

## Review on manuscript

### Public Education Expenditure, Institution Development, and Regional Innovations: An Empirical Evidence from China

The objective of this paper is to study the determinants of innovation. The authors pay particular attention to the impact of education expenditure, intellectual property protection and “financial market development” on innovation activities. The empirical analysis is based on panel data at provincial level (31 provinces, 1998-2014). The authors adopt the number of applications for international patents or patent grants as a proxy for innovation activities. This is a relevant measure, which has been widely used in previous studies. The results show that government education budget, R&D, and some other factors can significantly promote innovation.

During the last decades, China has experienced rapid economic expansion. However, China’s ability to innovate in science and technology has lagged behind the rate of economic growth. The topic of this study is undoubtedly relevant and original. Moreover, the author successfully presents research context, literature review, data, and results of econometric estimation. The presentation is relatively complete.

However, the paper has some shortcomings, making the results less convincing. It accordingly requires some major corrections. Here are my comments and suggestions.

- The authors apply the fixed effects model to panel data without explaining the reasons. In the analysis of panel data, it is necessary to perform two tests: Breusch-Pagan Lagrange multiplier test and Hausman test. The first serves to determine whether there are individual effects: pooling model or panel model; and the second allows to determine the most appropriate model: fixed effects or random effects. The authors should firstly justify the relevance of their analysis strategy.
- The introduction of explanatory variables is generally related to the objectives and hypotheses of the research; in other words, the explanatory variables are supposed to influence the explained variable. Thus, the authors should give a detailed presentation of their model specification. For example, why do they introduce “Output of First Industry as a share of GDP”? What is its anticipated effect? What is the role of “Number of enterprises above designated size” in innovation activities? In my opinion, it would be more interesting to examine the effect of firm size on innovation: small and medium-sized enterprises are often more dynamic in innovation activities; while large enterprises have a greater capacity to manage the risks resulting from innovation. In addition, it would be necessary to introduce some variables that characterize the level of regional infrastructure.
- In Section 4.1, “*The independent variable and all control variables are lagged by 1 year to account for the time lag between educational spending and changes in innovative activities.*” In my opinion, the variables lagged by 1 year are not enough to avoid

potential endogeneity. I suggest to redo the estimation using variables lagged by 3 or 5 years to check for the robustness of the results.

- The authors present the complete results only for baseline regression in Table 2. For other estimations, only the coefficients of key variables are presented. As potential multicollinearity could lead to odd results, it would be necessary to present full results in all estimations; otherwise, it is difficult to judge the relevance of the model and the results.
- Intellectual Property Protection is a hotly debated topic in China. The motivation for innovation requires effective protection of intellectual property. It is relevant and important to introduce this factor in the analysis of innovation (Section 5.2). However, this section seems unclear to me. I do not understand the following sentence: “*FIN(protect)* demotes the level of financial market development or Intellectual Property region of province *i* while the coefficient of the interaction term ... measures the effect of each of the institutional factor on the effect of education spending.” What are the definitions of “financial market development” and “Intellectual Property”? Without more detailed presentation, this section makes no sense.

Two minor comments are listed in below:

- There is a systematic error throughout the text on the interpretation of results: “*This paper endeavors to present an empirical survey of the correlations between education spending and innovation on a provincial level...*”. “*...our analysis shows a strong positive correlation between R&D spending and innovation...*” Correlation means that two phenomena or two random variables have a common evolution. It is possible that there is an interaction or causality between these two variables. In the regression, only dependent variable is random, and it is always independent variable that influences dependent variable; the last one cannot play on the first one.
- In the tables, it would be better to use the full name of the variables instead of their abbreviation. For each abbreviation, it is necessary to specify the full name when it is used for the first time, e.x. EDF, EDP in Table 5.

Due to problems mentioned above, I don't recommend this paper for publication at this stage.