

*What Criteria Might Be Used to
Effectively Measure Research and Innovation in
Post-Secondary Environments?*

Prepared by

Canadian Council on Learning

For

British Columbia Ministry of Advanced Education

October 23, 2006

Ottawa

215-50 O'Connor
Ottawa, ON, Canada K1P 6L2
t. 613.782.2959 f. 613.782.2956

Vancouver

1805-701 West Georgia Ouest
P.O. Box/C.P. 10132, Vancouver, BC, Canada V7Y 1C6
t. 604.662.3101 f. 604.662.3168

www.ccl-cca.ca | info@ccl-cca.ca

1. BACKGROUND	5
1.1 The Question	5
1.2 Framing the Question and Defining the Terms	6
1.2.1 <i>Backgrounds and Definitions</i>	6
1.2.2 <i>Some Notes on Econometrics</i>	7
1.3 Historic Overview of the Study of Innovation.....	8
1.4 Theoretical Framework	8
1.5 Research and Innovation: The Canadian Context.....	9
1.6 Using Research to Inform Policy	10
1.7 Review Question and Approach.....	10
1.7.1 <i>Being Systematic</i>	11
1.7.2 <i>Stages of the Systematic Review</i>	11
2. METHODS.....	14
2.1 Search Strategy and Inclusion/Exclusion.....	14
2.2 Identification of Studies	15
2.3 Secondary Inclusion/Exclusion.....	15
2.4 Review and Analysis.....	16
2.5 Cross-Study Synthesis	16
3. MAPPING THE RESULTS: IDENTIFICATION AND DESCRIPTION	17
3.1 The Flow of Literature.....	17
3.2 Study Characteristics.....	18
3.2.1 <i>Descriptive</i>	18
3.2.2 <i>Analytical/Critical</i>	18
3.2.3 <i>Empirical</i>	23
3.3 Discussion of Indicators.....	29
3.3.1 <i>Grouping of Indicators</i>	30
3.3.2 <i>Trends in Studies Exploring Innovation</i>	31
4. ANALYSIS	34
4.1 Characteristics of the Studies	34
4.2 Evaluating Methodological Quality.....	43
4.2.1 <i>Empirical studies</i>	49

4.2.2 <i>Cross-Section Studies</i>	50
4.2.3 <i>Time Series</i>	53
4.2.3 <i>Panel</i>	54
4.3 Summary Table of Methodological Quality.....	56
4.4 Strengths and Limitations of the Three Most Common Indicators	60
4.4.1 <i>Publications</i>	60
4.4.2 <i>Patents</i>	61
4.4.3 <i>Investment in R&D</i>	61
4.4.4 <i>Summary</i>	62
4.5 Measurement Techniques: Strengths and Limitations	63
4.5.1 <i>Citation Analysis</i>	63
4.5.2 <i>Innovation Surveys</i>	64
4.6 Barriers to Innovation	65
5. RESULTS AND DISCUSSION	66
5.1 Summary of Principal Findings.....	66
5.2 Summary of Effective Measuring Techniques	67
5.3 Strengths and Limitations of This Review.....	68
6. CONCLUSIONS AND FURTHER CONSIDERATIONS.....	70
BIBLIOGRAPHY.....	72
Included Articles	72
Background Articles	74
GLOSSARY OF TERMS.....	77

EXECUTIVE SUMMARY

This paper systematically reviews and evaluates studies of research and innovation (R&I) to determine criteria that might be used to measure R&I in post-secondary environments. This review applies thorough and transparent procedures to the collection and analysis of evidence from research-oriented articles obtained from a variety of academic and fugitive sources disseminated between 1985 and the present.

The field of R&I is a complex one that presents many challenges for those who try to study and understand it. For example, the field lacks clear definitions of “research” and “innovation” and relies heavily on proxies or indicators. To address these challenges, the studies captured by this review are divided into three categories, based on their treatment of innovation indicators. The three categories are descriptive, analytical/critical and empirical. The focus of this review is on the empirical category because the articles included in this category tested the validity and reliability of the innovation indicators using statistical methods. Research trends are identified, based on frequency analysis of indicators. The methodological quality is assessed by an econometrician.

Based on this analysis, conclusions may be drawn regarding the study of R&I, including, but not limited to, the following:

- Scientific production is often confused or combined with innovation. This leads to an underlying assumption that if a study measures scientific production, it will measure innovation. As a result, the majority of studies measure innovative production and fail to measure innovative processes (Salazar and Holbrook 2004).
- Collaboration seems to be positively associated with innovation (Belderbos, Carree and Lokshin 2004; Gulbrandsen and Smeby 2005; Medda, Piga and Siegel 2005).
- R&I seems to be positively associated with labour productivity and research production (Belderbos, Carree and Lokshin 2004; Gulbrandsen and Smeby 2005; Medda, Piga and Siegel 2005).
- The majority of studies measured innovation using three indicators: research and development (R&D), patents and publications. In fact, 60% of the studies failed to examine indicators other than those included in the research and commercialization category. This result suggests that other aspects of innovation are being under-utilized.
- One aspect of innovation that has received particularly little attention is the innovation environment. This category includes indicators such as tax revenue and internet connectivity. Including this category in the study of innovation may have an effect on the results.
- Eight out of 17 studies were determined to be of good methodological quality, and the remaining nine studies were determined to be of fair quality. Not a single study received a poor rating. As a result, it appears that the overall methodological quality of studies examining R&I is reasonably good.

1. BACKGROUND

1.1 The Question

The twenty-first century is often characterized as the century in which the economy is knowledge-based; some of the most valued commodities are human capital and knowledge creation (Coulombe and Tremblay 2005). The creation, dissemination and exchange of information and technology keep countries competitive and economically viable in this global market. The retention of a skilled and productive workforce, therefore, is an important component of a country's economic growth and stability. Authorities are aware of the "brain drain" phenomenon, in which a region's brightest and most ingenious workers are attracted to other countries or regions because they have more amenable innovative environments. Invention and advancement are social and economic goods; therefore, the creation of an innovative environment must be an important priority for both the public and private sectors if we want to remain both strong and competitive (Science Council of British Columbia 2002).

Post-secondary systems are a major facilitator of research and innovation. Post-secondary institutions are a primary source of infrastructure (research laboratories), human capital (graduates and research personnel) and funding (Abbott and Doucouliagos 2004). In fact, the higher education sector provides 45% of the total funding for R&D in Canada (Statistics Canada 2006). As Canada has the highest proportion of post-secondary educated citizens in the OECD, one might suspect then that Canada also ranks extremely high on the OECD innovation scale. Not so.

Although Canada's workforce is the most educated in the world, our innovative capacity ranks near the bottom of all G8 countries on all performance measures (OECD 2001). Perhaps as a result of Canada's history of poor rankings in innovation, the federal government committed to reaching four key innovation targets by 2010: to rank among the top five countries in the world in terms of R&D performance, to at least double the Government of Canada's current investments in R&D, to rank among world leaders in the share of private sector sales from new innovation and to raise venture capital investments per capita to prevailing U.S. levels. This commitment appears to be working. In the past few years, dramatic increases have occurred in "the number of workers devoted to research and development (R&D), in external patent applications and in business expenditures on R&D" (Government of Canada 2003, 6).

Building innovative capacity often requires the investment of scarce public resources, and with government spending comes accountability. But how do governments measure ingenuity and innovation? What proxies might be used, and how might those proxies or indicators be measured? The necessity to make choices about how and where to allocate resources to improve the innovation environment has increased the demand for evidence-based research in this area.

This review seeks to answer such questions using research evidence. It systematically reviews research devoted to identifying effective measures of innovation. Nine academic databases were searched, as well as 22 journals and 47 fugitive literature sites. In total, 2,683 articles were initially captured. This number was eventually reduced to 28, using systematic and

transparent processes. The following report analyzes and synthesizes the findings of these 28 studies in an effort to produce a list of common indicators of innovation and reliable methods of measuring those indicators as determined by the evidence-based research in the field.

1.2 Framing the Question and Defining the Terms

1.2.1 Backgrounds and Definitions

In his discussion of Nation Systems of Innovation (NSI), Holbrook (1997) draws attention to Schumpeter's seminal book about innovation, published in 1937, in which Schumpeter defines five types of innovation: "(1) the introduction of a new product or qualitative change in an existing product; (2) process innovation new to an industry; (3) the opening of a new market; (4) the development of new sources of supply for raw material or other inputs; and (5) changes in industrial organization" (cited in Holbrook 1997, 6). Similarly, the Oslo Manual (the international standard for measuring innovation) defines innovation in terms of technological product and process. The above definitions clearly focus on the production of goods, whether these goods are new or simply improvements on existing ones. However, much of the research we located that specifically relates to post-secondary education tends to define innovation in much broader terms.

Existing research on R&I appears to lack a cohesive definition of the terms "research" and "innovation." Often the terms are used interchangeably (European Commission 2002). During the course of our review, we found that many authors use R&D as an indicator of R&I. We followed this trend, identifying R&D expenditure as an indicator of innovation.

In order to define research, we made a distinction between basic research and applied research. Basic research is generally understood to be "research for research's sake," that is, research that is exploratory in nature and does not necessarily seek to scientifically disprove an existing theory. Rather, basic research seeks to acquire further knowledge about a given subject or topic. Applied research seeks to answer a specific question or find a solution to a specific problem. Some studies make a distinction between basic and applied research, but others do not. The studies exploring research in this review are primarily concerned with applied research and innovation. For the purpose of this review, we will define research as any effort to expand knowledge and develop novel products and processes. This definition includes research directed toward a specific need (applied research) and "research for research's sake" (basic research).

Innovation is also a difficult concept to measure, partly because it too lacks a universal definition. The definition used by the Conference Board of Canada captures the multidimensional nature of the term: Innovation is defined as a "process through which economic or social value is extracted from knowledge, through the generation, development, and implementation of ideas to produce new or improved products, processes, and services" (Conference Board of Canada 2002, 3). The OECD's Oslo Manual (2005b), which is considered the scientific community's guide to conducting research on R&I (Salazar and Holbrook 2004), provides some additional components of innovation from a business-oriented perspective. Innovation includes uncertainty, investment, spillover, utilization of

new knowledge, a combination of existing knowledge and attempts to gain a competitive advantage (OECD 2005b). These characteristics were used in conjunction with the Conference Board of Canada's definition as the framework for determining which articles addressed the topic of innovation and should therefore be included in this review. We defined innovation as the transformation of ideas or knowledge into a new or improved activity, process or product. Innovation may result from basic or applied research, collaboration and transmission.

Not only are there no universal definitions of research and innovation, but there is also no consistent approach to constructing a framework for investigating these concepts. Some studies use "innovation" as a proxy for research production, while others use proxies of research production (for example, the number of journal publications) as a measure of innovation. For the purposes of this review, studies that identify "research production" as a proxy for innovation are included, while studies that examine research production in isolation (and did not mention it in conjunction with innovation) were excluded.

The above-mentioned definitions for research and innovation were used in conjunction with inclusion/exclusion criteria (described in detail in section 2.1) to determine which articles should be included in this review.

1.2.2 Some Notes on Econometrics

Having established definitions, it remains necessary to discuss econometrics because the concept of innovation is, to a certain extent, rooted in economics. As a result, a large percentage of the studies included in this review rely on econometrics.

Econometrics is the quantitative measurement and analysis of economics. Literally, the word means "measurement in economics." A generally accepted definition for econometrics does not exist. The common perception is that econometrics involves a combination of statistics, mathematics and economics, but remains distinct from each of them. Two definitions of econometrics are:

the application of statistical and mathematical methods to the analysis of economic data, with a purpose of giving empirical content to economic theories and verifying them or refuting them (Maddala 1988, 1)

the integration of economics, mathematics and statistics for the purpose of providing numerical values for the parameters of economic relationships and verifying economic theories (Koutsoyiannis 1973, 3)

Regression analysis is the method used by econometricians to make quantitative estimates of relationships suggested by economic theories. Although both econometricians and statisticians use regression analysis to estimate relationships, there are major differences between the two applications. First, econometricians conduct little controlled experimentation, because their data are mostly obtained from observations of what is going on in the real world. Due to this lack of controlled experimentation, econometricians often violate the standard assumptions for regression analysis adopted by statisticians. Econometricians apply different assumptions and statistical techniques to regression analysis.

1.3 Historic Overview of the Study of Innovation

The study of research and innovation is a relatively new phenomenon. In 1963, the OECD published the first manual standardizing the study of research and experimentation. The Frascati manual provided a standard for the collection of R&D investment. The report was based on the linear model (see section 1.4) and focussed on the measurement of inputs and their transformation into scientific and technical outputs (Godin 1996b). This manual allowed for benchmarking and cross-country comparison through the use of indicators. The manual has since undergone numerous editions and revisions.

In the 1960s and 70s, there was an emergence of science and technology policies. It was believed that the measurement of R&D would reveal the areas in which public funds were needed. In the 1980s, various Western countries and the OECD began to recognize the complexity of science and technology, and, as a result, the limitations of relying on simple R&D indicators. Governments were also becoming more interested in the innovative component of science and technology. In 1992, this interest led to the development of the Oslo Manual (Godin 1996b). The Oslo Manual, perhaps the most prominent work on innovation today, provides a methodological template for examining innovation. This manual warrants further discussion because it is cited often and frequently used as a tool for developing research design. We used this report to inform our understanding of the study of innovation, and to determine the strengths and limitations of methodologies and indicators.

The Oslo Manual provides information on the measurement and interpretation of data related to science, technology and innovation. It provides guidelines for collecting and interpreting data on innovation to allow for international comparison. The manual has two goals. The first is to study innovation and economic growth. The second is to develop indicators for benchmarking national performance. The manual also includes a discussion of the barriers to innovation and guidelines for designing survey questions and establishing survey procedures (OECD 2005b).

In the Oslo Manual, innovation is divided into four categories: product innovations, process innovations, organizational innovations and market innovations. The manual examines innovative activities of the business sector, the scope of these activities and the outputs and effects of innovation. The surveys are designed to represent numerous industries, so the results can be compared among industries and countries (OECD 2005b).

Around the same time as the introduction of the Oslo Manual, governments recognized the need to measure research outputs. This fuelled a search for output and impact indicators. Although several countries were already using their own output indicators, no standard measurement existed. To meet this demand, the OECD developed the technology economy productivity program (TEP). The primary goal of this program was to develop new innovation and output indicators. It also set itself the task of operationalizing a national system of innovation; this task would shape the future study of innovation (Godin 1996b).

1.4 Theoretical Framework

Several theoretical approaches have contributed to the history of innovation and the state of the field today. As noted, one highly influential work on innovation was by Joseph

Schumpeter. Schumpeter postulates that economic development is driven by innovation and that innovation is a process of creative destruction. This process is characterized as dynamic; old technologies are replaced by new technologies. Schumpeter maintains that radical innovations result in large disruptive changes; incremental innovation is preferred because it continually advances the process of change (cited in OECD 2005b).

Schumpeter's work provides the theoretical underpinnings of research on R&I. The most prevalent theoretical approach in the literature is the linear model. The linear model is based on Schumpeter's view that innovation drives economic growth. This theoretical framework provided the basis for the development of the Frascati Manual. Linear theory postulates that science and research lead to new technologies and that these new technologies satisfy market needs. More specifically, innovation is viewed as a process that moves through a series of stages: basic research, applied research, design, development and production. Therefore, scientific inputs have a direct effect on innovative outputs. According to this view, public policy can affect R&I outputs simply by increasing inputs (Salazar and Holbrook 2004).

The development of the OECD TEP program resulted in the abandonment of the linear model in favour of a systems approach to innovation (Godin 1996b). The systems model emphasizes the demand side of technology and examines the influence of external forces on firms and other institutions (Salazar and Holbrook 2004). An underlying assumption is that science undergoes constant evolution within the system (Arnold 2004). Systems theory has two main areas of concentration: the transfer and diffusion of skills, ideas and knowledge and the interaction between firms and the application of knowledge. The interaction process also must take into account the role of government, and its effect on markets through monitoring and regulating the innovative process. There are some common characteristics of systems of innovation: the firms are part of a network; there are linkages between firms and institutions; intellectual resources flow between and within institutions; and knowledge and learning are considered to be key economic resources (Salazar and Holbrook 2004).

The application of systems theory to innovation shows that innovation is not merely an input-output equation; instead, it is composed of regularly interacting stakeholders and interdependent relationships among public, private, industrial and academic sectors, to name only a few.

1.5 Research and Innovation: The Canadian Context

Since the 1950s, the federal government, by means of three granting councils, has been responsible for funding the direct costs of sponsoring research. In contrast, the provincial government has been responsible for the operating costs necessary to support research and the costs associated with infrastructure and equipment. In the late 1970s and the early 1980s, there was a shift from cost funding to block funding. This shift was accompanied by a gradual imposition of spending limits, which resulted in serious constraints on the fiscal resources provided to provinces for post-secondary education. In the 1980s, the provinces began to fund applied research in an attempt to fill the gap left by the federal government funding reductions (Wolfe 1998b). Currently, the provincial and federal governments fund R&I in Canada.

In 1992, the government launched an innovation strategy with the mandate of becoming one of the top five innovative countries. In an attempt to meet this goal, the government made an effort to increase R&D expenditure. In 2004/2005, higher education R&D expenditure increased to 9 billion dollars; this represents an increase of 10.5% from 2003/2004. The government portion of this increase in R&D expenditure was 7.1% above 2003/2004 levels (Statistics Canada 2006).

In spite of these efforts, Canada currently ranks sixteenth on the international innovation ranking scale. This represents a three-position fall from thirteenth in 2005. The Canadian government has recognized the importance of R&I in securing Canada's place in the global economy, wants to improve its position relative to other countries and is allocating more resources to facilitate R&I.

1.6 Using Research to Inform Policy

If the Canadian government is to meet its aim to become one of the top five innovative countries, increased importance will need to be placed on mechanisms that facilitate R&I. Public policy is an important mechanism for facilitating R&I activities and securing long-term scientific viability. Unfortunately, the abstract nature of the concepts of research and innovation make it difficult for public policy officials to employ research in this field. In an attempt to organize the literature in a meaningful way, we will focus on the indicators used to measure innovation and conduct an analysis of the quality of the empirical research within this field.

This type of analysis allows conclusions to be drawn regarding the effectiveness of the indicators, the extent of innovation within an institution or country, the identification of underutilized indicators, the criteria for the effective measurement of R&I and the directions for future study. This information allows policy decisions to be grounded in empirical research that can be replicated and validated.

1.7 Review Question and Approach

This review evolved from an earlier, more limited survey of research designed to determine the feasibility of conducting an in-depth review on a given topic. The tool used in this survey is known as a Question Scan.¹ Question Scans survey the literature devoted to a given topic using a sample of relevant databases and fugitive literature. The purpose of such a tool is to determine the practicability of carrying out a much larger, more comprehensive, systematic review of evidence on the question at hand.

After completing the Question Scan, the decision was made to proceed with a systematic review of the question: what criteria might be used to effectively measure research and innovation in post-secondary environments? Through discussions with ministerial staff, the scope of the question was elucidated. The review was to address, where possible, the strengths and limitations of the various methods of evaluating R&I, the indicators of innovation and the commercial value of R&I.

¹ CCL would like to acknowledge EEPI as the developer of the original version of Question Scan.

A variety of databases across disciplines were searched. The databases, which included social science, education and economic disciplines, are identified in greater detail in the methodology section. Furthermore, key systematic review databases were also searched to determine the existence of other reviews of research evidence on the topic and to locate any empirical studies not located within the original search. Finally, a list of professional organizations with reputations for conducting research in the field was also searched.

Inclusion/exclusion criteria were developed by the review team and approved by the ministerial staff. The retrieval of documents identified for secondary review was conducted using internet sources and hand searches at the University of British Columbia library. Final inclusion criteria were selected by the review team and approved by ministerial staff.

Because the majority of the studies included for review used econometric methods, an econometrician was added to the team to carry out the analyses of the studies and evaluate their quality. A scoring rubric was included for transparency and accountability purposes.²

Graphs and tables are included where possible to represent the summative findings, and add clarity and ease of reference through visual representations.

1.7.1 Being Systematic

Documentation and transparency are essential components of a systematic review. CCL's approach to conducting reviews of evidence is in keeping with well-established review protocols: thorough and transparent procedures are applied to the collection and analysis of evidence-based research in education. The CCL research and review team uses a variety of documentation templates when conducting Question Scans and systematic reviews to ensure consistency, transparency and accountability.

The use of the above-mentioned templates allows for the replication of our reviews. Each review conducted by CCL includes the use of a search diary completed by the information-retrieval specialist who conducts the searches. In the same fashion, all inclusion/exclusion criteria are documented, as are reasons for the exclusion of studies. A document control sheet is used to track and monitor the movement of studies from their initial capture to their final inclusion and analysis. A flow chart describing the movement of studies through the review is also provided for the reader as a visual representation of the process.

1.7.2 Stages of the Systematic Review

The following list provides a short description of each stage of the system review process. The results of each stage and their subsequent application are described in greater detail in the methodology section.

Identification of the research question(s) and conceptual framing

² See Table E

Consultation with the client to identify the key concepts and emphases of the research is the first step of the review process. At this stage, the terms used to formulate the question are clarified and defined, and relevant databases and websites are identified with the assistance of experts and practitioners in the field.

Development of search strategies

CCL created a search strategy development process that includes the creation of a comprehensive list of both free terms and controlled vocabulary. This process ensures both a broad review and a focussed selection of literature devoted to a particular question. Free terms are generally provided by experts or practitioners in the field, and are used to identify matching and/or complementing controlled vocabulary using database thesauruses. A facet analysis (which involves grouping like terms according to category) is performed for all terms. Search terms are then linked using truncation and Boolean terms, resulting in a single search strategy for each category of terms (for example, population/location, issue and method). All category search strings are then combined to create the final search strategy. Search strategies differ slightly according to the capacity of the database.

Searching

CCL's search procedures include the use of a search diary. The diary is used to document all searches, their results and any decisions to adapt the search. Search results are imported to a citation management program and sorted for duplicate references.

Searching for fugitive literature (research not published in journals) requires the collaboration of experts and practitioners in the field. Key websites are gathered and reviewed for relevant publications, references and links to other pertinent sites.

These websites are then used to locate similar websites using the advanced search function of Google, "Find pages similar to the page." The reference lists of included articles are also reviewed for relevant studies.

The initial review of results

Once all the search results have been imported to the citation management program and sorted for duplicates, inclusion and exclusion criteria are developed to remove irrelevant material from the review process. The criteria are developed collaboratively by the review team and sent to the client for editing and approval.

The initial inclusion/exclusion criteria are applied using titles and abstracts only. Reviewers are instructed to include studies in cases where it is difficult to decide whether they meet the criteria so as not to eliminate relevant studies in error. Exclusion decisions are recorded in the inclusion/exclusion spreadsheet.

Document retrieval

Once the initial inclusion/exclusion stage is complete, the full texts of all included articles are retrieved for the second stage of review. The majority of the studies are obtained via the internet. The reviewer saves an electronic copy of each retrieved study. Studies that are not

available via the internet are located through university libraries whenever possible. The information retrieval specialist tracks the retrieval of the studies using the document control sheet. Each study is tagged with identification of its origin, which may be a database link, the Web (typically Google Scholar), the library or inter-library loan.

Secondary review

Secondary inclusion criteria are typically more detailed than the initial criteria because the reviewer has acquired a general sense of the body of literature through the initial review of the titles and abstracts. At this stage, more studies are excluded due to inappropriate research designs or a lack of focus on the specific question at hand. Each included article must meet specific inclusion criteria.

Coding

After the initial studies are read in greater detail and the determination of included articles is finalized, the review team meets to develop coding criteria in accordance with the question. The secondary review process also includes the development of category codes for data extraction. Reviewers code all the articles according to the data extraction categories, using the data extraction coding spreadsheet.

Summaries and quality analyses

Each study is summarized by the reviewer. In the case of this review, an econometrician was brought in to review and analyze the quality of the included studies. A rubric of quality was developed, applied to each study and included in the report.

Synthesized findings

Studies are synthesized using a variety of techniques and methods appropriate to the question and the resulting capture of studies. Using the coding sheet, studies are grouped and synthesized according to focus, results and study quality. General statements of finding are made accordingly.

2. METHODS

2.1 Search Strategy and Inclusion/Exclusion

The search strategies for this review were developed by the research team in consultation with experts in the field, as well as with the British Columbia Ministry of Advanced Education. The general process of searching, identifying and reviewing relevant articles is outlined in Section 1.5. Prior to creating our search strategies, the research team reviewed key background articles, which informed our understanding of the issues related to R&I. Our initial search was structured broadly in order to maximize the capture of relevant articles. The search strategy was based on three main search categories:

- population: colleges, universities or research institutions; professors, faculty or teachers
- issue: research and development, educational productivity, research utilization, publication
- method: empirical studies using quantitative methods, including surveys, data analyses, longitudinal observations, statistical analyses, comparative analyses; meta-analyses and systematic reviews of such studies

Using these categories, experts were able to provide appropriate free terms, which were then used to identify controlled vocabulary and develop the final search strategy, as outlined in section 1.7. We applied our final search strategy to nine databases: ERIC, EconLit, Academic Search Premier, Business Source Premier, Theses Canada Portal, Education Index Full Text, CBCA Education and CBCA Business. Any adjustments to the search as a result of database limitations were documented.

We limited the scope of our search by confining the language of publication to English and French, and by imposing the date restriction of 1990 to the present.

In addition to database searches, we attempted to retrieve applicable reports and studies by hand searching pertinent journals. A total of 22 journals were hand searched. We also searched fugitive literature, focussing on sites that could provide research relevant to the Canadian context. A total of 47 fugitive literature sites were searched: Google Scholar; Merrill Advanced Studies Centre; Canadian Psychiatric Association; National Bureau of Economic Research; HKUST Business School; Vienna University of Economics and Business Administration; University of California, Los Angeles Higher Education Research Institute; The Campbell Collaboration; EPPI Centre; Stanford Institute for Higher Education Research; Centre for the Study of Higher and Post Secondary Education; Learning Alliance for Higher Education; National Centre for Research on Evaluation, Standards and Student Testing; Association of American Colleges and Universities; American Association of University Professors; Association for Institutional Research; Rand Corporation; Centre for Higher Education Policy Studies; National Centre for Public Policy and Higher Education; Centre for the Study of Higher Education; Australian Government Department of Education; Higher Education Policy Institute; EuryDice; Educational Policy Institute; Canadian Millennium Scholarship Foundation; Canadian Association University Teachers;

Associations of Universities and Colleges in Canada; Council of Ministers of Higher Education; Federation of Post Secondary Educators of B.C.; Ontario Confederation of University Faculty Associations; Council of Ontario Universities; B.C. Innovation Council; Canada Foundation for Innovation; British Columbia Knowledge Development Fund; Leading Edge Endowment Fund; Human Resource Development Canada; Canadian Policy Research Network; ERIC Centre for Research on Innovation and Competition; Research Unit for Innovative Economics; Max Planck Institute of Economics; Eindhoven Centre for Innovative Studies; NIFU STEP Centre Innovation for Research; Trend Chart Innovative Policy in Europe; OECD; Conference Board of B.C. and Statistics Canada. Relevant articles identified in the bibliographies of articles found through hand and fugitive literature searches were also collected.

Articles were included if they met several criteria. The article must examine post-secondary institutions/departments, whether public or private, industry/university collaborations or university research hospitals. Included articles focussed on R&I or R&D, identified criteria for the measurement of R&I or used R&I as a criteria for the measurement of quality in PSE.

Methodologically, articles were included if they used empirical measurement, comparative analysis or survey design in conjunction with data analysis. Articles were excluded if the country of origin was non-OECD or the primary focus of the research was on a professor's innovative instruction or an institution's innovative programs or instructional practices. Additionally, articles were excluded if they did not contain measurement and analysis of R&I or if they were position pieces, books or articles in newspapers or magazines.

2.2 Identification of Studies

The initial database searches in conjunction with fugitive literature searching and hand searching resulted in 2,683 articles. These articles were subsequently searched for duplicates and 41 articles were removed. The application of the first inclusion/exclusion criteria resulted in the exclusion of 2,379 articles, so 263 articles remained for further analysis.

The next step in the process was document retrieval. All but 84 articles were retrieved via the internet and the UBC library. Time and resource constraints did not permit the retrieval of documents through specific journals or from authors themselves. A total of 179 articles were collected for further analysis.

2.3 Secondary Inclusion/Exclusion

Secondary inclusion/exclusion criteria were developed in consultation with the British Columbia Ministry of Advanced Education. These criteria were applied to the full text of all 179 retrieved articles; we selected only studies that applied bibliometrics, surveys or data analysis using statistical methods.

We excluded studies if:

- The study did not address R&I or R&D by or in collaboration with post-secondary institutions.
- The study did not identify an indicator of R&I or R&D quality, productivity or outcome.
- The study did not use or identify a mode of analysis to establish the reliability of the indicator.

Articles that did not meet the inclusion criteria but provided relevant background on R&I were also earmarked at this time. These included relevant reviews of literature, meta-analyses and reviews that elaborated on the history and theory of R&I. The application of the secondary inclusion criteria resulted in the removal of 118 articles. Among the remaining 61 articles, 28 were deemed to be included articles and 33 were considered relevant background articles.

All studies in the second inclusion were summarized and identified according to the central purpose of the study, the study design and method of analysis, the Dataset, the indicators used for R&I and, where possible, the limitations of the study as identified by the authors and reviewers.

2.4 Review and Analysis

An econometrician completed a subsequent review of the methodological quality of all the included studies. A rubric for evaluating the quality of each study according to econometric analyses was developed and applied. Each study received a rating between 11 and 33. Studies were thus judged to be of good, fair and poor methodological quality. The results of this review are elaborated in section 4.

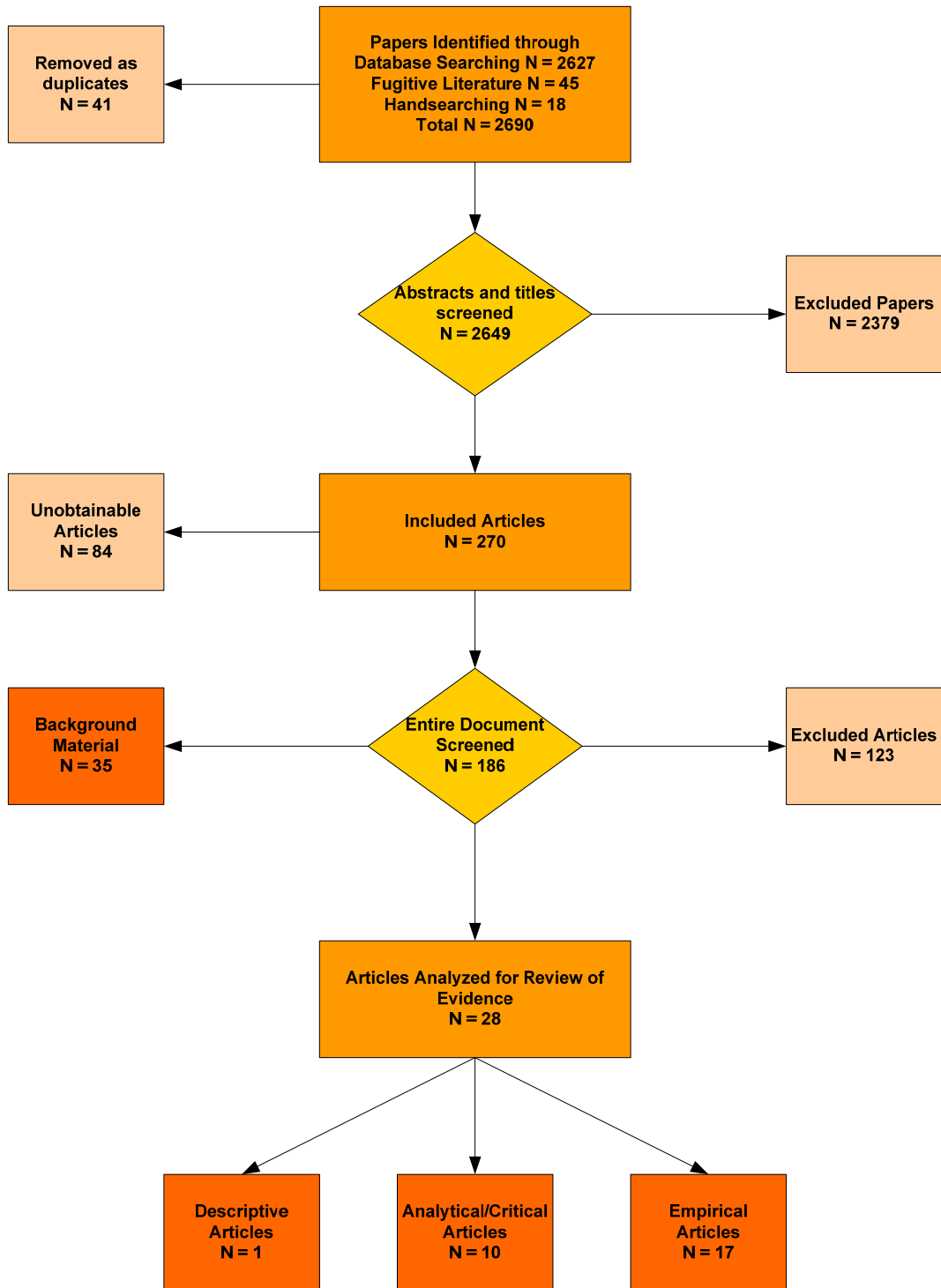
A coding matrix was developed to organize the results of all studies and allow for cross-study comparisons. For each study, the author, the country and year of publication, the population, the design and the analysis methods were specified.

2.5 Cross-Study Synthesis

This report will identify the common indicators explored in each included study. This information will be used to determine which of the indicators are used most frequently, which indicators are insufficiently examined and the extent to which multiple indicators are explored in each study.

3. MAPPING THE RESULTS: IDENTIFICATION AND DESCRIPTION

3.1 The Flow of Literature



3.2 Study Characteristics

In this section, we will summarize each included article. The studies are organized according to methodological design, as follows: descriptive, analytical/critical and empirical.

Descriptive articles describe indicators used as proxies for innovation but fail to engage in a critical discussion of the strengths and limitations of the indicators or to empirically analyze the relationships between the indicators and innovation. Analytical/critical articles include some critical discussion of the indicators and their use, but fail to conduct empirical analysis as well. Empirical articles discuss the indicators and engage in some empirical analysis of these indicators and their reliability as proxies for innovation. Typically, these indicators can be categorized in accordance with the four SCBCIC (2003) indicator categories: research and commercialization, human resources, innovation environment and innovation outcome (see section 3.3.1). Each article will be summarized, highlighting the aim of the study, the design and analysis, the indicators and the results.

3.2.1 Descriptive

Van den Berghe *et al.* (1998)

The authors of this study suggest that scientific progress is accomplished through basic research and the transfer and availability of research findings. The authors assume that scientific literature is a good proxy for scientific activity and progress. The study aims to use bibliometric techniques to compare research productivity in medicine, science and pharmaceutical science faculties at three universities in Flanders.

Indicators examined in this study fall under the research and commercialization category. The authors measure research production, research impact and research capacity between 1980 and 1991. Research production was determined by the number of publications produced by each department; research impact was determined by the average number of times cited per department as compared against the world average; and research capacity was determined by the number of full-time equivalents spent on research. Each department was subsequently ranked on all three indicators.

Although this article provides a measure of scientific activity and progress, it fails to conduct any statistical analysis of the indicators. The article provides a description of the scientific outputs for the respective universities and departments. Statistical analysis is necessary to determine the relationship between these indicators and whether or not the indicators are a good measure of scientific production and innovation.

3.2.2 Analytical/Critical

Tijssen (2003)

The article by Tijssen examines research excellence and the methodological challenges in its evaluation. This article also includes a discussion of some of the strengths and limitations involved in using single indicators and the measurement challenges that result. These strengths and limitations will be discussed in section 4 of this report. Tijssen proposes a scoreboard approach to cope with some of the difficulties associated with measuring

research excellence. A scoreboard is an analytical tool, based on collections of quantitative data, for comparing countries and institutions.

Tijssen defines research excellence as a multidimensional concept that includes the successful transfer of research to inventions and innovations (94). He suggests that research excellence has several different components that require different techniques of measurement. The first component is originality. The author suggests this component should be measured by peer review because originality is not easily quantifiable. The second component, coherence and cumulative character, is best measured by an expert from within the field of study, while the third component, the quality of the scientific output, should be measured using citation analysis. In addition to distinguishing and measuring these three components of research excellence, the scoreboard approach would also include indicators of scientific output, measured using bibliometrics. All of the proposed indicators identified in this study fall under the research and commercialization category.

Tijssen concludes that the use of a scoreboard allows for the combination of multiple approaches and multiple indicators. This study appears to support a systems model of innovation, seeing innovation as a process of ongoing evolution rather than a single input-output equation. The author suggests that the scoreboard approach is most appropriate for indicator selection, comparative analysis and benchmarking.

Bessette (2003)

University research is considered critical to technological development and innovation. While many researchers recognize that university research impacts the economy, they do not agree on the best method to measure this economic impact. This article proposes a model for measuring research outputs where economic impacts are reported as a return on investment (ROI). This is done using an ROI mathematical formula:

$$R/100=(O-(C-I)/I)-1$$

The return on investment (R) is calculated as total economic output (O), minus cash operating cost (C) less investment (I), divided by investment (I), minus one. The formula includes multiple indicators, but fails to measure or analyze them.

The indicators used by Bessette include new products, licence agreements, number of graduates, GDP and the costs associated with technological development. All of these proxies for innovation fall under the categories of research and commercialization, human resources, innovation environment and innovation outcome. The author suggests that the ROI has two key advantages. First, ROI is a commonly understood measure and is standardized, which allows different technologies to be compared; second, the model is dynamic, flexible and capable of measuring divergent outputs.

Martin (1996)

This article examines the multidimensional nature of basic research, based on the Big Science Project, an English study that was one of the first to use multidimensional indicators of research. The author also discusses the indicators used to evaluate research and what they

measure. This article is based on raw data from an exploration of articles published in the journal *Scientometrics*. The author hopes to determine the common indicators used to evaluate research. This study examines indicators from the research and commercialization category. The author discusses three indicators: scientific activity/progress and production, scientific impact and scientific quality. Scientific activity/progress and production are measured by publication counts. Scientific impact is measured by citation analysis, and scientific quality is measured by peer review.

The author concludes that research evaluating scientific production relies heavily on the simplest indicator, publications counts. Additionally, the research tends to rely on one or two indicators. The author also found that peer review is generally used infrequently, primarily to conduct evaluations at the level of the individual. Peer review is rarely used at the department or organizational level. The author concludes by suggesting the majority of researchers recognize the importance of multiple indicators and favour this approach over single indicators. The author believes most studies fail to use three or more indicators because it is too costly.

European Commission (2002)

This report was designed to evaluate the strengths and limitations of methodological approaches and indicators used to measure scientific and technological production. The commission compares European Union (EU) countries with the United States and Japan. The indicators used in this report can be grouped into research and commercialization, human resources, innovation environment and innovation outcome. This study includes multiple indicators for each category, which will not be listed here due to the volume of indicators. The main focus is on scientific productivity outputs, the quality of these outputs and funding/expenditure. Publications and patents are used as proxies for scientific productivity outputs and citations are used as a proxy for quality. The proxies used for funding/expenditure include numerous R&D measures, such as GERD (gross expenditure on R&D) and HERD (higher education R&D). The report suggests that each indicator needs to be matched with the appropriate R&D input, for example, high-tech industries with business R&D. The report also examines workers in natural and applied sciences, as well as aspects of manufacturing and service sector R&D.

The results of this report are twofold. The first relates to benchmarking scientific productivity by country based on the indicators. The commission found that the EU has the largest scientific output of scientific publications in terms of absolute numbers, but ranks second, behind the United States, in scientific output of publications on a per-inhabitant basis. The EU is also a world leader in scientific output of publications in terms of dollars spent on research universities. However, the United States is the leader in patent output in terms of BERD (business expenditure on R&D) and citation output.

The second result relates to the strengths and limitations of each indicator and the methodological approaches to their measurement. (The strengths and limitations of the indicators are discussed in more detail in section 4.) The report suggests an analysis of relative trends for each country over time may be more appropriate for measuring and benchmarking scientific output, and that indicators need to be normalized and lagged appropriately to allow for a proper interpretation of the indicator.

Dosi, Llerena and Labini (2005)

Dosi, Llerena and Labini (2005) aim to evaluate and compare the innovative performance of the United States and the EU. The design of the study is comparative, and the authors examine indicators. The categories of indicators explored in this study are research and commercialization, human resources and innovation outcome. Specifically, the article examines scientific output, technological output, technological innovation and systems of higher education. Scientific output is determined by publication counts and citation analysis. Technological output is determined by the number of patents. Technological innovation is determined by several measures of R&D as a percentage of total GDP. Lastly, systems of higher education are determined by HERD (higher education R&D) and the number of science and engineer graduates.

The results of this investigation suggest that if total population is considered when calculating scientific productivity, the United States produces more publications, top one-percent publications, and citations. The authors suggest that contrary to popular opinion, the EU does not spend more on publicly funded R&D than the United States. The United States also produces more biotechnology patents than the EU. Based on these findings, the authors conclude that the EU lags behind the United States in both scientific output and technological innovation (as measured by R&D).

Langford, Hall, Josty, Matos and Jacobson (2005)

This paper examines the Canadian government's efforts to foster and evaluate university knowledge commercialization. Specifically, the paper explores the effectiveness of indicators, and the authors suggest implications for further measurement through the improvement of existing indicators and the development of additional indicators. The authors also conducted a survey of university faculty at the University of Calgary. The study design is a survey/case study. The study's indicators are from the research and commercialization, human resources and innovation outcome categories. Specific indicators include publications, patents, citations for both publications and patents, bachelor degrees in science and engineering, and high-tech sector employment.

The study's critique of the indicators is based on the indicators' ability to measure the five pathways of knowledge migration from university to industry: co-operative research and development, sale and spin-off of intellectual property, technical assistance, information exchanges and hiring skilled people. Langford *et al.* suggest that current indicators focus on patents and spin-offs and fail to measure several paths of knowledge flow.

The survey results found that the majority of respondents thought publications in open literature and collaboration with external organizations were the most efficient way to commercialize research. The authors concluded that measurement techniques used for the indicators rely too heavily on aggregate data and fail to capture the path-dependent nature of innovation. In particular, it is suggested that current indicators fail to capture the flow of human capital from universities into the workforce. The authors suggest that more indicators need to be developed to capture the multidimensional nature of innovation.

Holbrook (1997)

The aim of this paper is to explore the national innovation system and examine indicators of innovation and technology. This paper includes a pilot study of B.C.'s system of innovation, as well as a survey of business enterprises. The indicators used in this study are from all four categories. The author is specifically interested in investment in knowledge, investment in knowledge training, the productive use of human engineers and scientists, and knowledge transfer. Investment in knowledge is determined by R&D in terms of B.C.'s GDP. Investment in knowledge training and productive use of trained scientists and engineers is determined by human capital statistics and trends in real economic growth. Knowledge transfer is determined by R&D tax subsidies and estimated high-technology trade flows.

The results of this investigation suggest that data need to be added on investments in intellectual property and technologies within high-technology industries. The author suggests there also needs to be an improvement in our ability to capture the transfer of knowledge from scientific disciplines to employment. The survey results indicate that 78% of B.C. enterprises reported introducing new products or processes over the last five years. However, only 47% of the same enterprises reported introducing a product that they considered unique to the industry.

Godin, Gingras and Davignon (1998)

This paper examines knowledge production in Canada in an attempt to identify knowledge flows within the Canadian research system. Specifically, it is concerned with inter-sectoral, interprovincial and international knowledge flows. The paper uses bibliometric analysis and indicators from the research and commercialization category. Publication numbers are used to compare country and provincial performance, as well as the degree of knowledge flow between Canadian institutions. Citation analysis is also used to determine the impact of the scientific publications.

The article suggests there are four major institutional players in Canada: universities, industries, governments and hospitals. The results determined that Canada ranks sixth among OECD countries in total volume of publications and that this share has remained stable in the last fifteen years. Within Canada, the publications are unevenly distributed among the provinces, with Ontario and Quebec producing the most, followed by B.C. and Alberta. Lastly, the university sector produces about two-thirds of all publications.

Gauthier (1998)

Gauthier provides a methodological guide for the implementation of bibliometric analysis of science and technology. The author provides a review of bibliometric analysis, its main functions and examples of empirical studies. This includes scientific and technological output and analysis of flows or trends. The author examines indicators from the research and commercialization category. Gauthier discusses publications and patents and suggests that citations are a good method for analyzing collaboration and knowledge flow.

The author concludes by suggesting that bibliometric analysis can be used for three purposes. The first is to measure output, knowledge flows and trends. The second is to

evaluate programs, impact and quality of indicators. Lastly, bibliometrics can be used to monitor state or national research performance.

Godin (1996a)

The author examines the state of science and technology indicators in OECD countries. This analysis includes defining and evaluating specific indicators based on two international statistics criteria: the inventory must cover a large group of indicators, and the indicators must be produced regularly, in a systematic manner. Four inventories meet these criteria: those of the OECD, the European Economic Community (EEC), the National Science Foundation (NSF) and the Observatoire des sciences et des techniques (OST). The indicators examined in this article are from the research and commercialization and human resources categories, and include indicators such as publications, patents, R&D and number of science and engineering graduates.

The author considers the strengths and limitations of each indicators (see the detailed discussion in section 4.1), leading to several determinations. The author suggests that small firms and the service sector in general are not well represented by current indicators. He argues that more disaggregated data are needed that focus on output indicators. The author also suggests that more R&D information is needed, in the form of more specific streams of R&D expenditure.

When one examines the analytical articles as a group, two themes begin to emerge. The first is that numerous articles suggest current indicators and their measurement techniques are inadequate (Holbrook 1997; Langford *et al.* 2005). The second is that multiple indicators need to be used, as there is an over-reliance on the simplest indicators (Langford *et al.* 2005; Martin 1996; Tijssen 2003).

3.2.3 Empirical

Bilbao-Osorio and Rodriguez-Pose (2004)

This study examines R&D and innovation and their relationship with economic growth. The authors assume that technology and its advancement are important components of both innovation and economic growth. Multiple indicators within the research and commercialization category are examined. The authors are interested in a number of areas: the amount of new technology produced, the stock of previous knowledge available in the region, investment, skills and economic indicators. New technology is measured as the number of patent applications per million population for each region. The stock of previous knowledge is measured as GDP per capita in each region. Investment is measured as R&D expenditure over GDP. Skills are measured by the percentage of the population between the ages of twenty-five and fifty-nine having only a primary education. Lastly, economic indicators are defined as the percentage of people employed and the percentage of people working in the high-tech sector.

In the first part of their analysis, the authors investigate the effect of regional socio-economic conditions, such as GDP, investment in R&D, employment rate, economic structure and the skill level of population, on the number of applications for patents, which

serves as a proxy of innovation capacity. The model distinguishes between private, public and higher-education R&D expenditures. The second part of their analysis investigates the effect of innovations, using the growth rate of patents, while controlling for the socio-economic factors, on economic growth. Cross-sectional data of 103 regions in the peripheral and non-peripheral areas of the EU are used in the [OLS regressions](#) with log-linear specifications.

The major finding of the first part of the analysis is a positive relationship between innovation and initial level of wealth, although the significance of the results varies between peripheral and non-peripheral (central European) regions. The results of the second part show a positive and significant relationship between innovation and economic growth in the peripheral area but not in the non-peripheral area.

Goldstein and Renault (2004)

This study attempts to determine the extent to which public institutions and research universities generate regional economic development and what factors help knowledge-producing organizations contribute to economic development. The indicator categories examined were research and commercialization and human resources. Within these broad categories, the authors examined regional economic development, the presence of universities and technological innovation.

Goldstein and Renault investigate whether research and development from universities contribute to regional economic development in the United States. Changes in average earnings serve as [proxy variables](#) for economic development. The four measures of research and development from universities include the presence of a top-50 research university in the metropolitan statistical area (MSA), the sum of total research expenditures across all universities within the MSA, the sum of all degrees awarded in all higher-education institutions within the MSA and the number of patents assigned to universities within the MSA. Their analysis consists of tests of mean differences and linear regressions using data from 312 MSAs. Although their results show significant contributions are made to regional economic development, the magnitude of the effect is relatively small compared to the control variables they considered.

Belderbos, Carree and Lokshin (2004)

This study considers the effects of four different types of R&D partners on the growth of labour productivity and productivity in innovation. The authors are interested in four types of R&D partnerships, with competitors, suppliers, customers, and universities and research institutions. The indicators used in this study can be categorized as research and commercialization. Specifically, the authors use a measure of R&D expenditure and two measures of firm performance. Labour productivity is measured as growth of value added per employee, while innovative sales productivity is measured as the growth of sales per employee from new products to the market.

Their analysis involves [two-stage least squares regression](#) estimation using a large sample of Dutch innovating firms from the Community Innovation Survey. The authors conclude that the different types of R&D co-operation exhibit heterogeneous goals of productivity type.

Jacobsson and Rickne (2004)

Jacobsson and Rickne (2004) analyze the weaknesses of the conventional indicators used to measure R&D and suggest additional criteria to be included in the indicators. The conventional indicators include expenditure on R&D in the higher-education sector as a percentage of GDP and the number of published scientific articles. The additional indicators pointed out by the authors, which fall under research and commercialization, and include R&D expenditure by governments and private non-profit research bodies and time expenditures on academic R&D.

The study involves very little econometrics. Only a small part of the authors' argument is supported by results from simple regressions, which use a single explanatory variable to explain the dependent variable. Their argument is mainly supported by analyzing the relevant statistics. They conclude by suggesting the number of person-years spent on R&D in relation to a country's population is a better indicator of the volume of R&D inputs. They also suggest that time may be a more appropriate indicator than R&D input.

Calderini and Scellato (2005)

Calderini and Scellato analyze the relationships between academic research and private sector innovation in the wireless telecommunication sector. The indicators used to measure innovations in this study are from the research and commercialization category. The authors examine the absolute number and the Herfindahl index of patenting and publications in the wireless sector and in telecommunication. They use [panel data](#) for 33 regions and 15 countries in Europe with [fixed effects](#) and [Poisson](#) and [negative binomial model specifications](#). They also estimate a lagged variables model to account for [auto-regression](#).

The authors argue that their results suggest a geographical determinant of co-specialization. Second, there is a weak link over time between public research and private output. Third, the innovative performance of the wireless sector in Europe demonstrates a high degree of variability. Lastly, the presence of a large innovative firm increases smaller firms' interest in university research: the main finding of the study is that local, smaller companies are more likely to do academic research if a large, established innovative company exists in the region.

Medda, Piga and Siegel (2005)

The study of Medda, Piga and Siegel analyzes the impact of internal and external (university and other firms) research on total factor productivity growth of Italian manufacturing firms. The ratio of R&D expenditure to sales is used as a proxy for the rate of R&D investment; this falls under the larger category of research and commercialization. Data for more than 2,000 Italian manufacturing firms are used in the estimating a [two-stage, treatment effects model](#), which involves a bivariate [probit](#) sample selection model in the first stage. Their results suggest external R&D generates significantly higher returns than internal R&D. However, the higher return is mainly observed in collaborative research with other firms, rather than with universities.

Niosi and Banik (2005)

This study looks at the growth and performance of regional systems of innovation for biotechnology. The data are related to about 90 Canadian-based publicly traded biotechnology firms. The authors examined the research and commercialization category, using the average number of patents per firm to measure the amount of innovation. Although the study does not use an econometric model, four hypotheses are posed and are tested by analyzing the graphical and statistics trends of the relevant data, such as initial public offerings (IPO), patents and university spinoffs.

Results suggest that the innovative performance of firms was higher in large provinces than in smaller provinces. The authors also determined that university spinoffs were more inventive (produced more patents) than other firms' innovations; however, the associated employment growth was smaller.

Gulbrandsen and Smeby (2005)

Universities and colleges are seen as key actors in the national system of innovation (Wolfe 1998a). These organizations provide infrastructure for research institutions in which many innovative activities take place. Gulbrandsen and Smeby investigate the relationship between commercialization of research and university professors' research performance.

The data are collected from the questionnaire responses of faculty members at four Norwegian universities. The indicators used to measure research and innovations performance include academic outputs, such as kinds of publication, and commercial outputs, such as patents, commercial products, establishment of firms and consulting contracts. These indicators are from the research and commercialization category. The bivariate relationships between research funding and the indicators are presented in figures and tables. [Logit regression](#) is also used to analyze the probability of reporting different types of commercial output.

The authors' main finding is the significant relationship between industry funding and research performance. They also find no relationship between academic publishing and entrepreneurial outputs.

Adams and Griliches (1996a)

Adams and Griliches investigate the relationship between research output and R&D using field- and university-level data from the United States. The authors state that universities account for 50% of basic research in the United States, and argue that basic research is one of the vital components of industrial innovation. The study is based on indicators from the research and commercialization category. Research output is measured by papers and citations related to R&D. Various estimation methods with different specifications are employed in the study, including linear regression, [non-linear least squares regression](#), and [seemingly unrelated regression](#). With the exceptions of mathematics and agriculture, they find that the growth rate of paper and citations and the growth rate of R&D are moving together. Furthermore, they find diminishing returns for R&D at the university- and field-level data.

Kocher, Luptáčík and Sutter (2001)

In this study of research productivity, Kocher, Luptáčík and Sutter measure and compare the relative research efficiency of 21 OECD countries. The authors use indicators from the research and commercialization category. Research output is measured by the number of articles published in 10 leading economic journals with the highest impact factor. Research input is measured by R&D expenditure. The ratio of research output to research input is used to measure efficiency. The empirical analysis involves data envelopment analysis, which is a derivation of linear programming. The authors derive two models from data envelopment analysis, one with constant return-to-scale and one with variable-return-to-scale. The results indicate that the research in the United States is the most efficient with constant-return-to-scale. All other sampled countries exhibit increasing-return-to-scale for research production, implying the possibility of improving efficiency by expanding research activities.

Jaffe and Trajtenberg (1996)

This study analyzes knowledge diffusion by modelling the process generating citations to patents in a double exponential function. Hence, the indicator category used is research and commercialization. Data from a random sample of patents registered in the United States Patent Office are applied to [non-linear least squares regression](#) for estimating the empirical results. The variable of interest is the probability of patent citation over time after a patent is granted. The effects of citing year, cited year, field, country, university-to-corporate differential and government-to-corporate differential are included in the model. The authors find that within-country citations are more numerous than across-country citations. In addition, they find that government patents are cited significantly less than corporate patents, and university patents are the most cited of the three.

Adams and Griliches (1996b)

In their study of U.S. R&D data, Adams and Griliches explore the problems with the existing measures of R&D expenditures, the number of papers published and the number of citations from different U.S. Datasets. All these indicators are from the research and commercialization category. The authors suggest that although it is far from perfect, the new Dataset INST100 assembled by ISI provides the best data available at the moment. After trend and statistical analysis of the data, the authors apply the “best” Dataset to estimate a production function using regression, incorporating time trend issues with log-linear specification. The main finding from the regression analysis is that U.S. academic science is exhibiting diminishing returns in terms of papers produced per dollar of R&D expenditure.

Rogers, Yin and Hoffmann (2000)

An important aspect of research and innovation is the ability to transfer information from post-secondary institutions to private companies. This study aims to develop a measure of technological transfer effectiveness and to determine the characteristics of institutions that score high on this measure. The authors define technology transfer effectiveness as “the degree to which research-based information is moved successfully from one individual or organization to another.” The six variables include the number of invention disclosures

received, the number of U.S. patents filed, the number of licences/options executed, the number of licences/options yielding income, the number of start-up companies and the gross licensing. These indicators fall under the research and commercialization category. The measure is then used to score the 131 universities in the sample. The results show that the effectiveness score is more correlated with the composite measure than the individual variables. In addition, the authors investigate 20 factors that influence the relative effectiveness of technology transfer from U.S. research universities to other organizations. From correlation analysis, five of the factors are shown to have significant correlations with the effectiveness score.

European Union (2005)

The 2005 European Innovation Scoreboard (EIS) of the European Union is a report that evaluates and compares the innovation performance of the 25 EU member states, as well as Bulgaria, Romania, Turkey, Iceland, Norway, Switzerland, the United States and Japan. The EIS consists mainly of trend analyses, with some correlation and regression analyses. The EIS uses an “input/output approach,” analyzing innovation in terms of transformation of innovation assets into innovation returns. Indicators related to innovation drivers, knowledge creation, and innovation and entrepreneurship are used as innovation inputs. Indicators related to application and intellectual property are used as innovation outputs. In total, 26 indicators are considered in the EIS. The indicators used in this report fall under all four of our broad categories: research and commercialization, human resources, innovation environment and innovation outcomes. The EIS distinguishes the sample countries into four types: leading countries, average performers, catching up and losing ground. It also indicates that many countries show signs of catching up. None of these countries, however, are expected to catch up to the average performers by 2010.

OECD (2005a)

This paper builds on the indicators included in the EIS and tries to determine which of these indicators are interrelated. The authors believe several of the EIS indicators are redundant, in that they measure the same phenomena, and should be eliminated. The design of this study is applied statistics, and the authors implement a [factor analysis](#). This study examines indicators from all four broad categories: research and commercialization, human resources, innovation environment and innovation outcome.

The specific indicators analyzed in the factor analysis are S&E graduates, population with tertiary education, participation in lifelong learning, employment in medium-high and high-tech manufacturing and high-tech service, public R&D, BERD, patent applications and patents. The factor analysis also examines various indicators of SMEs, sales of new-to-market products and internet access.

Factor analysis resulted in the elimination of indicators that measured the same phenomena. The remaining indicators were grouped into two categories. The first is knowledge development, a category that includes the indicators of population with tertiary education, participation in lifelong learning, employment in high-tech services, public R&D expenditure, business expenditure on R&D, EPO high-tech patent applications, USPTO high-tech patents granted, SMEs involved in innovation co-operation, share of early stage

venture capital in GDP, and share of manufacturing value added in high-tech sectors. The second category is knowledge application, a category that includes sales of new-to-firm, but not new-to-market, products; innovative expenditure; SMEs using non-technical change; and SMEs innovating in-house. These two categories explained 66% of the variance of innovation. Indicators were excluded if they did not help explain a significant portion of the variance in innovation.

Dietz and Bozeman (2005)

This paper focuses primarily on innovation through industrial experience with universities. The authors explore career patterns within industry, academia and government by examining the publication and patent productivity of scientists and engineers working in university research centres. The design of the study is econometric, using [Poisson](#) and [Tobit](#) regression analyses.³ The broad indicator category explored in this study is research and commercialization. The indicators used are publication and patent counts. Publication counts and patent counts are determined by the number of publications or patents per career-year after the completion of a doctoral degree. The data are collected from the curriculum vitae of scientists and engineers and from the USA Patent and Trademark Office. The main finding is that beginning in the early 1980s, increasing “career transformation” between universities and industry and between academic areas enhances higher publication and patent productivity.⁴

Kim, Lee and Marschke (2005)

This paper examines the influence of university research on industrial innovation, and more specifically, the role of researchers in the diffusion and spillover of ideas from universities to the private sector. The design of this study is econometric, and the statistical analysis is [logit](#) and [probit](#) regression. The indicators examined in this study fall under the broad category of research and commercialization. In particular, the authors are interested in patents, patent citations and dissertation abstracts. These data are combined, using an inventor matching procedure, to determine the inventor’s life cycle productivity in patenting. The authors find that the pharmaceutical industry makes greater use of inventors with university experience than the semiconductor industry, but both industries show an increasing employment of inventors with university experience.

3.3 Discussion of Indicators

Research and innovation are difficult concepts to measure and, as demonstrated by the literature reviewed here, inconsistently defined. The terms are abstract and involve both products and processes. As a result, those who attempt to measure R&I are forced to use what they deem to be indicators of, or proxies for, research and innovation. This section groups the indicators into categories based on those identified by the SCBCIC (2003):

³ Tobit regression is used when the data are truncated above or below specific values.

⁴ Career transformation between universities and industry (between academic areas) means that the person is leaving one job at university (at one academic area) and entering another job at industry (at another academic area) or works concurrently between universities and industry (between academic areas).

research and commercialization, human resources, innovation environment and innovation outcome.

3.3.1 Grouping of Indicators

Identified indicators of R&I are grouped into four main categories, each including common indicators identified by the Science Council of British Columbia and Industry Canada (2003). The Science Council of British Columbia and Industry Canada (SCBCIC) determined an indicator to be common if it was used in three or more of the seven studies they reviewed. It should be noted that some of the studies included in this review use indicators that do not fall under any of the four main categories.

SCBCIC (2003) indicator categories were used for two reasons. First, the SCBCIC report aggregated innovation indicators from several large reports into a single universal set of common indicators in a cohesive and coherent manner. Second, the seven reports used by SCBCIC (2003) are often cited, cover a number of international comparisons and sectors, are relevant to Canada and represent how Canadian provinces measure innovation.

The first indicator category, research and commercialization, includes, but is not limited to, such indicators as total R&D; collaboration between business, higher education and the public sector; publications; patents; university technology commercialization; and total venture capital investment. The human resources category includes education attainment, bachelor degrees in science and engineering, and research personnel. The third category, innovation environment, includes corporate income tax rates, fixed business investment and internet connectivity. The final category, innovation outcome, includes such indicators as employment, pay/earnings and labour productivity. A complete list of the indicators included in each category is listed in Table A below.

Table A

Category	Indicator
Research and Commercialization	Total R&D
	Sectoral R&D performance
	Collaboration between business, higher education and public sector
	Scientific publications; also broken down by sector
	Patents
	University technology commercialization
	Total venture capital investment
	Innovative manufacturing firms

Category	Indicator
	Entrepreneurship
Human Resources	Educational attainment
	Bachelor degrees in science and engineering
	Educational spending
	Student performance in science and math
	Research personnel
	Workers in natural and applied science-related occupations
	Education levels of immigrants
Innovation Environment	Individual tax rate
	Corporate income tax rate
	Cost of R&D/tax subsidies
	Fixed business investment
	Internet connectivity
Innovation Outcome	GDP
	Employment
	Pay/earnings
	Trade
	Labour productivity

3.3.2 Trends in Studies Exploring Innovation

Frequency analysis is used to determine trends in research examining innovation. First we will examine broad category trends. Of the 28 articles included, all examined indicators from the research and commercialization category, nine examined indicators falling under human resources, five examined indicators from innovation environment, and seven examined indicators under the heading of innovation outcome (see Table B). Nineteen articles examined only a single category, in each case research and commercialization. Two articles examined indicators from the same two categories: research and commercialization and human resources. Two articles examined three categories, and of the 28 articles that

examined indicators from the common categories, five examined all four categories (see Table C).

Table B

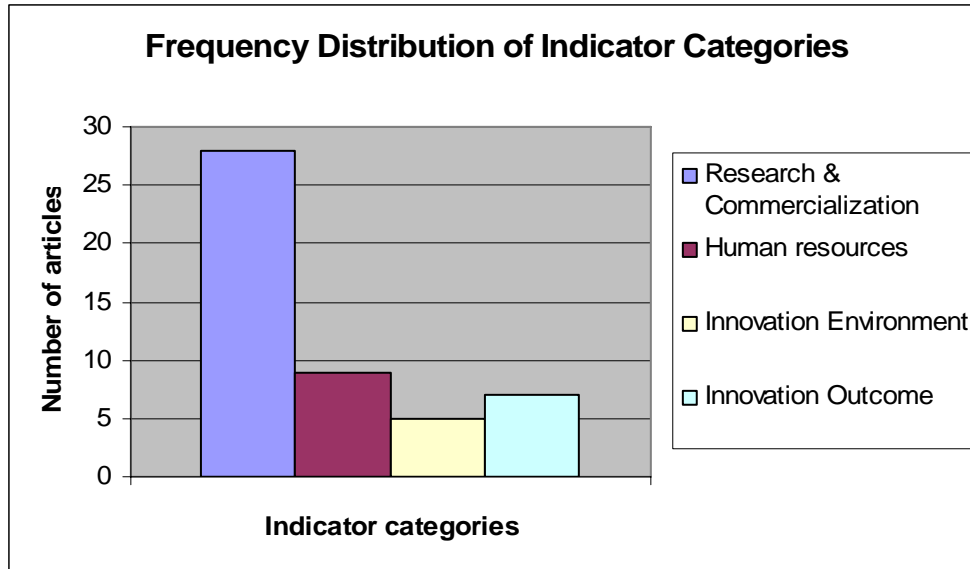
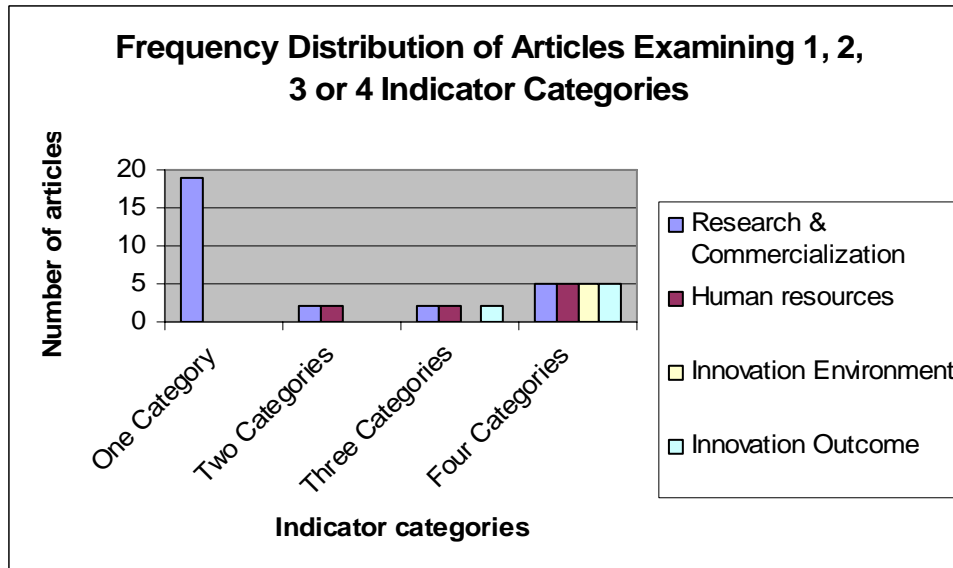


Table C

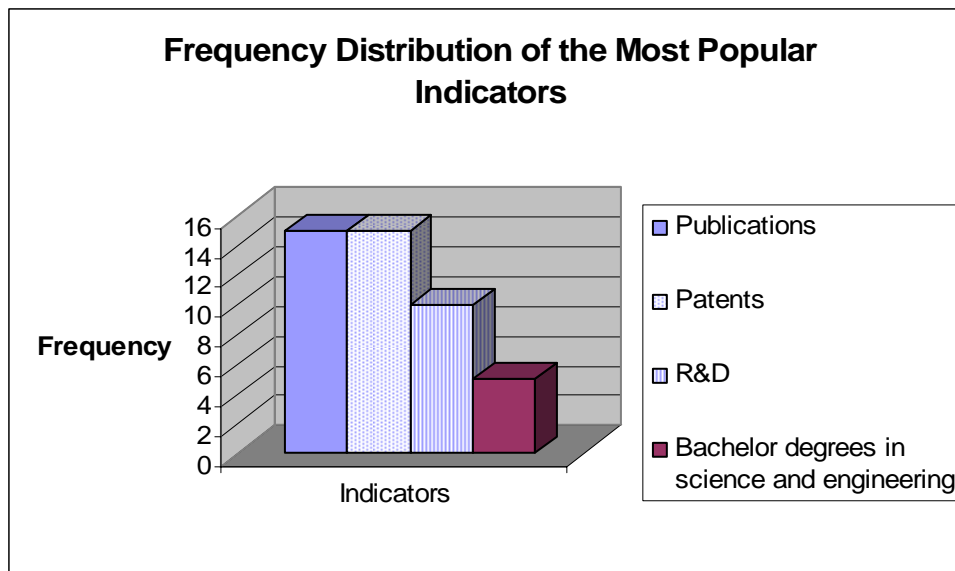


The prominence of the research and commercialization category warrants further discussion. All 28 articles examined at least one indicator from the research and commercialization category. The most common indicators from this category were publications, patents and R&D expenditure. (Among all the articles included in this review, not one examined any of the other indicator categories without also examining the research and commercialization category.) This category appears to be the foundation of research in innovation.

One category of indicators, while used less frequently by researchers, represents a significant trend in this field of study: innovation environment, addressed by only five of the 28 articles reviewed. This category includes indicators such as fixed business investment, internet connectivity and various tax rates. One explanation for the under-utilization of this category might be that tax rates have a high degree of variability. As a result, this indicator may be difficult to compare across different regions and time periods. The internet is also a relatively new technology, and researchers may not recognize that internet access, use or connectivity can be used as a proxy for innovation.

It is also useful to examine which indicators were used most frequently, regardless of category. The top four indicators are represented in Table D. This graph makes it clear that publications and patents are the most frequently used proxies for innovation. R&D expenditure is the third most popular indicator; 10 articles used this proxy.

Table D



4. ANALYSIS

Only studies using empirical methodologies will be evaluated in this section. Those studies considered above as descriptive or analytical/critical, while interesting and informative on the issue of measurements of innovation, do not make empirically grounded statements as to the reliability and ability of indicators to actually measure innovation.

The empirical articles, on the other hand, can be evaluated with regard to a number of dimensions of methodological quality. The majority of these studies are econometric, while only three use other empirical means.

4.1 Characteristics of the Studies

Table E summarizes the characteristics of each of the empirical studies, including the indicators measured; the data source, type and level; and the model specifications and model estimations.

Table E

Characteristics of Methodologies

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Adams and Griliches	1996b	Research and commercialization: Number of scientific papers originating in U.S. universities Number of papers published Number of citations to the published papers	CASPAR Institute for Scientific Information Bureau of Economic Analysis	Panel	Field University by field	Trend analysis Statistical analysis Log-linear regression
Adams and Griliches	1996a	Research and commercialization: Total papers per year; total citations to papers over five years	CASPAR Institute for Scientific Information	Panel	University	Trend analysis Linear regression Non-linear least square (NLLS)
Belderbos, Carree and Lokshin	2004	Research and commercialization: Total innovation expenditure divided by sales	Community Innovation Surveys	Cross-sectional	Firm	Semi-log linear Two-stage least square regressions

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Bilbao-Osorio and Rodriguez-Pose	2004	Research and commercialization: R&D expenditures of private, public and higher-education sector as percentages of GDP Number and the growth rate of patents applications	New Cronos Eurostats	Cross-sectional	Region	Log-linear OLS regressions
Calderini and Scellato	2005	Research and commercialization: Number of patents, number of publications Herfindahl index for patenting Herfindahl index for publications	European Patent Office Institute for Scientific Information	Panel	Country Region	Fixed effects Poisson Negative binomial
Dietz and Bozeman	2005	Research and commercialization: Publications and Patents	Curriculum vitae of 1,200 research scientists and engineers supported by DOE, Department of Defense, and NSR research centres U.S. Patent and Trademark Office	Cross-sectional	Individual	Descriptive analysis Tobit Poisson

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
European Union	2005	<p>Human resources: S&E graduates, population with tertiary education, participation in lifelong learning, youth education attainment level</p> <p>Research and commercialization: Knowledge creation (public R&D expenditure, business R&D expenditure, share of medium-high and high-tech R&D, share of enterprises receiving public funding for innovation, share of university R&D expenditures financed by business sector), intellectual property (EPO patents, USPTO patents, Triadic patent families, new community trademarks, new community designs), SMEs innovating in-house, innovative SMEs co-operating with others, innovation expenditures, early-stage venture capital, ICT expenditures, SMEs using non-technological change</p> <p>Innovation environment: Broadband penetration rate</p> <p>Innovation outcome: Employment in high-tech services, exports of high-tech products as a share of total exports, sales of new-to-market products, sales of new-to-firm but not new-to-market products, employment in medium high and high-tech manufacturing</p>	Eurostat OECD OHIM	Panel	Country	Trend analysis

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Goldstein and Renault	2004	Presence of a top-50 ranked research university Research and commercialization: Number of patents assigned to universities, sum of total research expenditures across all universities Human resources: Sum of all degrees awarded in all higher education institutions	Census CASPAR database	Cross-sectional	Metropolitan statistical area	Test of mean difference Linear OLS regressions
Gulbrandsen and Smeby	2005	Research and commercialization: Academic output (publication of articles in journals, books, textbooks and conference proceedings), commercial output (patents, commercial products, establishment of firms and consulting contracts)	Questionnaire responses	Cross-sectional	Faculty members	Analysis with graphs Logistic
Jacobsson and Rickne	2004	Research and commercialization: Expenditure on R&D in the higher education sector as a percentage of GDP, number of published scientific articles, R&D expenditure in government and private non-profit research bodies Time-expenditure on academic R&D	OECD National Science Foundation	Cross-sectional	Country	Simple regression Correlation
Jaffe and Trajtenberg	1996	Research and commercialization: Probability of patent citation over time	United States Patent Office	Time series	Patents	Double exponential function Non-linear least squares

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Kim, Lee and Marschke	2005	Research and commercialization: Fraction of patents that involve inventors who have university patenting experience, fraction of patents that have inventors on them who have advanced degrees, fraction of patents citing past university-assigned patents	US Patent and Trademark Office ProQuest Digital Dissertation Abstract Securities and Exchange Commission S&P's Annual Guide Thomas Register, Mergent and Corpotech NBER Patent-Citations	Panel	Individual Firm	Trend analysis Logit Random-effects probit
Kocher, Luptáčík and Sutter	2001	Research and commercialization: Number of articles published in 10 leading economic journals (research output) Ratio of research output to R&D expenditure	Journal Citation Reports Der Fischer Weltalmanach Frankfurt am Main Institutes and Research Centres	Cross-sectional	Country	Data envelopment analysis (linear programming)

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Medda, Piga and Siegel	2005	Research and commercialization: Ratio of R&D expenditure to sales	Mediocredito Centrale Survey	Cross-sectional	Firm	Two-stage treatment effect model First stage: bivariate probit sample selection model
Niosi and Banik	2005	Research and commercialization: Average number of patents per firm	Canadian Biotechnology CanCorp SEDAR	Cross-sectional Time series	Firm	Analysis with graphs and statistics

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
OECD	2005	<p>Human resources: S&E graduates, population with tertiary education, participation in lifelong learning, employment in medium-high and high-tech manufacturing, employment in high-tech services</p> <p>Research and commercialization: Public R&D expenditures, business expenditure on R&D, EPO high-tech patent applications, USPTO high-tech patents granted, EPO patent application, USPTO high-tech patents granted, EPO patent applications, USPTO patents granted, SMEs innovating in-house, SMEs involved in innovation co-operation, innovation expenditures, SMEs using non-technological change</p> <p>Innovation environment: Internet access</p> <p>Innovation outcome: Output and markets (share of high-tech venture capital investment, share of early-stage venture capital in GDP, sales of new-to-market products, sales of new-to-firm but not new-to-market products, ICT expenditures, share of manufacturing value added in high-tech sectors)</p>	EUROSTAT EVCA	Cross-sectional	Country	Factor analysis

Author	Year	Indicator(s) for measuring R&I	Data source	Data type	Data level	Model specification(s) and Estimation method(s)
Rogers, Yin and Hoffman	2000	Research and commercialization: Number of invention disclosures received, number of U.S. patents filed, number of licences/options executed, number of licences/options yielding income, number of start-up companies, gross licensing income received	AUTM Licensing Survey for FY 1996	Cross-sectional	University	Statistical analysis Zero-order correlation Multiple correlation Linear regression

4.2 Evaluating Methodological Quality

In this review, we evaluate the quality of an empirical study in three dimensions: the quality of data, the quality of the model and the quality of the result. There are four criteria for data quality, four for model quality, and three for result quality. The data quality criteria are data source, data completeness, representative sample and data description. The four criteria for model quality are type of analysis, model assumption, model specification and selection of variables. The three criteria for result quality are statistical significance, estimation bias and overall objectivity.

Although some of the criteria are related to each other, we believe each one is essential for producing a good study with meaningful results. Therefore, the same scoring weight is assigned to all criteria. A score of “3” indicates a study possesses good quality with reference to a particular criterion. A score of “2” indicates fair quality, and a score of “1” indicates poor quality. In order to determine the studies with the highest qualities, the scoring is downward biased: that is, if a study is judged to be between fair and poor for a particular criterion, it is scored as poor for that criterion.

Table F shows the rubric used by the econometrician to assess the methodological quality of the 17 empirical studies. With 11 criteria, the possible scores range from 11 to 33. A study with a score of 28 or above is considered to be a good study, with a high quality of data, model and result. The stated findings from a good study should be reliable and relevant for policy analysis. A study with a score between 22 and 27 is considered to be a fair study. A fair study only possesses good quality in relation to some of our criteria. While some minor methodological flaws may exist in a fair study, the results still provide some insight into the research question. However, the stated findings should be interpreted with caution for policy analysis. A study with a score of 21 or below is considered to be a poor study. Since a poor study typically has major methodological flaws, the stated findings are unreliable and should not be used for policy analysis.

Table F

Criteria for Assessing the Quality of Data

Facet of Study	Study scored 1 if	Study Scored 2 if	Study scored 3 if
Quality of Data			
<i>Data Source</i>	<p>The data is obtained from surveys/questionnaires conducted by the researcher(s).</p> <p>All sources of data are not documented.</p>	<p>The data set is obtained from a single, external source.</p> <p>Some sources of data are not documented.</p>	<p>The data set is obtained from multiple, external sources.</p> <p>The sources of all data used in the study are clearly documented.</p>
<i>Data Completeness</i>	<p>A substantial amount of data is missing.</p> <p>The missing data seriously affects the study results.</p>	<p>The researcher(s) provides a reasonable explanation for missing data.</p> <p>The missing data is not important to the study.</p> <p>The missing data does not seriously affect the study result.</p>	<p>There are no missing data.</p>
<i>Representative Sample</i>	<p>The chosen sample is a poor representation of the population of interest.</p>	<p>It is uncertain whether the chosen sample could serve as a good representation of the population of interest.</p>	<p>The chosen sample serves as a good representation of the population of interest.</p>

Facet of Study	Study scored 1 if	Study Scored 2 if	Study scored 3 if
<i>Data Description</i>	The researcher(s) does not describe the unit or the definition of the variables.	The unit or definition of the variables is described but not clear.	The unit and the definition of the variables are clearly described.
Quality of Model			
<i>Type of Analysis</i>	The study does not employ any econometric methods and relies solely on qualitative methods (trend analysis and correlation analysis).	The study only uses econometric methods for estimating the results.	The study is a mix of quantitative and qualitative analyses. The researcher(s) mainly uses econometric methods for estimating the results. Some qualitative analyses are provided for enhancement.
<i>Model Assumptions</i>	Assumptions are unreasonable. Assumptions are made without any explanation.	Assumptions are not relevant to the study. Assumptions are non-intuitive. The explanation made by the researcher(s) may not be very convincing.	Assumptions are intuitive in our opinion. Assumptions are used in other relevant studies. Assumptions are necessary and important for the study, and the researcher(s) have provided a reasonable explanation.

Facet of Study	Study scored 1 if	Study Scored 2 if	Study scored 3 if
<i>Model Specification</i>	<p>The model is unable to address the central research question.</p> <p>The specification is uncommon and the researcher(s) does not provide any statistical test.</p> <p>The specification is uncommon and the researcher(s) either does not provide any explanation or provides poor explanation.</p> <p>The chosen specification does not account for the issues regarding the type of data used.</p>	<p>The model is able to address the central research question.</p> <p>Although the researcher(s) does not justify or test the specification, it is common in relevant studies.</p> <p>The specification is consistent with the type of data used by the researcher(s).</p>	<p>The model appropriately addresses the central research question.</p> <p>The researcher(s) tests the validity of the functional form specification.</p> <p>The researcher(s) justifies the specification with reliable references.</p> <p>The specification is a good match of the type of data used by the researcher(s).</p>
<i>Choice of Variables</i>	<p>The model does not include many of the important influential factors.</p> <p>There are no control variables.</p> <p>Proxy variables, if any, are not relevant to their underlying factors.</p> <p>Instrumental variables, if any, are weak.</p>	<p>The model includes many of the important influential factors.</p> <p>Some control variables are missing.</p> <p>Proxy variables, if any, are relevant to their underlying factors.</p> <p>The strength of instrumental variables, if any, is adequate.</p>	<p>The model includes all of the important influential factors.</p> <p>Proper control variables are used.</p> <p>Proxy variables, if any, are highly relevant to their underlying factors.</p> <p>Instrumental variables, if any, are strong.</p>

Facet of Study	Study scored 1 if	Study Scored 2 if	Study scored 3 if
Quality of Results			
<i>Statistical Significance</i>	Estimates that capture statistical significance are not reported. Results are not discussed in terms of statistical significance.	Estimates that capture statistical significance are reported, but the researcher(s) does not discuss the results in terms of statistical significance.	Estimates that capture statistical significance are reported. Results are discussed in terms of statistical significance.
<i>Estimation Bias</i>	The results are biased.	The results may be biased, but the direction of the effects should be reliable.	The results are unbiased.
<i>Objectivity of the Discussion</i>	The researcher(s) discusses the results in a subjective manner. Implications and inferences are made beyond the estimated results. The discussion substantially overstates the estimated results.	The discussion slightly overstates the estimated results.	The researcher(s) discusses the results in an objective manner, such that implications and inferences are made on the basis of the estimated results.

The methodological quality of the 17 studies is summarized below. The studies are organized according to data type and data level.

Non-Econometric Studies

- Niosi and Banik 2005 (firm level cross-sectional data with graphs and statistics)
- OECD 2005a (factor analysis, cross-country data)
- Rogers, Yin and Hoffman 2000 (trend, statistics, correlation, index composition)

Cross-section

Individual

- Gulbrandsen and Smeby 2005 (graph, tables, logistic)
- Dietz and Bozeman 2005 (Tobit, Poisson)

Region

- Bilbao-Osorio and Rodriguez-Pose 2004 (log-linear, OLS)

Metropolitan statistical area

- Goldstein and Renault 2004 (test of mean difference, OLS)

Firm

- Belderbos, Carree and Lokshin 2004 (semi-log linear, 2SLS)
- Medda, Piga and Siegel 2005 (two-stage treatment effects model, bivariate probit sample selection model)

Country

- Jacobsson and Rickne 2004 (simple regression)
- Kocher, Luptáčík and Sutter 2001 (data envelopment analysis)

Time Series

- Jaffe and Trajtenberg 1996 (double exponential function, NLLS)

Panel

Firm

- Kim, Lee and Marschke 2005 (trend, logit, random effect probit)

Region and country

- Calderini and Scellato 2005 (fixed effect, Poisson, negative binomial)

Country

- European Union (trend, statistics, linear regression, index composition)

University and field

- Adams and Griliches 1996a (trend, linear, NLLS, SUR)
- Adams and Griliches 1996b (trend, statistics, linear)

The detailed scores received by each study are included in Table G, below the following summaries.

4.2.1 Empirical studies

Non-econometric studies

Niosi and Banik (2005)

Although the research questions are clearly stated through the four specific hypotheses in the study, the methodology used may not be legitimate for testing these hypotheses. The study does not use regression analysis, which enables one to isolate the effect from all other factors but the factors included in the model. It is possible that the two variables actually correlate with a third variable, rather than with each other: therefore, if one were to isolate the effect of the third variable, the two variables become uncorrelated. Furthermore, the authors only consider the mean of a variable in the statistical analysis. The observed difference between the sample means may be due to sample variation or difference in the true population mean. The authors should also include the standard deviation and conduct statistical tests by [testing the difference of means](#).

As a result, one needs to be sceptical about whether the results from the analysis of trends and statistics support the authors' conclusion; the findings may be overstated. Furthermore, although the sources of data are provided, the authors do not clearly state what data are obtained from which source. Despite these issues, the execution of the study is clear and apparent.

OECD (2005a)

The 2005 OECD Economic Survey of the Netherlands uses [factor analysis](#) to determine 20 variables in the European Innovation Scoreboard Summary Index are “related to the same underlying phenomena.” In this instance, 45% and 21% of the common variation shared by the 20 variables can be explained by the first and the second factors, respectively. The 10

variables correlated with the first factor for greater than 0.7 are loaded onto the first factor. These 10 variables are mostly under knowledge development. The four variables correlated with the second factor for greater than 0.7 are loaded onto the second factor. These four variables are mostly under knowledge application. The 20 OECD countries are ranked according to the first and the second factors. Sometimes it is a concern that factor analysis is entirely empirical: although it can find correlation patterns in any Dataset, it provides no explanation of the relationships. Such a concern does not apply in this case: these 20 variables are obtained from the European Innovation Scoreboard, which uses trend and regression analysis; that already explains the underlying relationships clearly.

Rogers, Yin and Hoffmann (2000)

The central research question is clearly stated by the authors. The authors pose three hypotheses that reflect the purpose of their study, and they provide extensive references from relevant studies. Nevertheless, the study design might not be able to test the hypotheses. Technically speaking, the study does not involve any regression analysis. Most arguments are supported by reference to other studies or by looking at statistics and correlations of relevant variables. As a result, the authors are unable to isolate the effects of other factors. Furthermore, the authors only report the [R-square](#) values of multiple regression, and not the estimated coefficients from the regression. It is unclear whether the statistically significant variables are economically significant. In other words, omitting the regression results strongly affects the [apparentness](#) of the study.

One should also be aware that the results may be overstated. Starting with 21 possible factors, the authors pick out the factors that have the significant correlation with the effectiveness scores and then calculate the multiple correlations of these chosen factors. Furthermore, the variables of the composite measure show high correlation with each other. If two variables are highly correlated, one can only gather a single piece of information from the two variables, because the two variables vary together. Making a six-variable composite measure may not be much different from using one or two of the variables. Lastly, the presentation of the correlation table is confusing. Because of these methodological concerns, the study of Rogers, Yin and Hoffman is only of fair quality.

4.2.2 Cross-Section Studies

Individual

Gulbrandsen and Smeby (2005)

This study is of high quality. The authors clearly state the five hypotheses they investigate. Their assumptions are always backed up by references to other relevant studies. Although the data are collected from questionnaire responses, the authors explain the sampling process, the context of the questions and the response rate. They are also fully aware of the possible sources of [bias](#), and the design of the questions can address the hypotheses. Furthermore, the authors acknowledge the [simultaneity](#) of funding and research, and avoid talking about “impacts” and “effects,” instead using concepts such as “relationships.” If data were available, another way to handle the simultaneity would be to use a more complex regression model, such as [three-stage least squares](#) with instrumental variables. The study

design appropriately incorporates both qualitative and quantitative analysis. When reporting the results, the authors emphasize statistical significance. The presentation of graph and statistics is clear and apparent. One consideration might improve the study: the study only concerns the source and the number of fundings; it is very likely that the size of fundings would also be of interest.

Dietz and Bozeman (2005)

The central research questions are clearly stated, and the study design fits the questions. The authors acknowledge the skewness of the data and use appropriate estimation models. The choice of variables is extensive. The assumptions are appropriate, as the authors often provide multiple references for their use from relevant studies. However, one might be concerned with whether the data came from a random sample. The authors do not describe the selection method for choosing the curriculum vitae of the 1,200 scientists and engineers, and the sample might not be representative, in which case the results could be biased. Nevertheless, the study is of good quality.

Region

Bilbao-Osorio and Rodriguez-Pose (2004)

This study utilized [cross-sectional data](#) on 103 regions in peripheral and non-peripheral areas of the EU and conducted [OLS regressions](#) with [log-linear specifications](#). The central research question is clearly formulated, and the study design fits the research question. However, data from a fairly large amount of regions are missing in the study, due to limited data availability. Furthermore, the authors are unclear about which years are the “beginning” and “ending” periods of analysis. Due to the substantial amount of [missing data](#), it is suspicious that different regions have different “beginning” and “ending” periods of analysis. In such a case, the cross-sectional OLS regression used in the study would not be a legitimate method to produce [unbiased](#) results. As a result, although the empirical models are appropriately constructed, with clear explanation, the quality of the study is much reduced by the poor data quality. Thus, the results should be interpreted with caution.

Metropolitan statistical area

Goldstein and Renault (2004)

The central research question is very clearly stated. The authors propose five specific hypotheses and clearly state the way to test the hypotheses, using the results from [tests of mean differences](#) and linear regressions. The data collection is fairly complete, although the data source is not clearly documented. Furthermore, the authors properly control for other regional variables that affect average earnings. Most importantly, the study design effectively isolates the hypothesized cause by including a dummy variable for identifying whether research universities exist in a MSA along with the control variables. However, the authors fail to explain the reason for segregating the full-sample period from 1969 to 1998 into two unequal sub-periods at 1986. The results would very likely be different if the analysis were made for the full-sample period. The procedure raises some suspicions of [data snooping](#), which in turn leads to bias, since a particular set of data is used deliberately to guide the

results to something the authors want. As a result, the findings should be interpreted with caution.

Firm

Belderbos, Carree and Lokshin (2004)

The central research question is clearly formulated. By including dummy variables for each type of R&D partner, the study design can address the central question. The authors are correct to acknowledge the [endogeneity](#) of spillovers and to use [two-stage least squares regressions](#). The equations for the first-stage regression, which are often omitted, are clearly shown in this study. However, the [apparentness](#) of the study would be improved if the results from the first-stage regression were included, so that the relevancy of the instrumental variables could be assessed. Although the authors provide explanations for the assumptions made, the assumptions are not intuitive and are not supported by references. The Dataset is not complete, but the authors provide a very clear explanation about the [missing data](#).

One should be cautious about survivorship bias in the “random” sample of firms used by the authors. About 6,000 firms in 1998 are matched with information on the firms from the 1996 survey, allowing the explanatory variable to lag two years. The population of firms has changed from 1996 to 1998; some firms are new and some firms have shut down. Consequently, firms that cannot be tracked are not random, and the firms that can be tracked (the ones in the sample) are not random as well. However, it is not clear that a possible survivorship bias would affect the estimated results. Despite the issues mentioned, this study is of high quality, as it is very apparent and well explained. The study design can answer the central research question posed by the authors.

Medda, Piga and Siegel (2005)

This study is of good quality in all aspects. The central research question is clearly formulated. The study design not only can address the question but also controls for potential sources of selection sample bias. The choice of explanatory and control variables is complete and extensive. In particular, the authors control for regional effects related to the location of firms. The description of the methodology and model construction is clear and [apparent](#). The assumptions made in the study are mostly referenced by multiple sources. The results are discussed in relation to statistical significance. In addition, the detailed Dataset and the model design enable the authors to segregate total R&D into external and internal R&D, which enriches the implications of the study’s results.

Only two minor points are not clearly explained. First, the authors do not justify including all firms with more than 500 employees and merely sampling “smaller” firms, with probability based on size. Second, the authors do not explain what is to be considered an “outlier” and to be eliminated from the study. These two issues might be sources of sample selection bias. In general, however, the results are fairly reliable and [unbiased](#).

Country

Jacobsson and Rickne (2004)

Although the central research question is clearly stated, the quality of the methodology used in this study is of poor quality. First, the data source and the [missing data](#) are not clearly explained in the body of the study. The time frame for the data used is not consistent throughout the study. Some tables use 1999 data and some tables use 2001 data. Second, the regression models are not clearly stated in the study. The interpretation for the regression line is misleading as well. The authors argue that Sweden is “placed well below” the regression line (regressing the number of person-years spent on R&D on R&D divided by GDP) by looking at the visual distance on a graph. This is misleading, because the authors do not conduct any statistical test. The regression models do not include any control variables. Furthermore, it is unclear which Dataset is used to estimate the regression models. Finally, the part of study that involves econometrics is quite poorly handled, which drastically affects the quality of the study.

Kocher, Luptáčík and Sutter (2001)

The central research question is clearly stated. The authors have developed a model that is able to overcome the weakness (measuring efficiency) that is common to other studies. As mentioned by the authors, the data envelopment analysis does not assume a specific functional form of the production function, but measures relative efficiency between countries. Although the model fits the research question, the study is only of fair quality. First, the quality of cross-country data is low in general. In this particular case, the authors use data from different sources for a single variable. For example, the data for the variable “manpower in economics” is collected from four different sources, a procedure that can be considered inappropriate. Given the possibility of methodology bias in collecting data, the ideal is to collect data for different variables from various sources. But since different sources use different methods to generate data, data for a single variable should be collected from a single source. Furthermore, the relevance of using the number of universities with at least one economic department as the proxy for the control variable “manpower in economics” is uncertain. Lastly, the authors put very little emphasis on the statistical significance of the results.

4.2.3 Time Series

Jaffe and Trajtenberg (1996)

The study is of good quality. The central research question is clearly stated and the theoretical model can address the question. The assumptions made by the authors are reasonable, although explanations and references are occasionally missing. For example, the data used contains information on the countries in which the inventors of patents reside. However, the authors aggregate the data into five countries (United States, Canada, European countries, Japan, and all other countries) without any explanation. Although aggregation is plausible, as it would effectively reduce the number of [dummy variables](#), the

authors should explain why they use the chosen aggregation pattern. The Dataset is of good quality, containing over 30 years' worth of data with extensive details, suggesting the data are fairly complete. The choice of explanatory variables is also extensive. Lastly, the findings are stated according to the estimated results and statistical significance.

4.2.3 Panel

Firm

Kim, Lee and Marschke (2005)

The central research question is clearly stated by the authors. The study contains both qualitative and quantitative analyses. The authors acknowledge the features of the data and apply the appropriate estimation model. The choice of variables is also suitable for answering the research question. Furthermore, the authors construct the [panel data](#) by matching six different data sources. The construction method is clearly stated, and appropriate assumptions are made concerning the matching method. However, there are still [missing data](#), since the authors are occasionally unable to match the data. Although the authors have considered the time trend in panel data, they do not conduct a proper statistical testing for the time trend. This affects the quality of model specification. Overall, however, the study is of good quality.

Region

Calderini and Scellato (2005)

The central research question is clearly formulated by the authors, and the design of the study can address the question. However, as pointed out by the authors, the major limitation of the study is that the results cannot be applied to other technological research and to innovations in general. In order to avoid selection bias, the authors define the population of interest in a specific area: the wireless telecommunication sector. Another weakness of the study might be counting some articles more than once, in cases where one or more of the authors of a multi-author article had affiliations in different geographical areas. This could be a possible source of bias, because articles with more than one author from different geographical areas would have larger weights than others. An improvement might be to include partial counts. For example, an article with two authors from different geographical areas could be counted as 0.5 for each area.

For model specification, the authors acknowledge the fact that the dependent variable, the number of wireless patents, is a discrete, countable variable, and they use [Poisson](#) and [negative binomial models](#) to address this. Because of the specific assumption (mean and variance are equal) required in using these count data models, the authors might consider using the less restrictive, continuous (normal) model. On the other hand, it is good to use [fixed effects](#) to account for the “region varying but time specific” and “time varying but region specific” effects. In addition, the authors use a solid approach to handle the existence of [autoregression](#) in the model by regressing with lagged variables. Although the data source is not very clearly documented (for example, what is ISI?), the description and explanation of

variables and model construction is extensive. Statistical significance is also addressed in the study. Therefore, the study is overall of fairly good quality, given the discussed limitations.

Country

European Union (2005)

Although only a relatively small portion of the EIS involves regression analysis, the study is of very good quality in all aspects. First, the purpose of the EIS is to compare innovation performance, rather than to investigate some casual effects. As a result, using mostly trend analysis is appropriate. Second, the EIS provides a good explanation of the methodology and emphasizes the validity of measurements and statistical significance. It clearly states the methods used to select the indicators. Furthermore, the data sources are reliable, and the variable description is extremely clear. Lastly, proper control variables are included in the regression analysis. In conclusion, the results of the EIS are reliable. The indicators considered in the EIS should provide useful information on measuring innovation.

University and field

Adams and Griliches (1996a)

The central research question is clearly stated, and the authors estimate the same model with various model specifications. Using too many specifications on the same model may raise suspicions of [data snooping](#)—that is, seeking which specifications generate “desirable” results that fit the authors’ hypotheses. Furthermore, one should be cautious about the representation of the sample. The 109 universities are the largest and have the most successful R&D programs. Therefore, it is not a random sample. As a result, the study results may be biased. On the other hand, the data collection is quite extensive, including 10 years’ worth of data. The analysis contains both trend and regression analysis. Therefore, considering both the strengths and the weaknesses, the study is of fair quality.

Adams and Griliches (1996b)

The central research question is not clearly stated, so it is uncertain if the study design can address the central question. On the other hand, although the arguments do not reference other studies, the discussion of the weaknesses of the existing Datasets is reasonable and convincing. However, the part of the study that involves estimating a production function using regression is unclear, although the authors incorporate the issues regarding time-series data into the regression model.

The authors do not provide a complete list of control variables. Only the descriptive statistics of the variables of interest are provided. The estimated coefficients for all control variables are omitted as well. Therefore, the [apparentness](#) of the study is low. Also, the authors do not report [R-square](#). The actual level of statistical significance used is not reported, and the quality of the regression results is uncertain. Therefore, the study is only of fair quality.

4.3 Summary Table of Methodological Quality

As noted, scores for methodological quality range between 11 and 33. A study with a score of 28 or above is considered to be a good study, with a high quality of data and model and reliable findings. A study with a score between 22 and 27 is considered to be a fair study, possessing high quality in only some of our criteria. Minor methodological flaws may exist in a fair study, but the results can still provide some insight on the research question; however, the stated findings should be interpreted with caution. A study with a score of 21 or below is considered to be a poor study. Since a poor study has major methodological flaws, the stated findings are unreliable.

All the studies included in this review were rated as either good or fair. Eight studies received a score of 28 or above, and nine studies received scores between 22 and 27.

Table G

Methodological Quality Assessment for Empirical Studies

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Adams and Griliches (1996a)	3	2	1	2	3	3	2	2	3	3	2	26	Fair
Adams and Griliches (1996b)	3	3	2	2	3	2	2	1	2	3	3	26	Fair
Belderbos, Carree and Lokshin	3	2	3	3	3	2	3	2	2	3	3	29	Good
Bilbao-Osorio & Rodriguez-Pose	3	1	1	2	2	2	3	2	3	2	3	24	Fair
Calderini and Scellato	2	2	2	3	3	2	2	3	3	2	3	27	Fair
Dietz and Bozeman	3	3	2	3	3	3	3	3	3	2	3	31	Good

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
European Union	3	3	3	3	3	3	3	3	3	3	3	33	Good
Goldstein and Renault	2	3	3	2	2	1	2	3	3	2	3	25	Fair
Gulbrandsen and Smeby	1	2	2	3	3	3	3	3	3	2	3	28	Good
Jacobsson and Rickne	2	2	2	3	2	2	1	2	2	1	2	22	Fair
Jaffe and Trajtenberg	3	3	3	2	3	2	3	2	3	3	3	30	Good
Kim, Lee and Marschke	3	2	2	3	3	3	2	3	3	3	3	30	Good
Kocher, Luptáčík and Sutter	3	2	1	3	3	3	2	2	1	2	2	24	Fair
Medda, Piga and Siegel	2	2	2	3	3	3	3	3	3	3	3	30	Good

Author	Data source	Data completeness	Representative sample	Data description	Type of analysis	Model assumption	Model specification	Choice of variables	Statistical significance	Estimation bias	Overall Objectivity	Total score	Quality
Niosi and Banik	2	2	3	3	2	3	2	3	2	2	2	26	Fair
OECD 2005a	3	3	3	3	3	2	2	3	2	2	3	29	Good
Rogers, Yin and Hoffman	2	2	3	3	1	2	2	2	2	2	2	23	Fair

4.4 Strengths and Limitations of the Three Most Common Indicators

Within the literature captured for this review, most studies did not identify the strengths and limitations of the indicators. Those that did tended to discuss a few of the most prevalent indicators. Bound by these limitations, this section evaluates the strengths and limitations of some of the most common innovation indicators: publications, patents and investment in research and development.

4.4.1 Publications

One common indicator of innovation is publications. The publication of research is seen as a measure of scientific output and, therefore, innovation. The tool used to measure publications is commonly referred to as citation analysis. Publications are usually measured at the individual level, but can also be measured by department or institution. A strength of this indicator is that it is measured objectively, which allows results to be quantified, compared and replicated (Campbell 2003).

Still, this indicator has numerous limitations and requires careful consideration when using and interpreting results. This indicator tends to favour the hard sciences, since these disciplines have, on average, higher publication rates. Other disciplines may have fewer publications due to differing reward structures. Researchers must be careful not to interpret the greater publication output of the hard sciences as an indication that these disciplines are more innovative than other disciplines.

Another limitation is that this indicator may be a superficial proxy for innovation. The underlying assumption is that publications are an index of scientific productivity, which in turn indicates innovation. This proxy may fail to sufficiently capture innovation, however, since publications are only one of the outputs researchers and scientists produce, and not all innovative ideas result in publication (Campbell 2003).

A final limitation of this indicator is that publications may not be the best proxy for scientific progress. The majority of publications make modest contributions to knowledge, while very few make major contributions. However, publication counts are unable to distinguish between the two. It may be that publication counts are a better proxy for university publication expectations than innovative activity (OECD 2005b). This has led to questions regarding this indicator's ability to capture the multitude of objectives that countries pursue through investment in innovation.

The interpretation of this indicator also raises further concerns. Some researchers may fail to consider the size of the institution and the level of funding when interpreting the results. These are important factors that need to be considered, since they can have a significant effect on publication outputs (European Commission 2004). If these factors are not considered, there is a danger of overstating the innovative production of large, established institutions and understating that of small institutions.

4.4.2 Patents

Among the alternative methods of measuring scientific production, analysis of patents has become particularly prominent. Patents are indicators used to measure technological knowledge production and innovation. This indicator has several strengths. The data is available in disaggregated levels, at individual units of analysis, and over an extended period of time. Patents are also considered to be closely linked to innovation, because they are generally used to protect new ideas and products. Moreover, patents allow for the measurement of novel outputs. Patent data is also a readily available, rich source of information that can cover a broad range of technologies (Khan and Dernis 2006). Another strength of patent indicators is that the information is standardized, which allows for ranking of countries and cross-country comparisons (European Commission 2002).

In spite of these advantages, this indicator has numerous limitations. First, patents may reflect corporations' business strategies of corporations as much, or even more, than they reflect scientific output and innovation. It may be that patents are a better indication of protectionist policies than they are of innovation, since the primary purpose of a patent is to protect the economic returns on an invention (European Commission 2002).

Second, this proxy fails to capture important differences between patents. For instance, some patents have a high technological or economic value, while others have no such value (OECD 2005b); accordingly, some patents have industrial application and others do not. Patent indicators fail to distinguish the two (European Commission 2002). Also, not all inventions are patented, so some innovations will not be captured by using this indicator. Lastly, changes in patent law make it difficult to analyze trends over time (Khan and Dernis 2006). Patents may become more or less regulated over time, and therefore more or less restrictive to potential inventors. As a result, using this indicator may capture degrees of innovation that vary over time. Patent regulation also makes it difficult to compare countries on this measure, because countries may have different restrictions on which inventions can be patented.

Additional limitations are present when patent indicators are used for cross-country comparisons. First, the propensity to patent varies by country and industry. This suggests that some countries' innovative production might be overstated, if average patent rates are not considered. Second, patent comparisons between countries have demonstrated a home advantage bias. This occurs when the propensity to patent is higher in the home country than in other countries (Khan and Dernis 2006). Third, there are differences in patent regulations that make it difficult to engage in cross-country comparisons. For example, a specific country's patent regulation may be more inclusive than another country's. This regulation may result in a larger number of patents being filed (European Commission 2002). In this instance, it is the regulation that determines the number of patents, not the degree of innovation.

4.4.3 Investment in R&D

R&D expenditure is used to measure investment in innovation. R&D is typically defined as a creative work undertaken to increase the stock of knowledge. This includes activities such as basic research, product development, process, marketing, publications, venture capital,

organizational, software development, linkages, and the construction and testing of prototypes (OECD 2005c).

The strength of this indicator is that it is the most common measure of the costs associated with innovation. This indicator is also quantifiable and allows for cross-country comparison. It is considered the primary mechanism by which policy-makers affect research and innovation (Salazar and Holbrook 2004).

However, R&D expenditure fails to capture the full extent of the investment in innovation (Holbrook 1997). Some of the aspects of innovation that are not captured by investment in R&D include later phases of development in preproduction, production and distribution, support activities such as training and development activities, acquisition of external knowledge and capital goods (OECD 2005a). Additionally, not all R&D investment is dedicated to basic research and practices fostering innovation. Some of the investment is dedicated to infrastructure and maintaining current mechanisms already in place (Munn-Venn 2006). Critics argue that R&D expenditure is related to technological change and innovation, but fails to measure it. It is suggested that R&D expenditure is one component of the multifaceted aspects that encompass innovation, but that it should not be used in the absence of other proxies (OECD 2005b).

R&D expenditure is also criticized for being biased, inappropriate and ill suited for developing nations. The argument is that developing nations may have low R&D investment, but high investment in other science-related activities. Consequently, some suggest that this indicator is best suited for developed nations (Holbrook 1997).

A final limitation of this indicator concerns its calculation. To examine the effect of investment in R&D, it is necessary to determine how long a period must pass between the initial investment and the observed innovation. Mathematically, this is represented as a time lag. However, researchers are not sure how much time is required to turn an investment in innovation into an innovation outcome (European Commission 2002). The time lag may also vary by sector or by country.

4.4.4 Summary

Several general statements can be made regarding the indicators examined above. A substantial amount of knowledge is not made public, and is therefore not captured by these indicators. Some studies fail to relate the output (publications, patents) with the number of researchers or the associated R&D investment. This may result in the overestimation of the scientific productivity of large, established firms or institutions. R&D investment should be normalized to avoid this problem. Normalizing will ensure that the size of the institution or country and the number of researchers are controlled for. This will allow the results to be comparable across a single standard of measurement.

As mentioned above, these indicators are limited by the comprehensiveness of the Datasets from which the information is derived. It may be beneficial to use multiple Datasets to counter this problem. Researchers and scientists have also expressed some concern that outcome proxies will become the goal of research, rather than a measurement of research.

Studies examining R&I must recognize that indicators are just proxies and should not be viewed as the desired goals (European Commission 2002).

Due to the numerous limitations of these common indicators, it is advisable for researchers examining innovation to use multiple indicators. This will help avoid the limitations associated with individual indicators. The use of multiple indicators also has the advantage of allowing researchers to determine the reliability of specific indicators. If each indicator is capturing innovation, one would expect that a country or institution would have similar output results for each indicator.

4.5 Measurement Techniques: Strengths and Limitations

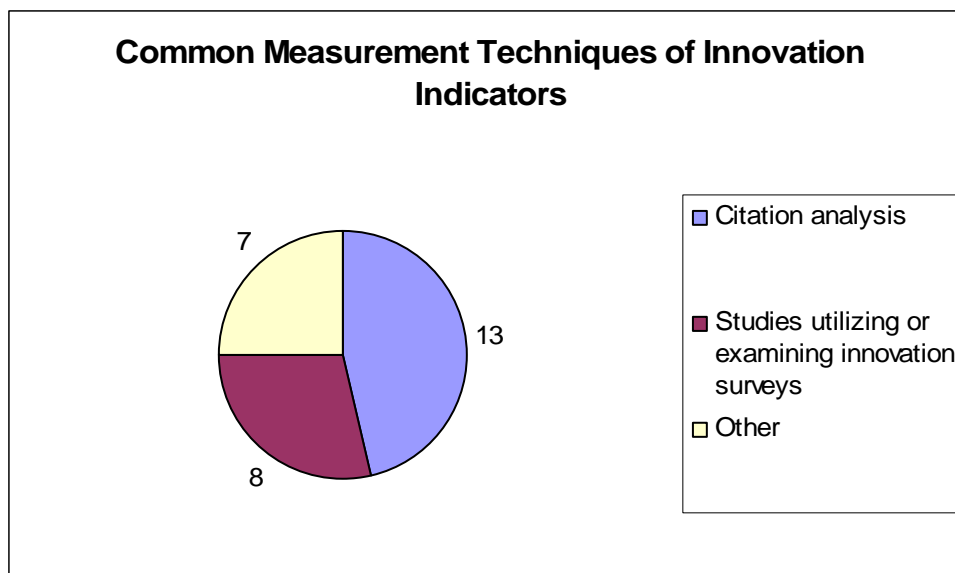
Many of the studies analyzed in this review used citation analysis or innovation surveys to measure innovation indicators. In this section, we discuss the strengths and limitations of these two common measures.

4.5.1 Citation Analysis

The count and quality of a scientific output is primarily determined by citation analysis. Citation analysis is a method used to determine the scientific quality, importance and impact of a scientific output. This technique involves an analysis of the number of publication citations or patent citations produced by a specific group or over a specific period of time (Martin 1996). Of the 28 studies reviewed, 13 employed citation analysis.

The technique is used to measure publication and patent indicators, and to determine a publication's or patent's contribution to scientific progress by counting the number of times the output has been cited in other publications or patents.

Table H



The technique has several strengths. Citation analysis is one of the only quantifiable measures of the quality, impact and importance of scientific output (Martin 1996). Second, the analysis is most useful at the meso or organizational unit of analysis; this allows measurement of the indicator to avoid the problem of magnifying the role of dominant disciplines (Langford 1999).

There are also several limitations: mistaken citations can occur, and publications may also fail to cite earlier scientific classics, although it is unclear to what extent this is the case. Another limitation is that citation rates vary across fields of study and types of publication; this can result in overstating citations in some disciplines. A final limitation is that the majority of citation analysis comes from a single source, the Science Citation Index (SCI). This Dataset includes only international journals. Thus, if a study has a significant impact within a country, but fails to be recognized by international journals, its impact will be missed by the SCI (Gauthier 1998). Further, since most citation analysis comes from a single source, it is subject to problems related to this Dataset. These may include journal selection bias, entry errors and biases that stem from the goals and objectives of the source organization (European Commission 2002).

To help alleviate some of these limitations, it may be necessary to define disciplines and conduct citation analysis subject to each discipline's average publication, patent and citation rates. It may also be beneficial to engage in extensive data cleaning, which will facilitate the elimination of existing errors within the SCI database (Langford 1999).

4.5.2 Innovation Surveys

Another technique for evaluating innovation indicators is surveys. Surveys are perhaps the most popular method of evaluating innovation for the purpose of cross-country comparison (see Table H). A survey is a data collection technique designed to collect standard information on a particular topic. Innovation surveys are conducted for two primary purposes: to benchmark a country's performance in innovation, and to further our understanding of the innovation process. As mentioned in section 1.3, the majority of innovation surveys are based on the Oslo Manual. As a result, innovation surveys have traditionally focused on the outcomes or products of innovation.

Innovation surveys are a popular method of exploring innovation, but they are not without limitations. Surveys based on the Oslo Manual fail to collect data on the dynamics of national or regional systems. This may be in part due to a lack of a single universal definition of a national innovation system. Some people have also questioned whether the firm is the best unit of analysis, given the new theoretical view of innovation as a system (Salazar and Holbrook 2004): if innovation occurs within a system, does it make sense to model surveys around the assumptions of a particular sector?

Some researchers have also questioned the indicators used within surveys to measure innovation. For example, Francoz and Patinson (2000) (as cited in Salazar and Holbrook 2004) suggest that innovation surveys are not producing comprehensive and reliable indicators that can be considered consistent between countries and over time. This result suggests that more indicators need to be utilized and new indicators need to be developed.

Another limitation of innovation surveys is the grouping of firms into innovators and non-innovators. This method fails to capture firms who are involved in the innovation process but have not produced an output (Salazar and Holbrook 2004). This results in the omission of many firms who are engaging in innovative activities that will eventually produce innovative products.

In addition to the above-mentioned criticisms, Salazar and Holbrook (2004) identify some areas not adequately covered by innovation surveys. The first is the diffusion of new technologies. The surveys have been unable to capture the paths or flow of new technologies to commercialization. The second aspect of innovation that has not been sufficiently covered is education, lifelong learning and training. These are key components of the innovation system that contribute to human capital and should be evaluated accordingly. Lastly, in recent years the global economy has undergone a shift from manufacturing industries to service industries, and as a result, innovation surveys have had to include measures of service sector innovation. In the future, innovation surveys will need to include indicators to capture the service sector (Salazar and Holbrook 2004).

4.6 Barriers to Innovation

Based on our discussion thus far, it is apparent that R&I is difficult to measure. Three barriers in particular make innovation difficult for researchers to evaluate effectively. First, it is costly for researchers to study and evaluate innovation. Researchers may be faced with limited resources and/or time constraints. This barrier may explain why few studies captured by this review examined all four indicator categories. And the most time- and resource-efficient approaches may not be the most empirically or methodologically robust.

A second barrier is the difficulty of measuring knowledge transfer. To begin, it is difficult to determine what proportion of research output is transferred to industry (Sobol and Newell 2003). Furthermore, it is a challenge to obtain reliable estimates regarding the nature and measure of research outputs from research institutions. Also, as discussed in the background section, it is difficult to make distinctions between different types of research and different types of knowledge flows. Finally, not all research has commercial potential, so it is important to make such a distinction before attempting to measure innovation (Wolfe 1998a).

The third barrier is the nature of the concept itself. Innovation is considered a dynamic process that consists of both products and processes. The process of innovation is inherently difficult to measure. How do we distinguish between activities that fostering innovation and those that do not? Are all activities that lead to a product part of the innovative process, and how do we quantifiable measure processes? In light of the complexity of these problems, most of the literature examining innovation has focussed on innovative products (Salazar and Holbrook 2004).

Based the findings of section 4.1 and section 4.2, it would appear that most researchers recognize the barriers facing the study of R&I. The next step is to determine how to address these barriers and expand the study of innovation. The section that follows will discuss the principle findings of this review, the criteria for the measurement of R&I and the policy implications.

5. RESULTS AND DISCUSSION

5.1 Summary of Principal Findings

This report explores research and innovation in the context of post-secondary environments. We outline the complexity of defining and measuring innovation and the theoretical underpinnings of research within this field. We explore the common indicator categories and summarize the findings of the studies included in this review. We also highlight some of the trends and gaps that exist in research on innovation research, the strengths and limitations of the common indicators and the barriers to effective measurement of innovation. Based on this analysis, our principal findings are as follows.

5.1.1 General Findings

- There is a lack of consensus on how to define the terms “research” and “innovation.”
- The majority of studies reviewed by this report view innovation as a multidimensional concept that should be treated as a system with linkages and interdependencies. As a result, the consensus in this field is that multiple measures are necessary to capture innovation (Langford *et al.* 2005; Martin 1996; Tijssen 2003).
- R&D is the most common input explored by researchers and is often relied on heavily (Arundel and Hollanders 2006; Colecchia 2005).
- Scientific production is often confused and combined with innovation. This leads to an underlying assumption that if a study measures scientific production it will measure innovation. As a result, the majority of studies measure innovative production and fail to measure innovative processes (Salazar and Holbrook 2004).
- Collaboration seems to be positively associated with innovation (Belderbos, Carree and Lokshin 2004; Gulbrandsen and Smeby 2005; Medda, Piga and Siegel 2005).
- R&I seems to be positively associated with labour productivity and research production (Belderbos, Carree and Lokshin 2004; Gulbrandsen and Smeby 2005; Medda, Piga and Siegel 2005).

5.1.2 Findings Regarding Indicator Use

- Every study reviewed here used at least one indicator from the research and commercialization category to measure innovation. This suggests that innovation is being determined primarily by three measures: R&D expenditure, publications and patents.
- Sixty percent of the studies failed to examine indicators beyond those included in the research and commercialization category. This suggests that other aspects of innovation are under-represented in studies.

- A large portion of studies examined the quality of the scientific output.
- The innovation environment category is under-utilized.

5.2 Summary of Effective Measuring Techniques

In addition to identifying the most frequently used proxies for innovation, this review also seeks to identify the most reliable measurements of those indicators. The results of our review indicate that the most common measurement techniques might not in fact be the most rigorous and robust. More likely, these techniques, such as citation analysis, may simply be the most cost effective. In research and analysis, measurement and method are of utmost importance: they can be the difference between reliable and unreliable statements of findings.

With this in mind, the following guiding questions are offered to those contemplating effective means for measuring research and innovation:

1. *Does the study provide a clear, concise definition for R&I?* Researchers should define what is meant by innovation and the measurement tool used to capture it. It is often difficult to distinguish between research examining scientific production and research examining innovation, because both studies may use the same indicator. If researchers provide a clear definition of innovation and how they plan on capturing it, the scientific study of innovation will be greatly improved.
2. *Does the study use multiple indicators of innovation?* It is clear from this review that a single measure, such as publications or patents, fails to capture all the aspects of innovation. If a study fails to capture the multiple aspects of innovation, the results may not be meaningful. It is possible that the unmeasured aspects of innovation could have significant effects on the outcomes. A more effective study would use multiple indicators of innovation in an attempt to capture the multifaceted nature of the topic. This would include multiple indicator categories and numerous indicators within each category. This approach would give greater confidence that innovation is being measured.
3. *How does the study use R&D, and how much emphasis is placed on this indicator?* Few would argue that R&D is an input related to R&I; however, it needs to be given the proper treatment in order to yield meaningful results. To do this, R&D can be normalized by the number of researchers or by aggregate levels for each region under examination. If R&D is not normalized, innovative performance may be overstated. There is currently a consensus within the scientific community that investment in R&D does not lead to immediate returns in innovation, and the relationship should therefore be lagged (European Commission 2002). However, scientists are unclear how long the lag should be. Further complicating the issue, some sectors or regions may take longer to turn R&D into innovative outcomes. Thus, some sectors may require short lags and others long lags. At this point, there appears to be no agreement on what lags to use, although it is evident that a lag must be used in studies of R&D and innovation. Ideally, studies should compute R&D with multiple lags to determine the effect that differing lags may have on the results.

It is also important to examine the weight that is placed on R&D expenditure. Evaluators should ask two questions: is R&D the only input, and is there a large emphasis on R&D? Although R&D is an important component of innovation, it is not the only component. Future research on innovation needs to move beyond an emphasis on R&D inputs (Arundel and Hollanders 2006). Additional indicators need to be considered in conjunction with R&D, such as the number of research personnel, the institutional or regional infrastructure and linkages with the private sector.

4. *Does the study include a measure of the quality of the scientific output?* Many studies simply provide data on the number of outputs and assume that all the outputs are equal. This is rarely, if ever, the case; some scientific outputs may be highly innovative and others may not be innovative at all. Measurements of quality are limited to two techniques, citation analysis and peer review. Citation analysis is the more prevalent, due to the high costs associated with peer review. However, both peer review and citation analysis techniques have limitations. In spite of these limitations, a study that uses a measure of quality in conjunction with a measure of scientific output is preferred to a study that simply measures scientific output.
5. *Does the study implement statistical analysis to examine the reliability of the indicators?* Some studies use bibliometrics to provide aggregate totals of science indicators, but fail to explore the relationship between the indicators. The effective measurement of R&I should involve more than simple counts of raw data. Numerous methods can be used, including a variety of regression techniques. Statistical analysis allows researchers to determine the relationship between multiple indicators, which is in line with the first criterion and the systems view of innovation.
6. *What aspects of innovation are being explored?* Does the study examine products and processes, or just products? And what unit of analysis is the study examining: is the study primarily concerned with firms or with individuals? These questions will help to determine which areas of innovation are being studied and whether the authors' conclusions are appropriate based on the indicators they used. For example, did the study use indicators for knowledge transfer, collaboration or production?

5.3 Strengths and Limitations of This Review

The largest challenge facing this review and other reviews examining R&I is determining how to define research and innovation. Studies investigating R&I lack a cohesive, universal definition for innovation. These studies also fail to agree on the activities that are encompassed within innovation. To resolve this issue, studies were only included if the authors explicitly stated they were attempting to measure R&I or R&D. This was at times a difficult distinction to make, as innovation is often assumed to be part of scientific production. Studies measuring scientific production were only included if they explicitly stated that innovation was an aspect of production. We are confident that this subtle distinction ensured this review captured only the most relevant literature.

It should also be mentioned that this review only captures literature created by OECD countries and produced in English and French, and that a number of articles were excluded because they could not be retrieved by CCL staff members, who searched the Web and the

University of British Columbia library. Due to time and resource constraints, we were unable to extend the search to other libraries or contact the authors themselves. However, we are confident that the articles captured in this review are representative of the articles we were unable to retrieve. As a result, we believe the captured articles are sufficient to determine effective criteria for the measurement of R&I.

In spite of the above-mentioned constraints, this review has several strengths. It provides an in-depth analysis of the quality of the scientific research obtained. To our knowledge, no other review exists that has examined R&I in this manner. This report also highlights trends and identifies gaps within the research. This information provides important insights into the state of research on innovation and offers directions for future study. Lastly, the research captured within this review is often grounded in complex econometric techniques, and this review presents the results in a concise manner, allowing the information to be understood by people outside this field of study.

6. CONCLUSIONS AND FURTHER CONSIDERATIONS

In recent years, public accountability, a scarcity of public resources and a knowledge-based economy have heightened the importance of evidence-based decision-making and have also led to concerns about the cost-effectiveness of research expenditure. Reliable and valid assessment of returns on investment in the context of research and innovation is therefore essential for jurisdictions seeking to determine whether or not scarce resources are being strategically spent.

Below is a list of further considerations, founded upon the evidence-based research reviewed here, for those contemplating how to measure research and innovation.

1. Encourage research that explores alternative indicators. The majority of studies seem to be exploring indicators from the research and commercialization category. Numerous studies have determined that these current indicators fail to capture some aspects of innovation (Adams and Griliches 1996b; Goldstein and Renault 2004). The identification and testing of alternative indicators such as infrastructure, the process of innovation, regulation and taxes may prove beneficial. Exploring alternative indicators is appropriate, given that research has identified innovation as a multidimensional concept (Langford *et al.* 2005; Martin 1996; Tijssen 2003).
2. Research examining R&I seems to have determined that innovation is best represented by a systems approach (Arnold 2004; Salazar and Holbrook 2004). Consequently, future research should examine multiple indicators. The use of multiple indicators will ensure that the multiple aspects of the system are captured. As a rule, innovation studies should investigate as many indicators for each category as possible. This will reduce the likelihood that some aspects are not being measured. The use of multiple indicators may also allow research to determine which indicators capture more aspects of innovation, and are therefore better proxies for innovation.
3. The use of indicators notwithstanding, the methodological quality assessment of this review determined that eight of 17 studies had good methodological quality. Four of these studies examined R&I and its relationship with labour or research production. Although it is difficult to draw general conclusions about the relationship between R&I and economic outputs, it would appear that R&I has a positive effect on production (Belderbos, Caree and Lokshin 2004; Medda, Piga and Siegel 2005; Gulbrandsen and Smeby 2005). However, the effect may vary depending on the type of production and the type of R&I indicators used. The remaining nine studies were ranked as fair, and not a single study was rated poor. Thus, as a whole, the R&I studies captured by this review exhibited reasonably good methodological quality, and therefore provide a sound empirical foundation for future innovation studies.
4. Of the 17 empirical studies, 14 used econometric analysis, and of the eight studies that received good quality rankings, seven used econometric analysis. This suggests that econometric analysis is one of the most robust methods for effectively measuring R&I. Future studies of innovation will likely continue to focus on econometric methods. An

understanding of this type of methodology is essential to ensure the proper interpretation of research examining R&I.

5. National criteria should be developed to help determine research funding in this area. These criteria could be used to encourage research designed to capture the multidimensional nature of innovation. The criteria would also facilitate research that targets aspects of innovation that are not currently being explored, including, for example, features of the innovation environment. Such an approach may face some challenges, due to the current system's structure, which involves both provincial and federal funding. This system makes it difficult to coordinate and implement national standards for funding research—a result symptomatic of a system characterized by a lack of clearly defined roles and institutional mechanisms (Wolfe 1998b). An alternative to national criteria would be to develop guidelines. The guidelines would be based on criteria, but would fall short of determining who received funding. The guidelines could be distributed to all funding organizations with the expectation that these organizations would consider the guidelines when allocating funding for research. Although strict criteria are preferable to guidelines, a guideline approach would have the benefit of raising awareness.

BIBLIOGRAPHY

Included Articles

- Adams, J. D., & Griliches, Z. (1996a). Research Productivity in a System of Universities. Working Paper No. 5833. Cambridge, MA: National Bureau of Economic Research.
- Adams, J., & Griliches, Z. (1996b). Measuring Science: An Exploration. *Proceedings of the National Academy of Sciences of the United States of America*, 93(23), 12664–12670.
- Belderbos, R., Carree, M., & Lokshin, B. (2004). Cooperative R&D and Firm Performance. *Research Policy*, 33(10), 1477–92.
- Bessette, R. W. (2003). Measuring the Economic Impact of University-Based Research. *Journal of Technology Transfer*, 28(3-4), 355–61.
- Bilbao-Osorio, B., & Rodriguez-Pose, A. (2004). From R&D to Innovation and Economic Growth in the EU. *Growth and Change*, 35(4), 434–55.
- Calderini, M., & Scellato, G. (2005). Academic Research, Technological Specialization and the Innovation Performance in European Regions: An Empirical Analysis in the Wireless Sector. *Industrial and Corporate Change*, 14(2), 279–305.
- Dietz, J., & Bozeman, B. (2005). *Academic Careers, Patents, and Productivity: Industry Experience as Scientific and Technical Human Capital*. Retrieved July 13, 2006, from http://www.rvm.gatech.edu/paperfiles/Dietz_Research_Policy.pdf
- Dosi, G., Llerena, P. & Labini, M. (2005). *Evaluating and Comparing the Innovation Performance of the United States and the European Union*. Retrieved July 12, 2006, from <http://trendchart.cordis.lu/scoreboards/scoreboard2005/pdf/EIS%202005%20EU%20versus%20US.pdf>
- European Commission. (2002). *Final Report of the Expert Group on "Benchmarking S & T Productivity"*. Retrieved July 11, 2006, from http://www.insme.org/documenti/bench_stp_0602.pdf.pdf
- European Union. (2005). *European Innovation Scoreboard 2005: Comparative Analysis of Innovation Performance*. Retrieved July 12, 2006, from <http://trendchart.cordis.lu/scoreboards/scoreboard2005/pdf/EIS%202005.pdf>
- Gauthier, E. (1998). *Bibliometric Analysis of Scientific and Technological Research: A Users Guide to the Methodology*. Retrieved September 21, 2006, from <http://www.statcan.ca/english/research/88F0006XIE/88F0006XIB1998008.pdf>
- Godin, B., Gingras, Y. & Davignon, L. (1998). *Knowledge Flows in Canada as Measured by Bibliometrics*. Retrieved September 21, 2006, from <http://www.statcan.ca/english/research/88F0006XIE/88F0006XIB1998010.pdf>

- Godin, B. (1996a). The State of Science and Technology Indicators in the OECD Countries. Retrieved September 26, 2006, from <http://www.statcan.ca/english/research/88F0017MIE/88F0017MIE1996001.pdf>
- Goldstein, H. A., & Renault, C. S. (2004). Contributions of Universities to Regional Economic Development: A Quasi-experimental Approach. *Regional Studies*, 38(7), 733–46.
- Gulbrandsen, M., & Smeby, J. (2005). Industry Funding and University Professors' Research Performance. *Research Policy*, 34(6), 932–50.
- Holbrook, J. (1997). *The Use of National Systems of Innovation Models to Develop Indicators of Innovation and Technological Capacity*. Retrieved July 11, 2006, from <http://www.sfu.ca/cprost/docs/9706.pdf>
- Jacobsson, S., & Rickne, A. (2004). How Large Is the Swedish “Academic” Sector Really? A Critical Analysis of the Use of Science and Technology Indicators. *Research Policy*, 33(9), 1355–72.
- Jaffe, A. B., & Trajtenberg, M. (1996). Flows of Knowledge from Universities and Federal Laboratories: Modeling the Flow of Patent Citations Over Time and Across Institutional and Geographic Boundaries. *Proceedings of the National Academy of Sciences of the United States of America*, 93(23), 12671–77.
- Kim, J., Lee, S., & Marschke, G. (2005). The Influence of University Research on Industrial Innovation. Retrieved June 13, 2006, from <http://papers.nber.org/papers/w11447.pdf>
- Kocher, M. G., Luptáčík, M., & Sutter, M. (2001). Measuring Productivity of Research in Economics. A Cross-Country Study Using DEA. Working Paper 77. Vienna: Vienna University of Economics.
- Langford, C., Hall, J., Josty, P., Matos, S., & Jacobson, A. (2005). *Outcomes of University Research in Canada: Innovation Policy and Indicators of Triple Helix Relationships*. Retrieved July 12, 2006, from <http://www.thecis.ca/working%20papers/LANGFORD%20Outcomes%20of%20University%20Research%20in%20Canada.pdf>
- Martin, B. R. (1996). The Use of Multiple Indicators in the Assessment of Basic Research. *Scientometrics*, 36(3), 343–362.
- Medda, G., Piga, C., & Siegel, D. S. (2005). University R&D and Firm Productivity: Evidence from Italy. *Journal of Technology Transfer*, 30(1-2), 199–205.
- Niosi, J., & Banik, M. (2005). The Evolution and Performance of Biotechnology Regional Systems of Innovation. *Cambridge Journal of Economics*, 29(3), 343–57.
- OECD. (2005a). *Economic Survey of the Netherlands 2005: Factor Analysis to Identify Inter-related EIS Innovation Indicators*. Retrieved July 12, 2006, from <http://www.oecd.org/dataoecd/57/55/35858453.pdf>

- Rogers, E. M., Yin, J., & Hoffmann, J. (2000). *Assessing the Effectiveness of Technology Transfer Offices at U.S. Research Universities*. Retrieved July 5, 2006, from <http://www.autm.net/pubs/journal/00/assessing.pdf>
- Tijssen, R. J. W. (2003). Scoreboards of Research Excellence. *Research Evaluation*, 12(2), 91–103.
- Van den Berghe, H., Houben, J. A., de Bruin, R. E., Moed, H. F., Kint, A., & Spruyt, E. H. J. (1998). Bibliometric Indicators of University Research Performance in Flanders. *Journal of the American Society for Information Science*, 49(1), 59–67.

Background Articles

- Abbott, M., & Doucouliagos, H. (2004). Research Output of Australian Universities. *Education Economics*, 12(3), 251–265.
- Arnold, E. (2004). Evaluating Research and Innovation Policy: A Systems World Needs Systems Evaluations. *Research Evaluation*, 13(1), 3–17.
- Arundel, A., & Hollanders, H. (2006). *Searching the Forest for the Trees: “Missing” Indicators of Innovation*. Retrieved July 12, 2006, from http://trendchart.cordis.lu/scoreboards/scoreboard2006/pdf/eis_2006_methodology_report_missing_indicators.pdf
- Campbell, D. F. J. (2003). The Evaluation of University Research in the United Kingdom and the Netherlands, Germany and Austria. In P. Shapira, Ed. *Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe* (98–131). Cheltenham, U.K.; Northampton, U.S.: Edward Elgar.
- Colecchia, A. (2005). *What Indicators for Science, Technology, and Innovation Policies in the 21st Century?* Retrieved July 12, 2006, from [http://www.statcan.ca/english/conferences/sciencetech2005/pdf/BSII-Background_En\(Web\).pdf](http://www.statcan.ca/english/conferences/sciencetech2005/pdf/BSII-Background_En(Web).pdf)
- Coulombe, S., & Tremblay, J. F. (2005). Public Investment in Skills: Are Canadian Governments Doing Enough? C. D. Howe Institute Commentary No. 217. Toronto: C.D. Howe Institute.
- Conference Board of Canada. (2002). *The Road to Global Best: Leading Innovation Through R & D Innovation Challenge Paper #2*. Retrieved July 26, 2006, from <http://www.conferenceboard.ca/Boardwiseii/LayoutAbstract.asp?DID=368>
- European Commission. (2004). *Scenarios for the Future of European Research and Innovation Policy*. Retrieved August 1, 2006, from ftp://ftp.cordis.europa.eu/pub/foresight/docs/strata_report_vd5_051004_out_ec.pdf
- Godin, B. (2002). *Are Statistics Really Useful? Myths and Politics of Science and Technology Indicators*. Retrieved July 28, 2006, from http://www.csiic.ca/PDF/Godin_20.pdf

- Godin, B. (1996b). *The Rise of Innovation Surveys: Measuring a Fuzzy Concept*. Retrieved July 28, 2006, from <http://www.thecis.ca/working%20papers/innovationsurveys.pdf>
- Government of Canada. (2003). *Achieving Excellence: Investing in People, Knowledge and Opportunity - Canada's Innovation Strategy*. Retrieved October 07, 2006, from www.innovation.gc.ca/gol/innovation/site.nsf/en/in02427.html
- Khan, M., & Dernis, H. (2006). *Global Overview of Innovative Activities from the Patent Indicators Perspective*. Retrieved July 12, 2006, from <http://lysander.sourceoecd.org/vl=1100215/cl=13/nw=1/rpsv/cgi-bin/wppdf?file=519psecs8bmzp.pdf>
- Koutsoyiannis, A. (1973). *Theory of Econometrics: An Introductory Exposition of Econometric Methods*. London: Macmillan.
- Langford, C. H. (1999). *The Evaluation of Research Done in Post Secondary Institutions*. Retrieved June 14, 2006, from <http://www.cmec.ca/postsec/evaluation.e.pdf>
- Maddala, G. S. (1988). *Introduction to Econometrics*. New York: Macmillan.
- Munn-Venn, T. (2006). *Lessons in Public-Private Research Collaboration: Improving Interactions Between Individuals*. Retrieved July 26, 2006, from <http://www.conferenceboard.ca/boardwiseii/temp/BoardWise2ELOFEFDBAPEIKAGHNLBACHOH2006726174725/002-07%20AIR%207%20RPT-web.pdf>
- OECD. (2005b). *Oslo Manual*. Retrieved July 12, 2006, from <http://213.253.134.29/oecd/pdfs/browseit/9205111E.PDF>
- OECD. (2001). *Main Science and Technology Indicators—Principaux indicateurs de la science et de la technologie, 2001(1)*.
- OECD. (2005c). *Science, Technology and Industry Scoreboard—Towards a knowledge-based economy*. Retrieved July 14, 2006, from <http://puck.sourceoecd.org/vl=3400723/cl=15/nw=1/rpsv/scoreboard/>
- Science Council of British Columbia. (2002). *British Columbia Innovation Benchmarks*. Retrieved July 5, 2006, from http://www.bcinnovationcouncil.com/database/img_4176c01b27168.pdf
- Science Council of British Columbia. (2003). *Project Report: Development of Indicators for Benchmarking Innovation in British Columbia*. Retrieved 7/5, 2006, from http://www.bcinnovationcouncil.com/database/img_4176bd3256680.pdf
- Salazar, M., & Holbrook, A. (2003). *A Debate on Innovation Surveys*. Retrieved 07/28, 2006, from http://www.sfu.ca/innocom/docs/innoCom2003Slides/Salazar_Holbrook.pdf
- Sobol, M. G., & Newell, M. K. (2003). Barriers to and Measurements of the Diffusion of Technology from the University to Industry. *Comparative Technology Transfer and Society*, 1(3), 255–78.

Statistics Canada (2006). *Science Statistics*, 30(5).

Wolfe, D. A. (1998a). *Quality and Accountability in PSE Research: The Measurement Challenge*. Retrieved June 14, 2006, from <http://www.cmec.ca/postsec/quality.e.pdf>

Wolfe, D. A. (1998b). *The Role of the Provinces in PSE Research Policy*. Retrieved June 14, 2006, from <http://www.cmec.ca/postsec/wolfe-e.pdf>

GLOSSARY OF TERMS

Apparentness

All regression results are reported in the study.

Autoregression

In a time series or panel data model, correlation between the errors in different time periods.

Bias

The difference between the expected and the population parameter values of an estimator.

Cross-sectional Dataset

A Dataset collected from a population at a given point in time.

Data mining/snooping

The practice of using the same data set to estimate numerous models in a search to find the “best” model.

Dummy variable

A variable that takes on the value zero or one.

Economic significance

The practical or economic importance of an estimate, which is measured by its sign and magnitude.

Endogeneity

A term used to describe the presence of an endogenous explanatory variable.

Factor analysis

Factor analysis is a statistical technique used to explain variability among observed random variables in terms of fewer unobserved random variables, called factors.

Fixed effects

In a panel data model, “fixed effects” refers to an unobserved variable in the error term that does not change over time. For cluster samples, “fixed effects” refer to an unobserved variable that is common to all units in the cluster.

Functional form misspecification

A problem that occurs when a model has omitted functions of the explanatory variables or uses the wrong functions of either the dependent variable or some explanatory variables.

Logit model

A model for binary response in which the response probability is the logit function evaluated at a linear function of the explanatory variables.

Log-linear specification

A specification that involves a linear relationship such that the dependent variable and all explanatory variables are transformed to logarithms.

Missing data

A data problem that occurs when we do not observe values on some variables for certain observations in the sample.

Negative binomial model specification

Used to estimate count models when the Poisson estimation is inappropriate due to overdispersion.

Non-linear least squares regression

A non-linear model is any model of the basic form in which the functional part of the model is not linear with respect to the unknown regression parameters and the method of least squares is used to estimate the values of the unknown parameters.

Nonrandom sample selection

A sample selection process that cannot be characterized as drawing randomly from the population of interest.

OLS (ordinary least squares)

A method for estimating the parameters of a multiple linear regression model by minimizing the sum of squared residuals.

Panel data

A data set constructed from repeated cross-sections over time.

Poisson model

A model for a count-dependent variable in which the dependent variable, conditional on the explanatory variable, is nominally assumed to have a Poisson distribution.

Probit model

A model for binary responses in which the response probability is the standard normal cumulative density function evaluated at a linear function of the explanatory variables.

Proxy variable

An observed variable that is related but not identical to an unobserved explanatory variable in multiple regression analysis.

R-squared

In a multiple regression model, the proportion of the total sample variation in the dependent variable that is explained by the independent variable.

Seemingly unrelated regression

A set of equations that has contemporaneous cross-equation error correlation.

Simultaneity

A term that means at least one explanatory variable in a multiple regression model is determined jointly with the dependent variable.

Test of mean differences

A hypothesis testing procedure that tests whether the sample means from two population are statistically different.

Time-series data

Data collected over time on one or more variables.

Tobit model

A model for a dependent variable that takes on the value zero with positive probability but is roughly continuously distributed over strictly positive values.

Treatment effects model

Used to estimate models when responses are heterogeneous from the first-stage regression and are segregated into different regressions in the second stage.

Three-stage least squares

Three-stage least squares is a combination of multivariate regression and two-stage least squares. It obtains instrumental variable estimates, taking into account the covariances across equation disturbances as well. The objective function for three-stage least squares is the sum of squared transformed fitted residuals.

Two-stage least squares

An instrumental variables estimator in which the instrument variable for an endogenous explanatory variable is obtained as the fitted value from regressing the endogenous explanatory variable on all exogenous variables

Unbiased

The sample estimates are expected to be equal to the population value.