Scientometrics
DOI 10.1007/s11192-014-1236-4

On the data and analysis of the research output of India and China: India has significantly fallen behind China

Rahul Panat

Received: 18 September 2013

Abstract An analysis of the number of research papers from India and China in the fields of sciences and engineering between the years 1975 and 2012 is presented. The results show that while Indian research output has increased steadily, the Chinese research output has been increasing at a rate far outpacing that of India. The research output of China has been increasing with distinct inflection points that show an acceleration in output growth. The research output for India shows periodic inflection points that show either an acceleration or deceleration in output growth. The possible reasons for the inflection points are discussed. Simple statistical analyses are used to analyze the trends in output. Although multiple factors affect a nation's research output, this paper highlights that the government programs targeted to increase the research output from universities may create inflection points resulting in a rapid increase in the research output. The article also highlights that India has fallen far behind China in terms of scientific and engineering research output, providing important clues for the future growth of the two countries.

Keywords India · China · World share · Publication productivity · Science indicators · Economic growth · Performance measures

Introduction

Scientific and technological innovation is necessary to sustain growth in any modern economy (David et al. 2008). Due to the importance of innovation, many developing countries are striving to catch up with the developed countries in terms of science and technology research. Prominent amongst them are India and China, where both aspire to become the superpowers of the next century. The science and technology research output

R. Panat (✉)
Mechanical Engineering, School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ 85287, USA
e-mail: rpanat@asu.edu

Published online: 01 February 2014

from these countries will determine the sustainability of their growth and whether they can achieve their stated goals. It is important to assess where they stand today and the trajectory they have been following for this critical metric. Such an analysis will also provide an idea to the policymakers regarding the specific actions needed to spur technological innovation.

A comparison between the economies of India and China has been a subject of intense investigation in the past; both with scholars (Bosworth and Colluns 2007; Huang and Khanna 2003; Madhan et al. 2010; Srinivasan 2006) and in popular media ("Chasing the dragon: How the Asian superpowers compare on various measures of development" 2011; Wright 2012). The Indian economy started to integrate with the global economy about 13 years after that of China and her current GDP in absolute terms is at \$1.9 trillion versus that of China at \$8.2 trillion in 2012 (IMF Statistics 2012). The Indian GDP growth has been typically below 8 %, while that of China exceeds about 10 % over the last decade. Although there is a vast difference in the current economic output between the two nations, their science and technology innovation would tell us if this continues in the future. In other words, if the economic growth is not accompanied by growth in scientific innovation, the types of industries that will grow can be copied elsewhere and the real wage growth will slow, eventually slowing the entire economy. Historically all the successful economies have had a strong backbone of university and industrial research, constantly producing new ideas that spur new industries and/or perfect the existing ones. Such high growth was seen in UK in the eighteenth and nineteenth centuries, in the US in the second half of twentieth century, and in Germany in the first half of the twentieth century.

Several scientometric studies have been conducted to compare China with India. King (2004) carried out a study of research output as measured by highly cited papers from different countries, including India and China. The output from the two countries was comparable in the period of 1993–1997. However, the Chinese output started to surpass that of India from 1997 to 2001. This work also established a correlation between wealth intensity of a nation and her scientific output as shown by the citation intensity. Other studies have compared specific research areas, viz., lasers (Garg 2002), computer science (Kumar and Garg 2005), diabetes (Arunachalam and Gunasekaran 2002), and fish sciences (Arunachalam and Balaji 2004). A recent report on research output from BRIC countries gives their total national research paper outputs and patent outputs and compares this metric to that of the world (Adams et al. 2013). The report shows a rapid rise of China as a global player in research. This work however, concentrates on analyzing the research output from the BRICs and no attempt is made to understand the differences in Indian and Chinese research outputs. Altenburg et al. (2008) found evidence that for both India and China, their research efforts rarely materialized in cutting-edge innovations but suggested that if capital accumulation proceeded at the then current pace, the innovation capabilities will rapidly build up in China and with a time lag in India. It will be interesting to know if this has indeed happened in the 5 years since this study was published.

In this paper, a simple analysis of the research output from India and China in sciences and technology since 1975 is presented. Although the number of papers is a subjective measure and differs in different sub-fields within Sciences and Engineering; a measure of the total number of papers published in peer reviewed journals can provide a fairly reasonable measure of the scientific progress. The results show that the Indian research output has fallen far behind that of China with a gap that has only increased in the recent times. Simple statistical tools are used to analyze the trends and identify the modulators for these trends. Although several complex factors can potentially affect the observed trends, it is shown that specific government programs and the funding for research in sciences and technology can be significant factors in determining the research output.

Methods

Thomson Reuters Web of Knowledge database was used to get the research articles coming out of India and China between 1975 and 2012. The total number of research papers is counted such that at least one author affiliation comes from these countries. The area is limited to sciences and engineering with review articles and articles from conference proceedings being included in the analysis. Only English language publications are considered as per the Thomson Reuters database. While plotting the research trends of India and China, several inflection points where the research output growth either accelerates or decelerates are discussed. The growth rate of the research output (slope of the number of research papers vs. time) is plotted as a function of time to get this trend. A trend in acceleration or deceleration is said to exist if the rate increases for four or more years steadily with no more than 1 year of slower growth (excluding the last year) in this time period. This definition for acceleration/deceleration is chosen to differentiate the true trend from local variations. The choice of four or more years is chosen by the author and is for interpretation purpose only. While analyzing the research trends, the research output is plotted as a function the gross domestic expenditure on research and development by China and India. Curve fitting is used to establish the correlation between funding and output with the highest possible correlation coefficient amongst linear, polynomial, exponential and moving average fits used to determine the relationship.

While analyzing the observed trends, the ‘difference in the difference’ (DID) approach is used in order to isolate the possible significant factors affecting research output. The difference in difference approach is a statistical tool to separate the effects of a particular treatment on an event from effects of multiple other parameters influencing the event at identical periods of time (Abadie 2005; Berry 2011). The effect of the treatment obtained by this method is the “Average Treatment Effect” or ATE. Note that the DID approach assumes that the control group identifies the outcome in absence of the treatment. As a result, this approach is approximate and can only be used as an estimate. In the simplest case where the effect of the control is known, the ATE can be obtained by simply subtracting its effect of the control from the effect of the treatment (i.e. as the difference in difference). The simplest example of DID is shown in Table 1. The ATE in this case is the $(T_2 - T_1) - (T_4 - T_3)$. Note that for multiple input parameters, the DID approach becomes semi-empirical and requires nonlinear analysis to isolate the effects of each of the parameters.

Results

The research paper output for India and China as a function of time between 1975 and 2012 is plotted in Fig. 1. The Fig. 1 also shows the ratio of China’s output to that of India for the same time period. Figure 2 shows the *rate of growth in research output* of China (top) and India (bottom) as a function of time. The slope for India and China is positive for all the years examined, indicating a growth in output. From Figs. 1 and 2(top), it is clear that the Chinese output was growing at a steady pace until 1995, when the growth rate showed a rather sudden jump. This is indicated as an inflection point in Fig. 1, where the output

Table 1 Example of difference-in-difference approach to obtain the effect of the treatment of interest (i.e. ATE) (Berry 2011)

	Before the event	After the event
Treatment of interest	T_1	T_2
Other events taken as control	T_3	T_4

The ATE of interest is the effect of the treatment ($T_2 - T_1$) minus the effect of the control ($T_4 - T_3$)

started to grow rapidly. Although the Chinese growth slowed between the years 2000 and 2003, it picked up rapidly thereafter and has continued this trend until today. Correspondingly, a second inflection point for China is indicated for the year 2003 in Fig. 1. In the case of India (Figs. 1 and 2 bottom), the growth rate has been varying such that it is seen to go higher and lower throughout the period examined. The output growth decelerated from 1978 to 1985, and in the 1990s, while the output accelerated between the years 2000 and 2006. The output has dropped significantly since then and the total output has plateaued out since 2009/2010. Another interesting point is indicated by the scale of the *y* axis (i.e., the growth in research output per year) in Fig. 2 for China and India. The Chinese growth rate since 2000s has been $>5\times$ that for India. From Fig. 1, the ratio of research articles per year from China to India has been increasing steadily all the time and does not appear to plateau out ($\sim 4\times$ in 2012). In other words, China's output is not only higher than that of India, but the gap between them is steadily widening. This is remarkable given that Chinese output was lower than that for India before 1995.

We next plot the regional comparison between the research output of India and China from 1990 to 2012 as the percentage increase in research output compared to 1990 (Fig. 3). The trends for the two largest city-states of China are similar to the trends for China shown in Fig. 1, i.e. acceleration in output is seen in the years 1995 and 2003 compared to previous years. For Indian research output, the states with largest GDP are plotted plus the national capital of New Delhi. The general trends for the regions of India are also somewhat similar to the national trends shown in Fig. 1, i.e., rapid growth in the first half of 2000s, with a slowing down in the output growth in the recent years. The plot is unremarkable and does not point to any particular region having led the increase in research output.

Figure 4 compares the research paper output from 1996 to 2007 as a function of the total Science the Technology gross domestic expenditure on R&D (GERD) by India and China (World Bank Database 2013). The total expenditure (*x* axis) is calculated by taking into account the purchasing power parity (PPP) of the nations in order to be comparable in US$ equivalent. The use of PPP is necessary since a similarly productive faculty will cost less in India versus China in absolute terms and hence the absolute numbers are not comparable. The measure of GERD takes into account the funding from public as well as private sectors. Note that the scale for the Chinese output and the Indian output are separated (scale on vertical axis on left vs. on right) so that they fit on the same graph. A polynomial fit results in the best correlation to the data for both the countries.

Discussion

The factors that lead to the changes in the research output including the points of acceleration and deceleration are complex and cannot be easily quantified. These

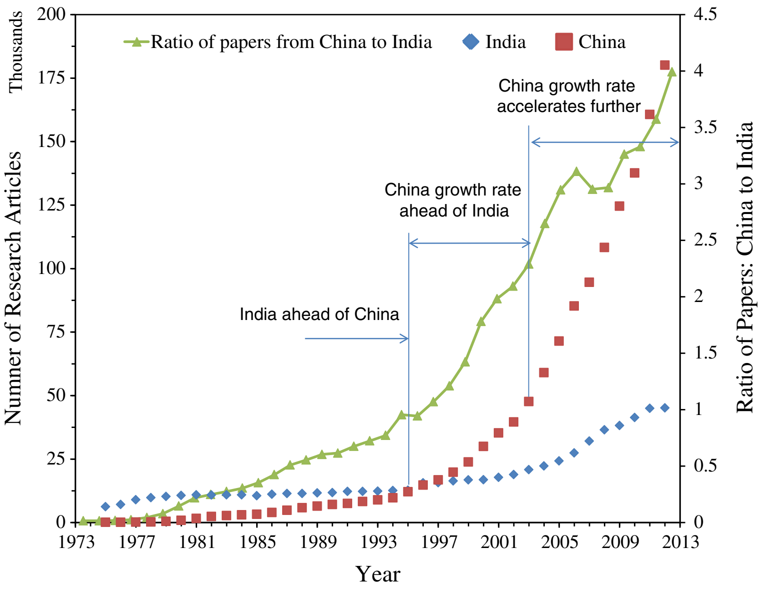

Fig. 1 Research articles from China and India since 1975. China's output crossed that for India in 1995 and has increased rapidly since then. The ratio of the research output of China to India has been growing steadily since 1977 with no sign of slowing down

factors may include (a) government policies including setting up of new academic institutions and increasing funding for the existing ones, (b) push by regional governments, (c) GDP growth driving a growth in industrial R&D causing an increase in the graduate school enrollment (due to availability of such jobs), and (d) overall increase in the positive attitude caused by economic growth that can attract academic talent including that from diaspora, (e) presence of infrastructure and traditions for excellence in research (e.g. the presence of universities of national importance even when research output was insignificant), and (f) regional factors such as political changes, research assessment, university structures etc. These factors are summarized in Table 2. Note that these factors are by no means exhaustive considering the dynamic nature of research output. We use simple statistics to assess if targeted government policies, regional research growth and/or the research funding in the area of science and technology have impacted the research output.

In case of China, the government enacted a policy of 'project 211' in 1995 with a 'Decision on Accelerating Scientific and Technological Progress'. This was a major push from the government infused funding into science and technology (S&T). This policy was a complete reversal from the isolationist policies enacted during the Cultural Revolution (1967–1977). The Chinese government followed through by giving large funding to S&T research. For a policy enacted in 1995, it is expected to have a minimum time lag of 4–5 years between the funding for research proposals and the research output (Carpenter et al. 1988; Cunningham 1997; May 1998, 1997). These lags are uncertain and variable and the interactions between institutional, national and international effects are very large. Although some of the existing research projects may get additional funding from the 'project 211' immediately after 1995, it is clear that

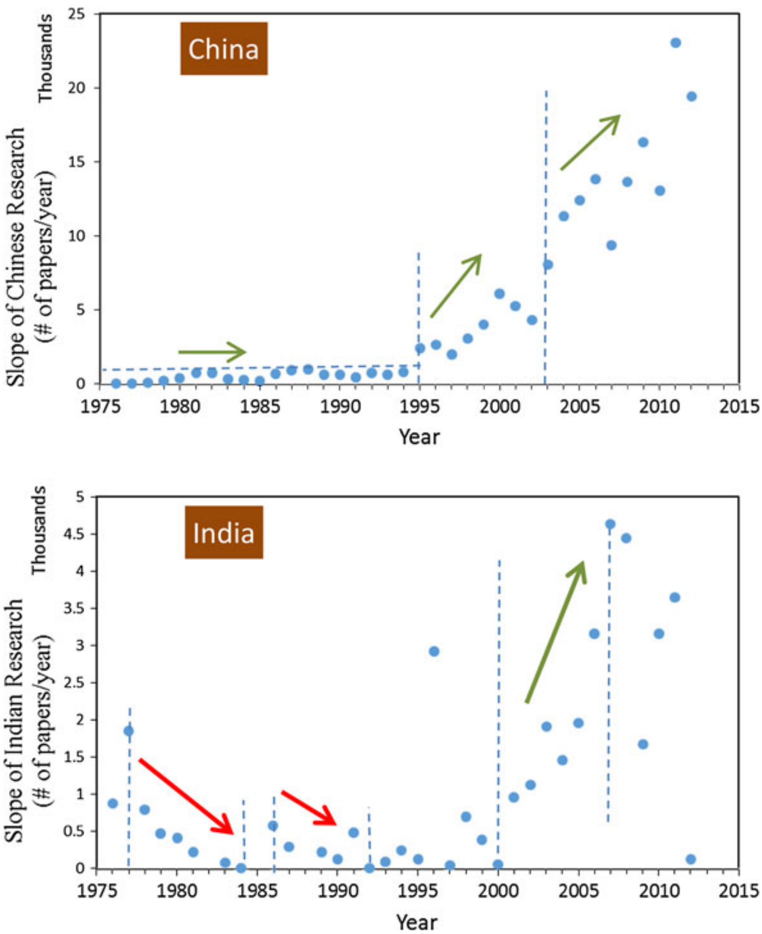

Fig. 2 The slope of the research output for China and India. The slope of the Chinese output shows distinct inflection points with significant increases in growth in 1995 and 2003. A period of slower growth between the years 2000 and 2003 is also noted. The Indian output growth rate has accelerated and decelerated consistently throughout the studied time period. An accelerated growth period from 2000 to 2007 can be observed. However, this time period is followed by significant slowing of the growth rates. Both countries show positive growth from 1975 to 2012. The *arrows* represent trends with the slope of the arrow approximating the slope of the data points. *Red arrows* indicate decrease in growth rate while *green arrows* indicate an increase in growth rate. Note that the scale of the *y* axis for China is an order of magnitude higher than that for India

factors other than ‘Project 211’ are responsible for the inflection point of 1995. In 1998, the Chinese government embarked upon ‘Project 985’ with the aim of promoting the development and reputation of the Chinese higher education system in general and certain universities in particular. The project involved the national and the local governments allocating large amounts of funding to build new research centers, improve facilities, attract world-renowned faculty and visiting scholars, etc. Considering the time of 4–5 years required for converting policy initiatives into increased research publications, the ‘Project 985’ could indeed be a significant factor in the inflection point observed in the Chinese research output in 2003. Note that this project also had follow-

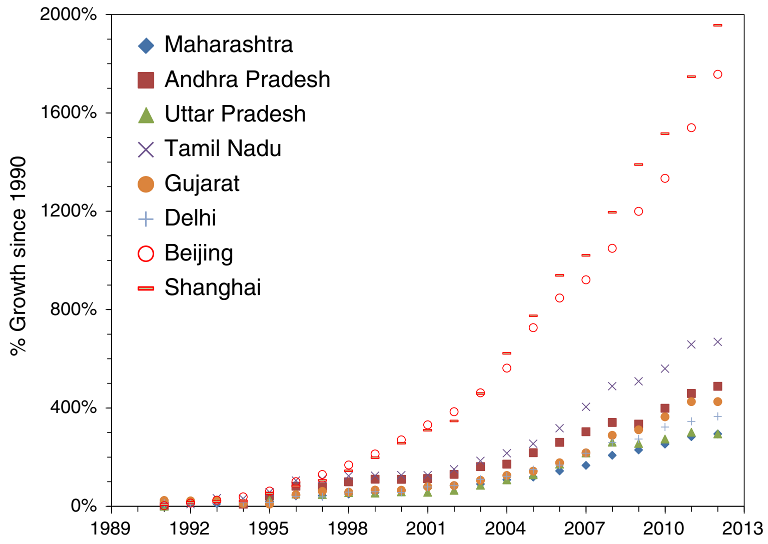

Fig. 3 Increase in research output from 1991 to 2012 for different regions of India and comparison with two major Chinese administrative regions. Chinese growthis starting to take off since 1995 with further acceleration seen since mid-2000s, mirroring the national trend. The research output of India is also seen to mirror the national trend including the recent plateauing out of the output

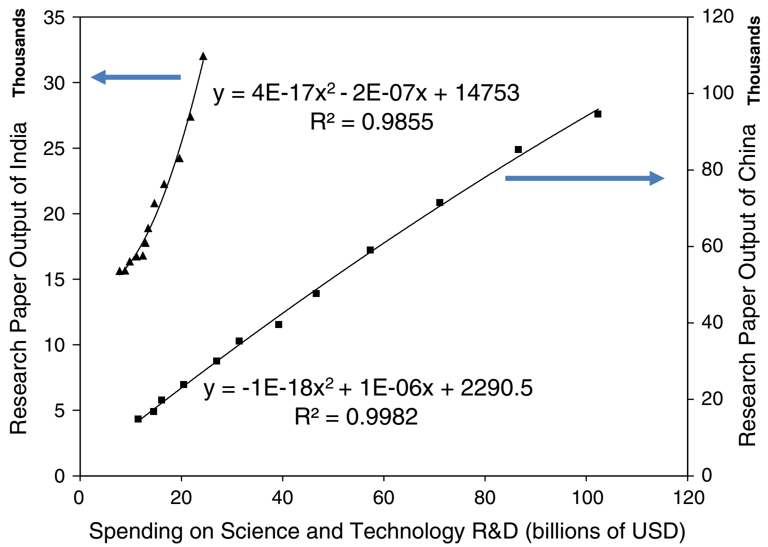

Fig. 4 Number of research papers as a function of the science and technology research funding for India and China (World Bank Database 2013). The data for the spending is calculated by taking into account the purchasing power parity of the nations for accurate comparison

on funding phases where additional universities were included. The Indian research output has specific inflection points in early 2000s when the output rate increased with time, however at a much smaller pace when compared to China.

Table 2 Possible important factors affecting national research paper output

Primary factor	Secondary factor	Result
Government policy	Setting up of new institutions Increased availability of funds for research	Increase in research output
Regional government policies	Same as above	Region specific increase in research output
Growth in economy	Growth in industry due to domestic factors and FDI (Tang and Hussler 2011) Increase in industrial R&D Increase in enrollment at Universities due to availability of research jobs Increase in funding for research in science and technology Return of trained diaspora	Increase in GDP and research output
Pre-existing infrastructure and tradition for excellence in research		Seed for good research institutes. Increase in research output
Region specific events	Political change (Braun and Schubert 1996), changes in research assessment (Currie 2008; Jiménez-Contreras et al. 2003; Moeh 2008), changes in university structures, and strategic changes in the balance between academic and research institutions (Philippe and Mustar 2001)	Changes in research output

Table 3 Number of research papers per year for China and India in 1998, 2003, and 2006

	# of papers per year 1998	# of papers per year 2003	# of papers per year 2006
Treatment—Project 985 (China)	19,822	47,652	85,282
Control (India)	16,370	20,811	27,398

The Project 985 was enacted by the Chinese government in 1998

In order to test the effect of ‘Project 985’, we use the DID approach described in the ‘Methods’ section. We use the premise that China went through a specific government program of ‘Project 985’, while Indian government did not have any such specific program. Further, the assumption is that the government initiative acted independent of the other factors. Table 3 shows the number of papers before the after treatment. The ATE for 1998–2003 (when the Project 985 is expected to have minimal impact) is 23389 papers/year. The ATE for 2003–2006 (when project 985 is expected to have high impact) is 31,043 papers/year. The Project 985 appears to be significant for the Chinese research output compared to the comparative government policies in India. We note that this approach is only approximate and ignores interdependencies. Further, this change happened during a period of significant structural and political changes in China.

In case of India, one of the speculations has been that since India started opening up of the economy about 13 years after that for China, the Indian output will show a lag with that of China by that time (Altenburg et al. 2008). It is clear from Figs. 1 and 2, that the rate of increase of research publications per annum from China seen after 1995 is higher than that for India until today. Further, other than in early 2000s, there seems to be no indication of an inflection point for India where growth accelerated similar to that seen by China in 1995 and 2003. Thus, we do not have an indication at this time that India will 'follow' China in terms of the research paper output in sciences and engineering. The growth for Indian research since 2000 could possibly be due to the general growth in economy resulting from the opening of the economy. However, further studies are needed to establish this link. The root cause for the deceleration in growth rate seen in the output growth since 2006 is less clear. It is interesting to know that India did not embark upon specific thrusts for science and technology research and instead largely relied upon general increase in research funding as the economy grew in size. In addition, the Indian government is shown to have neglected the Indian 'State University' system over the latter decades of the twentieth century (Kanhere et al. 2009). Clearly, this *laissez-faire* approach adopted by India has not helped in boosting the research output to the levels seen in China.

In the case of economic output of China and India, it was the special economic zones of China (e.g. Shenzhen) and the software and hardware hub (Bangalore) of India that led the economic growth before the rest of the country showed similar growth rates (Leong 2013; Pandit 2012). From Fig. 3, we see no indication that the local governments are 'leading the way' in accelerating the research output. This plot is rather unremarkable since the regional trends are visibly following the national trend. Since the main objective of this paper is to analyze the national research outputs of China and India, analyzing the detailed causes of regional output variations will be an objective of a different investigation. Further, Fig. 4 confirms that the funding for R&D is a significant factor to the total research output for these countries. The rate at which the research output increases with funding (constants in the polynomial fit) is different indicating the complex economic and socio-cultural factors that are not accounted for in this simplified analysis.

Note that the number of patents granted can potentially be taken as an important measure of scientific progress. Bhaumik et al. (2009) carried out such a study in 2009 comparing the patent output of India and China. This study showed an increase in Indian patent output with an even higher increase in output of Chinese patents. A close look, however, revealed that the patents grants originating from India were dominated by Chemistry area, particularly in the fields related to pharmaceuticals and biotechnology. This patent trend mirrored the strong growth in India's pharmaceutical R&D outsourcing business.

Limitations

We note that the analysis presented in this paper is only approximate and has several limitations. First, since some papers can have multiple authors with affiliations from both China and India on the same papers. Such papers are counted for both the countries. However, the author believes that this does not significantly affect the results. The analysis also ignores the interdependencies between different factors affecting research output (bias in the ATE obtained from the difference-in-difference approach). The socio-cultural conditions of each nation is different and can have an entirely different effect of the same input parameters; we ignore this factor since it is difficult to quantify. Lastly, the Chinese language papers are not included in this study. Clearly, an extensive cause-effect analysis needs to be undertaken to fully understand the observed trends. Similar to other scientometric studies, the time trends of a significant phenomenon (in this case the research output from the two most populous countries) is plotted with the investigation of the related important factors that show a correlation with these trends. To the author's opinion, the current analysis provides useful conclusions and trends that can help enact specific government programs and increase in the overall funding for science and technology research to potentially boost the research output and hence improve the competitiveness of her economy. The study also shows that from the perspective of the research paper output, India has fallen behind significantly behind that of China with no sign of catching up.

Conclusions

An analysis is presented comparing the research outputs from India and China since 1975. The Chinese research output has increased at a rate far outpacing that of India. Chinese research output is seen to accelerate at two distinct points; first in 1995 and then again in 2003. Although the actual research paper output is a result of several interrelated factors; simple statistical analyses shows that targeted government programs and overall funding for sciences and technology research are significant factors in increasing the research output. The finding in the paper also points to the fact that Indian output is not a mere time lag when compared to that for China as speculated in literature. These findings provide future indicators for the growth of these two countries.

Acknowledgments This paper is dedicated to the memory of author's father, late Prof. Padmakar V. Panat (1943–2009) who deeply cared about the state of the higher education in India. The author would like to thank the anonymous reviewers for helpful comments that greatly improved the manuscript.

References

Abadie, A. (2005). Semiparametric difference-in-differences estimators. *Review of Economic Studies, 72*, 1–19.

Adams, J., Pendlebury, D., & Stembridge, B. (2013). Building bricks: Exploring the global research and innovation impact of Brazil, Russia, India, China and South Korea. *Global Research Digests*, Thomson Reuters.

Altenburg, T., Schmitz, H., & Stamm, A. (2008). Breakthrough? China's and India's transition from production to innovation. *World Development, 36*, 244–325.

Arunachalam, S., & Balaji, J. (2004). Fish science research in China: How does it compare with fish research in India? *Scientometrics, 52*, 13–28.

Arunachalam, S., & Gunasekaran, S. (2002). Diabetes research in India and China today: From literature-based mapping to health-care policy. *Current Science, 82*, 1086–1097.

Berry, C. (2011). *Fixed effects*. Chicago: Harris School of Public Policy.

Bhaumik, P., Chakrabarti, A., & Makinen, S. (2009). Technology development in China and India: A comparative evaluation. *Journal of Indian Business Research, 1*, 213–237.

Bosworth, B., & Colluns, S. (2007). Accounting for growth: Comparing China and India. *National Bureau of Economic Growth (Working Paper Number 12943)*. Cambridge: National Bureau of Economic Growth.

Braun, T., & Schubert, A. (1996). Indicators of research output in the sciences from 5 Central European countries, 1990–1994. *Scientometrics, 36*, 145–165.

Carpenter, M. P., Gibb, F., Harris, M., Irvine, J., Martin, B. R., & Narin, F. (1988). Bibliometric profiles for British academic institutions: An experiment to develop research output indicators. *Scientometrics, 14*, 213–233.

Chasing the Dragon: How the Asian superpowers compare on various measures of development. *The Economist* (2011).

Cunningham, P. (1997). The evaluation of European programs and the future of scientometrics. *Scientometrics, 38*, 71–85.

Currie, J. (2008). The research assessment exercise in Hong Kong: Positive and negative consequences. *International Education Journal: Comparative Perspectives, 9*, 47–58.

David, P., Aghion, P., & Foray, D. (2008). Science, technology and innovation for economic growth: Linking policy research and practice in "STIG Systems". *Paper No. MPRA Paper 12096*. Munich: University Library of Munich.

Garg, K. (2002). Scientometrics of laser research in India and China. *Scientometrics, 55*, 71–85.

Huang, Y., & Khanna, T. (2003). Can India overtake China? *Foreign Policy, 137*, 74–81.

IMF. (2012). *IMF statistics*. Washington, DC: IMF.

Jiménez-Contreras, E., de Moya Anegón, F., & Delgado López-Cózar, E. (2003). The evolution of research activity in Spain: The impact of the National Commission for the Evaluation of Research Activity (CNEAI). *Research Policy, 32*, 123–142.

Kanhere, D., Arjunwadkar, M., & Vichare, A. (2009). Rise and decline of India's state university system: Neglect, design, or neglect by design? *Current Science, 97*, 1013–1021.

King, D. (2004). The scientific impact of nations. *Nature, 430*, 311–316.

Kumar, S., & Garg, K. (2005). Scientometrics of computer science research in India and China. *Scientometrics, 64*, 121–132.

Leong, C. (2013). Special economic zones and growth in China and India: An empirical investigation. *International Economics and Economic Policy, 10*, 549–567.

Madhan, M., Chandrasekar, G., & Arunachalam, S. (2010). Highly cited papers from India and China. *Current Science, 99*, 1–12.

May, R. M. (1997). The scientific wealth of nations. *Science, 275*, 793–796.

May, R. M. (1998). The scientific investments of nations. *Science, 281*, 48–51.

Moeh, H. (2008). UK research assessment exercises: Informed judgments on research quality or quantity? *Scientometrics, 74*, 153–162.

Pandit, K. (2012). The Indian landscape after two decades of liberalization: An introduction. *Eurasian Geography and Economics, 53*, 1–6.

Philippe, L., & Mustar, P. (2001). French research and innovation policy: Two decades of transformation. In E. Elgar (Ed.), *Research and innovation policies in the new global economy: An international comparative analysis* (pp. 447–496). Oxford: Oxford University Press.

Srinivasan, T. (2006). China, India and the world economy. *Economic and Political Weekly, 41*, 3716–3727.

Tang, M., & Hussler, C. (2011). Betting on indigenous innovation or relying on FDI: The Chinese strategy for catching-up. *Technology in Society, 33*, 23–35.

World Bank Database (2013). Washington, DC.

Wright, T. (2012). India far from overtaking China, report says. *Wall Street Journal*.