Innovation Survey Practice in China

Yongxu DENG, Zhikang JIAO, Yin LI, Feng ZHEN

Paper prepared for the 16th Conference of IAOS
OECD Headquarters, Paris, France, 19-21 September 2018

Session 2.D.1, Day 1, 19/09, 14h30: From R&D statistics to innovation – China

Yongxu DENG
dengyongxu@stats.gov.cn
National Beaurau of Statistics of China

Zhikang JIAO
jiaozk@stats.gov.cn
National Beaurau of Statistics of China

Yin LI
liyin@stats.gov.cn
National Beaurau of Statistics of China

Feng ZHEN *
zhen@ruc.edu.cn
Renmin University of China

Innovation Survey Practice in China

DRAFT VERSION 30/08/2018
PLEASE DO NOT CITE

Prepared for the 16th Conference of the
International Association of Official Statisticians (IAOS)
OECD Headquarters, Paris, France, 19-21 September 2018

ABSTRACT

Operated by Ministry of Science and Technology and National Bureau of Statistics, China formally funded its National Innovation Survey System since April 2017. Firm based innovation survey is one of the two main parts of the system. This paper will introduce the practice of firm innovation survey in China since the first version of *Oslo Manual* published, compare the survey organization programme with Community Innovation Survey in Europe, and summarize the problems and interesting directions like measuring R&D efficiency, joining R&D and innovation survey data with other official data recourses, and opening of micro data. It will give an overview of Chinese practice on innovation survey based on international statistical standards, and the common problems facing by both China and developed economies on innovation survey and its data application.

Keywords: R&D Statistics; Innovation Survey; China

1. INTRODUCTION

Innovation survey is an important step to push hard data based R&D statistics to a new stage with understanding of the mechanism of innovation, i.e. innovation process from R&D input to innovation output and its efficiency, other types of innovation besides technique innovation, R&D activities and its organization, cooperation and effects. *Oslo manual*, published by Organization for Economic Cooperation and Development (OECD) and European Commission, gives international standards and principles of innovation survey, and Community Innovation Survey (CIS), operated by Eurostat and member national statistics institutions (NSIs), gives a good practice of collecting and disseminating innovation survey data. Based on *Oslo manual* and the experience of CIS practice, China started to investigate the possibility of innovation survey from 1990s, organized innovation survey in 2007 and 2014 for large sample size of practice, and formally built National Innovation Survey System in April 2017.

The National Innovation Survey System includes two main parts. The first part is innovation survey, operated by National Bureau of Statistics of China (NBS China), which will be a yearly firm based micro innovation survey. The second part is innovation capacity monitoring and evaluation report organized by Ministry of Science and Technology of China (MoST China). The first part will give a very large size of sample of firm data with similar structure like CIS, and it will also partly support the evaluation in the second part. The MoST operated part will be a macro national or regional based indictor system to find and compare changes and characters of innovation like the European Innovation Scoreboard.

This paper will focus on the build up of innovation survey system in China. Since it spends around twenty years from the first version of *Oslo Manual* to the build up of innovation survey system of China, there must be some practices, discussion, changes according with China's reform and opening. The paper will organize like follows. The first part will introduce the practice of innovation survey in China since 1990s, to give background information of the exploration of the new area. Take 2014 innovation survey programme as an example, the second part will introduce how to organize the innovation survey and compare the difference with CIS in European countries. The third part will focus on problems or directions of improvement like measuring R&D efficiency, joining R&D and innovation survey data with other official data recourses, and opening of micro data. The last part will be a summary comment of China's practice on firm innovation survey.

2. FROM R&D STATISTICS TO INNOVATION STATISTICS

The R&D statistics in China started from mid-1980s. It developed into an international standard based national wide statistical programme no later than 2000. After that, several key progresses were made up with the status change of measuring R&D based on economic development and social progress. Gao et al. (2018), the first paper of the session will give a summary of different stages of development of R&D statistics in China.

After the first version of *Oslo Manual* was published in 1992, NBS China published the Chinese version in the following year. It gave a new viewpoint of understanding R&D activity. Innovation survey, as a brand new concept, came to the view of research institutions and government. Two stages of testing survey were made since then on. The first stage of testing surveys organized in 1994 and 1996 independently with relatively small size of firm sample. The second stage of testing survey with large sample size organized in

2007 and 2014. Then as important part of planed national innovation survey system, NBS China started yearly firm level innovation survey from 2016.

The two rounds of innovation survey in 1990s are based on the interests of researching and testing. Development Research Center of the State Council and NBS China organized the 1994 survey, but collected only 2,000 samples from two provinces, i.e. Fujian and Gansu. In 1996, National Science and Technology Commission (later MoST) and NBS China collected around 5000 answers from large and medium size firms in six provinces and municipalities include Beijing, Shanghai, Jiangsu, Shandong, Liaoning and Harbin. These two rounds were only testing of the survey, but gave basic experience of firm level innovation survey in China. In other words, they were the best samples reflecting the spirit of reform in China, i.e. across the river by feeling the stones.

The following main step is the 2007 survey, more than ten years later. Organized by NBS China, the 2007 innovation survey made great progress. It is the first time following strict sampling method and international statistical principles and standards in this area. Besides the main reference of CIS system, the survey collected 67,000 answers from industrial firms national wide, with answers representing the time periods from 2004 to 2006 for qualitative data. Disadvantages are also obvious. Firstly, the survey included only industrial firms, but not included firms from other sectors like construction and service, which may do had active innovation activities. Secondly, it asked activities of product innovation and process innovation, but did not include organizational innovation and marketing innovation. Thirdly, it was not included in the yearly official statistical programmes, but organized separately, which increased the cost of survey quite a lot. Considering about the size of industrial firms above designated scale was nearly 300,000 firms at that time, the cost was an important indictor influencing the necessity of innovation survey.

The 2014 round of innovation survey improved a lot comparing with the 2007 round. More or less, the disadvantages shown above were solved. This round of survey is an important transfer from testing to the formal standing official survey programme. Firstly, this round of survey included not only industrial sectors, but also construction sectors and part of service sectors, of which with active innovation activities. Secondly, the questionnaire for the first time included all the four kinds of innovation categories, i.e. product innovation, process innovation, organizational innovation and marketing innovation. Thirdly, it was synchronously organized with the general firm survey programmes, which made it a formal programme of survey in NBS China. The sample size was 440,000 and 378,000 were industrial firms. The number of sampling based firms above designed scale[1] in service sectors, not including financial sectors, was around 52,000[2].

The 2016 and 2017 rounds of survey followed the basic principles of 2014 round, but with some significant changes. Firstly, the design of the survey started to based on per one year instead of two years, which means the questionnaire will ask for the innovation situation of firms not in two years but in the exact survey year. Secondly, the sector structure and sample size changed. The industrial sectors kept

[1] The designed scale is based on revenue from principal operations and partly on number of employees, but different standard in different sectors, and the standard changed upon years. Take industrial sectors and construction sectors as an example, firms above the designed scale means their revenues from principal operations are more than 20 million Renminbi, around 2.52 million Euro in September 2018.

[2] No less then 20% of total number of firms above designed scale in selected sectors, which was around 260,000 in 2014.

stable in both structure and sample size. The construction sectors and most of the service sectors enlarged sample size, but the whole financial sectors was removed due to sampling frame, response rate and some other problems in the experience of 2014. The total sample size of firms above the designed scale went to 726,000, but the number of industrial firms kept stable in 379,000[3]. For the sampling of firms under the designed scale, the sample size was 54,000.

After 20 years practice and thinking, innovation survey came into a standardized formal official statistics programme. The reasons why using such a long time include at least the follows. The first is as a relatively new concept and phenomenon, innovation and its soft survey data need to be understood with the progress of market economy and change of social background. Step by step, researchers and government realized that the organization and mechanism of innovation activities, the interaction among government, firms and research institutions, and many other soft environment of innovation are important to R&D and innovation activities. The second is the good foundation of traditional R&D statistics made it seems measuring R&D activities are sufficient to understand innovation, and the high budget of a new national wide survey made it not easy to make the final decision. Finally, in the deep changing world economy and fast developing digital and internet age in the new century, the central government has more and more strong view of promoting sustainable development by the strategy of constructing innovation-oriented country.

3. DESIGN OF INNOVATION SURVEY IN CHINA

Innovation survey in China based on the common principle from *Oslo Manual*, and the questionnaire is comparable with CIS in Europe. In order to express the diversity of Chinese firms, there are some changes. Take 2014 survey as an example[4], this part will introduce the design of the survey, especially the difference comparing with CIS.

3.1 Two sampling systems among sectors

The 2014 round innovation survey added part of service sectors in the sample, but the sampling methods were different comparing with the industrial sectors.

For industrial sectors[5], especially manufacturing sectors, based on its importance in national economy system and its good traditional data collection system, the survey included all the firms above designated scale, which means it was a census for large and medium sized firms. The census method was also used in the construction sectors[6] and financial sectors, but only for large firms.

[3] Firms above the designed scale slightly reduced in 2016 from 383,000 in 2015, which made the number of firms changed not too much in 2016 comparing with 2014.

[4] The 2014 survey is a turning point to a standard survey programme. The 2016 and 2017 survey design followed the principle of 2014 survey in general, changed slightly in questionnaires, but enlarged the sample quite a lot in as introduced in section 2.

[5] These sectors may approximately compare with sections B, C, D and E in ISIC v4 published by United Nations (2008).

[6] These sectors may approximately compare with section F in ISIC v4.

Financial sectors[7] are special ones, for the general official data in this area is collected by not NBS China but the central bank and regulatory institutions. There are two reasons to do the innovation survey by census of firms in and above province level. One is NBS China do not have the basic units and sampling frame of these sectors. The other is the total number of firms in and above province level were around 8,000 at that time, which was not a big number if separate them into each province. These sectors was reduced since 2016 round of survey.

For the other selected service sectors[8], a strategic sampling method was used to find the cases in across provinces and sectors. The strategic variable was number of employee. The error of sampling was controlled in 5% level, and the theoretical sample size was around 52,000 firms. A sampling programme was designed and iteration method was used to find exact sample size in across each provinces and sectors. Based on this sample, the estimation could only be done in province level, but not in regional level under province .

After 2016 round of survey, the two sampling systems of "census" and sampling kept, however they were separated by not sectors anymore but size of firms. No matter industrial sectors or not, firms above designed scales were all involved in the survey, but firms under the designed scales were based on a sampling survey, no matter what sectors they were in.

3.2 Two sets of questionnaires in three kinds of sector groups

The 2014 survey used two sets of questionnaires. One was prepared for the selected firm, and the other for the general manager or owner of the firm. The firm questionnaire was similar like CIS. The one for the general manager had two main parts besides the background questions. The first part was to ask the effects of four kinds of innovation methods, and the second part was to ask the actions and relations of firms in certain policies and innovation activities.

The questionnaires were separated into three groups based on sectors, i.e. the industrial sectors, the construction sectors, and the service sectors. The questionnaires in each sector groups are similar, but changed a little bit of either the order of questions, or the words in describing or explaining the questions. For example, in the construction sectors, the questionnaire for firms asked the questions about process innovation first, and then product innovation, for the main innovation activities in these sectors are not production innovation but process innovation. The change of the order may encourage firms to continue the answers.

Then in total, there were six questionnaires for three sector-groups and two sets in each group.

3.3 Quality control and dissemination

Quality control, including the evaluation of the survey afterward, is a key step of official statistics. NBS China has a certain set of documents and mechanism to control the process of data collection. The innovation survey is under control during the process of designing, sampling, training, data collecting,

[7] These sectors may approximately compare with section K in ISIC v4.
[8] These sectors may approximately compare with sections G, H, J and M in ISIC v4.

aggregating, and disseminating. However, the sampling and non-sampling error will always happen in any sampling survey. NBS China will give an evaluation like what Eurostat (2012) done in its Synthesis of Quality Reports, but the document is not open to public till now.

The main results of 2007 round survey were published online for public reference after the survey. The 2014 and 2016 rounds with detailed aggregated data in sectors, ownership, size, province and so on was published to public, and several research series by NBS research institution were published in NBS open journal.

4. PROBLEMS AND INTERESTS IN PRACTICE

The budget of innovation survey in China is very high due to the large sample size, which makes how to use the results effectively is important. However it will never avoid the dilemma of micro data mining and data protection at the same time. In the following contents, we will firstly discuss the interaction between theory development and practice, list several valuable directions for application, and then discuss about micro data protection and opening, and practice of combining innovation data with other data resources.

4.1 Interaction between theory and practice

Innovation is a good topic to understand the interaction between theory and practice. On the one hand, the development of theory and quantitative research will give new directions of how to understand and develop the questionnaire. On the other hand, a good design of questionnaire and micro data opening will push new chances of understanding the practice and promoting or even adjusting the theory.

The research on productivity and effect of innovation (R&D) is a good example. The productivity ratio between input and output has been a classical area of study since the Cobb-Douglas production function was proposed in 1928. Since then on, thousands of discussions, ameliorations and empirical studies have been contributed to this academic area, together with the remarkable improvements by Tinbergen (1942), Solow (1957), and Jorgenson (1988). Griliches (1979) develops the knowledge production function and gives innovation criteria a new position in the equation.

The CIS micro data gave a new chance to the research. By using CIS firm data, Crépon et al. (1998) propose a new system of measuring innovation to productivity not in one step but in three steps to find the mechanism of innovation clearly. The first step is to understand why firms select to do R&D and what influence their expenditure. The second step explains innovation output of both product and process from R&D input by using knowledge production function. The third step is an extended production function to analyze the effect of innovation outputs to productivity.

Further more, since innovation input like R&D will partly get no return, Mohnen et al. (2006) work on CIS firm data to compare 7 European countries and develop the measure of innovativity, which emphasized the efficiency of innovation input. In the new century, sustainable development came in to research area, which makes “green” innovation a new topic of theory framework. Leeuwen & Mohnen (2014) gave some exploration in this area.

The theoretical analysis gives policy recommendations in new viewpoints, which improves CIS and innovation statistics at the same time to find more valuable indictors to measure. This example suggests data opening and micro research is important, but must be based on a good protection of individual privacy of each sampling unit.

4.2 How to evaluate the efficiency of innovation

Innovation is not always the first choice in development. The good example is the postponement of R&D ratio to 3% in European from 2010 to 2020 due to the influence of 2008 world financial crisis. There must be some other important areas need to increase public budget and firm investment urgently in the case, which suggests any spending of public or private revenue should be evaluated the efficiency of the input.

Innovativity, or the output of innovation comparing with its input, gave a good thinking of how to evaluate the efficiency of innovation. Before that, policy and research got similar results that R&D was important to promote development of economy, which was always a key evaluation of success of governance in no matter developed or developing countries. However, innovation input like R&D expenditure is part of revenue in both public and private sectors, and a certain part of R&D input may get nothing. How to evaluate the efficiency of R&D input or other innovation activities should be considered at the same time of emphasizing innovation input. It should be considered as a key direction to develop research and practice in innovation area.

4.3 How to evaluate innovation output

Innovation input is relatively easy to measure for R&D is the main part and it can be calculated based on accounting forms. Then how to evaluate innovation output besides new product, for CIS can only give yes or no in process innovation and other three areas. Data of research papers, patents, trademarks and copyrights may be a good way to consider about. However, the following new question is how to joint the data in different areas to understand the direct output and indirect output of innovation in the same framework but with different units and effects.

The official statistics do have most of the data mentioned above. The question is how to join them together. U.S. Patent and Trademark Office (2016) and European Patent Office & the European Union Intellectual Property Office (2016) gave some efforts. A more deep and clear framework is still waiting for establishing. The third paper of the session will give a introduction of the effort in China.

4.4 R&D capitalization

How to practice R&D capitalization in the new SNA system is another important topic based on innovation statistics. Gao et al. (2018), the first paper of the session will give a discussion in its section 4.

4.5 Opening of micro data

In the big data environment, if data or information is a mineral, then the census or random sampling based official data will be a rich mineral. However, it will always be a dilemma to open micro data and protect the individual privacy at the same time.

European Union had some leading work on micro data opening. On the one hand, it gave strict protection of data by both legal framework and administrative practice. Manual on Disclosure Control Methods (Eurostat, 1999) is a good sample of standardizing data opening. On the other hand, some research like Nordholt (2003) discussed Micro Aggregated Data and other methods to protect micro data. It found out that the empirical results before or after the moving average of micro data were stable, which means it might be a beneficial way to solve the problem in technical mechanism.

Based on similar rules, NBS China started to consider limited micro data opening in 2017. Two experimental ports were promised in Tsinghua University and Renmin University of China[9]. These ports will joint selected NBS micro data and open to public. However, there is still a lot of work to do to make it into reality in both hardware and software construction.

4.6 Joining data in different areas

Joining official statistical data with administrative data is a good way of promoting research in both policy and theory topics.

Take innovation as an example, the R&D statistics and innovation survey generally operated by the national statistics institutions, but data like patent, trademark and copyright may be operated in other institution(s). The R&D and innovation data is collected by firms and institutions, but the data of intellectual property rights is collected by registration. In theory, these different resources of data can be easily joined by using the bridge of firm id in firm level[10]. However there are still a lot of non-sampling errors in the data sets. NBS China and State Intellectual Property Office of China (SIPO China) are trying to do the work in practice. The third applied paper in this session will introduce an example.

Doing the similar thing in developed countries may also have difficulties. By doing the green innovation research in Leeuwen & Mohnen (2014), it tried to join the innovation data with energy data in the Netherlands. The authors structure from Statistics Netherlands and leading research institution made the work relatively easy, but there were still barriers of doing so.

4.7 National Innovation Survey System

In April 2017, the State Council approved the National Innovation Survey System. MoST and NBS China are the two leading and coordinating ministries. Together with other 15 ministries or ministry level institutions, the system will organize national innovation survey, monitoring and evaluation. NBS is responsible for the yearly innovation survey and other micro data collection to measure innovation activities. Besides data collection, NBS is also responsible for the mechanism of administrative data

[9] The Tsinghua port was opened in June 2018.
[10] It is only part of the data in firm level for intellectual property rights may register by private person or other institutions.

sharing system among central government institutions. MoST is responsible for the monitoring and evaluation of innovation capacities, and issuing regular reports.

In the area of official statistics, the system has at lease two important meanings. The first is that NBS became the only resource of official innovation data publisher, which will solve the problem that the same criteria may have different values from different official data collector. The second is organizing and combining official administrative data from different ministries will be more smoothing and less resistance.

5. SUMMARY

Based on R&D statistics and innovation survey, China has built the modern innovation statistics system. It follows the common international standards and principles, and the data has a good base to compare with other leading countries. However, there are still some challenges in data collection, data mining, and application. They may be solved or improved in practice together with the efforts of NBS China, researchers and users of the data from China and abroad.

REFERENCE

[1] Crépon, B., E. Duguet, and J. Mairesse, 1998, Research and Development, Innovation and Productivity: An Econometric Analysis at the Firm Level, Economics of Innovation and New Technology, 7(2), 115-158.

[2] Economics & Statistics Administration, U.S. Patent and Trademark Office, 2016, Intellectual Property and the U.S. Economy: 2016 Update.

[3] European Patent Office & the European Union Intellectual Property Office, 2016, Intellectual Property Rights Intensive Industries and Economic Performance in the European Union.

[4] EUROSTAT. 1999. Manual on Disclosure Control Methods, Luxemburg.

[5] EUROSTAT. 2012. Community Innovation Survey 2008 Synthesis of Quality Reports. http://ec.europa.eu/eurostat/cache/metadata/Annexes/inn_cis2_esms_an3.pdf.

[6] Griliches, Z., 1979, Issues in Assessing the Contribution of Research and Development to Productivity Growth, Bell Journal of Economics, 10(1), pp.92-116.

[7] Jorgenson, Dale W., 1988, Productivity and Postwar U.S. Economic Growth, The Journal of Economic Perspectives, Vol. 2, No. 4, pp. 23-41.

[8] Leeuwen George van & Pierre Mohnen. 2014. Green Innovation: Assessing the Porter Hypothesis for the Netherlands. Miro.

[9] Minxue Gao, Xiaojing Guan, Peng Zhang and Jingping Li, 2018, The Status Quo and Issues of R&D Statistics in China, draft for the IAOS 2018 conference.

[10] Mohnen, Pierre, Jacques Mariesse and Marcel Dagenais, 2006, Innovativity: A Comparison across Seven European Countries, Economics of Innovation and New Technologies, 15(4/5), 391-413.

[11] Nordholt, Eric Schulte, 2003, Applications of statistical disclosure control methods, Statistics Canada International Symposium Series.

[12] Schumpeter, Joseph, 1934, The Theory of Economic Development[M], Cambridge, MA: Harvard University Press.

[13] Solow, Robert M., 1957, Technical Change and the Aggregate Production Function, The Review of Economics and Statistics, Vol. 39, No. 3, pp. 312-320.

[14] Tinbergen, J., 1942, Professor Douglas' Production Function, Review of the International Statistical Institute, Vol. 10, No. 1/2, pp. 37-48.

[15] United Nations. 2008. International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4. New York: United Nations Publication.