

CDDRL

WORKING PAPERS

Number 127
April 2013

Republic of Government Vis-À-Vis Republic of Science: Analyzing China's Scientific Knowledge Production

Devin Fensterheim
Massachusetts Institute of Technology

Yasheng Huang
Massachusetts Institute of Technology

Fiona Murray
Massachusetts Institute of Technology

Center on Democracy, Development, and The Rule of Law
Freeman Spogli Institute for International Studies



Additional working papers appear on CDDRL's website: <http://cddrl.stanford.edu>.

Center on Democracy, Development,
and The Rule of Law
Freeman Spogli Institute for International Studies
Stanford University
Encina Hall
Stanford, CA 94305
Phone: 650-724-7197
Fax: 650-724-2996
<http://cddrl.stanford.edu/>

About the Center on Democracy, Development and the Rule of Law (CDDRL)

CDDRL was founded by a generous grant from the Bill and Flora Hewlett Foundation in October in 2002 as part of the Stanford Institute for International Studies at Stanford University. The Center supports analytic studies, policy relevant research, training and outreach activities to assist developing countries in the design and implementation of policies to foster growth, democracy, and the rule of law.

China is increasingly considered to be an essential participant in the global knowledge economy (King 2004). Recent studies have highlighted the rising number of scientists and engineers educated in Chinese institutions of higher education, and the growing funding allocated to the production of knowledge. By 2005, China ranked fifth amongst nations in total scientific publications produced; in engineering it ranked second, lagging only the United States in the production of new engineering publications (NSF 2007). China has meanwhile embraced major scientific projects: in genomics it was the only developing nation to participate in the Human Genome Project; in space science it was the third nation to develop a successful manned space program. According to a science policy expert at the Chinese Academy of Sciences, by investing heavily in major scientific projects China intends to “transform from the largest developing country to a world powerhouse.” (Xin and Yidong 2006).

While the rapid increase Chinese scientific knowledge production is generally undisputed, much greater uncertainty surrounds assessments of whether and to what extent China is producing scientific knowledge at the global frontier. Are China’s investments in scientific work paying off with knowledge that is globally competitive of high quality as well as quantity and if so, in which fields? This paper provides the first comprehensive enquiry, of which we are aware, into scientific knowledge production (as measured through the lens of scientific publications) in China across a wide range of disciplines over the period 1980 – 2008. In particular having established the widely reported increase in the absolute number of publications published by scientists affiliated with China’s institutions, we explore three key puzzles in the literature: (1) Whether the composition of China’s publications have changed or remained stable across disciplines and across the basic-applied continuum; (2) Whether the quality of China’s publications remains uniformly low or whether there are specific areas of leading edge

knowledge that reach the intellectual frontier; (3) Whether the distribution of China's research activities and publications is centered on a small number of research institutions or is widely distributed across the country and across a variety of collaborations. Before turning to our quantitative analysis we provide some detailed background into China's scientific institutional arrangements describing the historical context, the modern commitment to significant spending on science, and then the scientific institutions that support (or hinder) high quality knowledge production (as defined by those institutions supporting freedom and openness).

HISTORICAL PERSPECTIVES ON SCIENCE IN CHINA

Recent models of endogenous economic growth have recently recognized the importance of knowledge production, and more importantly, the powerful role played by the accumulation of scientific knowledge [Romer, 1990; Grossman and Helpman, 1991; Jones, 1995] and step-by-step technical progress [Scotchmer, 1991; Gallini and Scotchmer, 2003]. However, scientific knowledge production can only serve as a foundation for growth when the research is cumulative – researchers “stand on the shoulders” of prior generations. This characteristic feature of scientific work, as Mokyr (2002) has noted, is not guaranteed by the mere production of knowledge. Instead, effective sharing and accumulation of scientific knowledge across researchers and over time must be provided by a complex and intertwined system of scientific institutions (as has been well documented in the case of the scientific institutions underpinning the expansion of science in seventeenth century England) (Merton 1953; Merton 1973; Mokyr 2002). The reliance on the institutional context for scientific knowledge production is as salient in China as it is in the West. Thus we start with a brief description of the historic development of

science in China. By describing the (limited) scientific institutions that form part of China's history, scientific work in contemporary China can be better understood.

China has a distinguished history of technical and scientific achievements. Needham noted the four great inventions of ancient China: the compass, gun-powder, paper and the printing press. However, in a puzzle that became known as Needham's paradox he also remarked on China's lack of scientific mastery and the rapid advance of Renaissance Europe.

[Hall,1968;Chang andLee,1998]. In a 1922 paper entitled Why China Has No Science, Yu-Lan Fung, a Chinese scholar studying at Columbia University postulated a cultural dissonance between traditional scientific theory, as observed in Europe, and traditional Chinese philosophy [Fung, 1922]. In addressing the China Paradox, Needham argued that an entrenched feudal system prevented the rise of a merchant class to expedite technological development.

Nonetheless in the early 20th century china would undergo a significant scientific transformation and the Academia Sinica in Beijing would form the scientific epicenter of scientific research. Between 1912 and 1936 the Republic of China would establish 53 research institutions, 42 periodical publications and Chinese scientists received international recognition particularly in the geosciences (Wang 1943).

With the formation of the People's Republic of China, the Academia Sinica was merged into the Chinese Academy of Sciences. It experience rapid growth, but was staffed by party officials who sought to mold research to meet the needs of industrial and agricultural production. At that time there were only about 1000 scientists in China with doctorate-level degrees (Suttmeier 1975) and the exclusive communities, so prevalent in the West among scientists (often referred to as "invisible colleges"), were perceived as antithetical to socialist theory.

Instead, a class-based theory of science was promoted to resist the bourgeois institutions of science.

With the onset of the Cultural Revolution in 1966 the Chinese scientific infrastructure was virtually eviscerated: Professional societies were abolished, the State Technology Council was eliminated, many research institutes were closed and scientists sent to work in factories and fields (Suttmeier 1980). A typical education stopped at primary school, cumulating in an entire generation lacking the education and training necessary for technical development. During this time, the obstacle to scientific research was also the subordination of science to political ideology. Science was deliberately sidelined to promote the revolutionary ideology that prized loyalty to the communism. It was not until 1975 that progressive scientific reform would find a place in Chinese policy.

CHINA'S COMMITMENT TO SCIENCE

Contemporary science in China emerged from the strictures of the Cultural Revolution in 1975 when Zhou Enlai, first Premier of the People's Republic of China, proclaimed the Four Modernizations. He outlined a vision for modern Chinese society that could be realized by the end of the twentieth century through technological progress. One year later, China entered a period of rapid economic, technological, and scientific growth sustained over three decades. According to a science policy expert at the Chinese Academy of Sciences, by investing heavily in major scientific projects China intends to "transform from the largest developing country to a world powerhouse." (CITE). Several policy goals illustrate this vision including keji xingguo (science and technology making the country prosperous).

The first systematic political statement outlining a science and technology strategy was the “Decision on the Reform of the Science and Technology Management System” in 1985. The first systematic political statement outlining a science and technology strategy was the National Science Conference in 1978, where party officials reversed Mao-era restrictions on scientific practice and declared a central role for science and technology in its vision to become a socialist economic power by the 21st century. The conference was followed in 1985 by the “Decision on the Reform of the Science and Technology Management System,” which linked national research and development to socioeconomic advancement [Frame & Narin, 1987] and provided increased autonomy to research bodies, acknowledging the state was “undertaking too much and exercising too rigid a control” and encouraging technology transfer within the economy through research organizations, production units, and agriculture. The policy introduced the commoditization of scientific information and economic incentives for technology transfer (Canada IDRC and SSTC, 1997, 29) as an alternative to the fiat mechanisms of the command economy. Later that year, China would reform its patent and trademark law to encourage innovation and stimulate imports of foreign technologies (Frame & Narin, 1987).

This commitment was linked to educational reforms beginning in 1985 that established compulsory education, decentralized control of education, and granted greater freedom and mobility to professors, scientists, and students. In 1986, China set up its National Natural Science Foundation to fund basic research. However, the main resource commitments to tertiary education and science came a decade later. In 1995, at the Third National Conference on Science and Technology, “The Decision on Accelerating Scientific and Technology Progress” described *zizhu chuangxin* or “indigenous innovation” as the source of China’s future development. In the

same year, the 211 Project was launched to strengthen 100 universities and develop key academic fields.

More recently, at the Fourth National Conference on Science and Technology held in January 2006, Premier Wen Jiabao called innovation “the soul of scientific and technological development and the engine behind national development”. The current leadership is thus no less wavering on science and technology compared with its predecessors despite the fact that there has been a rhetorical shift toward “social harmony,” income equality and rural welfare.

Following these policy pronouncements, the Chinese government has conferred increasing resources on science and technology. Starting in the late 1970s, China has made substantial commitments to science and technology, measured, for example, by massive levels of resource allocation to scientific research. The funding levels of science in China are particularly striking for a country of China’s per capita GDP. The country’s emphasis on science and technology is unusual among developing countries at China’s level of per capita GDP, as illustrated in Figure X (King 2005) which shows that relationship between R&D spending and GDP per capita.

The 1995 emphasize on indigenous sources of science and technology was coupled to a huge rise in R&D expenditure. The 10th Five Year Plan (2001-05) identified targets of raising R&D funding to more than 1.5% of GDP and increasing enrollment in higher institutions to over 15 percent.¹ In 1998 the “Action Plan for Invigorating Education in the 21st Century (1998-2002)” was issued by the Ministry of Education together with the 1999 “Decision on Deepening Educational Reform and Promoting Quality Education” which stipulated goals of increasing higher education attendance to 11% in 2000 and 15% by 2010, increasing central government expenditures on education by 1% every year during the five year period 1998-2002, and

developing first-rate institutions and global centers of excellence.²³ The 985 Project was initiated in 1998 to provide significant funding to ten leading universities with the aim of creating world-class universities and was expanded in 2004 to 30 universities. The 111 Plan launched in September 2005 aims to recruit 1,000 scientists from the world's top universities in an effort to make Chinese universities innovation centers.

These massive resource commitments are the most visible signal of a country poised to emerge as a major global player in science and technology. The policy targets outlined above have been followed with rising R&D spending, reaching 1.3% of GDP in 2003, closely linked to blueprint laid out in the 10th Five Year Plan (2001-05) which identified targets of raising R&D funding to more than 1.5% of GDP by 2005. Moreover, some economists have pronounced China to be on the verge of “science and technology takeoff” (Gao & Jefferson 2005) on the basis of the historical relationship between R&D and GDP. Documenting that “a country's R&D spending approaches one percent of GDP abruptly accelerates to the vicinity of two percent, and then levels off in the range of two to three percent of GDP” the authors argue that with China's 2003 R&D spending reaching 1.3% GDP up from 0.6% in 1996, the country is set to take off. Since 2003 China has indeed continued to increase its R&D spending. In 2006 China spend \$136 billion on R&D overtaking the \$130 billion spend by Japan and reaching about 40% of the United States spending levels (\$330 billion in 2006). This has been paralleled by China's education expenditures which, as a percentage of GDP, have increased from 2.55% in 1998 to 3.41% in 2002, with the 11th Five Year Plan stipulated to increase this to 4% by 2010. Investment in colleges and universities has doubled between 1998 and 2004 to \$12 billion. University facilities have also been upgraded with teaching and experimental equipment having

¹ 11th Five Year Plan website document

² <http://www.moe.edu.cn/edoas/website18/level3.jsp?tablename=208&infoid=3337>

doubled in the past five years.⁴ Moreover, post-graduate enrollment has grown from 70,000 in 1998 to 365,000 in 2006, of which doctoral enrollment is 208,000.^{5 6}

This sustained and high-level commitment to science and technology represents a rare consensus view among three generations of Chinese policy makers who otherwise disagreed about many issues.

CONTEMPORARY SCIENTIFIC INSTITUTIONS IN CHINA

While China's considerable allocation of resources for science is not subject to dispute, the issue is whether the institutional foundations for scientific progress in China support the rapid disclosure, exchange and evaluation of novel scientific ideas (Merton 1973; Dasgupta and David 1994, Aghion Dewatripont and Stein 2008).

An emerging literature has sought to analyze the institutional environment supporting the so-called 'republic of science' (Dasgupta and David 1994; Mokyr 2002). This system of closely intertwined institutional is embedded in the broader national system of innovation (Nelson 2002) but specifically includes the recognition of scientific priority and a system of public (or coordinated) expenditures to reward those who contribute to cumulative scientific knowledge production over the long term (Merton, 1973: Dasgupta and David, 1994). Thus it is clear that at least two elements are critical to the effective functioning of the scientific system and the accumulation of knowledge. These elements can be characterized as supporting openness and freedom.

³ <http://www.moe.edu.cn/edoas/website18/level3.jsp?tablename=208&infoid=3314>

⁴ Ministry of Education report October 12, 2006

⁵ China's colleges to enroll 5 percent more students in 2007. Xinhua News Agency. January 24, 2007

⁶ Ministry of Education report (October 2007)

The focus on openness follows the description of the Mertonian norms pursued by sociologist of science who sought to understand a previously ignored but critical element in institutional life (Merton 1973). Key elements include an effective system for the disclosure of scientific information, its validation and rewards for disclosure as provided by peer-reviewed journals and the tenure-based reward system that rewards cumulative research contributions disclosed in the public domain. Another key element is grounded in a variety of organizations, such as Biological Resource Centers, that collect, certify, and distribute key research materials including biological organisms, such as cell lines, microorganisms, and DNA material. The significance of such institutions to the life sciences derives from the importance of using certified research materials in life sciences and beyond. The ability to build upon existing knowledge often depends on access to the cells, cultures, and specimens used in prior research as well as certainty about the fidelity of those materials. More broadly, databases, software code repositories and other such facilities provide the institutional foundations for effective knowledge accumulation [Hunter-Cevera, 1996; OECD, 2001; Smith, 2003; Stern, 2004].

In contrast, the notion of freedom as provided by the institutions of science is an often taken for granted aspect of the scientific system that has been less well characterized (de Solla Price 1965). Recent work by Aghion and co-authors (Aghion, Dewatripont and Stein 2008) argues that while rewards for openness and disclosure constitute a critical aspect of the academic system, freedom to pursue a chosen research agenda is also central to the contract between academics and those who fund them. By building institutions that credibly commit to freedom, academia endogenously emerges as the most effective set of institutional arrangements for highly uncertain but potentially important early-stage research. This is particularly salient today when the traditional arguments for funding academic research on the basis of underinvestment due to

appropriability concerns has faded with expanding patenting of academic research (Nelson 1959, Mowery et al. 2005). Thus Aghion et al. (2008) suggest that a system which places strong control rights with scientists will be particularly salient in ensuring that high quality, basic research is undertaken.

Against the contours of the scientific institutions laid out above, China's scientific system is poorly documented, particularly as it shapes the daily practices of scientists (rather than their idealized configuration) (Latour & Woolgar 1979). Certainly recent work narrowly focused on the area of stem cell science undertaken by one of the authors suggests that China's institutions continue to stifle openness, free flow of information, collaboration, and peer-review (Murray & Spar 2007). Taking a broader, less fine-grained approach, we examine the stylized features of China's institutional system.

INSTITUTIONAL FOUNDATIONS OF ACADEMIC FREEDOM IN CHINA

With respect to freedom to pursue specific research projects, China has a strongly top-down system for managing resource allocation, goal setting and implementation. Unlike the United States, the Ministry of Education (under the 211/985 programs) together with local governments not only funds the tertiary education and provides long-term educational and research goals but has also been actively involved in the day-to-day management of the Chinese universities (e.g., Qiping and White 1994). It appoints all the presidents of the universities. The literature suggests this changed in about 1993. Specifically, the 1993 Program for Education Reform and Development in China stimulated a transition from a state-controlled machinery to a state-supervised one, where "university presidents would become responsible for their own institutional policies and long-term development plans." (Yang et al., "Dancing in a Cage:

Changing Autonomy in Chinese Education” High Educ (2007) 54:575-592, pg. 580); also Yin and White (1994). Nonetheless, faculty, unless they occupy administrative positions, have almost no say over the running of a university. The Ministry of Education has a department that certifies whether a faculty member at any university is qualified to supervise PhD students. It is in the business of approving whether a particular university can roll out a new degree program (such as an EMBA program by a business school).

Beyond directly controlling academic institutions, the role of the government also encompasses operating and managing all the public sources of scientific research. The key mechanism for funding allocation arises through the so-called 973 Program, formerly the “National Basic Research Program” (initiated in the 1980s) which focuses on basic research in “strategic” industries. The 973 Program was said to have been personally endorsed by Zhu Rongji, China’s vice premier in charge of economy (later the premier 1991-2002). Basic research funded under the 973 Program received 10 to 15 times the funding normally allocated through the National Natural Science Foundation. The definition and the specification of “strategic” industries and interest would rest with the government. Other major initiatives directing and controlling scientific research were guided after 1998 by the National Steering Group for Science, Technology and Education. This group was chaired by Zhu Rongji, China’s premier, the group consisted of State Council officials and it had broad control over the entire science and technology sector, including firms. All the major producers of scientific research, such as National Academies and universities reported to this group.

As part of its mission, this steering group laid out long-term scientific and technological “visions”. For example, the Middle and Long-Term Program for Science and Technology Development 2006-2020 outlined a vision of China based on measures of patents and academic

citations, increasing R&D four-fold to 900 billion Yuan (estimated to be 2.5% of GDP in 2020), increasing science and technology's contribution to 60% of the country's development, and reducing foreign technology to less than 30%. The plan was then grounded in a specific research agenda that focused on “frontier” and breakthrough technologies in biology, information technology, and nanotechnology, but also on addressing China's energy, water, and human organ shortages and develop technologies to explore the seas, oceans, and space.⁷ More specifically, several waves of programs of substantial government funding of science and technology have controlled research resources through a strongly guided program of resource allocation that appears (at least on the surface) to severely limit academic freedom as constituted by the ability of researchers to follow their chosen research agenda. The list of the programs is long and complex and includes: The National High Tech R&D Program or 863 Program (1986), the Spark Program, the Torch Program, the National Key Technologies Program, Project 211 (1996), Project for Funding World-Class Universities (Project 863, 1998), and Project 111 (2006). The cumulative impact of these programs is not clear but as we will discuss in our analysis it suggests the careful allocation of research resources to targeted scientific fields rather than across the board.

It is through these programs that the Chinese government controls the balance of research spending between basic versus applied research. In the United States, 58% of R&D expenditures are allocated to development research, 23% to applied research, and 19% to basic research.⁸ In China, of the R&D expenditures in 2005, 77% went to developmental research, 18% to applied research, and only 5% to basic research.⁹ The OECD argues that the lack of basic and applied

⁷ China issues guidelines on sci-tech development program, GOV.cn Thursday, February 09, 2006

⁸ NSF Science and Engineer Statistics 2006.

⁹ China Statistical Yearbook 2006. 21-38 Basic Statistics on Scientific and Technological Activities

research will hamper innovation efforts.¹⁰ That said, few developing countries – which have to struggle with poverty and complicated politics – have the Chinese level commitment to science and technology.

INSTITUTIONAL FOUNDATIONS OF ACADEMIC RESEARCH IN CHINA¹¹:

REPUBLIC OF GOVERNMENT

Commitments to openness with the Western university system are typically characterized as being grounded in the incentive mechanisms introduced to reward research productivity and disclosure through publication. By premising career rewards (such as tenure) on disclosure through publication, these institutional aspects of science leverage the public goods nature of research and are intended to promote disclosure and subsequent cumulative innovation (Dasgupta & David 1994; Murray & Stern 2007). Within China, however, the micro mechanism shaping how science is performed, disclosed and accessed each day, are poorly characterized. Our preliminary enquiries suggest that in contrast to the United States, China's incentives to reward research productivity are akin to piece production in manufacturing.

The American academia operates on a tenure system that rewards the cumulative achievements of a scholar during a specified period. Moreover, the system relies exclusively on peer review rather than the top-down judgment of deans or university presidents (let alone a bureaucrat in Washington DC). The Chinese university system is entirely different and more akin to systems in widespread use in Europe in the 1950s and 1960s (Mairesse 2008). In effect, China grants scholars tenure on day one and then uses salaries, bonuses, and promotion to incrementally reward scholarly achievements and punish underperformance.

¹⁰ OECD Reviews of Innovation Policy – China Synthesis Report, August 2007.

¹¹ This part of the discussion is based on Chinese media reports and private discussions with Chinese academics.

One effect of this system is that it relies heavily on specific incentive mechanisms to reward scholarly achievements. These mechanisms are exclusively financial in nature and are based on highly quantitative metrics. A case in point is the widespread renown of Eugene Garfield, the founder of Science Citation Index (SCI). Mention his name of to a US academic and you are likely to get a blank stare. In China he is widely known because the Chinese academic incentive mechanisms are linked to SCI in a highly mechanical and transparent fashion. Publishing in a journal covered by SCI leads to a specific monetary reward; publishing in many of the SCI journals leads to promotions (such as directorships of labs).

In some respects, Chinese academia resembles Wall Street in its compensation design: The base pay is low but (beginning in the early 1990s) the government introduced a bonus system that was rolled out on a massive scale in the mid-1990s. Although we do not have systematic evidence, conversations with Chinese academics reveal that the bonuses are a significant share of the total compensation package and among those who are considered as “star” academics the bonuses can be several multiples of the base salary. (We will return to this issue later in the paper.) The bonus system requires regular performance reviews and the basis of the performance review prevailing in almost all the academic institutions in China today is the practice of linking bonus pay to publications included in SCI. This system was introduced in the early 1990s by Nanjing University. The rule was that the university would pay RMB 1,000 to researchers who published in journals indexed by SCI. Subsequently, Nanjing University would also require graduate students to publish in SCI journals to fulfill their graduation requirements. Currently, SCI covers 6,000 journals of which 74 are published in China (although all 6,000 are of course open to Chinese scholars as venues for their scholarly research). After a Chinese study

revealed that Nanjing University ranked as No. 1 institution in SCI publications between 1992 and 1998, this SCI-based system was widely adopted by other Chinese universities (CITE).

The financial incentives linked to SCI publications escalated enormously since the late 1990s. While there no systematic data exist, a number of reports in the Chinese media reveal that SCI-linked pay ranges from a few thousand to tens of thousands of Yuan. This SCI-linked pay system has become both elaborate and specific: At many universities the bonus is linked with specific SCI journals. For example, Beijing Normal University would pay RMB 50,000 to researchers who succeeded in publishing in *Nature*. Another university promised to pay RMB 500,000 to researchers who published in *Science*. Responding to this incentive system, one professor in Wuhan published 65 SCI papers in single year (2003). He received RMB 10,000 per publication and RMB 650,000 (about \$95,000) in total. Another academic couple in the northeast region of the country published an average of one SCI paper per day. Perhaps the most famous “SCI master” is Professor Zheng Yueqin, a professor of chemistry at Ningbo University in Zhejiang province (a third-tier university in China). Professor Zheng gained fame because of his seeming research productivity having been wooed back by the Ningbo government with a grant of RMB 1 million in 1998 to open his own laboratory. He was heavily promoted as an academic star and he was invited to join the government in an advisory capacity. He won substantial grant money for his university and for Ningbo. His case is very telling both of the top-down management of the Chinese academia and as well as of the massive distortions introduced by the SCI-indexed compensation.

By publishing 24 SCI papers in 2004 alone under the compensation rules at Ningbo University he received the handsome sum of RMB 168,000 (about \$24,000) – a supplement that more than doubled his base salary (estimated to be at least RMB 100,000). Upon further scrutiny,

it appears that he published heavily in extremely low impact SCI journals. (Some Chinese dub the phenomenon “garbage publications”.) He published 39 papers in a journal called New Crystal Structures with a Journal Impact Factor (JIF) of 0.349 (i.e. the average the paper in this journal is cited 0.349 times during the life of the publication). His best publication was in a journal with a JIF of 3.389, of which he was the 5th author. His research publications are documented in Table 1.

Table 1: The Sci Publications by Professor Zheng Yueqin

Journal	Journal impact factor	# of papers by Prof. Zheng
New Crystal Structures	0.349	39
Journal of Chemical Crystallography	0.4	2
Acta Chimica Sinica	0.643	1
Journal of Chemical Sciences	0.729	3
Journal of Coordination Chemistry	0.841	7
Journal of Molecular Structure	1.021	3
Allgemeine Chemie	1.127	17
Solid State Sciences	1.327	1
Journal of Solid State Chemistry	1.413	3
Inorganic Chemistry Communications	1.513	1
Polyhedron	1.584	2
Tetrahedron Letters	2.326	1
Crystal Growth & Design	2.742	1
Inorganic Chemistry	2.289	1

Source: Chinese Media Reports

Beyond rewards for disclosure in scientific journals, openness in scientific research is also supported by scientific institutions that allow for the effective sharing of materials and information and by institutions whose incentives allow for widespread, democratic access to key knowledge inputs (Von Hippel 2006, Furman and Stern 2009). While access to materials and other inputs has not been the subject of widespread analysis in China (as it has in the United

States e.g. Campbell et al. 1995), anecdotal evidence suggests that politics still intrudes when scientific research intersects with a public policy issues and with the perceived need on the part of rulers to maintain political stability. Nowhere is this conflict between science and politics more clear than the case of SARS in 2003. Below is a lengthy quote from an article published in Science (July 18, 2003):

In a well-equipped lab in southern Beijing, a group of virologists had already discovered a new virus in samples from some of the earliest patients... By the first week of March, the group had tentative evidence that the new virus might indeed be linked to the epidemic. There was just one problem. They didn't dare tell the world... A call or an e-mail to Stöhr [coordinator of the WHO network] might also have ensured Yang and his colleagues a more prominent place in the history of the disease..... That failure, many note, stems in part from systemic problems in Chinese science: a lack of coordination and collaboration, stifling political influence, hesitation to challenge authorities, and isolation from the rest of the world. Teams in Stöhr's network, which at the time didn't include anyone from mainland China, started holding daily teleconferences, posting their findings on a secure Web site, and sending each other samples and reagents by overnight delivery.

The above case is about an active suppression of scientific information in the interest of what was viewed as the need for political stability, but politics intrudes in other ways and at times when a public policy exigency is absent. This has to do directly with the top-down management of Chinese science. The top-down management system has two related features. One is that information flows up and down rather than sideways; the other is that bureaucratic

barriers exist to hamper horizontal flows of information. The latter effect arises because the complexity and the multiplicity of any subject matter are such that often different bureaucracies are involved in the day-to-day management. Below is an admission of this problem by the Chinese themselves, again with respect to the SARS outbreak (People's Daily 10-23-2003):

“Li [Director of the Chinese Centre for Disease Control and Prevention] admitted there are still some problems with the management of specimens of the SARS virus. Many laboratories hold their own specimens and need to co-ordinate their research into SARS to make their work safer and more efficient.”

This problem with information hoarding is apparently pervasive in the Chinese system, as documented in the case of stem cell research (Murray & Spar 2007). An interview with the Chinese director of Pasteur Institute in China conducted by one of the authors of this paper reveals that information sharing is hampered routinely and it is not just politically-sensitive information that is affected. The Pasteur Institute in China works formally with the Shanghai branch of the Chinese Academy of Sciences (CAS) but its research requires access to samples controlled by Ministry of Agriculture and Chinese CDC, which are separate bureaucracies from CAS. With its relatively limited influence over the Chinese CDC, the Institute became entangled in a variety of challenges regarding their samples all of which hampered their ability to undertake experiments in a timely way. In another setting, a Hong Kong scientist was criticized severely by the Chinese government when he used strains he had collected for research on avian flu rather than the strains collected by the Ministry of Agriculture. Similarly, stem cell

researchers report a greater willingness to share samples and techniques with foreign colleagues than with colleagues at other institutions within China.

PUZZLES FOR CHINA'S SCIENTIFIC KNOWLEDGE PRODUCTION

It is generally undisputed that China is producing a vast quantity of scientific knowledge (see Figure 1). But are the institutional foundations of China's scientific system conducive to producing innovative, leading-edge scientific research? What is the role of the quantity-driven and financial incentive-based system in shaping both the quantity and the quality of Chinese scientific research?

Figure One: China's Role in Global Knowledge Production

PRODUCTIVITY OF NATIONS BASED ON THEIR PUBLICATION RECORD					
Country	Country classification	No. of publications (1997-2001)	GDP per capita	Publications per GDP per capita	Publications per GDP per capita per year
India	IDC	77,201	487	158.52	31.7
China	IDC	115,339	989	116.62	23.32
United States	G8, OECD	1,265,808	36,006	35.16	7.0
Brazil	IDC	43,971	2,593	16.96	3.39
Germany	G8, OECD	318,286	24,051	13.23	2.64
United Kingdom	G8, OECD	342,535	26,445	12.95	2.59
Japan	G8, OECD	336,858	31,407	10.73	2.15
South Africa	IDC	18,123	2,299	7.88	1.58
Canada	G8, OECD	166,216	22,777	7.30	1.46
Italy	G8, OECD	147,023	20,528	7.16	1.43
PRODUCTIVITY OF NATIONS BASED ON THEIR CITATION RECORD					
Country	Country classification	No. of citations (1997-2001)	GDP per capita	Citations per GDP per capita	Citations per GDP per capita per year
India	IDC	188,481	487	387.02	77.40
China	IDC	341,519	989	345.32	69.06
United States	G8, OECD	10,850,549	36,006	301.35	67.27
United Kingdom	G8, OECD	2,500,035	26,445	94.54	18.91
Germany	G8, OECD	2,199,617	24,051	91.46	18.29
Brazil	IDC	155,357	2,593	59.91	11.98
Japan	G8, OECD	1,852,271	31,407	58.98	11.79
Canada	G8, OECD	1,164,450	22,777	51.12	10.22
Italy	G8, OECD	964,164	20,528	46.97	9.39
Australia	OECD	623,636	20,822	29.95	5.99

Source: D. A. King *Nature* 430, 311-316 (2004). IDC, internationally developing country; OECD, Organisation for Economic Co-operation and Development.

This study will begin to probe into some of these issues. More specifically, the scholarly literature and the popular discourse on China's place in the global engine that produces scientific knowledge hold a number of puzzles. These relate to the composition, quality and distribution of scientific work.

Composition: First, with regards composition, China's shift towards a publication-based incentive system grounded in increasing the number of ISI-publications suggests that scientific publications are likely to have increased across the board. The policy is not a nuanced incentive system designed to elicit particular types of scientific effort focused but instead one that is designed to increase effort overall. On the other hand, specific policy interventions of the type described in the 973 program and the focus on specific areas of expertise appear designed to shape the composition of scientific effort along particular composition margins. Moreover, funding allocation suggests a shift towards applied research and away from basic research rather than a stable compositional pattern. Thus it remains a critical puzzle to understand whether the rise in Chinese scientific knowledge has taken place disproportionately along key margins specifically: particular research fields and the basic-applied continuum.

Quality: With regards to the quality of scientific research, China is increasingly viewed as a serious participant in the global knowledge economy (CITES). Recent studies have highlighted the rising number of scientists and engineers educated in Chinese institutions of higher education, and the growing funding allocated to the production of knowledge. By 2005, China ranked fifth amongst nations in total scientific publications produced; in engineering, second, behind only the United States. China has meanwhile embraced major scientific projects in genomics, becoming the only developing nation to participate in the Human Genome Project

in the 1990s; in space science, becoming the third nation to develop a successful manned space program; and more recently, in environmental and energy research through partnerships with the European Union. On the other hand, there is widespread concern that research quality is generally low not least because of the SCI-18 indexed compensation system. It is certainly plausible to argue that this practice has a disproportionate quantity effect at the expense of the quality of the research. Indeed, Chinese academics are already beginning to rethink about the SCI-based system. One prominent Chinese computer scientist and a member of CAS, Li Guojie, dubbed the SCI system a “stupid Chinese idea.” His institute, the Computer Research Institute under CAS entered into a collaborative project with Tsinghua University and one of the first decisions Mr. Li made was to abolish the SCI-indexed compensation system.

Distribution: The third “puzzle” in the debate over science in China pertains to the distributional nature of knowledge production. While little has been written about the changing organization of knowledge production in China at least two perspectives emerge. First, the expansion of higher education to an increasing number of students suggests that a growing number of faculty members are participating in the educational process. If we extent this view to our understanding of knowledge production, similar trends would lead us to expect that knowledge production would be expanding to encompass an ever growing number of institutions. An alternative view holds, however, when we consider the importance of a small number of highly qualified scientists typically trained abroad and then lured back to the country through the 111 program and others. Most of the anecdotal stories of such individuals locate them within either the Chinese Academy of Sciences system or a small number of elite universities. If this is the case, then we might expect that a small number of institutions (continue to) dominate knowledge production, particularly when we take quality ratings into consideration.

A related distributional topic highlights the role of international collaboration in shaping knowledge production in China. Again, the prevailing view is that much of China's rise in production is closely tied to participation in global scientific projects. In contrast, however, the perspective of "indigenous innovation" argues that the real rise in science has been predominantly local in nature, not closely tied to international collaboration but instead through local upgrading of laboratories, technology and the institutional environment for scholarly work.

In the analysis that follows we develop a comprehensive analysis of China's publications (as captured in SCI) to examine these puzzles.

PUBLICATION ANALYSIS

While publications only capture one facet of scientific knowledge production, they provide a particularly important measure for the analysis of academic scientific research in the late twentieth century and early twenty first century because they have become the central mechanism used by academics to disclose their research findings and disseminate them to the global scholarly community (Cole 2000). Conferences, on-line sites and even private communications continue to serve as key elements of scientific life, nonetheless publications dominate knowledge production and form a key pillar in the broader set of scientific institutions that shape freedom and openness as articulated above.

METHODOLOGY

Scientific indices such as the Science Citation Index (SCI) and Engineering Index index the bibliographic content of articles printed in a set of journals. The data provided for an article varies by index, but typically includes such information as the names and institutional affiliations

of authors, pertinent keywords, and the year and source of publication. For the purposes of this analysis we use the expanded Science Citation Index (SCI-Expanded), an index of approximately 8,000 scientific journals published by Thomson Reuters. As noted above, SCI is widely used in China. It is also a de facto standard for scientific scientometric studies (e.g. Zhou and Leydesdorff) and is used by both NSF and Most (MOST 2007)¹² as a research indicator. SCI is a proprietary database, with access through the Web of Science (WoS), a web-based portal operated by Thomson Reuters. The service allows queries to be executed against the database, and for the full records of the corresponding articles to be downloaded in tabular format.

In 2008 and 2009, the full records for about 800,000 Chinese publications published between 1979 and 2008 were downloaded from Web of Science. A publication was defined as “Chinese” if the term “Peoples R China” appeared in the address field of the publication. In other words, this captures any and all articles indexed by SCI that include an authors whose affiliation is listed as being located in China. While publications are useful in providing simple count of publication volume over time, in order to address the various puzzles laid out above, we developed a series of post processing methods to categorize publications along a variety of margins. Specifically these included: field, basic-applied, quality, and distribution.

COMPOSITION – FIELD ANALYSIS: With refard to field, SCI categorizes journals into zero or more subject categories (SCs) that identify the field of articles appearing in the journal, with the vast majority of journals (about 99.4% of unique journals in the dataset and 99.98% of articles) assigned to at least one category. The dataset contains 246 unique SCs, primarily in scientific disciplines. SCs are not hierarchal, and all combinations are valid. However, while useful for some purposes, the entire set os SCs is too unwieldy to use for an

¹² Although the databook sums the SCI, EI, and ISTP (Index to Scientific and Technical Proceedings) indices to obtain total papers, the different indices are not mutually exclusive.

analysis of field-level variations in scientific knowledge production. Instead, we used the 13 subject fields developed by the National Science and Engineering Indicators (published by the United States National Science Foundation). NSF explicitly defines 126 subfields as mapping to specific fields. Of these, 107 map directly to specific subject categories, and are mapped as explicit. These map almost directly to the SCI subject categories and were used to categorize journals (with one exception -- the S&E field astronomy was merged with physics since the corresponding SC entry, Astronomy and Astrophysics, could not distinguish between astronomical and astrophysical applications). Additionally, two fields were added: Multidisciplinary, corresponding to the SC field Multidisciplinary Sciences, for inherently-multidisciplinary journals such as Science; and Non-S&E for humanitarian subject categories. These top-level fields are summarized in Table 2.

TABLE 2: Science & Engineering Indicators Subfields

Agricultural sciences	Geosciences	Other life sciences
Biological sciences	Mathematics	Physics[1]
Chemistry	Medical sciences	Psychology
Computer sciences	Multidisciplinary	Social sciences
Engineering	Non-S&E	

(1) Includes astronomy, a separate field in Science and Engineering Indicators

The multidisciplinary categorization is assigned to a journal, but specific articles may be classified differently where citation analysis suggests the applicability of a more specific field. Since the multidisciplinary classification is then published at a higher scope (at the level of the journal rather than the level of the publication), any more specific subject takes precedence. Certain fields do not correspond directly with a single category. In some cases, the subject field may be divided between explicitly-defined fields; for instance, Agricultural Engineering may

correspond to agricultural sciences or engineering; or “Geriatrics and gerontology” where NSF classifies geriatrics as Medical Sciences and gerontology as Social Sciences. Other fields, such as linguistics and architecture, do not easily correspond to a specific field and are also classified at this level. Of the articles in the dataset, 90.5% correspond with exactly one field. Where an article corresponds to multiple top-level fields at the highest priority, the field corresponding to the first SC at that priority is selected.

COMPOSITION - BASICNESS: CHI Research, Inc. (now the Patent Board) has classified SCI-indexed journals on a four-level scale based upon the subject category and content of the journal. Level 4, the most basic, is defined as basic research; level 3 is defined as applied research; level 2 for engineering and technical science; and level 1 to applied technology. Most journals analyzed by CHI specialize in basic or applied research. Journals established after 1996, when the CHI analysis was performed, are not indexed.

JOURNAL IMPACT FACTOR: ISI assigns to each indexed journal a journal impact factor (JIF), calculated as the number of citations to the journal divided by the number of articles (excluding front-matter material such as editorial content) over a two-year period. The JIF can be estimated as the expected number of citations to the average article two years after publication. Scientometric literature often uses the JIF as a proxy for the quality of the journal, implicitly inferring that many citations to a journal article is indicative of a high-quality article. Such analysis requires caution: the metric can be manipulated by, for instance, publishing review articles that contain no original research yet are known to be highly-cited, encouraging self-citations, manipulating the online accessibility of articles, and encouraging strategic use of front-matter material (which contributes to the numerator, but not the denominator, in calculating the

JIF). Additionally, citation practices differ amongst fields, preventing a meaningful quality comparison between fields.

While an imperfect measure, the JIF can be used as a proxy for quality when controlling for its deficiencies by, for instance, not using the JIF to ascertain the quality of individual authors or articles, or to compare the quality of discrete fields. The analysis presented here uses the JIF only to discriminate amongst journals in the same field, making no assertion of the quality of individual articles or authors published therein.

DISTRIBUTION: SCI includes the institutional affiliations of all authors affiliated with each publication, but does not include the type of each institution. During post-processing, the institution names were parsed with each institution classified into one of four categories: institute of higher education, government (including research of the national academies), industry, and unknown. Institutes are classified by keywords within the name, such as “University,” “inc.,” and “institute”. Using this methodology, slightly over 90% of the authoring institutions could be automatically classified. Each article was then assigned to one or more institutional types.

RESULTS

Our data show that the number of Chinese scientific publications increased exponentially since the mid-1990s with this increase coincided with rising research and educational expenditures (especially on university education), increasing university enrollments and more graduates in science and technology.

References

- Adams, J. D. and Z. Griliches. 1998. "Research Productivity in a System of Universities." *Annals of INSEE*, 127-162.
- Aghion, P., M. Dewatripont, J. Stein. 2008. "Academic Freedom, Private Sector Focus & the Process of Innovation." NBER Working Paper No. W11542 and forthcoming.
- Allison Paul D. and Stewart, John A. 1974. "Productivity Differences among Scientists: Evidence for Accumulative Advantage." *American Sociological Review* 39: 596–606.
- Campbell, E. G., B. R. Clarridge, M. Gokhale, L. Birenbaum, S. Hilgartner, N.A. Holtzman, D. Blumenthal. 2002. "Data withholding in academic genetics: Evidence from a national survey." *Journal of the American Medical Association*, 287(4): 473 - 479.
- Canada IDRC and SSTC, 1997. *A Decade of Reform: Science and Technology Policy in China*. International Development Research Centre, Ottawa, Canada. Jointly authored by the International Development Research Centre of Canada and the State Science and Technology Commission of the People's Republic of China.
- Cole, J. 2000. "A short history of the use of citations as a measure of the impact of scientific and scholarly work." In E. Garfield, B. Cronin & H. B. Atkins (eds.), *The web of knowledge: A festschrift in honor of Eugene Garfield*. Information Today, Medford, NJ, 281 -300.
- Cole, Jonathan R. & Stephen Cole. 1973. *Social Stratification in Science*. Chicago, IL: University of Chicago Press.
- Collins, H. M. 1974. "The TEA Set: Tacit knowledge and scientific networks." *Science Studies*, 4: 165-186.
- Crane, D. 1969. "Social structure in a group of scientists: A test of the invisible college hypothesis." *AMERICAN Sociological Review*, 34(3): 335-352.
- Dasgupta, P., P.A. David. 1994. "Towards a new economics of science." *Research Policy*, 23: 487-521.
- David, Paul. 2003. "Can 'Open Science' be protected from the evolving regime of IPR protections?" Working Paper #03-011, Stanford.
- DiMaggio, P. J., W. W. Powell. 1983. "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields." *American Sociological Review*, 48: 147-160.

- Frame, J. D. and Narin, F., 1987. "The Growth of Chinese Scientific Research, 1973-84." *Scientometrics*, 12(1):135-144. 25
- Friedland, Roger, Robert, R. Alford. 1991. "Bringing Society Back In: Symbols, Practices, and Institutional Contradictions. In Walter W. Powell and Paul J. DiMaggio (eds.), *The New Institutionalism in Organizational Analysis*, 232-263. Chicago: University of Chicago Press.
- Garfield, E. 1979. *Citation Indexing*. Philadelphia, PA: ISI.
- Hagström, W. O. 1965. *The Scientific Community*. Basic Books, New York.
- Hunter-Cevera, J.C. (1996) "The importance of culture collections in industrial microbiology and biotechnology, in *Culture Collections to Improve the Quality of Life*," eds. R. A. Samson, J. A. Stalpers, D. van der Mei et al., Centraalbureau voor Schimmelcultures, Baarn (Netherlands), 158-62.
- Latour, B. S. Woolgar. 1979. *Laboratory Life: The Construction of Scientific Facts*. Sage Publications, Beverly Hills.
- Levin, S., P. Stephan. 1991. "Research Productivity Over the Life Cycle: Evidence for Academic Scientists." *American Economic Review*, 81(1): 114-132.
- Merton, R. K. 1973. "The normative structure of science." *The sociology of science: Theoretical and empirical investigations*, 267-280. The University of Chicago Press: Chicago, IL.
- Merton, R. K. 1988. "The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property." *Isis*, 79(4): 606-623.
- Meyer, J. W., B. Rowan. 1977. "Institutionalized Organizations: Formal Structures as Myth and Ceremony." *American Journal of Sociology*, 83: 340-363.
- Mokyr, J. 2002. *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press, Princeton.
- MOST [Ministry of Science and Technology of the People's Republic of China], 2007. "2007 China Science Technology and Statistics Data Book." Available: <http://www.most.gov.cn/eng/statistics/2007>.
- Murray, F., Aghion, P., M. Dewatripont, J. Koley, S. Stern. 2008. "Of Mice and Academics: The effects of openness on scientific research." MIT Sloan Working Paper.
- Murray, F., S. Stern. 2007. "Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis." *Journal of Economic Behavior and Organization*, 63(4): 648-685.

- National Institutes of Health (NIH). 2003. Final NIH statement on sharing research data [NOT-OD-03-032]. Available: <http://grants.nih.gov/grants/guide/notice-files/not-od-03-032.html>.
- Nelson, R. 1993. *National Innovation Systems: A Comparative Analysis*. New York (NY): Oxford.
- North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives*, 5(1): 97-112.
- OECD (2001) *Biological Resource Centers: Underpinning the Future of Life Sciences and Biotechnology*, Paris (FR): OECD, Directorate for Science, Technology, and Industry.
- Romer, Paul M. 1990. "Endogenous technical change". *Journal of Political Economy* 98(5): S71-102.
- Schubert, A., T. Braun. 1993. "Standards for citation-based assessments." *Scientometrics*, 26: 21-35.
- Scotchmer, S. 2004. *Innovation and Incentives*. MIT Press, Cambridge, MA.
- Scott, W. R. 2001. *Institutions and Organizations* 2nd ed. Thousand Oaks, CA: Sage.
- Smith, D. (2003) "Culture collections over the world," *International Microbiology*, 6(2), 95-100.
- Stern, S. (2004) *Biological Resource Centers: Knowledge Hubs for the Life Sciences*, Washington (DC): Brookings Institution Press.
- Xin, H., G. Yidong 2006. "China bets big on big science." *Science*, 311:1548-1549.
- Zhou, P., L. Leydesdorff 2006. "The emergence of China as a leading nation in science." *Research Policy*, 35:83-104. 27