The Influence of Population Quality Competitiveness to Regional Innovation: the China Case

Yihui Xie[†]
Center for Applied Statistics, Renmin University of China

Abstract
The human resource is playing a more and more important role for a nation's innovation, therefore it is necessary to make further analysis on different population qualities and examine how they influence the capability of innovation respectively. This paper is mainly focused on the research of relationship between the population quality competitiveness and regional innovation and aims to find out the most important factors that lead to stronger competitiveness for China.

Keywords: population quality, regional innovation, competitiveness

1. Introduction

Nowadays the world economy is changing from the industrial economy to the knowledge-based economy, and the modern international competition has been turned into the competition of science, knowledge and intellectuals, which is, in fact, the competition of human resources, especially of talents. Among all the factors that influence the productivity, human resource is the most active one, and it is the most valuable resource compared with other sorts of resources, playing a leading role in the activities of a nation's innovation.

After more than 20 years' reformation and implementation of the opening-up policy, China's economy has made rapid development. But China is still a developing country; the human resources need to be further rationalized and strengthened substantially in many aspects, such as the size, structure, utilization, mobility and quality and so on. From an international viewpoint, the adjustment of the industrial structure and the pace of the labor transfer are being accelerated, the skills of workers are being constantly improved, and the quality of national life is being continuously promoted, and international competition is becoming increasingly fierce. Facing the new challenges, we have to expand the scale of the human resource, adjust its structure, enhance people's knowledge levels and labor skills, improve the quality of life, increase the Chinese population quality, and ultimately promote the innovation capability of the whole nation.

This paper is based on the population data of the "China Statistical Yearbook 2007", and gives an overview of the features of China's population and an exploration of the relationship between human capital and regional innovation from a perspective of population qualities.

[†] Address: School of Statistics, Room 1037, Mingde Main Building, Renmin University of China, Beijing, 100872, China. Email: xieyihui@gmail.com

2. Overview of China's Population Qualities

(1) Historical Development

Based on the data from the last five times of national population census in history, there is a clear increasing trend in the population quality of China. Figure 1~4 has displayed a variety of changing characteristics of China's population ever since 1953. (1) Age structure: strengthened labor force and less burden for supporting families, as the proportion of the young and the old is decreasing; (2) Education level: educational quality is also increasing; (3) Life structure: higher average life expectancy has indicated a healthier population; (4) Region structure: increasing proportion of urban population just shows better living conditions.

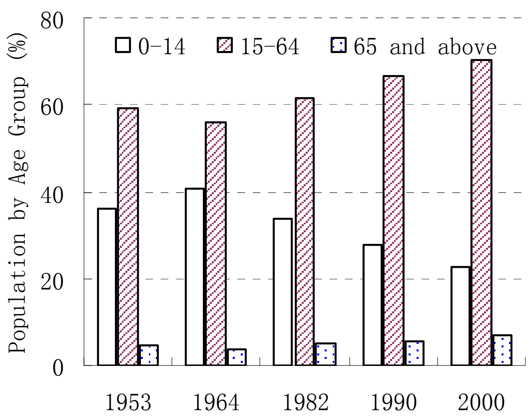

Figure 1 Population age structure: 1953~2000

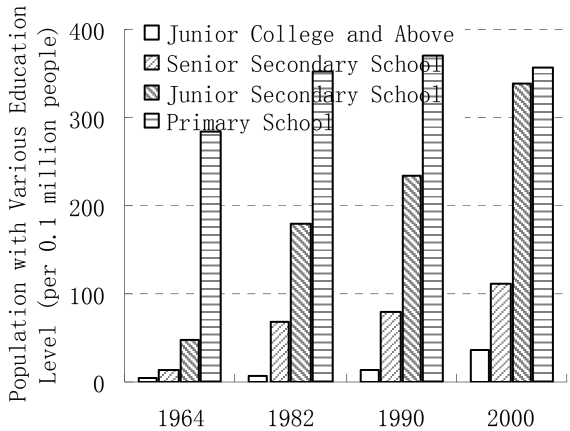

Figure 2 Education level structure: 1964~2000

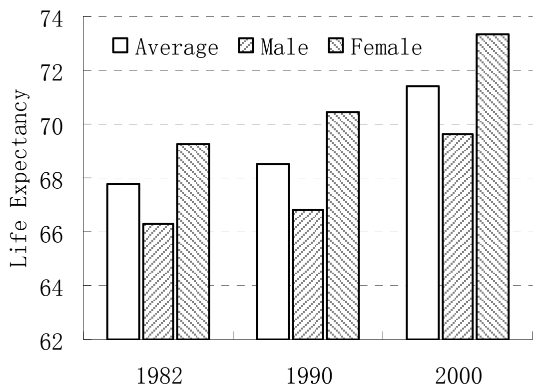

Figure 3 Gender and life structure: 1953~2000

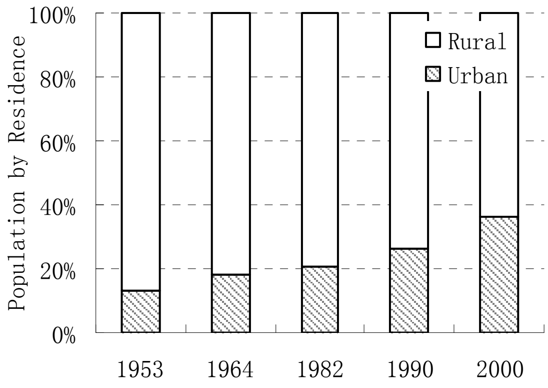

Figure 4 Region structure: 1964~2000

In all, the population qualities are increasing year by year, and we hope to make a further exploration to the relationship between population qualities and innovation (especially regional innovation), so we give a more specific description of population qualities for the 31 regions next.

(2) Regional Distribution

We will examine five aspects of the population in this section: gross quantity of the population, growth rate, health quality, educational quality and living standard. Figure 5 is an overview of China's population qualities distributed in the 31 regions. The width of the thermometers is proportional to the number of people in each region; the height is proportional to the growth rate; the "degree of temperature" has indicated the proportion of urban people in a region.

The two box-and-whisker plots alongside of the axes denote the distributions of the number of people with higher education and life expectancy respectively. The contour lines have shown the 2D density based the two variables. Obviously all the 31 regions in China are clustered from the viewpoint of population qualities: eastern regions are more developed and with better population qualities, while the middle and western regions are less developed and with weaker population qualities consequently. Figure 6 is a perspective plot corresponding to the contour lines in Figure 5, which might help us observe the clustered characteristic more clearly (there are obviously 3 "peaks").

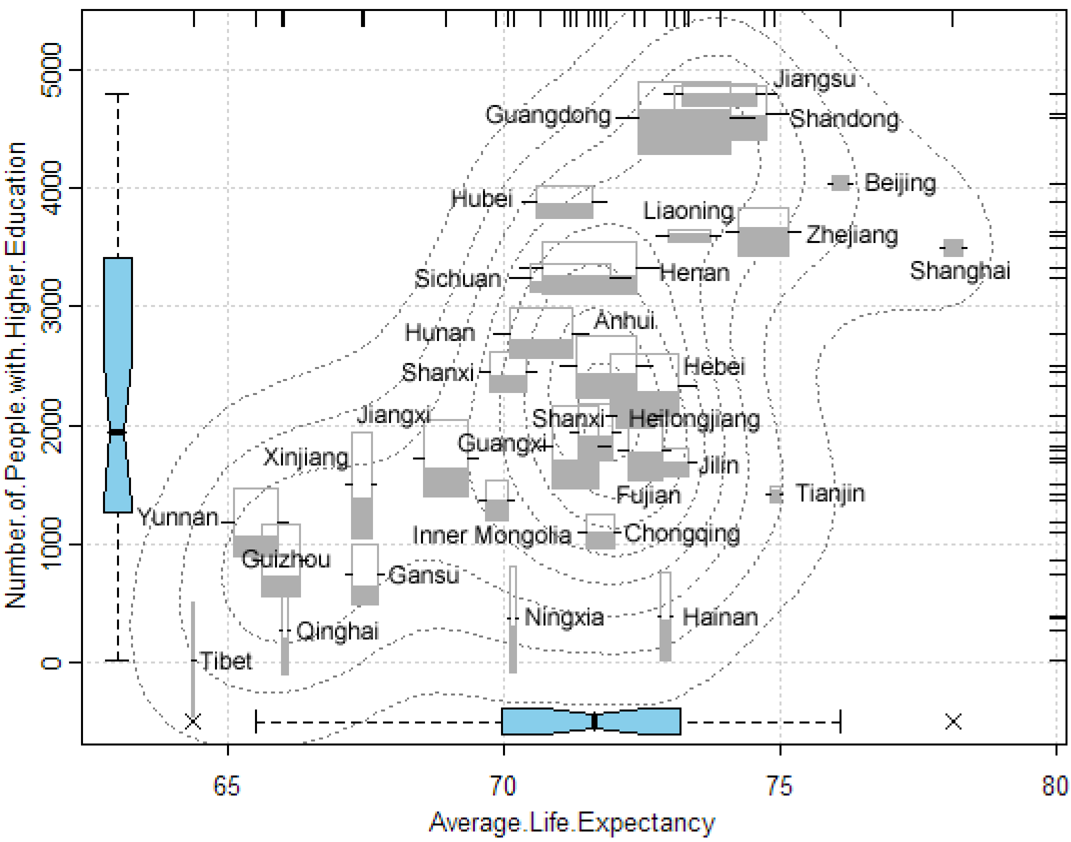

Figure 5 Overview of population qualities in 31 regions: gross quantity of human resource, growth rate, health quality, educational quality and living standard

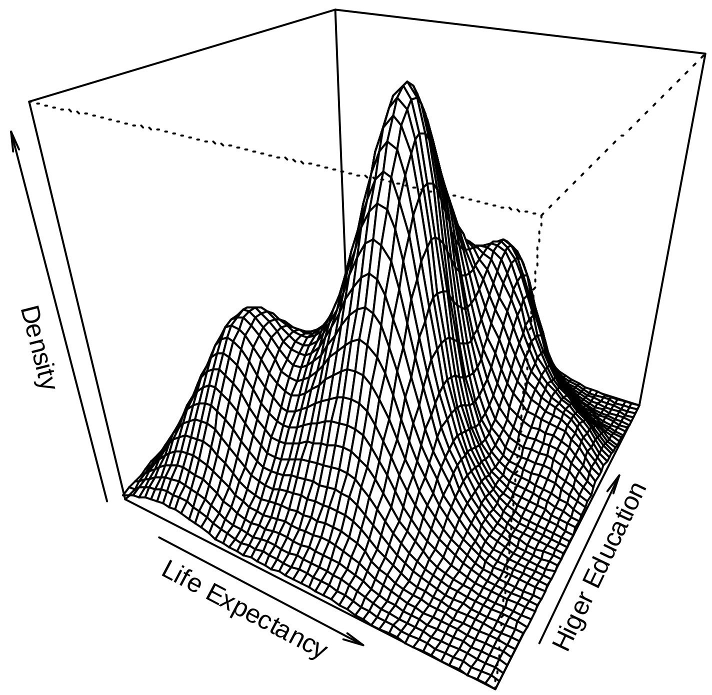

Figure 6 The joint distribution of life expectancy and higher education

3. Population Qualities and Regional Innovation

(1) Indicators of Innovation

Patent is the most common indicator to evaluate the output of innovation, e.g. Griliches (1990) and Acs *et al* (1989) have justified that it is reasonable to use the patent data to make research on the output of innovation. Compared with other indicators, patent as a measurement of innovation has three apparent advantages: (1) the statistical data for patents is available from the current system of statistics; (2) the actual meaning of patent is closely related to innovation; (3) the standard for granting patents is relatively objective and the changes are rather slow. Therefore, we also employ this indictor as the main variable in measuring the capability of innovation. At the same time, we also want to know the eventual value of innovation, thus another indicator "the average transaction value in technical market" is adopted in our framework too.

Here we may first of all check the relationship between GDP per capita, the average transaction value in technical market and the number of patents (accepted and granted respectively) using a scatter plot matrix in Figure 7.

Apparently the number of patents has a strong influence on the economy, so does the average transaction value in technical market. As the numbers of patents accepted and granted are nearly collinear, we only use the number of patents granted as the sole indicator in following sections.

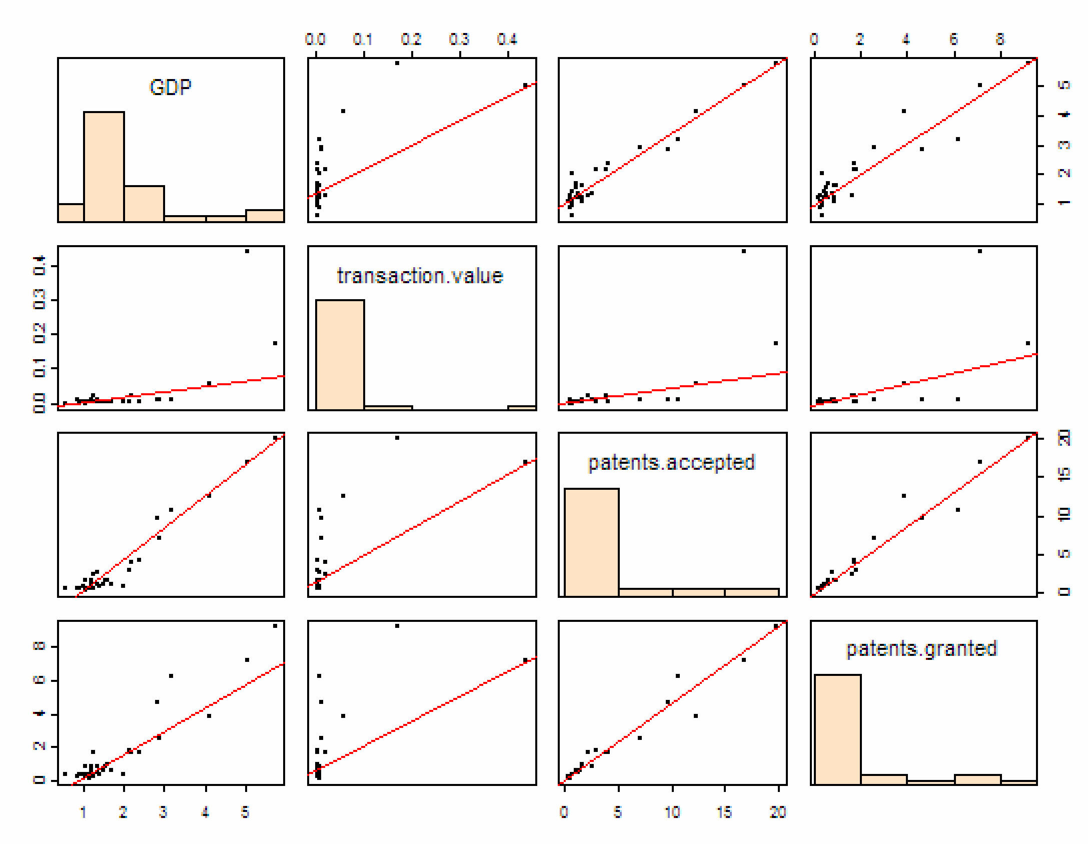

Figure 7 The relationship between GDP per capita, the average transaction value in technical market and the number of patents per capita.

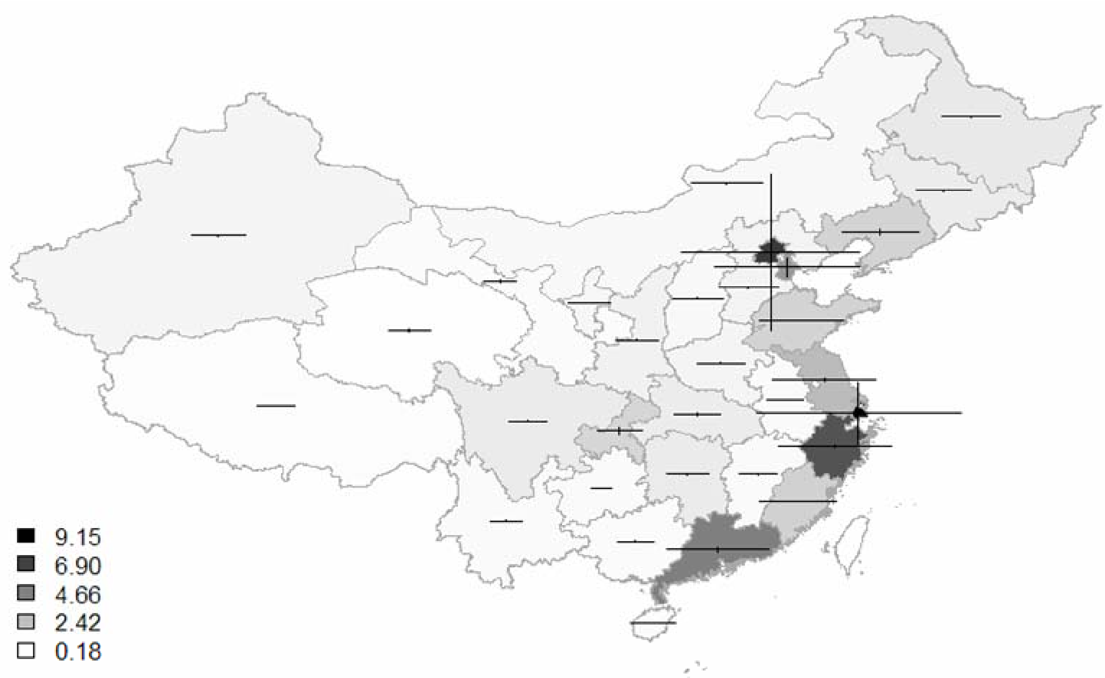

Figure 8 Regional distribution of innovation output of China's 31 regions

Figure 8 displays the regional distribution of innovation output denoted by the number of patents. Darker regions mean they have more patents per capita; the length of horizontal lines is proportional to GDP per capita, while the height of vertical lines is proportional to the average transaction value per capita in technical market. It is clear that Beijing, Tianjin, Shanghai, Guangdong and Zhejiang are the top 5 regions in the innovation output. And again, there is a phenomenon of clustering geometrically – provinces in the southeast have stronger competitiveness in innovation output. Naturally we will make a further analysis to explore the reasons for this clustering phenomenon.

(2) Influence of Population Qualities on Innovation

In this section we consider five indicators of population qualities: (1) Living condition ("Life"): denoted by the proportion of urban people – larger proportion can indicate better living conditions; (2) Health quality ("Health"): denoted by life expectancy – longer lives mean better health; (3) Population mobility rate ("Mobility"): computed from the sampling survey data – it is the proportion of people who move from their hometowns to other regions; (4) Educational quality ("Education"): denoted by the proportion of people with higher education; (5) Employment situation ("Employment"): denoted by the registered unemployment rate in urban areas. Figure 9 shows the influence of these indicators on the innovation output by a scatter plot matrix with smoothed lines.

All these five indicators except the last one have positive influence on the innovation output, and this is quite reasonable. Among these indicators, special attention should be paid to the "mobility", which reveals a fact that higher mobility of the population tends to bring higher innovation output. This preliminary conclusion can be heuristic for further statistical modeling.

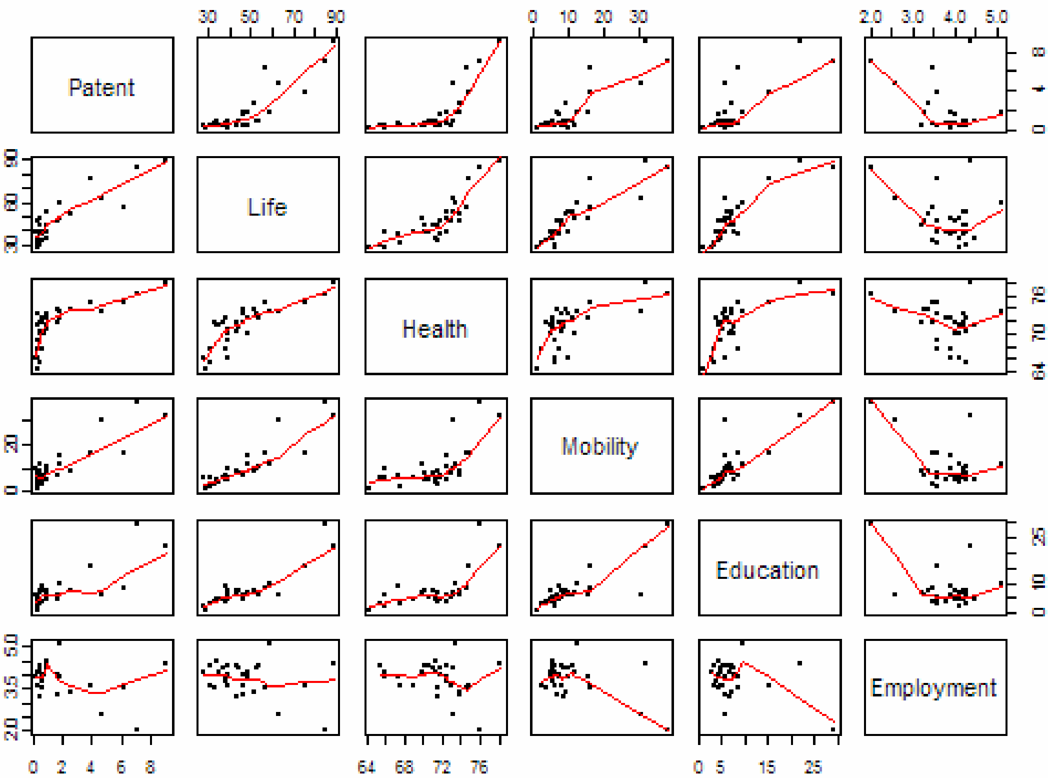

Figure 9 Relationship of population qualities and innovation output

We employ the traditional multivariate regression first to explore relationships between the five exploratory variables and the dependent variable, and the result is summarized in Table 1. It is acknowledged that sometimes the traditional multivariate regression model has disadvantages, for example, the existence of multicollinearity will lead to difficult explanations of regression coefficients. In Table 1, the coefficients of "proportion of urban people" and "unemployment rate" seem to be unreasonable judging from common senses, which is probably caused by severe multicollinearity among the exploratory variables.

Table 1 Multivariate regression on the innovation output against population qualities

	Coefficient	Standard error	t	Pr(>|t|)
Intercept	-16.0396	7.65277	-2.096	0.0468
Proportion of urban people	-0.00124	0.04618	-0.027	0.9788
Life expectancy	0.21005	0.1165	1.803	0.084
Proportion of floating population	0.16287	0.05692	2.861	0.0086
Proportion of people with higher education	0.05962	0.0695	0.858	0.3995
Registered unemployment rate	0.15628	0.38655	0.404	0.6896

Next we examine how the population qualities influence the innovation output from the perspective of "Classification and Regression Tree" (CART, Breiman *et al* 1984), which can both reveal the importance of independent variables to the response and the path for this influence.

In order to make a more thorough research on the direct and indirect output of innovation, we define the response variable "innovation output" in four situations in Table 2:

Table 2 Different definitions of the response variable in the regression tree

	number of patents	output value of new products
averaged by all the people	(1)	(2)
averaged by employed people	(3)	(4)

Below is a summary for the regression tree in model (1). There are 31 regions at the root of the tree, and the first variable used to split the tree is "Mobility". Definitely speaking, if the proportion of people who move from other regions is larger than 11.98%, the average number of patents will be 4.94 per capita (averaged by all the people in the region), which is strikingly high for all the 31 regions; then if "Mobility" is smaller than 11.98%, the rest 24 regions will be classified into two groups: if the proportion of urban people is higher than 43.41%, the average number of patents will be 1.16 per capita, otherwise 0.70 per capita.

```
n= 31
node),  split,                  n,      deviance,       yval
     * denotes terminal node
1)      root                    31      149.7414000     1.6576640
  2)    Mobility< 11.98255  24      7.1607880       0.7013894
    4)  Life< 43.405        16      0.6270062       0.4742679 *
    5)  Life>=43.405        8       4.0577410       1.1556320 *
  3)    Mobility>=11.98255  7       45.3865700      4.9363180 *
```

The regression tree in model (1) has classified all the 31 regions into three groups according to two exploratory variables "Mobility" and "Life", see Table 3.

Table 3 Three innovation groups defined by the regression tree in model (1)

Group	Members
1	Beijing, Tianjin, Liaoning, Shanghai, Zhejiang, Fujian, Guangdong
2	Inner Mongolia, Jilin, Heilongjiang, Jiangsu, Shandong, Hubei, Hainan, Chongqing
3	Hebei, Shanxi, Anhui, Henan, Hunan, Guangxi, Sichuan, Guizhou, Yunnan, Tibet, Shanxi, Gansu, Qinghai, Ningxia, Xinjiang

The above results can be illustrated in a tree structure as Figure 10; it has shown how 31 regions are divided into 3 groups with nodes and leaves in a tree. The boxplots just display the distribution of the number of patents for each group. Obviously the Group 1 has the strongest competitiveness in innovation, and the corresponding members are in the east of China. This is consistent with the conclusions in Section 2.2 and 3.1.

Similarly we may build other regression tree models, and the results are in Figure 10 and 11.

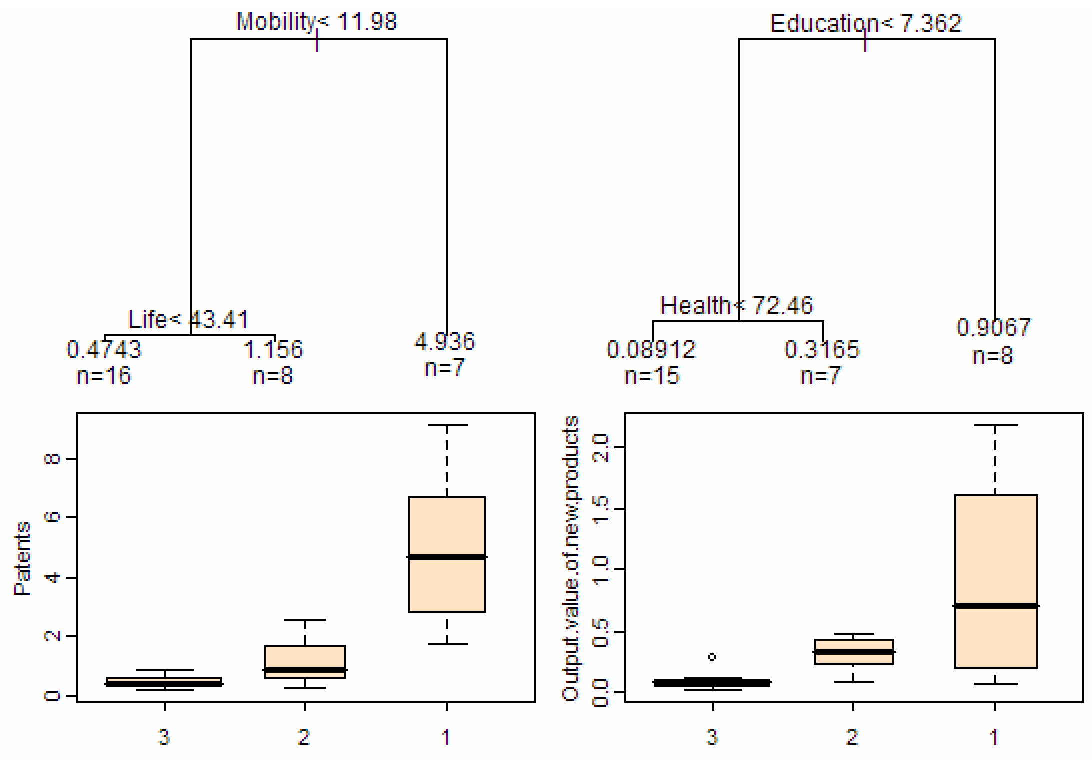

Figure 10 Regression trees from model (1) and (2)

After an examination of the group members for these tree models, we find that the membership is quite stable for all the models, and the clustering phenomenon is clear geometrically.

From the splitting variables in these four trees, we can conclude that the four factors "Mobility", "Education", "Health", and "Life standard" play different roles in determining the capability of innovation. Higher population mobility and healthier life can give birth to more innovation output for all the population, whereas the educational level can decide the eventual output value averaged by the employed people.

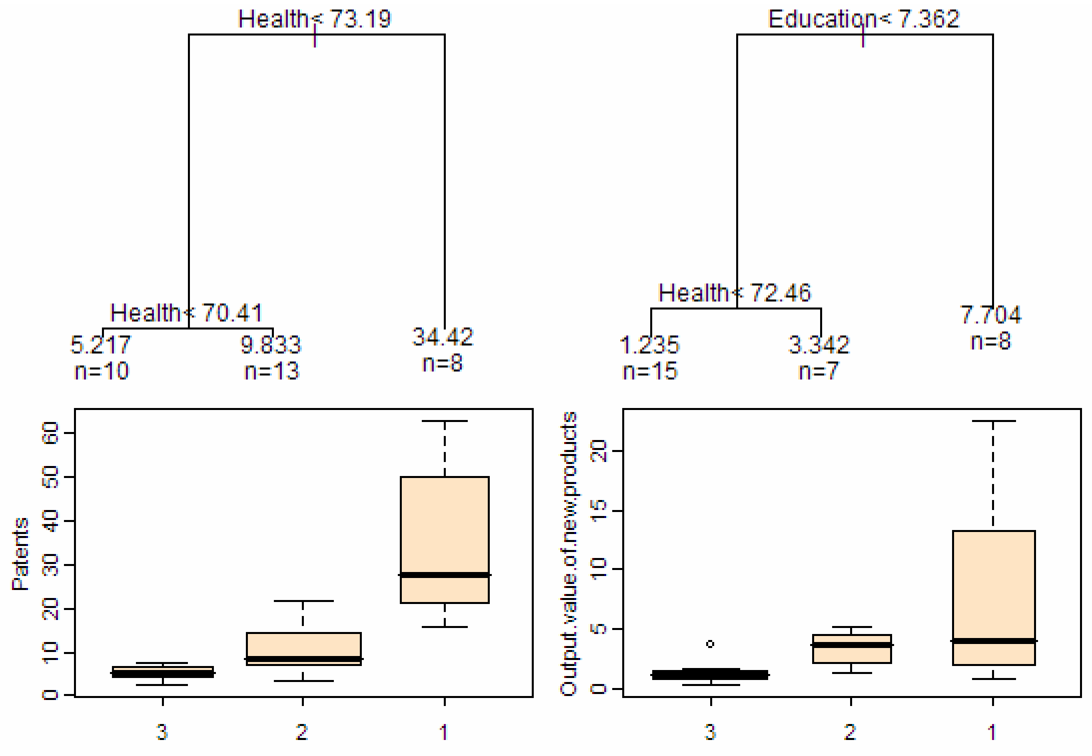

Figure 11 Regress trees from model (3) and (4)

4. Conclusions

Based on the above description and analysis about the regional population qualities and capability of innovation, it is not difficult to draw some clear conclusions as follows:

(1) Judging from the data of national population census in history, the overall national competitiveness of China's population has kept on increasing with the fast development of the economy. However, the population quality competitiveness is rather unbalanced for all the 31 regions, and the characteristic of spatial clustering of population qualities is obvious too. The clustering of population qualities has also led to the clustering of regional innovation. Further details can be found in Zhao *et al* (2008).

(2) From the perspective of patents as the indicator of innovation output, we discovered that the mobility of population is the most important factor in regional innovation. Higher population mobility can bring stronger capability of innovation. The mobility of population may both increase the number of labor forces and attract more investment at the same time. On the other hand, regions with high capability of innovation will also attract more talents and investment. Therefore the population mobility and the innovation are mutually dependent.

(3) From another perspective of the "output value of new products", the educational level plays the most critical role in regional innovation. This means the capability of converting products of innovation into market values depends on the educational quality of a region.

(4) The other two factors "health quality" and "living condition" can discriminate regions with middle and low levels of innovation, i.e. they are the secondary factors which can decide the capability of innovation for less developed regions.

In summary, we have analyzed the relationship between the population quality competitiveness and the capability of regional innovation; the importance must be attached to population qualities (especially the mobility, educational quality and health conditions) when we hope to increase the innovation capability for a region. The path of influence has also been analyzed via regression trees, which can be adopted as a practical way to promote regional innovation.

Reference

Acs Z. J., Audretsch D. B. *Patents as a Measure of Innovative Activity*. Kyklos 1989, 42 (2), 171-180.

Breiman, Friedman, Olshen, and Stone. (1984) *Classification and Regression Trees*. Wadsworth.

Griliches, Z. *Patent statistics as economic indicators: a survey*. Journal of Economic Literature, 1990, 28 (12): 1661-1707.

National Bureau of Statistics. *China Statistical Yearbook 2007*. China Statistics Press. 2007.

Yanyun Zhao, Nailin Feng, Yuncheng Zhao, Yihui Xie, Ning Xiao. *Research on the National Quality Competitiveness of China*. Statistical Research, Vol 25 (3). 2008.