



Same est. $\rho$ for all cross-sections	0.88	0.88	0.88
BUSE [1973] $R^2$	0.91	0.91	0.91
BUSE raw moment $R^2$	0.92	0.92	0.92
Variance of the estimate- $\sigma^2$	0.03	0.03	0.03
Standard error of the estimate- $\sigma$	0.18	0.18	0.18
Sum of squared errors (SSE)	9.00	9.01	9.03
Mean of dependant variable	-0.85	-0.85	-0.85
Log of the likelihood function	74.16	74.08	73.87

---

Source: Authors' calculations.

As the final step in our modelling procedure, robustness test is conducted by adding some possibly relevant variables into the selected Log-Lin model. We added both Pop and CPI, and added each separately. The results are presented in Table 4. Clearly, our selected model is fairly robust. In passing, it is noted that the commonly used double log model is rejected here thus earlier studies adopting this functional form are likely to suffer from misspecification errors.

#### **4.2 Regression-based inequality decomposition**

Based on the selected regression model, we now conduct the decomposition exercise to quantify contributions of relevant factors to the inequality of innovation capability in China. It is known that different measures of inequality often produce different results, which may carry over to inequality decomposition. Consequently, we consider the four of the five inequality measures defined earlier in this paper. These include  $GE(0)$ ,  $GE(1)$ , the squared coefficient of variation ( $CV^2$ ), and the Gini coefficient.

Follow Wan et al. (2007), we collect all regional dummy variables terms, and name it *Loc* to represent location, the  $DI * Edu$  is merged with *Edu* and *Edu2* to form *HC*, human capital. Naturally, *Opent* and *Openf* are combined to represent globalization denoted by *Open*. All other variables retain their original definitions. To measure the inequality of innovation, not inequality of its logarithm, we solve the estimated model for *Y* and then proceed to inequality decomposition using the technique developed by Wan (2004). To briefly explain, the contribution of the residual represents those made by variables not included in our model. All remaining inequalities are explainable by variables included in our regression. See Wan (2004), Wan and Zhou (2005) and Wan et al. (2007). The decomposition results are tabulated in Table 5.

Several findings deserve particular discussion. First, location is found to contribute the most to the inequality (around 30-40 per cent). This is not surprising as location captures the effects of culture, tradition and even policy biases. For example, Beijing, Tianjin, Guangdong, and Shanghai are China's traditional and emerging innovation centres. As the national political and manufacturing centres of the Huabei Area, Beijing and Tianjin have enjoyed more higher education institutions, national research institutes, and national industrial bases. Guangdong rose to a local power house economically due to the economic reform and open policy in the 1980s. Shanghai has gained ever more rapid development since Pudong New Area was established in the 1990s. It is worth mentioning that all these locations belong to east China. The location variable may also capture other socio-economic or institutional factors that are critical to innovation but were not included in our model.

Table 5a: Result of decomposition (1995-2000)

1995	GE 0	GE 1	CV	Gini	1996	GE 0	GE 1	CV	Gini
Inequality	0.45	0.55	1.44	0.52	Inequality	0.44	0.51	1.31	0.51
Contribution (%)					Contribution (%)				
Human Capital	28	25	27	29	Human Capital	28	26	28	29
Open	-9	-8	-10	-3	Open	-8	-7	-8	-2
Urban	11	12	17	11	Urban	12	14	20	11
Structure	30	28	34	30	Structure	33	33	41	31
GDP	-2	-2	-2	0	GDP	-1	-1	-2	0
Location	34	36	46	30	Location	37	42	56	31
High-Tech	1	1	1	1	High-Tech	1	1	2	1
Subtotal	94	91	113	97	Subtotal	102	107	139	101
Residue	6	9	-13	3	Residue	-2	-7	-39	-1
1997	GE 0	GE 1	CV	Gini	1998	GE 0	GE 1	CV	Gini
Inequality	0.43	0.49	1.25	0.51	Inequality	0.41	0.46	1.17	0.50
Contribution (%)					Contribution (%)				
Human Capital	27	25	25	28	Human Capital	30	28	29	30
Open	-10	-9	-10	-4	Open	-10	-9	-10	-4
Urban	11	13	19	11	Urban	12	14	21	11
Structure	33	33	42	31	Structure	33	34	43	31
GDP	-1	-1	-2	0	GDP	-1	-2	-2	0
Location	36	41	56	31	Location	39	46	62	32
High-Tech	1	1	2	1	High-Tech	1	2	3	1
Subtotal	97	102	131	98	Subtotal	104	114	147	102
Residue	3	-2	-31	2	Residue	-4	-14	-47	-2
1999	GE 0	GE 1	CV	Gini	2000	GE 0	GE 1	CV	Gini
Inequality	0.41	0.46	1.17	0.50	Inequality	0.38	0.43	1.12	0.48
Contribution (%)					Contribution (%)				
Human Capital	30	28	27	30	Human Capital	29	26	23	30
Open	-10	-10	-11	-4	Open	-8	-8	-10	-3
Urban	12	14	20	11	Urban	12	14	19	11
Structure	32	33	41	30	Structure	32	32	38	30
GDP	-1	-1	-2	0	GDP	-1	-2	-2	0
Location	38	44	57	32	Location	40	46	57	33
High-Tech	131	138	70	97	High-Tech	4	5	7	3
Subtotal	104	111	136	102	Subtotal	107	113	134	103
Residue	-4	-11	-36	-2	Residue	-7	-13	-34	-3

Source: Authors' calculations.

Table 5b: Result of decomposition (2001-04)

2001	GE 0	GE 1	CV	Gini	2002	GE 0	GE 1	CV	Gini
Inequality	0.46	0.51	1.22	0.52	Inequality	0.54	0.56	1.26	0.56
Contribution (%)					Contribution (%)				
Human Capital	26	24	24	29	Human Capital	21	20	20	27
Open	-6	-5	-7	-3	Open	-5	-5	-6	-3
Urban	12	13	20	12	Urban	9	10	16	10
Structure	25	26	34	26	Structure	20	22	30	23
GDP	-1	-2	-3	0	GDP	-1	-2	-3	0
Location	34	40	55	31	Location	28	34	50	28
High-Tech	4	5	8	3	High-Tech	3	4	6	3
Subtotal	95	101	132	98	Subtotal	74	83	114	89
Residue	5	-1	-32	2	Residue	26	17	-14	11
<hr/>									
2003	GE 0	GE 1	CV	Gini	2004	GE 0	GE 1	CV	Gini
Inequality	0.68	0.73	1.55	0.61	Inequality	0.61	0.61	1.32	0.58
Contribution (%)					Contribution (%)				
Human Capital	17	16	16	24	Human Capital	19	18	19	26
Open	-3	-3	-3	-2	Open	-4	-4	-5	-2
Urban	7	9	15	9	Urban	8	10	16	9
Structure	16	17	25	20	Structure	17	19	28	20
GDP	-1	-1	-2	0	GDP	0	-1	-2	1
Location	24	29	45	27	Location	26	34	53	28
High-Tech	4	5	10	4	High-Tech	6	8	14	6
Subtotal	65	71	107	83	Subtotal	72	84	123	88
Residue	35	29	-7	17	Residue	28	16	-23	12

Source: Authors' calculations.

Second, economic structure (share of agricultural GDP) has a significant impact on inequality. This essentially reflects the impact of different pace of industrialization on innovation inequality. A positive contribution of this variable echoes well the significant contribution of the urbanization variable. Both of these findings support our earlier argument that China's certified patent are generated mainly by industrial/service sectors, which are mostly located in urban area. Although China is advanced in agriculture research and technology, most innovations in the agricultural sector are created by public research institutes in Beijing and those provinces which do not have a high share of agricultural GDP.

Third, human capital is found to be the third most important factor in driving the inequality in China's innovation capability. This finding is consistent with a priori expectation that human resource is an important input of knowledge production (Audretsch and Feldman 2004). Silicon Valley and the Boston metropolitan area are

typical examples to illustrate the crucial importance of human capital in affecting knowledge-based industries. They are clustered by computing and biomedical industries which take advantage of the rich human resources that the regions can supply through prominent universities and research institutes. Similarly, in China, high-tech firms, regardless domestic or multinational, tend to locate in places such as Beijing, Shanghai, and coastal provinces where there exists a large skilled labour force.

Fourth, the negative contribution of openness implies that trade and foreign direct investment (FDI) can help mitigate disparity in innovation. Our regression model indicates that both trade and FDI are negatively associated with innovation, confirming the crowding-out effects frequently discussed in the literature. Finally, we observe that the contribution of high-tech development (per capita revenue from high-tech parks) to innovation disparity increased from 1 per cent in 1995 to 6-10 per cent in 2004. The increase demonstrates that high-tech parks, initiated by the national government, have successfully stimulated high-tech industrial development in the respective regions, leading to higher regional inequality in innovative capability.

It seems pertinent to placing results from the regression and decomposition analysis in the context of the changing national innovation system (NIS) in China. More engagement from the business sectors, especially large- and medium-enterprises, represents the most prominent change of China's NIS. As a result, industrial enterprises generated most of utility model and design patents. Though universities and colleges are major owners of invention patents, the role of industrial enterprises has increased significantly over time (Fan and Watanabe 2006). In fact, enterprises invested heavily in innovation activities, reflected by its share of S&T funding and R&D expenditure. While China's S&T funding increased to 258 billion yuan in 2001, more than six times of that in 1991, the contribution of enterprises grew more than ten times, reaching 56 per cent of the total in 2001. Similarly, large- and medium-enterprises increased their R&D spending from 14 billion Yuan in 1995 to 42 billion yuan in 2001, contributing to 42 per cent of the national R&D expenditure (Fan and Watanabe 2006). Furthermore, building innovation capability has become the first priority of some domestic high-tech companies. A study on these firms, such as Huawei, ZTE, DTT, and GDT, found that innovation capability and self-developed technologies have been the key to the leading domestic firms' catching up with the multinational corporations (Fan 2006).

Our research findings are at odd with that of Sun (2000), who found that R&D and agglomeration are not significant factors for innovations in China. This inconsistency can be largely attributed to the inefficiency in China's innovation system in the early period examined by Sun. Our results suggest that China has improved its effectiveness of R&D activities in the past decade.

## 5 Summary

This paper uses several indexes to measure the inequality in innovation capability in China from 1995 to 2004. It reveals that the east-central-west inequality has increased over time, whereas the inter-provincial inequality showed a V-pattern. Major factors driving these inequality trends are location, industrialization and urbanization, human capital, and openness (foreign direct investment). Unbalanced development in high-tech parks is found to play an increasing role in causing innovation disparity in China. As innovation capability plays a vital role for growth, the found increasing inequality can seriously affect lagging region's catching up in economic development. Accordingly, policy makers in China should focus on promoting enterprises' involvement in innovation and nurturing domestic high-tech companies in the inland provinces. Also, our results appeal for a more equalized approach in human capital investment.

## References

- Acs, Z.J., D.B. Audretsch, and M.P. Feldman (1992). 'Real Effects of Academic Research', *American Economic Review* 82: 363-7.
- Audretsch, D., and M. Feldman (2004). 'Knowledge Spillovers and the Geography of Innovation', in J.V. Henderson and J.F. Thisse (eds), *Handbook of Regional and Urban Economics* Vol. 4, Elsevier: Amsterdam.
- Barsberg, B.L. (1987). 'Patents and the Measurement of Technological Change: A Survey of the Literature', *Research Policy* 16: 131-41.
- Brandt, L., and C. Holz (2006). 'Spatial Price Differentiation in China: Estimates and Implications', *Economic Development and Cultural Change* 55: 43-86.
- Fan, P. (2006). 'Catching Up Through Developing Innovation Capability: Evidence From China's Telecom Equipment Industry', *Technovation* 26: 359-68.
- Fan, P., and C. Watanabe (2006). 'Promoting Industrial Development Through Technology Policy: Lessons From Japan and China', *Technology in Society* 28(3): 303-20.
- Fargerberg, J. (1994). 'Technology and International Differences in Growth Rates', *Journal of Economic Literature* 31(1): 147-75.
- Fargerberg, J., B. Verspagen, and M. Caniels (1997). 'Technology, Growth and Unemployment across European Regions', *Regional Studies* 31: 457-66.
- Feldman, M. (1994). *The Geography of Innovation*, Kluwer Academic: Boston MA.
- Griliches, Z. (1990). 'Patent Statistics as Economic Indicators: A Survey', *Journal of Economic Literature* 28: 1661-707.
- Guerrero, D.C. and M.A. Sero (1997). 'Spatial Distribution of Patents in Spain: Determining Factors and Consequences on Regional Development', *Regional Studies* 31: 381-90.

- Hagedoorn, J., and M. Cloudt (2003). 'Measuring Innovative Performance: Is There an Advantage in Using Multiple Indicators?', *Research Policy* 32: 1365-79.
- Jaffe, A.B. (1989). 'Real Effects of Academic Research', *American Economic Review* 79: 967-70.
- Malecki, E.J. (1987). 'Hope or Hyperbole? High Tech and Economic Development', *Technology Review* 90: 44-51.
- Malecki, E.J. (1997). *Technology and Economic Development*, Longman: London.
- Mansfield, E. (1986). 'Patents and Innovation: An Empirical Study', *Management Science* 32: 173-81.
- National Bureau of Statistics (1994-2005a). *China Statistic Year Book*, China National Bureau of Statistics: Beijing.
- National Bureau of Statistics (1995-2005b). *China Statistical Yearbook on Science and Technology*, China National Bureau of Statistics: Beijing.
- OECD (Organization for Economic Co-operation and Development) (1988). *OECD Economic Studies*, OECD: Paris.
- OECD (Organization for Economic Co-operation and Development) (2001). *European Competitive Report*, OECD: Paris.
- Schumpeter, J. (1942). *Capitalism, Socialism, and Democracy*, Harpercollins: London.
- Sun, Y. (2000). 'Spatial Distribution of Patents in China', *Regional Studies* 34(5): 441-54.
- Sun, Y. (2003). 'Geographic Patterns of Industrial Innovation in China During the 1990s', *Journal of Economic and Social Geography* 94(3): 376-89.
- Wan, G.H. (2002). 'Regression-based Inequality Decomposition: Pitfalls and a Solution Procedure', *Discussion Paper 2002/101*, UNU-WIDER: Helsinki.
- Wan, G.H. (2004). 'Accounting for Income Inequality in Rural China: A Regression-based Approach', *Journal of Comparative Economics* 32: 348-63.
- Wan, G.H., and Z.Y. Zhou (2005). 'Income Inequality in Rural China: Regression-based Decomposition Using Household Data', *Review of Development Economics* 9:107-20.
- Wan, G.H., M. Lu, and Z. Chen (2006). 'The Inequality–Growth Nexus in the Short and Long Run: Empirical Evidence from China', *Journal of Comparative Economics* 34(4): 654-67.
- Wan, G.H., M. Lu, and Z. Chen (forthcoming 2007). 'Globalization and Regional Inequality: Evidence from within China', *Review of Income and Wealth*.