8. Methodology and Notes on Data Series

International Data

UNESCO Data

Dr. Morton Brown of the Division of Statistics of the United Nations Educational, Scientific, and Cultural Organization (UNESCO) provided the National Science Foundation (NSF) with available data on enrollments and graduates in six fields of science and engineering, by gender, for more than 150 countries for the time period 1975–88. The fields of study (i.e., a student's main area of specialization) were as follows:

- social and behavioral science
- natural science
- math and computer science
- engineering
- agricultural science
- medical science (medical science data were not used in this report)

Dr. Brown provided data in these fields by level of programs. The definitions used for program levels are based on the International Standard Classification of Education (ISCED) and are as follows:

Level 5: Programs leading to an award not equivalent to a first university degree. Programs of this type are usually practical in orientation in that they are designed to prepare students for particular vocational fields in which they can qualify as high-level technicians, teachers, nurses, production supervisors, and the like.

Level 6: Programs leading to a first university degree or equivalent qualification. Programs of this type comprise those leading to typical first university degrees, such as a bachelor's degree or a license, or to first professional degrees such as doctorates awarded after completion of studies in medicine, engineering, or law.

Level 7: Programs leading to a postgraduate university degree or equivalent qualification. Programs of this type generally require a first university degree or equivalent qualification for admission. They are intended to reflect specialization within a given subject area.

ISCED program levels do not guarantee comparability of degrees across countries.

The degree data were verified country by country through national education statistics. See National Data for each country below or the bibliography for source documents. Educational statistics for each country were reconfigured to the NSF classification of fields of science. American and foreign scholars familiar with higher education in each country translated documents and reconfigured university departments' statistics to the NSF classification of fields of study in the Integrated Postsecondary Education Data System (IPEDS).

Demographic Data

UNESCO's Division of Statistics provided demographic data by age segments for all countries.

Dr. My T. Vu, a demographer with the World Bank, provided data files for China, India, Japan, Singapore, Korea, and Hong Kong from the forthcoming 1992 edition of *East Asia and Pacific Region, South Asia Region Population Projections* (Bos, Vu, and Levin 1992).

Mr. Charles Peng provided demographic data on Taiwan from the *Statistical Yearbook of the Republic of China* (Government of the Republic of China, National Science Council 1989).

GDP Data

National accounts time-series data for gross national product (GNP), gross domestic product (GDP), and deflator and exchange rates for 1975-90 were obtained from the International Monetary Fund's *International Financial Statistics Yearbook* for 1990, 1991, and 1992 (International Monetary Fund 1990, 1991, 1992).

University of Pennsylvania "Penn World Tables"

Current national currencies were converted to constant 1987 national currencies and then converted to 1987 constant purchasing power parities ($PPPs) using the University of Pennsylvania World Tables of International Price Comparisons, developed by Dr. Alan Heston and Dr. Robert Summers.

The "Penn World Tables" are based on the United Nations International Comparison Program (ICP) empirical benchmark studies conducted in 1970, 1975, 1980, and 1985. The ICP studies are pricing exercises of identically specified goods and services to estimate price parities for each country's currency and an overall purchasing power parity that makes real comparisons across countries possible. For countries that are not part of the ICP studies, Heston uses post allowance estimations of the Department of State for comparative prices.

The underlying work for $PPPs has been for particular outputs, e.g., food, clothing, and services. No $PPPs exist for inputs; neither is there a specific $PPP for research and development (R&D). Using the overall $PPP will not be perfect, but it is much preferable to using the official exchange rate to compare R&D across countries because the latter measure is volatile and would be very misleading. Most international organizations (e.g., the European Community, the Organisation for Economic Co-operation and Development [OECD], and the United Nations Development Program) and some governmental organizations (such as the Agency for International Development) now extensively use $PPP-based comparisons. Planned economies are not formal members of the ICP project. Price samples are from other sources.

National Data

China

Education data. UNESCO enrollment and graduate data were verified with the estimates made by Dr. Leo Orleans, a China scholar at the Library of Congress. It is Dr. Orleans' opinion that the UNESCO figures for Chinese students and degrees are conservative, not inflated. No attempt was made to derive estimates for the early missing years (1975–81) because the quality of the worker-peasant-soldier-scholars admitted to Chinese universities in the 1970s would not be comparable to students in other countries.

Known uncertainties in R&D data. The first R&D survey of 1985 showed how hard it was to capture R&D expenses from the extensive system of R&D institutes built in the 1950s. With the "Great Leap" of 1958, scientists were encouraged to learn the science, technology, and production wisdom of the masses, and R&D was decentralized. Scientists and engineers in R&D came under the control of "work units."

For the 1985 R&D survey, then, central and local governments had to cooperate to have data come up from these work units and there is a lack of consistency in the concept of R&D at each level. The State Education Commission, which has a strong desire to produce accurate, internationally comparable statistical data, uses a strict definition of R&D in its survey of R&D in the universities. The State Science and Technology Commission collects data from Government research institutes in the natural and social sciences (county, local, ministerial, the Chinese Academy of Science, etc.). Its reports, entitled Science and Technology Statistics, are not yet fully comparable to international standards. The Science and Technology Department of the State Statistical Bureau collects R&D information on large and medium-sized enterprises. By 1988, the annual White

Paper incorporated data from all three of these sources.

China would like to capture R&D expenditures in its surveys and then corroborate these numbers through national budget data. There are many problems to overcome. There are major sections in the national budget for health, education, science, and culture, but no one knows what the R&D component is. The State Statistical Bureau uses the concept of technological development, which can include science and technology (S&T) administration projects and S&T services.

R&D data on China were obtained from Richard P. Suttmeier (1990b) and from Mr. Zhao Yuhai of the Division of Statistics, Department of Planning, State Science and Technology Commission, in a seminar he gave at NSF in April 1991 and in follow-on correspondence in May 1991.

GDP data. Dr. Jeffrey Taylor, formerly of the Bureau of Census, International Statistical Programs Center, and currently with the China Branch of the International Monetary Fund (IMF), recommended computing GDP and deflators from the original source, the *Statistical Yearbook of China* (Government of the People's Republic of China 1991), instead of from the IMF figures. GDP was computed from the values given for the primary, secondary, and tertiary sectors of the Chinese economy. The deflator was rebased from 1978 to 1987, and the Penn World Tables were used to convert constant yuan to constant $PPPs.

India

Education data. UNESCO had received only 2 years of enrollment and graduate data on India: 1974 and 1978. Dr. A.R. Rajeswari of the Department of Science and Technology in India provided data for 1947–83 from the Survey of the Council of Scientific and Industrial Research. More recent data—for 1974, 1979, and 1984–86—were reported in the *Pocket Data Book, 1991* (Government of India, Department of Science and Technology 1991). At NSF's request, the Department of Science and Technology supplied 1987 degree data. Social science data are not included and math and computer science are included in natural science.

Known uncertainties in R&D data. The higher education sector has traditionally not been surveyed in R&D statistics that have come out since 1973. There are 8,000 colleges in India and it was difficult for academic staff to apportion time between research and teaching to classify clearly the human resources employed for R&D and auxiliary activities. It is also considered too difficult to provide data on R&D expenditures. Recently, however, India's Institute of Applied Manpower Research (1991) has published its first survey, sponsored by the Government's Department of Science and Technology. Data on social science research are not collected, social science research is not included in the overall R&D figures for India, and social scientists are not included in S&T human resources statistics.

GDP data. These data were obtained from the Government of India's Central Statistical Organization (1991).

Japan

Education data. Mr. Toshio Kudo, a Fellow in the Japan Program of the NSF Division of International Programs, provided line by line translation of all fields of science included in the six categories, requested and received a detailed written explanation of how Monbusho data are reported to UNESCO, and requested and received the detailed *Monbusho Survey* of higher education for the years 1975-90 (Government of Japan, Ministry of Education, Science, and Culture 1975–90).

Japanese enrollment and graduate data are submitted to UNESCO from the *Monbusho Survey*. UNESCO requests that the Ministry disaggregate math and computer science from natural science, so Japan reports them separately to UNESCO but aggregates them for its national statistics. UNESCO also requests that agricultural engineering be included in engineering. For national statistics, the Government reports it as an agricultural science. NSF made one further adjustment to the Japanese data: Agricultural

economics was subtracted from agricultural science and added to social science.

R&D data. R&D data for Japan were obtained from the Government of Japan's Statistics Bureau, Management and Coordination Agency (1991) and Science and Technology Agency (1991a).

Adjusted data in R&D. The Organisation for Economic Co-operation and Development (OECD 1992) provided adjusted Japanese data (line 26 in each OECD table) below the reported Japanese submission (line 13). The adjustments convert head counts of research scientists and engineers (RSEs) to full-time equivalents (FTEs). The adjusted data still appear in OECD 1992 with the notation (L), "Overestimate," because the head count reported by industry has not yet been converted to FTEs.

The OECD adjustments cause a reduction in Japanese overall R&D of 8 percent and a reduction of overall RSEs by 20 percent. Industrial reporting of FTEs and industrial R&D is not adjusted. University-performed R&D is reduced by 45 percent. In other NSF reports, the numbers used for Japanese R&D scientists and engineers have been adjusted by using only natural scientists and engineers. Therefore, the data differ somewhat.

Singapore

Education statistics. Dr. S. Gopinathan, Director of the Comparative Education Center, National University of Singapore, provided University Annual reports on graduates.

R&D data. R&D data were obtained from the Government of Singapore's Science Council of Singapore (1984–85 and 1987–88) and National Science and Technology Board (1990).

South Korea

Education data. South Korean data submitted to UNESCO cover the years 1975–77 and 1980–89. They were verified against Korean Ministry of Education data, with the assistance of the Education Attache. Education data on enrollments and graduates are contained in the *Yearbook of Educational Statistics* (Government of the Republic of Korea 1976–77 and 1983–90). The data disaggregate junior colleges, colleges and universities, and graduate schools. Advanced degrees disaggregate master's and doctoral degrees, and all levels disaggregate gender.

Data reconfiguration was possible for South Korea and Japan because of the level of detail available. Reconfiguration of South Korean data included the following adjustments for comparability with U.S. data:

- In the social sciences, 11 departments out of the 60 listed in the *Yearbook of Educational Statistics* are in fields of science in IPEDS. Anthropology from the humanities list and agricultural economics from the agricultural scientists list are added to the social science total.

- Five departments of math and computer science are subtracted from the natural science total.

- The above five departments from the natural science list and five departments from the engineering list make up the math and computer science list.

- For engineering, seven inappropriate departments (architecture and computer science courses) are subtracted and agricultural engineering is added from agricultural science.

Known uncertainties in enrollment and graduate data. A classification expert, Dr. Stephen Hunt of the National Center for Educational Statistics, assisted in reconfiguring the Korean education data and selected departments comparable to U.S. definitions of these fields of science. However, similar labels for departments may not mean a similar education. For example, one cannot be sure whether industrial engineering in South Korea is engineering technology or a broader and more advanced engineering program. Does South Korea's food science and technology program have a curriculum equivalent to a bachelor's engineering degree in the United States? Does the information processing department have a

high or low level of curriculum? Are East Asian Studies, or any area studies program, closer to humanities or social science?

R&D data. South Korean R&D data were obtained from the Government of the Republic of Korea's Ministry of Science and Technology (1990). Missing data on overall R&D and breakdown by type, by source of funds, and by performer were requested and received from the Ministry of Science and Technology through the Korean Science Attache in Washington.

Additional industrially funded R&D information was obtained from Dr. Denis Simon of the Fletcher School of Law and Diplomacy, Tufts University.

Taiwan

Education data. Since no UNESCO data are available on Taiwan, degree data were derived from three documents of the Ministry of Education Statistics, each with different levels of detail on departments of science and engineering. The most detailed information by department and subdiscipline was contained in the *Educational Statistics of the Republic of China* (Government of the Republic of China, Ministry of Education 1982–90) for the years 1987–91. Because this detail was available, Taiwan's science and engineering fields were reconfigured to standard classification of fields of science. More summary degree data were available in annual reports of the Ministry of Education for the years 1982–86, but no standardization of the data was possible. A summary document was used for the years 1976–81 containing department totals only. These documents were obtained from the Coordination Council for North American Affairs for 1982-91. Mr. Charles C.T. Peng provided the translation of fields of science in these educational statistics and obtained the earlier missing years, 1975-81, from the Ministry of Education in Taipei.

R&D data. Mr. Peng also provided the R&D data series from the Government of the Republic of China's National Science Council (1988, 1989).

United States

Education data. U.S. education data were obtained from the Institute of International Education (1991a, 1991b); the NSF (1991, 1992b); and the National Research Council (1992), including special tabulations on postgraduation plans of doctorate recipients in 1980, 1990, and 1991.

R&D data. R&D data for the United States were obtained from the NSF (1992a).

Regression Data

Tables 11 through 13 illustrate the factors used to compute growth in R&D, industrially funded R&D, and GDP.

Table 11. Regressions to compute growth in research and development in selected Asian countries

Current currency			Constant 1987 currency		
India (rupees)					
Constant		–258	Constant		–93
Std Err of Y Est		0.022	Std Err of Y Est		0.028
R Squared		0.99	R Squared		0.91
No. of Observations		5	No. of Observations		5
Degrees of Freedom		3	Degrees of Freedom		3
Growth Rate	14.5%		Growth Rate	5.31%	
X Coefficient(s)	0.135		X Coefficient(s)	0.0517	
Std Err of Coef.	0.007		Std Err of Coef.	0.009	
Japan (1981–89 yen)					
Constant		–151	Constant		–136.065305
Std Err of Y Est		0.000	Std Err of Y Est		0.028
R Squared		0.99	R Squared		0.99
No. of Observations		10	No. of Observations		16
Degrees of Freedom		8	Degrees of Freedom		14
Growth Rate	8.8%		Growth Rate	7.96%	
X Coefficient(s)	0.083		X Coefficient(s)	0.076	
Std Err of Coef.	0.000		Std Err of Coef.	0.001	
Singapore (Singapore dollars)					
Constant		–119603	Constant		–109781.72
Std Err of Y Est		31	Std Err of Y Est		20.95
R Squared		0.97	R Squared		0.98
No. of Observations		10	No. of Observations		10
Degrees of Freedom		8	Degrees of Freedom		8
Growth Rate	25.5%		Growth Rate	21.8%	
X Coefficient(s)	60.40		X Coefficient(s)	55.443	
Std Err of Coef.	3.367		Std Err of Coef.	2.306	
South Korea (won)					
Constant		–563449095	Constant		–497231047
Std Err of Y Est		189567	Std Err of Y Est		123397
R Squared		0.95	R Squared		0.98
No. of Observations		10	No. of Observations		11
Degrees of Freedom		8	Degrees of Freedom		9
Growth Rate	33.9%		Growth Rate	23.3%	
X Coefficient(s)	284531.02		X Coefficient(s)	251173.82	
Std Err of Coef.	20870.66		Std Err of Coef.	11765.44	

Table 11. Regressions to compute growth in research and development in selected Asian countries

Current currency			Constant 1987 currency		
Taiwan (New Taiwanese dollars)					
Constant		–315	Constant		–334.120701
Std Err of Y Est		0	Std Err of Y Est		0.08
R Squared		0.98	R Squared		0.97
No. of Observations		11	No. of Observations		9
Degrees of Freedom		9	Degrees of Freedom		7
Growth Rate	17.8%		Growth Rate	15.9%	
X Coefficient(s)	0.16		X Coefficient(s)	0.171	
Std Err of Coef.	0.01		Std Err of Coef.	0.0107	
Total Asian combined R&D ($PPPs [1])					
			Constant		–161.96
			Std Err of Y Est		0.019
			R Squared		0.99
			No. of Observations		11
			Degrees of Freedom		9
			Growth Rate	9.3%	
			X Coefficient(s)	0.087	
			Std Err of Coef.	0.001	

[1] $PPPs = purchasing power parity dollars

Table 12. Regressions to compute growth in industrially funded research and development in selected Asian countries

[In constant $PPPs [1]]

India		
Constant		–203
Std Err of Y Est		0.079
R Squared		0.97
No. of Observations		16
Degrees of Freedom		14
Growth Rate	11.2%	
X Coefficient(s)	0.105	
Std Err of Coef.	0.004	
Japan		
Constant		–176.630
Std Err of Y Est		0.042
R Squared		0.99
No. of Observations		16
Degrees of Freedom		14
Growth Rate	9.9%	
X Coefficient(s)	0.094	
Std Err of Coef.	0.002	
Singapore		
Constant		–448.723149
Std Err of Y Est		0.114
R Squared		0.98
No. of Observations		12
Degrees of Freedom		10
Growth Rate	25.6%	
X Coefficient(s)	0.228	
Std Err of Coef.	0.009	

Table 12. Regressions to compute growth in industrially funded research and development in selected Asian countries

[In constant $PPPs [(1)]]

South Korea		
Constant		–937882.546
Std Err of Y Est		84.78
R Squared		0.99
No. of Observations		9
Degrees of Freedom		7
Growth Rate	26.2%	
X Coefficient(s)	473.47	
Std Err of Coef.	10.94	
Taiwan		
Constant		-406.40
Std Err of Y Est		0.097
R Squared		0.96
No. of Observations		8
Degrees of Freedom		6
Growth Rate	23.1%	
X Coefficient(s)	0.207	
Std Err of Coef.	0.015	
Total Asian		
Constant		–193
Std Err of Y Est		0.041
R Squared		0.99
No. of Observations		14
Degrees of Freedom		12
Growth Rate	10.8%	
X Coefficient(s)	0.102	
Std Err of Coef.	0.002	

(1) $PPPs = purchasing power parity dollars

Table 13. Regressions to compute growth in gross domestic product in selected Asian countries

Current Currency		Constant Currency	
China (yuan)			
Constant	–261.98	Constant	–81.355
Std Err of Y Est	0.070	Std Err of Y Est	0.05
R Squared	0.98	R Squared	0.92
No. of Observations	13	No. of Observations	13
Degrees of Freedom	11	Degrees of Freedom	11
Growth Rate	14.9%	Growth Rate	4.9%
X Coefficient(s)	0.138	X Coefficient(s)	0.047
Std Err of Coef.	0.005	Std Err of Coef.	0.004
India (rupees)			
Constant	–240.633	Constant	–79.5313
Std Err of Y Est	0.035	Std Err of Y Est	0.035
R Squared	0.99	R Squared	0.97
No. of Observations	16	No. of Observations	16
Degrees of Freedom	14	Degrees of Freedom	14
Growth Rate	13.7%	Growth Rate	4.9%
X Coefficient(s)	0.128	X Coefficient(s)	0.047
Std Err of Coef.	0.001	Std Err of Coef.	0.001
Japan (yen)			
Constant	–113.18	Constant	–69.83
Std Err of Y Est	0.047	Std Err of Y Est	0.011
R Squared	0.97	R Squared	0.99
No. of Observations	16	No. of Observations	16
Degrees of Freedom	14	Degrees of Freedom	14
Growth Rate	6.6%	Growth Rate	4.2%
X Coefficient(s)	0.063	X Coefficient(s)	0.041
Std Err of Coef.	0.002	Std Err of Coef.	0.0006
Singapore (Singapore dollars)			
Constant	–6113521	Constant	–4711346
Std Err of Y Est	3199	Std Err of Y Est	2330.94
R Squared	0.96	R Squared	0.96
No. of Observations	16	No. of Observations	16
Degrees of Freedom	14	Degrees of Freedom	14
Growth Rate	10.7%	Growth Rate	7.5%
X Coefficient(s)	3100.61	X Coefficient(s)	2393.96
Std Err of Coef.	173.51	Std Err of Coef.	126.41

Table 13. Regressions to compute growth in gross domestic product in selected Asian countries

Current Currency			Constant Currency		
South Korea (won)					
Constant		–346.69	Constant		–147.176
Std Err of Y Est		0.137	Std Err of Y Est		0.048
R Squared		0.97	R Squared		0.98
No. of Observations		15	No. of Observations		15
Degrees of Freedom		13	Degrees of Freedom		13
Growth Rate	19.8%		Growth Rate	8.3%	
X Coefficient(s)	0.180		X Coefficient(s)	0.08	
Std Err of Coef.	0.008		Std Err of Coef.	0.002	
Taiwan (New Taiwanese dollars)					
Constant		–241	Constant		–159
Std Err of Y Est		0.098	Std Err of Y Est		0.076
R Squared		0.97	R Squared		0.97
No. of Observations		16	No. of Observations		16
Degrees of Freedom		14	Degrees of Freedom		14
Growth Rate	13.8%		Growth Rate	9.1%	
X Coefficient(s)	0.13		X Coefficient(s)	0.09	
Std Err of Coef.	0.005		Std Err of Coef.	0.004	
Total Asian ($PPPs [1])					
			Constant		–79.651
			Std Err of Y Est		0.033
			R Squared		0.97
			No. of Observations		13
			Degrees of Freedom		11
			Growth Rate	4.9%	
			X Coefficient(s)	0.047	
			Std Err of Coef.	0.002	
United States (dollars)					
			Constant		–42
			Std Err of Y Est		0.028
			R Squared		0.962
			No. of Observations		16
			Degrees of Freedom		14
			Growth Rate	2.9%	
			X Coefficient(s)	0.028	
			Std Err of Coef.	0.001	

[1] $PPPs = purchasing power parity dollars