ASSESSING COMPARABILITY OF AVAILABLE DATA ON CHARACTERISTICS OF THE SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS WORKFORCE

CANADA, UNITED KINGDOM, AND UNITED STATES

By **Marlene A. Lee** and **Nadwa Mossaad**

July 31, 2010

POPULATION REFERENCE BUREAU

1875 Connecticut Ave., NW
Suite 520
Washington, DC 20009-5728
USA

202 483 1100 **PHONE**
202-328 3937 **FAX**
www.prb.org **WEB**
popref@prb.org **E-MAIL**

ASSESSING COMPARABILITY OF AVAILABLE DATA ON CHARACTERISTICS OF THE SCIENCE, TECHNOLOGY, ENGINEERING, AND MATHEMATICS WORKFORCE:

CANADA, UNITED KINGDOM, AND UNITED STATES

ABOUT THE AUTHORS

Marlene Lee is a senior research associate in Domestic Programs at the Population Reference Bureau. **Nadwa Mossaad** is a research associate in Domestic Programs at the Population Reference Bureau.

ACKNOWLEDGMENTS

The authors wish to thank Carl Haub, Lisa Frehill, William Butz, Linda Jacobsen, and an anonymous reviewer for their comments. Each of these readers had knowledge in one or more areas key to our assessment of data for comparing the STEM workforce in Canada, the United Kingdom, and the United States, including STEM workforce issues, U.S. data sources, immigration data for all three countries, World Bank, OECD, and UN data sources. We would also like to express our appreciation to PRB administrative staff who assisted in preparing this report.

INTRODUCTION

Purpose

Despite efforts (OECD 1995, 2002) to harmonize data on stocks and flows of highly skilled labor in science, technology, engineering, and mathematics (STEM), these data still lack international comparability. Additional efforts at making data comparable are necessary in order for countries to better understand patterns and trends in the supply and mobility of STEM workers. Data comparability is particularly an issue in studying the factors that contribute to international movements of highly skilled STEM workers. These factors include pay, career advancement, access to better research funding, higher-quality research infrastructure, the opportunity to work with well-known scientists, and intellectual freedom.

In order to increase understanding of the current and future global supply of human resources in STEM fields, there needs to be greater cross-national comparability of data about: STEM employment, geographic mobility of STEM workers, and the cross-national flow of R&D money and talent in multinational corporations.[1] However, several limitations impede cross-national research on STEM workers: insufficient data, underutilization of existing data sources, and insufficient linkages across existing datasets (NSF 2006).

This project investigates the feasibility of obtaining comparable estimates from existing national and international data sources and the extent to which relatively standard methods, such as interpolation, may be used to formulate comparable indicators from existing data sources. We limit the scope of the project to three countries, Canada, the United Kingdom, and the United States. These countries have relatively centralized national statistical agencies whose data are generally accessible, often for free or at relatively low cost. In addition, these countries also participate in many of the same economic agreements and/or organizations, making it more likely that the data collected in each one's national interest may also already be designed to respond to international reporting requirements for multilateral agreements. For example, Canada, the United States, and Mexico have a common industrial classification system to facilitate country-to-country comparison with these partners in the North American Free Trade Agreement (NAFTA).

A list of potential indicators that helped to guide our data review is included as Appendix A. In addition to this report, the Population Reference Bureau features on its website (www.prb.org) a guide to accessible sources of data on the STEM workforce. Some information in the guide overlaps with information presented in this report, but is not identical as its focus is broader and highlights online data resources. Neither this report nor the data guide is an exhaustive compendium of data resources. Both emphasize official national data sources.

[1] Science and Technology (S&T); Scientists, Engineering, and Technology (SET); and Scientists, Technology, Engineering, and Mathematics (STEM) are broad terms usually used to capture the technical labor force. They are often used interchangeably although STEM may well be the broadest usage and, at times, STEM is defined to incorporate individuals with postsecondary degrees in the humanities. Here we use these terms to capture a broad set of workers with advanced education and training. The precise industries, occupations, and fields of study used in these definitions may need to be refined based on the objectives of any particular analysis. For example, analyses of U.S. immigration policies may want to exclude degrees in social science and humanities. (See Regets 2010 for a breakdown by occupation and further discussion of relevant definitions.) For the purposes of this report, we use STEM to encompass a broad definition of the technical labor force.

Reasons for Lack of Comparable Data

The factors that reduce data comparability can usually be grouped into three general categories: (1) the original purpose for which the data were collected; (2) factors related to the analytical methods used such as the definition of occupational or industry grouping; and (3) factors related to data collection; for example, whether data collection is part of an administrative process.

However, the first factor—the original intent for the data—drives decisions related to the other two factors, including the best or most efficient means of obtaining the desired data, the key variables of interest, and how to report the data collected. For example, education data in the United Kingdom are collected primarily at the institution-level to guide assessment of education policy and institution performance, even though data may be collected via records on individual students. This rich source of data on students exists primarily because many objectives of the higher education systems require information on individual outcomes. On the other hand, visa data collected for border control require only limited information on individuals. The data's original intent is also important with respect to maintaining continuity in data that are collected at regular intervals or in longitudinal studies. Information on new industries, occupations, and fields of study may be difficult to obtain because data collection inevitably lags behind social and economic changes. This is the case for at least two reasons. First, emerging fields and trends may not apply to enough individuals or institutions to warrant a change in data collection initially. Second, there is a trade-off between the ability to do historical analysis (continuity of a series) and the ability to capture information on emerging industries or occupations.

The third category, the nature of the data source, is another important contributor to data comparability. Data may be collected as part of administrative processes, such as a visa application or a university's administrative records. Surveys of the population or of a sample are another means of collecting data. Even within one country, estimates taken from two different data sources may differ. For example, estimates of the number of international students in the United States can be obtained both from official immigration data and from administrative data provided by universities. These two sources yield slightly different estimates of international students. Much of the variations in cross-national estimates may be due to differences in survey concepts, coverage, and the reference date to which the data refer (for example, last week vs. last month). Other data collection factors that substantially affect comparability of cross-national data include the geographical coverage of surveys or definition of the populations of interest, such as age limits used for measuring the economically active population. Also, the method of data collection (phone interviews, in-persons interviews, and online surveys), unit of analysis, and a respondent's ability to answer questions will affect results.

In this report, we discuss potential data sources, describe how they stack up with respect to key factors affecting comparability, and assess how these data might be made comparable. Tables 1-4 present an overview of data sources for Canada, the United Kingdom, and the United States for some key indicators and identify main comparability issues. Table 5 reviews international data sources, some of which use modeling to produce comparable cross-national estimates for Organisation for Economic Co-Operation and Development (OECD) countries and their trading partners.[2]

[2] Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey United Kingdom, and United States

DATA SOURCES FOR SCIENCE, ENGINEERING AND TECHNOLOGY WORKFORCE CHARACTERISTICS

Human resources play a critical role in economic and technological development, particularly in knowledge-based societies. For this reason, understanding the demand for and availability of STEM workers has become increasingly important. Information about the mobility of this workforce has also become important as globalization makes it possible for more knowledge-based work to be conducted virtually anywhere. Data on the characteristics of STEM workers are available from a variety of national surveys and administrative data including labor force surveys, population censuses or annual population surveys, and official immigration data. Labor force surveys and population surveys allow the examination of more variables than administrative data or employer/establishment surveys.

Each of these sources has limitations. While a census usually provides information on a sufficiently large number of people to allow more detailed examination of employment by industry and occupation, the data are usually collected less frequently. Sample sizes for labor force surveys are smaller, but multiple surveys may be pooled to allow for analysis at a greater level of detail. (See Table 1 for a list of some key workforce characteristics of interest.)

The main obstacles to data comparability are differences across countries in the classification of education, industry, and occupations, despite internationally harmonized standards: the International Standard Industrial Classification (ISIC), the International Standard Classification of Education (ISCED), and the International Standard Classification of Occupation (ISCO). When applied, these international classification standards make data generally comparable, but because education systems and occupations differ substantially across countries, differences still remain (UNESCO 1997; UNSTATS; ILO). These standards are also not particularly useful in comparing detailed levels of occupations. Also, even international standards have undergone changes over time, requiring care in the construction of time series.

Eurostat, UNESCO, and OECD essentially limit comparisons across countries and over time to the highest levels of education—tertiary. The ISCO-88 groupings recommended for examining innovation in the *Canberra Manual* are professionals (group 2) and technicians and associate professionals (group 3). These occupations usually require the equivalent of a formal tertiary education—a university degree or rigorous postsecondary professional training.

Canada

The Census of Population conducted every five years, provides data on the foreign-born and native-born workforce, their share of the overall workforce and the STEM workforce, their distribution by occupation and industry, and their average wages. However, confidentiality requires suppression of more detailed occupations.

The Labor Force Survey (LFS) is a monthly household survey of a sample of individuals who are representative of the civilian, noninstitutionalized population ages 15 and older. The survey excludes about 2 percent of this population: persons living on reserves and other Aboriginal settlements in the provinces, full-time members of the Canadian Armed Forces, and the institutionalized population. The LFS allows identification of the country of birth of the respondent, whether or not the respondent was a “landed immigrant,” the month and year he/she became a landed immigrant, and the country where the respondent received his/her highest level of education—making the data on immigrants comparable to those in the Census. The LFS also provides employment estimates by industry, occupation, hours worked, and much more, all cross-classified by a variety of personal characteristics. Information on industry and occupation, and reason for leaving last job are also available for persons currently unemployed or not in the labor force but with recent labor force involvement. The LFS follows a rotating panel sample design, in which households remain in the sample for six consecutive months. This results in a five-sixths month-to-month sample overlap, which makes the design efficient for estimating month-to-month

changes. Data collection for the LFS is carried out each month during the week following the LFS reference week. The reference week is normally the week containing the 15th day of the month.

Citizenship and Immigration Canada (CIC) provides data on immigrants granted permanent and temporary residency. The main annual publication (*Facts and Figures: Immigration Overview: Permanent and Temporary Residents*) presents statistical information derived from CIC's administrative data files, using the individual (not visas) as the unit of analysis. Data on permanent residents cover the three main categories of immigration—family class, economic immigrants, and refugees—as well as data for other immigrants who do not qualify in any of these categories. Statistics for temporary residents cover temporary foreign workers, foreign students, humanitarian (including refugee claimants), and other temporary residents. Selected sociodemographic characteristics are presented for all permanent residents and include information on immigrant category, country of origin, intended destination in Canada, age at landing, level of education, marital status, language ability, occupational skill level, and intention to work. Information pertaining to marital status, level of education, occupational skill level, and intention to work is presented only for the permanent resident population 15 years of age or older. In addition, supplementary tables for the permanent resident population show the number of years of schooling for children younger than 15, the major source countries, the level of education for new workers, and the occupational skill level for skilled workers. The reference period for the data is the calendar year.

United Kingdom

Because the U.K. Census questionnaire is completed by the householder, it is kept as simple as possible, thus limiting the labor market information that can be collected. Consequently, the economic activity questions do not allow categories to be as accurately defined as in the U.K. Labour Force Survey (LFS), but they are as close as possible to the International Labour Organization (ILO) definitions used in the LFS.

The Annual Population Survey (APS) is a combined survey of households in Great Britain. Its purpose is to provide information on key socioeconomic variables between the decennial censuses. The APS is comprised of the Quarterly Labour Force Survey (LFS) (waves one and five for the United Kingdom funded by the Office for National Statistics), plus data from the Annual Local (Area) Labour Force Survey (LLFS) Boosts for England, Scotland, and Wales. A further supplement, the APS Boost, was collected from January 2004 to December 2005 for England only and was funded by ONS. The APS Boost was discontinued in January 2006 due to withdrawal of ONS funding so is only present in the five datasets released between June 2005 and June 2006.

The main dissemination route for labor market data from the APS is Nomis. Other data from the APS, for example, on housing, education, employment, ethnicity, and health are available from the Neighbourhood Statistics and National Statistics websites. Datasets are also available at the U.K. Data Archive and from the LFS Customer Service (SOS) team in ONS. Annual Population Survey datasets are produced quarterly with each dataset containing 12 months of data. There are approximately 170,000 households and 360,000 persons per dataset.

The Labour Force Survey, conducted every three months, is based on a random sample throughout the United Kingdom. The survey collects information about the personal circumstances and work of everyone living in these households. Labor information is collected on members of the household ages 16 and older. The reference period is the week before the interview. The latest three months' average data are published each month. For comparisons over time, the U.K. Office of National Statistics recommends that, for seasonally adjusted data, the preceding nonoverlapping quarter should be used. For nonseasonally adjusted data, comparisons should be made with the same period a year ago.

The Long-Term International Migration (MN) series tables provide the most inclusive information on flows of international long-term migrants to and from the United Kingdom These data are derived from the International Passenger Survey (IPS); Home Office data; estimates of flows between the United Kingdom and Irish Republic from Central Statistics Office, Dublin (up to 2007); and estimates of flows to and from Northern Ireland from the Northern Ireland Research and Statistics Agency (2008 onward). Adjustments are made to reflect people who change plans for their length of stay. Long-Term International Migration (LTIM) tables include estimates by citizenship; by country of last or next residence, by country of birth; by main reason for migration; by usual occupation; and by area of destination or origin within the United Kingdom, length of stay, age, sex, and marital status. For the years 1991 to 2008, there are also cross-tabulations of some variables derived from IPS data only.

The IPS data used in the LTIM series is drawn from a survey that was primarily intended to provide information on international tourism and to compile the travel account of the U.K. balance of payments. Passengers on all major routes into and out of the United Kingdom are included in the survey, accounting for roughly 90 per cent of all travelers entering and leaving the United Kingdom. Stratification of the sample ensures that it is representative by mode of travel, route and time of day. Interviews are conducted throughout the year. The IPS provides data on the flow of people into the United Kingdom, including those from who may enter without a visa.

The U.K. Home Office is responsible for the control of its borders, including granting visa requests. Reports from this agency use administrative data on visa applications and grants of settlement (permanent residency). Because citizens of most countries in Europe have the right to enter and live in the United Kingdom without a visa, these people would not be represented in administrative data drawn from visa applications. Nationals of countries in the European Economic Area and Switzerland also have the right to work in the United Kingdom without applying for permission from the U.K. Border Agency. Administrative data from visas granted also have a drawback because individuals may change their minds about entering the United Kingdom after they have been granted a visa. The types of visas granted provide information on skill levels of immigrants.

United States

The U.S. Decennial Census after 2000 does not include labor force questions. Currently, the American Community Survey (ACS) is an ongoing survey sent to a sample of U.S. housing units every month. It includes socioeconomic questions that used to be asked in the Census, including questions about place of birth, citizenship, employment, occupation, and industry. In 2009, a question on field of degree was added. The ACS publishes single-year data at the national level for the United States and for all areas with populations of 65,000 or more. The reference period for employment activity is the week prior to the survey.

The Current Population Survey (CPS) is a monthly survey of households conducted by the U.S. Census Bureau for the Bureau of Labor Statistics. It provides a comprehensive body of data on labor force, employment, unemployment, persons not in the labor force, hours of work, earnings, and other demographic and labor force characteristics. Basic labor force data are gathered monthly; data on special topics are gathered in periodic supplements. The survey covers the civilian noninstitutional population ages 16 and older. The reference period for employment data is the calendar week (Sunday through Saturday), which includes the 12th day of the month.

The Department of Homeland Security *Yearbook of Immigration Statistics* provides data tables on foreign nationals who were, during the fiscal year, either: (1) granted lawful permanent residence (for example, admitted as immigrants or became legal permanent residents); (2) admitted into the United States on a temporary basis (such as tourists, students, or workers); (3) applied for asylum or refugee status; or (4) were naturalized. Tables on nonimmigrants (temporary workers on HB-1 visas; foreign students on F, J, and M visas) present admission data, not person-level data, so an individual who traveled more than once in the year is counted multiple times.

Permanent residence tables, in contrast, show the number of people who received green cards in a given fiscal year.

The U.S. State Department publishes the *Report of the Visa Office* with tables on the number of visas issued by country of nationality and visa type. The table shows the number of visas issued per fiscal year by type issued to nationals of various countries. This does not provide an exact measure of the inflow because some people might postpone or cancel their entry to the United States after they get their visas. The table shows the number of visas by type issued to nationals of various countries. No further information (such as reasons for migration, usual occupations as in the U.K. Long-Term International Migration Series) is available.

Assessment

The International Labor Office (ILO) notes that labor force surveys have a coherent framework and are a comprehensive data source for measuring the economically active population. Accordingly, the labor force surveys in Canada and the United Kingdom and the Current Population Survey in United States can provide data on many of the characteristics of the STEM workforce as a whole at the national level, including characteristics of the foreign-born. However, the sample sizes of these surveys do not allow analysis of refined industry-occupation categories. Even pooling surveys may not provide sufficient sample size to study emerging occupations or industries. These surveys have national coverage of all economic activities and similar reference groups and reference periods for employment information. The frequency of the surveys varies, but methods similar to the ILO's for producing comparable employment and unemployment annual estimates might be applied after ensuring that there is similar treatment of self-employed workers, unpaid family workers, and immigrant status and after grouping industry and occupations similarly. (See ILO averaging and interpolation methods in Appendix B.) Canada and the United States share a similar industry classification scheme, the North American Industry Classification System (NAICS); and the U.K. system of classification is similar to systems used by Eurostat and the UN ISIC. The occupational classification system used in the United Kingdom is more difficult to map to the ISCO than are the systems used in Canada or the United States.

All three countries produce similar administrative data on permanent immigration. However, the definitions of immigrants vary slightly. Data on permission to settle permanently in the United Kingdom ("grants of settlement") exclude nationals of European Economic Area (EEA) and Switzerland. Under the right of residence under the European law, nationals of EEA and Switzerland do not need to apply to enter, live, or work in the United Kingdom (unless they explicitly choose to do so); thus no data are collected on them. Their family members, if they are not nationals of the EEA/Switzerland, do need to apply for permanent residence status, and would appear in the "grants of settlement" tables.

DATA ON EDUCATION AND PRODUCTION OF STEM WORKERS

Workforce education and training are central to sustainable economic growth and higher paid jobs, particularly in the global knowledge-based economy. Data on the development and mobility of STEM human resources and on the career paths of doctorate holders contribute to understanding the national and international labor market for these workers. These data may also help identify policy measures to increase the supply of STEM workers, in particular the participation of women and other underrepresented groups in science and science-related careers.

OECD has concluded that education statistics are the most robust source of data for analyzing international mobility and tracking doctorate holders. Although registers and/or administrative data include many indicators, these sources are available in only a few countries and do not always have the individual as the unit of analysis. Labor force surveys are more widely available and for many countries follow established rules for collecting internationally comparable data across key labor force indicators, but sample sizes in these surveys limit cross-classification of variables at a very detailed level. In many countries, censuses are the richest source of data but are administered infrequently, and the data may be outdated. Education statistics seem to be the most robust source of data, but they present only the availability to the labor market, not actual engagement in the labor market. (See Table 2 for a list of key indicators for the education and production of STEM workers.)

At the macro level, for cross-national comparison, the OECD recommends using seven broad fields of study in the definition of STEM education: natural sciences, engineering and technology, medical sciences, agricultural sciences, social sciences, humanities, and other fields. Data on the highest level of education in these fields correspond to ISCED-97 levels (OECD 2007):

5a: Theory-based programs designed to provide sufficient qualifications for entry to advanced research programs and professions with high skill requirements, such as medicine, dentistry or architecture.

5b: Typically shorter than those of in 5a, focusing on practical, technical, or occupational skills for direct entry into the labor market and have a minimum duration of two years full-time equivalent at the tertiary level.

6: Programs that lead directly to the award of an advanced research qualification, such as a Ph.D. The theoretical duration of these programs is three years full-time in most countries (for a cumulative total of at least seven years full-time at the tertiary level).

Canada

Data on qualifications continue to be available from the annual Labor Force Survey data, as mentioned above. Some data on international students are also available in the immigration publications mentioned in the previous section. In this section, we focus on education statistics.

Adult Education and Training Survey (AETS) is Canada's most comprehensive source of data on individual participation in formal adult education and training. The survey has been administered intermittently as a supplement to the Labor Force Survey in 1992, 1994, 1998, and 2003. The survey collects detailed information about the skill development efforts of the entire adult Canadian population for the calendar year prior to the survey year. It also provides information about the main subject of training activities, their provider, duration, and the sources and types of support for training. In addition, the AETS includes socioeconomic and demographic profiles of both training participants and nonparticipants.

The National Graduates Survey was designed to answer questions about the postgraduation employment experience of graduates of postsecondary programs including: whether they became

employed, the relationship between their programs of study and subsequent employment, job and career satisfaction, rates of underemployment and unemployment, and the influence of postsecondary education on occupational achievement. The survey covers graduates of public postsecondary institutions and excludes graduates of private postsecondary and graduates from postsecondary vocational programs lasting less than three months. Only residents of Canada at the time of the survey are included. The reference period for the survey is the calendar year. The survey has some undercoverage for graduates of colleges in some provinces, notably Ontario and Alberta. No weighting adjustment has been made to correct for this undercoverage.

The Survey of Earned Doctorates is an annual census of doctorate recipients in Canada. It collects data about graduates' postsecondary academic path, funding sources, field of study, and their immediate postgraduate plans. The survey also includes information on graduates' educational history and socioeconomic background. The survey covers all degree-granting postsecondary institutions in the reference academic year but excludes those not awarding doctoral degrees in the reference year.

The Survey of Colleges and Institutions is a census that has been used to update the annual OECD data request and the Pan-Canadian Education Indicators Program (PCEIP), but is currently inactive. The survey collects full-time aggregate public college and institute enrollment and graduate data by age, sex, and field of degree/certificate/diploma. The survey covers publicly funded colleges and institutes only, and excludes universities. The institutions covered offer certificate, diploma, and transfer or continuing education programs, as well as professional development programs. The response rate for this survey is 100 percent. The reference period is the academic year—August 1 of the previous year to July 31 of the current year.

Post-Secondary School Information System (PSIS) is an information system based on a national survey that when fully implemented will replace other enrollment and graduate surveys such as University Student Information System (USIS), Community College Student Information System (CCSIS), and Trade Vocational Student Survey (TVOC). The PSIS is intended to inventory all Canadian postsecondary education institutions and the programs and courses they offer. The survey covers all postsecondary education institutions, collecting detailed information about the programs and courses offered, as well as about the students themselves and the program(s) and courses in which they were registered or from which they have graduated. The reference period for the survey is the academic year—start and end dates vary according to whether the institution has classes over the summer. Student data include demographic and other descriptive information such as birth date, gender, previous education, visible minority status, and activity limitation.

United Kingdom

Data on educational achievement are from the annual Labour Force Survey/Annual Population Survey as mentioned above. Some data on international students are also available in published migration data mentioned above in previous sections of this report. In this section, we focus on education statistics.

The Higher Education Statistics Agency (HESA) publishes a set of reference volumes that cover higher education students, qualifiers, destinations of leavers, and staff and finances in higher education institutions. There are comprehensive student data built from student records completed by higher education institutions. These data include student domicile or usual residence, level and mode of study, subjects of study, qualifications obtained, and ethnic minorities. In addition, student data are aggregated to provide performance indicators in higher education. Among the indicators reported are measures of participation for underrepresented groups, as specified by socioeconomic status and neighborhoods. Online tables make some survey results available to the public. The Higher Education Information Database for Institutions (Heidi) provides data access to registered institutions. Data are collected on student enrollments at each publicly funded higher education institution in the United Kingdom, using the academic

year as the reference period. The online database also includes data collected by agencies other than HESA.

The Destinations of Leavers from Higher Education (DLHE) cross-section survey contains all students who usually reside in the United Kingdom or the European Union and who are reported to HESA as obtaining relevant qualifications and whose study was full-time or part-time in the academic year (between August 1 and July 31). The survey is conducted approximately six months after students leave their higher education institution. Relevant qualifications include doctorate and masters degrees, and other postgraduate qualifications obtained primarily through supervised research. The longitudinal DLHE follows a sample of those who participated in the cross-sectional survey. The follow-up survey is conducted approximately three years after the first interview. The sample is intentionally skewed toward foundation degree leavers, those who completed a master's or doctoral degree, and nonwhite leavers. Topics covered in the follow-up include employment, additional study, engagement in other activities, and satisfaction with their degree and career.

United States

The Integrated Postsecondary Education Data System (IPEDS) is the core program that the U.S. Department of Education's National Center for Education Statistics (NCES) uses for collecting data on postsecondary education. IPEDS is a single, comprehensive system of annual interrelated surveys that encompasses all identified institutions whose primary purpose is to provide postsecondary education. IPEDS gathers information from every college, university, and technical and vocational institution that participates in the federal student financial aid programs. IPEDS consists of institution-level data that can be used to describe trends in postsecondary education at the institution, state, and national levels. The data include enrollments of undergraduates, first-time freshmen, and graduate and first-professional students by completions (awards) by type of program, level of award, race and ethnicity, sex, and state of residence for first-time students. Before IPEDS, some of the same information was collected through the Higher Education General Information Survey (HEGIS).

Student and Exchange Visitor Information System (SEVIS), a database for the U.S. Immigration and Customs Enforcement, Student Exchange and Visitor Program, tracks schools and programs, students, and exchange visitors and their dependents throughout the duration of approved participation within the U.S. education system. SEVIS counts reflect a moment in time and include secondary schools, vocational schools, and other institutions outside of higher education. In addition, it includes the 13 categories of the Exchange Visitor Program: student (university/college and secondary); physician; au pair; camp counselor; summer work/travel; trainee; government visitor; international visitor; professor; research scholar; short-term scholar; specialist; and teacher.

Institute of International Education (IIE) Annual Census of International Students in the United States is another source of data on international students. IIE defines an international student as an individual who is enrolled for courses at a higher education institution in the United States on a temporary visa, and who is not an immigrant (permanent resident with an I-151 or "Green Card"), a citizen, an illegal alien (undocumented immigrant), or a refugee. This voluntary survey was administered in 2007-2008 to 2,657 institutions, excluding closed institutions and long-term nonrespondents. Results are similar to those of SEVIS, a mandatory survey.

The National Survey of College Graduates (NSCG) provides data on the number and characteristics of individuals with education and/or employment in science and engineering or related fields in the United States. The survey is conducted every 10 years after the U.S. Decennial Census. The sample frame for the 2003 survey was drawn from the 2000 Decennial-Census long-form responses. Individuals who responded to the long form were eligible for the sampling frame if, as of April 2000, they were living in a housing unit or noninstitutionalized group quarters; had received a bachelor's degree or higher; resided in the 50 states, the District of

Columbia, Puerto Rico, or the other outlying U.S. territories; and age 75 or younger as of the 2003 survey reference date of Oct. 1. The NSCG has been directly affected by the decision to eliminate the long form questions from the Decennial Census. With successful implementation of the ACS, the Census Bureau has agreed that the National Science Foundation may use the ACS as a sample frame for the NSCG in the future.

The National Survey of Recent College Graduates (NSRCG) provides information about individuals who recently obtained bachelor's or master's degrees in a science, engineering, or health field. This survey is also a component of the Scientists and Engineers Statistical Data System (see below). The NSRCG respondents are individuals under age 76 who received bachelor's or master's degrees in a science, engineering, or health field from a U.S. institution in the prior three years and who were living in the U.S. during the survey reference week (April 1 for the 2006 survey). Under an agreement with the NSF Division of Science Resources Statistics, the Census Bureau collects the survey data from graduates of the institutions sampled for this study. Data are currently available for 2001, 2003, and 2006.

The Survey of Earned Doctorates (SED) collects data continuously on the number and characteristics of individuals receiving research doctoral degrees from all accredited U.S. institutions in an academic year (July 1 to June 30). The SED collects information on the individual's education, characteristics (including race/ethnicity), and postgraduation plans. The SED collects a complete college education history. Because one-third of doctorate recipients from U.S. universities are citizens of foreign countries, the U.S. Department of Education developed a coding manual for foreign institutions of higher education (Hunt 1996). This coding system is used to code the baccalaureate and/or master's degree origins of U.S. doctorate recipients who earned earlier degrees in foreign countries.

The Survey of Doctorate Recipients (SDR) gathers information from individuals who have obtained a doctoral degree in a science, engineering or health field. The SDR is a longitudinal survey administered approximately every two years and follows recipients of research doctorates from U.S. institutions until age 76. The sample for this survey is drawn from the same sampling frame as the SED above. Key variables collected from this survey include information on citizenship, country of origin, education history, race/ethnicity, and employment.

The Scientists and Engineers Statistical Data System combine data from NSF surveys of scientists and engineers—the National Survey of College Graduates, National Survey of Recent College Graduates, and the Survey of Doctoral Recipients. The three surveys share the same reference date and nearly identical questionnaires.

Assessment

Official immigration statistics and surveys of postsecondary students offer the most comparable coverage of international students across Canada, the United Kingdom, and the United States. However, official immigration statistics are not comparable in identifying level of study of international students (such as enrollment in postsecondary institutions). College enrollment by program level and highly aggregated field of study is comparable in the surveys of postsecondary institutions. More general measures of education face comparability issues discussed above under characteristics of STEM workers, as these data would rely on the data discussed above. Finally, existing international resources such as *Eurostat* and *OECD Education at a Glance* provide comparable data on education. However, more specialized analysis of existing data available in individual countries might yield more detail relevant to assessment of trends in the production and mobility of scientists and engineers.

DATA ON MULTINATIONAL CORPORATIONS

Multinational companies are among the most important influences on employment practices in Canada, the United Kingdom, and the United States. Multinationals also exert a great deal of influence on policymaking because of their influence on national economies. In the United Kingdom and the United States, multinationals are among the leading sources of inward and outward flows of foreign direct investment (Edwards et al. 2007). Their demand for international labor also exerts influence over rule setting for visas.

International guidelines for the compilation of balance of payments and international investment position statistics appear in the International Monetary Fund's Balance of Payments Manual and the OECD's Benchmark Definition of Foreign Direct Investment (FDI). This body of recommendations provides comprehensive and detailed international standards for recording both positions and flows related to FDI. The United Nations World Investment Report provides summary profiles of FDI flows and of the operations of multinational/transnational corporations, using information supplied or published by individual countries. Despite international guidelines, the coverage of published data varies widely across countries. (See Table 3 for indicators relevant to the effects of multinationals and foreign direct investment.)

Canada

Statistics Canada is in the process of setting up statistics on foreign affiliates' trade in services (FATS). Information on the sales and employment of nonbank majority-owned affiliates of Canadian transnational corporations (TNCs) abroad were published, for the first time, in April 2002, for the reference year 1999. Data by geographical destination refer to the immediate host country and region. Variables are attributed entirely to a single country or region and not factored down by ownership shares. Data by industry and sector represent the economic activity of the foreign affiliate according to the NAICS classification. Data on the activities of Canadian TNCs, their foreign affiliates, and foreign affiliates in Canada have been compiled by Statistics Canada for the purpose of business register statistics, but these data are not yet available.

United Kingdom

Data on the operations of TNCs and their foreign affiliates are collected by the Office of National Statistics (ONS). Only inward data on majority-owned foreign affiliates in industrial production are available. These data are published in the annual *Business Monitor PA 1002*: *Production and Construction Inquiry* and are reported with a breakdown of geographic origin (according to the ultimate beneficial owner) as well as with a breakdown by economic activity of the foreign affiliate. The ONS also collects data on R&D expenditure and on employment in R&D through an annual survey that covers majority foreign-owned enterprises in agriculture and industrial activities in the United Kingdom.

The Production and Construction Inquiry is part of the Annual Business Inquiry (ABI). ABI estimates cover all U.K. businesses registered for Value Added Tax (VAT) and/or Pay As You Earn (PAYE). The ABI obtains details on these businesses from the ONS Inter Departmental Business Register (IDBR), which consists of companies, partnerships, sole proprietorships, public authorities, central government departments, local authorities, and nonprofit bodies. Each legal unit is classified to a single activity, whether it is wholly or mainly engaged in that activity.

The nature of a business can change with time, possibly because another business has been absorbed in a takeover. Some businesses have a very significant secondary activity, perhaps completely different from their main activity, and a small change of direction can lead to a new main activity and to the reclassification of the business. Other changes may arise from improvements to classification data held on the IDBR as a result of new information received about individual businesses. The reclassification of businesses for reasons such as these contributes to the annual changes in the figures published. Northern Ireland and Scotland are

sampled and estimated separately, while England and Wales are sampled separately but are combined for the estimation procedure.

United States

U.S. data on foreign direct investment (FDI) are prepared by the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce. The BEA collects FDI flow data in the balance-of-payments framework. Data describe the operations of transnational corporations (multinationals) and their foreign affiliates, as well as foreign affiliates in the United States.

Benchmark surveys conducted every five years, complemented by the annual surveys that cover a sample of companies (used to produce a universe estimate), are the source of these data. The results are published in the annual *Foreign Direct Investment in the United States: Operations of U.S. Affiliates of Foreign Companies* for inward data and *U.S. Direct Investment Abroad: Operations of U.S. Parent Companies and Their Foreign Affiliates* on outward data. The survey on FDI in the United States is conducted two years before the survey on United States FDI abroad. Quarterly sample surveys update the results of the benchmark surveys.

Assessment

Differences in coverage and rules for attributing geographic origin to foreign-owned multinationals operating in Canada, the United Kingdom, and the United States render little of the data published with geographic attribution comparable. In addition, the industry classification based on primary activity is not strictly adhered to and can be an unstable classification over time, even within one country. However, broad measures that are considered determinants of a nation's share of multinationals—real GDP growth, growth in labor productivity, and investments of multinationals as a percent of the difference between inward and outward foreign investment flows—are internally consistent relative measures and comparable across countries.

DATA ON RESEARCH AND DEVELOPMENT ACTIVITIES

The quality of opportunities to engage in research and development (R&D) activities may influence not only the extent to which young people pursue STEM careers and degrees, but also the extent to which these workers move to a particular country. Expenditures on research and development are a gauge of such opportunities in a country. Other indicators might also include research productivity (possibly measured by patents or publications) and the reputation of research institutions. Of course, these latter indicators may well be highly positively associated with levels of expenditure on R&D.

Differences exist in the extent to which data on small firms/agencies are included in national estimates of R&D expenditures and personnel. In addition, lack of comparable R&D data arises from differences in how data for diversified enterprises (companies) are allocated (OECD ANBERD methodology). Following international guidelines set in the *Frascati Manual* (OECD 2002), an enterprise's research and development are supposed to be allocated to the industrial class of the primary activity. However, some countries disaggregate the activities of large diversified firms, allocating their R&D across a number of industry classifications. For some organizations, there is ambiguity as to whether their R&D should be included as business enterprise activity, government, or nonprofit. Finally, in some countries, some industrial activities are concentrated in a few firms. For reasons of confidentiality, data for such activities are suppressed or combined with that of another industry. These problems are all compounded in time series data because decisions related to coverage and handling of data suppression may vary from year to year and because R&D surveys are not conducted every year in some countries. (See Table 4 for indicators relevant to research and development activities.)

The OECD ANBERD database was developed to provide analysts with a comprehensive and internationally comparable data set on industrial R&D expenditures. ANBERD is based on official business enterprise R&D data but substitutes estimates if there are: (1) significant problems associated with the enterprise basis of the survey or with borderline institutions; (2) important deviations from the standard BERD industrial classification; (3) significant adjustments required due to incomplete survey coverage; (4) discontinuities or breaks in series due to a change in industrial classification or survey techniques; or (5) missing data for entire years.

In addition to surveys of business R&D activities, there are also surveys estimating expenditures and employment in R&D for government, higher education, and nonprofit organizations. International data are also available on patents (OECD Patent Database), but differences in the patenting process in each country make it difficult to assess the comparability of these data. Although rankings of universities across countries also exist, differences in education systems make it difficult to develop comparable performance-based indicators across countries. The *U.S. News and World Report* ranking of world universities, for example, relies heavily on surveys of employers and academic peers.

Canada

Research and Development in Canadian Industry (RDCI) includes data collected directly from survey respondents and extracted from administrative files. An annual mail survey is administered to all companies in Canadian industries known to be performing or funding more than $1.5 million in research and development. The data for the small performers and funders are taken directly from the Canada Revenue Agency (CRA). Although a complete enumeration is carried out of known and suspected R&D performers and funders, records received from the administrative data do not provide as much information as responses to the mail survey. Based on patterns found in the survey responses, information missing from administrative records is imputed by industry.

United Kingdom

As of 1993, the ONS has conducted an annual small-scale, stand-alone sample survey based on a continually updated register of R&D performers. Estimates are made for the R&D activity of unsampled and nonresponding businesses. This excludes government organizations, higher education establishments, and charities. Government R&D data are obtained from an annual census of all government departments, excluding local authorities and a small area of central government. Higher education funding councils for England, Scotland, Wales, and the Department of Education for Northern Ireland provide estimates of higher education R&D. Private Non-Profit R&D estimates are currently based on a number of sources, including the total reported as other spending by government, including Research Councils.

United States

The U.S. Census Bureau conducts an annual survey of approximately 25,000 companies selected from the Business Register (BR), Survey of Industrial Research and Development (SIRD). All companies with $5 million or more in R&D and more than five employees are included, while other companies are assigned probabilities of selection. Companies with five or fewer employees in the prior year are not included. This survey will be replaced by the Business R&D and Innovation Survey (BRDIS), an annual survey of companies with more than five employees. Data on U.S. government federal expenditures on R&D are collected by the Survey of Federal Funds for Research and Development for the federal fiscal year. Data on R&D expenditures in higher education are collected by Survey of Research and Development Expenditures at Universities and Colleges. The Survey of Research and Development Funding and Performance by Nonprofit Organizations is no longer active. The most recent version (1996 and 1997) collected information on the STEM research and development activities of nonprofit organizations. Surveys of R&D activities are listed at www.nsf.gov/statistics/survey.cfm.

Assessment

Coverage of R&D activities varies between the United States, Canada, and the United Kingdom. However, all three countries estimate that their surveys capture 90 percent or more of all business enterprise R&D activity and 100 percent of all federal government spending on R&D. The surveys cover the fiscal year, which may differ across businesses and governments. For example, the U.S. federal government fiscal year runs from October 1 to September 30, Canada's from April 1 to March 31, and the United Kingdom's from April 6 to April 5. For cross-sectional analysis, estimates of business and government R&D personnel and expenditures are comparable, with adjustments for differences in the reference period necessary for some types of analyses. Combining data from two consecutive years to estimate a weighted average for the same reference period would be one means of reconciling the different time periods represented by the fiscal year. This method provides smoothing where there is reason to expect that fluctuations would have been captured by more frequent measurement. Similar comparability may be expected for OECD countries that follow the *Frascati Manual* suggestions.

CONCLUSION AND RECOMMENDATIONS

Although most data available for all three countries cannot provide comparable detailed information by industry or occupation, there is comparable information (or it can be derived) for broad STEM industry and occupation groupings, high-skilled workers, production of STEM workers, immigrants, and the flow of direct investments that would be useful in formulating policy and assessing the relative strength of broad initiatives. More information on the stock of science and technology workers by detailed industry, occupation, or field of degree might be useful to employers, although sample sizes are small.

Education statistics are the most robust source of data for analyzing international mobility and tracking doctorate holders but only represent availability to the labor market. Not included in these statistics are persons who have worked in STEM occupations and associated industries in the labor market previously. Labor force surveys are widely available and for many countries follow established rules for collecting internationally comparable data across key labor force indicators, but sample sizes in these surveys limit cross-classification of variables at a very detailed level. Table 5 and Appendix B provide information on cross-national databases as well as examples of estimation methods used in achieving comparability of indicators across countries. Censuses are the richest source of data, but in these countries are administered infrequently, and the data may be outdated. Data on multinationals and R&D activities have more limited availability and comparability. Model-based databases developed by OECD have great promise for researchers but may not provide details specific to particular research questions.

Continued documentation of national data resources can make an important contribution both to informing researchers and to promoting nongovernment use of these resources. Such uses could in turn identify new measures for cross-national comparison among groups of countries competing for STEM labor. A natural expansion of the current study would be an extension to competing English-speaking countries such as Australia and New Zealand.

REFERENCES

Tony Edwards et al., "Charting the Contours of Multinationals in Britain: Methodological Challenges Arising in Survey-based Research," *Warwick Papers in Industrial Relations* 86 (2007).

Stephen Hunt, *Mapping the World of Education: The Comparative Database System* (Washington, DC: Government Printing Office, 1996), accessed at www.nsf.gov/statistics/mapping/contents.htm#top.

International Labour Organisation (ILO), "International Standard Industrial Classification of All Economic Activities," accessed at www.ilo.org/public/english/bureau/stat/class/isic.htm

ILO Laborsta*, ILO-Comparable Annual Employment and Unemployment Estimates*, accessed at http://laborsta.ilo.org/applv8/data/iloce.pdf.

National Science Foundation (NSF), *Advancing Measures of Innovation: Knowledge Flows, Business Metrics, and Measurement Strategies, Workshop Summary Report* (2006), accessed at www.nsf.gov/statistics/workshop/innovation06/.

NSF, SRS Survey Descriptions, accessed at www.nsf.gov/statistics/survey.cfm.

Organisation for Economic Co-Operation and Development (OECD), *OECD Education at a Glance* (Paris: OECD, 2009).

OECD, *The Global Competition for Talent: Mobility of the Highly Skilled* (Paris: OECD, 2008).

OECD, *Frascati Manual 2002: Proposed Standard Practice for Surveys on Research and Experimental Development* (Paris: OECD, 2002).

OECD, *The Measurement of Human Resources Devoted to S&T*, *Canberra Manual* (Paris: OECD, 1995).

Mark Regets, "Employment in Science and Engineering Occupations Reached 5.8 Million in 2008," *Infobrief SRS, National Science Foundation Directorate for Social Science, Behavioral and Economic Sciences* (April 2010), accessed at www.nsf.gov/statistics/infbrief/nsf10315/nsf10315.pdf.

UNESCO, *International Standard Classification of Education: ISCED 1997* (Geneva: UNESCO, 1997), accessed at www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm.

UNSTATS, ISIC Rev. 3, accessed at http://unstats.un.org/unsd/cr/registry/regcst.asp?Cl=2.

Table 1. Workforce Characteristics				
Indicators	*Comparability Notes*	*Availability (Source)*		
A. Foreign-born Workers		U.S.	CA	U.K.
Distribution by visa status (temporary, guest worker, student)	Broad distinctions between permanent and temporary residents are available for all three countries. The U.S. data provide this information based on immigration status at the point of entry and not in census or population-based survey data. Canada and United Kingdom make distinctions in survey data between those with permanent residency and temporary residency. U.K. tables present estimates with main reason for migration and usual occupation (see LTIM tables 2.04 and 2.05 or IPS 1.02) but the categories published are extremely broad, such as professional/managerial vs. manual/clerical. Canada tables include educational and occupational skill level as well as student status for temporary immigrants.	6	7	4,5
Share of total workforce	U.S. and Canadian definition of foreign-born are similar in a census or a population-based survey. Caution needs to be exercised in use of statistics about "immigrants" as definitions differ across surveys and countries.	2	1,8	3
Share of STEM workforce	See information below on industry and occupations	2	1,8	3
Distribution by STEM occupation	U.S. uses the Standard Occupation Classification System (www.bls.gov/soc/) and releases crosswalks between this classification system and systems used by other international agencies; Canada use National Occupational Classification. United Kingdom uses its own Standard Occupational Classification. It is possible to map these classifications into the broad categories of the ILO's International Standard Classification of Occupations (www.ilo.org/public/english/bureau/stat/isco/intro.htm). However, it is more difficult to do so for the United Kingdom.	2	1,8	3
Distribution by industry	U.S. and Canada use North American Industry Classification System (NAICS); the United Kingdom uses a Standard Industrial Classification (SIC) of economic activities comparable to the Eurostat NACE system at the four-digit level and to the UN ISIC system at the two-digit level (www.statistics.gov.uk/statbase/Product.asp?vlnk=14012)	2	1,8	3
Average wage and salary by STEM occupation	See information above on industry and occupations.	2	1,8	3

B. Native-born/Citizen Workers		U.S.	CA	U.K.
Share of total workforce	All three countries collect socioeconomic data on native-born and foreign-born. The U.S. and Canadian data include information on citizenship. The U.K. data include information on nationality.	2	1,8	3
Share of STEM workforce	See information under Foreign-born Workers.	2	1,8	3
Distribution by STEM occupation	See information under Foreign-born Workers.	2	1,8	3
Distribution by industry	See information under Foreign-born Workers.	2	1,8	3
Average salary by STEM occupation	See information under Foreign-born Workers.	2	1,8	3
C. Characteristics of the Science, Engineering, and Technology Workforce		U.S.	CA	U.K.
Numbers of STEM workers by occupation	Comparable occupational groups for relatively broad categories, matching of U.S. and Canadian classification systems is easier than matching with the United Kingdom, which has a more distinct national classification system and requires a crosswalk or conversion table.	2	1,8	3
Share of industry groups	U.S. and Canadian NAICS correspondence to ISIC classification is available from U.S. Census Bureau. U.K. industrial classification system does not require correspondence table for ISIC classification	2	1,8	3
Median age of STEM workers by occupation	Limited correspondence of occupations across countries. See above.	2	1,8	3
Average salary of STEM worker by occupation	Limited correspondence of occupations across countries. See above.	2	1,8	3
Share of STEM with tertiary education by occupation	Limited correspondence of occupations across countries. See above.	2	1,8	3
D. Other Indicators		U.S.	CA	U.K.
Median wages of STEM workers by occupation	Limited correspondence of occupations across countries. See above.	2	1,8	3

Notes: (1) Statistics Canada, 2006 Census of Population. Statistics Canada catalogue no. 97-564-XCB2006008.
(2) U.S. Census Bureau, Department of Commerce, American Community Survey. Tabulations for the United States as a whole may be made from Public Use Microdata. (3) Annual Population Survey, Office of National Statistics, United Kingdom. User-defined tabulations can be requested from ONS's tabulation service at lfs.dataservice@ons.gov.uk.

(4) Long-term International Migration (MN series): www.statistics.gov.uk/StatBase/Product.asp?vlnk=507&More=Y. (5) International Passenger Survey, Office of National Statistics, United Kingdom. (6) Yearbook of Immigration Statistics, U.S. Department of Homeland Security, www.dhs.gov/files/statistics/. (7) Facts and Figures: Immigration Overview-Permanent and Temporary Residents, Citizenship and Immigration Canada: www.cic.gc.ca/english/pdf/research-stats/facts2008.pdf. Additional tables available upon request. (8) Canada Labour Force Survey, http://www40.statcan.gc.ca/l01/cst01/other/lfs/lfsintro-eng.htm

Table 2. Education and Production of Students				
Indicators	*Comparability Notes*	*Availability (Source)*		
A. General Education		U.S.	CA	U.K.
STEM graduates per population ages 20-29	The ratio may be constructed from university graduation data and population estimates. This measure is included in OECD database. U.K. data on students are aligned with the International Standard Classification of Education (ISCED) (www.hesa.ac.uk/index.php?option=com_studrec&task=show_file&Itemid=233&mnl=08051&href=a^_^COURSEAIM.html). Population-based surveys do not provide these characteristics of individuals at the level of detail needed. However, special tabulations may be possible. United Kingdom provides information on qualifications of the economically active population (detail to classify as STEM does not appear) in public release). U.S. to add field of study (detail to classify as STEM may not be readily available to the public).	6,2	9,5	4,3
Share of working-age population with some college/tertiary education	Censuses and population-based surveys provide data for these estimates. Also available in OECD and Eurostat databases.	2	1	3
Share of working-age population in life-long learning	Definitions vary with policy perspective. Indicators may include literacy, numeracy, STEM students per population ages 20-29, or share of population 25-64 participating in education and training. The last is also available for the U.K. in Eurostat, Labor Force Survey. Also available in OECD Education Indicators at a Glance (www.oecd.org/document/24/0,3343,en_2649_39263238_43586328_1_1_1_1,00.html).	7, 2	10	3
Share of youth graduated from high school/with upper level credentials	Available from population-based surveys.	2	11	3
Share of adults ages 20-24 with secondary education	Available in labor force and other population-based surveys.	7, 2	8	3

B. Student Production		U.S.	CA	U.K.
Number of international students	U.S. total number of students by country of citizenship, visa status, level of education, and course of study based on administrative data from visa issuing agency as well as data from surveys of postsecondary institutions. Canadian number of international students in postsecondary institutions as well as program of study and year in program (administrative data) — country of citizenship data available in data file. U.K. student data include years in program, usual domicile, and does not include data on students completing entire program out of the country.	12,13,14	9	4
Sending countries of international students	See above.	12,13,14	9	4
International students by level of studies	See above (concordance of level of study may be possible through crosswalk to ISED but not readily available).	12,13,14	9	4
College enrollments by program level and field of study	Broad categories available in public release data for United States, Canada, and United Kingdom.	14	9	4

Notes: (1) Census, Statistics Canada. (2) American Community Survey, U.S. Census Bureau. (3) Labour Force Survey, Office of National Statistics, United Kingdom (4) Student Data, Higher Education Statistics Agency, United Kingdom (www.hesa.ac.uk/index.php?option=com_content&task=view&id=1197&Itemid=266). (5) Survey of Earned Doctorates, Statistics Canada (www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=3126&lang=en&db=imdb&adm=8&dis=2).
(6) Survey of Earned Doctorates, National Science Foundation, United States (www.nsf.gov/statistics/srvydoctorates/). (7) Current Population Survey, October Supplement, U.S. Census Bureau (www.census.gov/population/www/socdemo/school.html). (8) Labor Force Survey, Statistics Canada (www.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&catno=71F0004X). (9) Post-Secondary School Information System, Statistics Canada. (10) Adult Education and Training Survey, Statistics Canada (www.statcan.gc.ca/dli-ild/data-donnees/ftp/aets-eefa-eng.htm). (11) Annual Population Survey, Office of National Statistics, United Kingdom (12) Student/Exchange Visitors, U.S. Immigration and Customs Enforcement, DHS (www.ice.gov/sevis/numbers/student/index.htm and www.ice.gov/sevis/numbers/). (13) Institute of International Education (IIE), Census of International Students (Open Doors). (14) National Center for Education Statistics, Integrated Postsecondary Education Data System (IPEDS) (http://nces.ed.gov/ipeds/). (15) National Learning Targets (www.statistics.gov.uk/StatBase/ssdataset.asp?vlnk=5048&More=Y).

Table 3. Multinationals and Foreign-Direct Investment				
Indicators	*Comparability Notes*	*Availability (Source)*		
A. Operations Data for Majority-Owned Foreign Affiliates of Nation(*) Multinational Corporations (MNC)		U.S.	CA	U.K.
Number of home-based MNC	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,4	1,2	n.a
Number of majority-owned foreign affiliates of home-based MNC abroad	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,4	1,2	n.a
Number of employees of majority-owned foreign affiliates of home-based MNC abroad	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,4	1,2	n.a.
Number of employees by industry of majority-owned foreign affiliates of home-based MNC abroad	Canada and U.S. confidentiality rules limits categories of industry for which data are reported.	1,4	1,2	n.a.
Number of employees by world region/country of majority-owned foreign affiliates abroad	Canada and U.S. confidentiality rules limits countries for which data are reported.	1,4	1,2	n.a.
Sales of majority-owned foreign affiliates of home-based MNC abroad	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,4	1,2	n.a.
Sales by industry of majority-owned foreign affiliates of home-based MNC abroad	Canada and U.S. confidentiality rules limits categories of industry for which data are reported.	1,4	1,2	n.a.
Sales by region/country of majority-owned foreign affiliates of home-based MNC abroad	Canada and U.S. confidentiality rules limits countries for which data are reported.	1,4	1,2	n.a.

A. Operations Data for Nation(*) Affiliates of Foreign MNC		U.S.	CA	U.K.
Number of foreign MNC affiliates in nation(*)	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,5	n.a.	n.a
Number of majority-owned foreign MNC affiliates in nation(*)	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,5	n.a.	1,8
Number of employees of majority-owned foreign MNC affiliates in nation(*)	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,5	n.a.	1,8
Number of employees by industry of majority-owned foreign MNC affiliates in nation(*)	Canada and U.S. confidentiality rules limits categories of industry for which data are reported.	1,5	n.a.	1,8
Number of employees by world region/country of majority-owned foreign MNC affiliate in nation(*)	Canada and U.S. confidentiality rules limit countries for which data are reported. Attribution of ownership in United Kingdom, United States, and Canada are based on immediate parent not on ultimate beneficial owner.	1,5	n.a.	1,8
Sales of majority-owned foreign MNC affiliates in nation(*)	Canada statistics on services does not include banks. U.S. data on banks in 2007.	1,5	n.a.	1,8
Sales by industry of majority-owned foreign MNC affiliates in nation(*)	Canada and U.S. confidentiality rules limits categories of industry for which data are reported.	1,5	n.a.	1,8
Sales by region/country of majority-owned foreign MNC affiliates in nation(*)	Canada and U.S. confidentiality rules limits countries for which data are reported.	1,5	n.a.	1,8

C. Investments of Affiliates of Foreign Companies in Nation (*)-inward FDI		U.S.	CA	U.K.
Stock of foreign-owned MNC in nation(*) as percent of GDP		1	1	1
Investment flows of foreign-owned MNC in nation(*) as percent of GFCF		1	1	1
D. General Determinants of Nation Share of MNC employment		U.S.	CA	U.K.
Growth rate of real GDP per capita		6	6	6
Labor productivity growth rate		7	7	7
Net flow of MNC investments as percent of GFCF (difference between inward and outward flow)	May be estimated from available data on inward and outward foreign-direct investment.	1	1	1

Notes: (*) Substitute relevant country name. FDI=foreign-direct investment. N.A. =not available. MNC=multinational corporation, also known as transnational corporation (TNC). GDP=gross domestic product. GFCF=gross fixed capital formation. Outward FDI is economic activity of the company abroad. Inward FDI is the economic activity of foreign company in host country.
(1) UNCTAD, WID (www.unctad.org/Templates/Page.asp?intItemID=1465) and FDI/TNC database (www.unctad.org/Templates/Page.asp?intItemID=1923&lang=1). (2) Foreign Affiliate Trade Statistics, Statistics Canada (www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=1539&lang=en&db=imdb&adm=8&dis=2).
(3) Foreign and Direct Investment in Canada (www.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=61-232-XIB&lang=eng).
(4) U.S. Direct Investment Abroad Financial and Operating Data for U.S. Multinational Companies, Bureau of Economic Analysis, U.S. Department of Commerce(www.bea.gov/international/di1usdop.htm). (5) Foreign Direct Investment in the United States, Financial and Operating Data (www.bea.gov/international/di1fdiop.htm). (6) Penn World Tables (http://pwt.econ.upenn.edu/). (7) OECD Statistics (http://stats.oecd.org/Index.aspx?DatasetCode=PDYGTH). (8) Annual Business Inquiry: Production and Construction, United Kingdom (www.statistics.gov.uk/abi/2007-archive/prod_const.asp).

Table 4. General R&D, Economic Development, and Relevant Demographic Indicators				
Indicators (for preliminary country set)	*Comparability Notes*	*Availability (Source)*		
Investment in R&D		U.S.	CA	U.K.
Total gross domestic expenditure on R&D	Reported in local currency	1,2	1,2,7	1,2
Total gross domestic expenditure on R&D as % of GDP		1,2	1,2,7	1,2
Government-financed R&D (% of total Expenditures)	Reported in local currency	1,2	1,2,7	1,2
Business-financed R&D (% of total Expenditures)	Reported in local currency. United States excludes capital expenditures and includes additional classes of funds.	1,2	1,2,7	1,2
Percent of business enterprise expenditure on R&D performed in the service sector		1,2	1,2,7	1,2
Number of R&D workers	From OECD, reported as full-time equivalent and head count.	1,2	1,2,7	1,2
Number of R&D workers with Ph.D.	From OECD, reported as full-time equivalent and head count.	1,2	1,2,7	1,2
R&D workers as percent of all workers	May be estimated from R&D worker head count and total employment.	2	2	2
Knowledge Utilization/Application		U.S.	CA	U.K.
Employment in high-tech services (% of total workforce)	Industry crosswalk needed for United Kingdom with United States and Canada.	3,4	3,5	3,6
Employment in medium-high and high-tech manufacturing (% of total workforce)	Industry crosswalk needed for United Kingdom with United States and Canada.	3,4	3,5	3,6
Export of high-tech products (% of total exports)	Industry crosswalk needed for United Kingdom with United States and Canada.	3,4	3,5	3,6

Economic Indicators		U.S.	CA	U.K.
Gross domestic product (local currency)		2	2	2
GDP growth rate		2	2	2
GDP per capita (PPP) US$		2	2	2
Annual population growth rate		2	2	2
Population (thousands)		2	2	2
Total employment (thousands)		2	2	2
Industrial employment (thousands)		2	2	2
Labor force (thousands)		2	2	2
Selected Demographic, Cultural, Labor Force, Educational, and Income Characteristics		U.S.	CA	U.K.
Race, ethnicity and gender of the labor force	Racial/ethnic categories differ.	5	4	6
Income by citizenship, gender, race and ethnicity	Racial/ethnic categories differ.	5	4	6
Education by citizenship, gender, race and ethnicity	Racial/ethnic categories differ.	5	4	6

Notes: (1) UNESCO Institute for Statistics, Data Centre, Science & Technology Tables (http://stats.uis.unesco.org/unesco/ReportFolders/ReportFolders.aspx). (2) OECD, Main Science and Technology Indicators (http://www.oecd.org/document/33/0,3343,en_2649_34451_1901082_1_1_1_1,00.html). (3) Structural Analysis Database (STAN), OECD (http://www.oecd.org/document/44/0,3343,en_2649_34445_40696318_1_1_1_1,00.html). (4) Statistics Canada, 2006 Census of Population, Statistics Canada catalogue no. 97-564-XCB2006008. (5) U.S. Census Bureau, Department of Commerce, American Community Survey. Tabulations for the United States as a whole may be made from Public Use Microdata. (6) Annual Population Survey, Office of National Statistics, United Kingdom. User-defined tabulations can be requested from ONS's tabulation service at lfs.dataservice@ons.gov.uk.
(7) Research and Development in Canadian Industry (RDCI), Statistics Canada (www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4201&lang=en&db=imdb&adm=8&dis=2; www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4204&lang=en&db=IMDB&dbg=f&adm=8&dis=2; www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4208&lang=en&db=IMDB&dbg=f&adm=8&dis=2; www.statcan.gc.ca/cgi-bin/imdb/p2SV.pl?Function=getSurvey&SDDS=4212&lang=en&db=IMDB&dbg=f&adm=8&dis=2.

Table 5. Cross-National Data Sources for Human Resources in STEM

Datasets	*Country/Region*	*Key Information*
Eurostat New Cronos	Over 30	Eurostat is the Statistical Office of the European Union. New Cronos is its principal database and contains high quality macroeconomic and social statistics data. Eurostat uses data collected from national statistical institutes to produce these statistics and harmonizes them according to a single methodology. The database covers 25 European Union (EU) member states, many of the central European countries, Japan, the United States, and the main economic partners of the EU. The Science and Technology indicators include: Employment in High Technology Sectors; Human Resources in Science & Technology (HRST); Research and Development expenditure and personnel; and other measures relevant to patents, national policies, and innovation. www.esds.ac.uk/international/support/user_guides/eurostat/cronos.asp
LABORSTA, ILO	Over 200 countries	LABORSTA is a database on labor statistics operated by the International Labour Organisation (ILO), Department of Statistics. The ILO-Comparable program presents 11 basic indicators related to employment and unemployment and drawn from national labor surveys. The available estimates are working-age population; total and civilian labor force; total employment by broad industry group and by age group; civilian employment; and unemployment by age group, labor force participation rates, and unemployment rates. Estimates by occupation and industry from other programs, such as SEGREGAT and the ILO October Inquiry, are not comparable across all countries for which estimates are available. http://laborsta.ilo.org/default.html; http://laborsta.ilo.org/applv8/data/iloce.pdf
ANSKILL, OECD		OECD estimates of skills by industry database, an internationally comparable time series that is forthcoming.

ANBERD, OECD	38	The Analytical Business Enterprise Research and Development database (ANBERD) provides a comprehensive and internationally comparable time series on industrial R&D expenditures. Many of the data points are estimated, so ANBERD data may differ from data published nationally. R-3 electronic database based on the International Standard Industrial Classification (ISIC) Revision 3 includes estimates for Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States, Chile, China, Israel, Romania, the Russian Federation, Singapore, Slovenia, South Africa, and Chinese Taipei. N.B. www.oecd.org/document/17/0,3343,en_2649_34451_1822033_1_1_1_1,00.html
Education At A Glance, OECD Indicators		This publication provides comparable and up-to-date indicators on the performance of education systems: who participates in education, what is spent on it, how education systems operate and the results achieved. Indicators on results include outcomes such as students' performance in key subject areas and the impact of education on earnings and on adults' chances of employment. www.oecd.org/edu/eag2009
OECD Research and Development Statistics		This electronic database provides annual data from 1981 onward on R&D expenditure and personnel broken down by industrial sector and field of science in OECD countries. www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=SUB-17201S1
Main Science and Technology Indicators, OECD	38	A set of indicators that reflect the level and structure of the efforts undertaken by states in the field of science and technology. These data include resources devoted to research and development, patent families, technology balance of payments, and international trade in R&D-intensive industries. Also presented are the underlying economic series used to calculate these indicators. Data are available for OECD member countries and nine nonmember economies (Argentina, China, Israel, Romania, Russian Federation, Singapore, Slovenia, South Africa, and Chinese Taipei). www.oecd.org/document/26/0,3343,en_2649_33703_1901082_1_1_1_1,00.html

STructural ANalysis Database, OECD	28 (more to become available)	The STructural ANalysis (STAN) database includes annual measures of output, labor input, investment, and international trade, allowing users to construct indicators relevant to areas such as productivity growth, competitiveness, and general structural change. Use of a standard industry list based on the International Standard Industrial Classification (ISIC) of economic activities make comparisons across countries possible. High-technology sectors can be identified. Countries covered are: Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, United Kingdom, and United States. www.oecd.org/document/44/0,3343,en_2649_34445_40696318_1_1_1_1,00.html
OECD Science, Technology and R&D Statistics	38	This electronic database combines the OECD Research and Development Statistics database and the Main Science and Technology Indicators database (see above). www.oecdbookshop.org/oecd/display.asp?sf1=identifiers&st1=SUB-17201S1
United Nations Conference on Trade and Development (UNCTAD)	Varies by data product	Statistics are drawn from existing national and international data sources as well as some estimates made by the UNCTAD secretariat. Foreign Direct Investment (FDI) data are available in an online database and in electronic publications. Detailed statistics on FDI and transnational corporations (TNCs), also known as multinational corporations (MNC), for 112 countries are available in the online World Investment Directory (WID). The World Investment Report includes FDI flows and stocks for 196 countries. These two measures appear to be consistent across countries. The WID country profiles provide most relevant information on data collection in order to determine if the indicators are measured in a comparable way. National data may provide the means of making estimates comparable across some countries for which comparable data are not reported in these databases. www.unctad.org/Templates/Page.asp?intItemID=1584&lang=1

UN Common Database	Varies by data product	UN statistical databases and some OECD data are now accessible through a single entry point. http://data.un.org/
World Development Indicators, World Bank	209	This database includes 331 indicators from the World Development Indicators (WDI) from 1960 to 2008 translated into Spanish, French, and Arabic. The science and technology indicators, listed under Infrastructure-Technology in the WDI are intended to shed light on countries' technology base: research and development, scientific and technical journal articles, high-technology exports, royalty and license fees, and patents and trademarks. Sources include the UNESCO Institute for Statistics, the U.S. National Science Board, the UN Statistics Division, the International Monetary Fund, and the World Intellectual Property Organization. http://data.worldbank.org/indicator

Appendix A: Potential Useful Variables for Analysis of Stock and Flow of STEM Workforce

The following variables are not necessarily available in extant data sources but helped to guide our review of available data.

Foreign-born Workers
Distribution by visa status (temporary, guest worker, student)
Share of total workforce
Share of Science, Technology, Engineering, and Mathematics (STEM) workforce
Distribution by STEM occupation
Distribution by industry
Average salary by STEM occupation

Native-born/Citizen Workers
Share of total workforce
Share of STEM workforce
Distribution by STEM occupation
Distribution by industry
Average salary by STEM occupation

Characteristics of the STEM Workforce
Numbers of STEM workers by occupation
Share of industry groups (STEM workers in industry/all workers in industry)
Median age of STEM workers by occupation
Average salary of STEM workers by occupation
Share of STEM workers with tertiary education by occupation

General Education
STEM graduates per population ages 20-29
Share of working-age population with some college/tertiary education
Share of youth graduated from high school/with upper level secondary education ages 20-24

Investment in R&D
Public R&D expenditures as % of GDP
Private (Business) R&D expenditures as % of GDP
Share of medium high/high tech R&D in manufacturing
Number of R&D workers
Number of R&D workers with Ph.D.

Knowledge Utilization/Application
Employment in high-tech services (% of total workforce)
Employment in medium-high and high-tech manufacturing (% of total workforce)
Export of high-tech products (% of total exports)

Multinational Corporations (MNC) Operations Data
U.S. parent companies
- Number of employees in the U.S.
- Sales ($)
- Capital spending ($)

U.S. affiliates of foreign companies (by country where possible)
- Number of employees in the U.S.
- Sales ($)
- Capital spending ($)

Foreign affiliates of U.S. companies (by country where possible)
- Number of employees outside the U.S.

Sales ($)
Capital spending ($)

Determinants of Change in U.S. share of U.S.-MNC employment
Differences in rate of economic growth, U.S. and abroad
Differences in rate of productivity growth, U.S. parent and foreign affiliates
Total amount U.S. parent companies outsource to foreign affiliated or foreign nonaffiliated

Production of Students Relative to Other Countries
U.S. share of worldwide 1st four-year degrees produced in 2004
G7 share of worldwide 1st four-year degrees produced in 2004

APPENDIX B: ILO METHODS FOR AVERAGING AND INTERPOLATION FOR COMPARABLE EMPLOYMENT STATISTICS

Methods Used to Calculate Adjusted Annual Averages

ILO #1
Method for calculating annual averages based on quarterly observations
1. Assuming that the quarterly observations are for March, June, September and December in year "t", they may be designated Xt,3,Xt,6,Xt,9 and Xt,12, respectively:

(a) estimates for January and February in year "t" will be:
Xt,1=1/3(Xt,3-Xt-1,12) + Xt-1,12
Xt,2=2/3(Xt,3-Xt-1,12) + Xt-1,12

(b) estimates for April and May in year "t" will be:
Xt,4=1/3(Xt,6-Xt,3) + Xt,3
Xt,5=2/3(Xt,6-Xt,3) + Xt,3

(c) estimates for July and August in year "t" will be:
Xt,7=1/3(Xt,9-Xt,6) + Xt,6
Xt,8=2/3(Xt,9-Xt,6) + Xt,6

(d) estimates for October and November in year "t" will be:
Xt,10=1/3(Xt,12-Xt,9) + Xt,9
Xt,11=2/3(Xt,12-Xt,9) + Xt,9

2. The annual average for year "t" is:
t X = 1/12(Xt,1 + Xt,2 + …. + Xt,12).

ILO-2
Method for calculating annual average estimates for annual employment from occasional, but regular observations

Let us use Xt to indicate annual employment estimates obtained from an occasional observation in year "t", Xt+1 for annual employment estimates obtained from the same source in year "t+1" and X for the annual average of the two consecutive years. The annual average for year “t” is thus the sum of the two years divided by 2.

APPENDIX B (continued)

OECD-1 (used in ILO program)

Method for interpolating between survey benchmarks

Given survey benchmarks X0 and X12, 12 months apart, and corresponding registration (administrative) data Y0, Y1 ... Y12, where Y0 and Y12 are for corresponding periods as X0 and X12 respectively:

- calculate A, the ratio between the change in the survey benchmarks and the registration data over the year where t=the month:

At = [(X12**/**X0)**/**(Y12**/**Y0)] x t/12 (t=1,2,3,...12)

- and correct the simple interpolation results by applying the factor At cumulatively.

This gives the monthly series:

X0 = X0
X1 = X0(Y1/Y0)A1
X2 = X0(Y2/Y0)A2
X3 = X0(Y3/Y0)A3
X4 = X0(Y4/Y0)A4
X5 = X0(Y5/Y0)A5
X6 = X0(Y6/Y0)A6
.
X12 = X0(Y12/Y0)A12 = X12

The use of the adjustment factor (A) ensures that the interpolated monthly series coincides exactly with the benchmark estimates.

Source: *ILO-Comparable Annual Employment and Unemployment Estimates*, accessible at http://laborsta.ilo.org/applv8/data/iloce.pdf.