

# Bridging the Business Data Divide: Insights into Primary and Secondary Data Use by Business Researchers

by Linda D. Lowry<sup>1</sup>

## Abstract

Academic librarians and data specialists use a variety of approaches to gain insight into how researcher data needs and practices vary by discipline, including surveys, focus groups, and interviews. Some published studies included small numbers of business school faculty and graduate students in their samples, but provided little, if any, insight into variations within the business discipline. Business researchers employ a variety of research designs and data collection methods and engage in quantitative and qualitative data analysis. The purpose of this paper is to provide deeper insight into primary and secondary data use by business graduate students at one Canadian university based on a content analysis of a corpus of 32 Master of Science in Management theses. This paper explores variations in research designs and data collection methods between and within business subfields (e.g., accounting, finance, operations and information systems, marketing, or organization studies) in order to better understand the extent to which these researchers collect and analyze primary data or secondary data sources, including commercial or open data sources. The results of this analysis will inform the work of data specialists and liaison librarians who provide research data management services for business school researchers..

**Keywords:** business, primary data, secondary data, graduate students, research data management

## Introduction

A bridge is an apt metaphor for the work of an academic liaison librarian, who acts as a boundary spanner between faculty, students, and the Library. Much of this boundary spanning activity is driven by traditional liaison responsibilities including reference service, information literacy instruction, and collection development. As Canadian academic libraries begin to develop new research data management (RDM) services, liaison librarians have been identified as 'crucial intermediaries between the library's services and its researcher community... [who] often have domain-specific expertise and a network of department-specific relationships' (Steeleworthy, 2014, p.7). Like many of its Canadian peers, Brock University Library has articulated a desire, through its most recent strategic planning exercise, to explore

---

**This study employs content analysis to investigate the research designs and data collection methods**

---

opportunities to support research data management and curation (Brock University Library, 2012). As the liaison librarian to the Goodman School of Business at Brock University, I was quite familiar with the challenges of working with complex, and often expensive, commercial sources of numeric business data such as Compustat and CRSP (Hong & Lowry, 2007), but less familiar with the data practices of business scholars who generated primary data as part of the research life cycle. In order to bridge the business data divide, I needed to acquire evidence-based

insight into business researchers in their dual roles as data producers and data consumers.

A key phase in the development of RDM services is the discovery phase, which documents and analyses current researcher data practices that may be shaped by a variety of factors such as discipline, funding source requirements, research team composition, and career stage (Whyte, 2014). An independent assessment of the management, business, and finance (MBF) research landscape in Canada, commissioned by the Social Sciences and Humanities Research Council (SSHRC), provides some insight into these factors (Council of Canadian Academies [CCA], 2009a). The number of business faculty in Canada was estimated at just over 2,900 individuals working at 58 different academic institutions (CCA, 2009a, p. 14). Business is diverse discipline comprised of many subfields, some of which are more research-intensive than others. A bibliometric analysis of Canadian MBF research output published between 1997 and 2006 found that the accounting subfield represented 14% of business school faculty but produced only 2% of the research output, while the organizational studies and human resources subfield represented 5% of total business school faculty but produced 11% of the research output (CCA, 2009a, p.22). An analysis of research grants administered by SSHRC between 2005 and 2008 calculated that just 1.7% of these grants went to MBF research (CCA, 2009a, p. 18). The Council of Canadian Academies also examined the level of collaborative activity among MBF researchers and found that: (a) 40% of all papers published between 1996 and 2007 were collaborative; and (b), among the top 25 Canadian universities, 45% of collaborative papers had an international co-author. Data management plans are not currently required for SSHRC-funded research, but researchers who collaborate internationally may find themselves subject to data management and sharing policies required by funding agencies in other countries (Corti et al., 2014).

This study employs content analysis to investigate the research designs and data collection methods found in one form of academic business research output, the master's thesis, in order to discover to what extent graduate student business researchers collect primary data, or rely on access to secondary data sources for their analysis, and to explore variations within and between business subfields. In order to distinguish between the terms research strategy (which was not considered in this study), research design, and research method, the following definitions were considered:

- 1 A research strategy refers to 'a general orientation to the conduct of social research' (Bryman et al., 2011, p.579). Commonly cited strategies are qualitative, quantitative, and mixed methods, while other terms used to describe research strategies include strategies of inquiry, traditions of inquiry, or methodologies (Creswell, 2003, p. 13).
- 2 A research design refers to 'a framework for the collection and analysis of data' (Bryman et al., 2011, p. 579). Examples of research designs described in standard accounting, business, and social science research textbooks include experimental, cross-sectional (survey), fieldwork, case study, and archival (secondary analysis) designs (Bryman et al., 2011; Neuman, 2003; and Smith, 2011).
- 3 A research method can be defined as 'simply a technique for collecting data' (Bryman et al., 2011, p. 77) such as self-completion questionnaires, structured interviewing, focus groups, structured observation, ethnography and participant observation, content analysis, and secondary analysis. For consistency's sake, the term 'data collection method' will be used in this study when discussing research methods.

Reliance on primary data collection has implications for the development of research data management services, while reliance on secondary data has implications for data reference support and collection development planning, particularly due to the high cost, proprietary nature, and complex interfaces of many business data sets. This study sets a baseline measurement for data practices at the master's level of business research, and can be used in future studies to compare current data practices at other career stages, or at other institutions, at the disciplinary or sub disciplinary level of analysis. This paper is structured as follows: section 2 reviews the literature related to methods of discovering researcher data practices; section 3 describes the purpose of the study and the research questions I will be exploring; section 4 describes the study's procedures including the setting, and methods of data collection; section 5 presents the findings of the content analysis; section 6 discusses the implications of the findings for research data management, reference support, and collection development; section 7 discusses the limitations of the study; and the final section presents suggestions for future research

## Literature Review

### *Surveys and Interviews*

Academic librarians and data specialists have used a variety of approaches to gain insight into current research data management practices such as case studies (e.g., Key Perspectives, 2010), campus-wide questionnaires (e.g., Parham, Bodnar & Fuchs, 2012), interviews (e.g., Carlson, 2012), and focus groups (e.g., McClure et al., 2014). One study revealed statistically significant differences in research data management practices and attitudes across four research domains (but did not consider discipline-specific distinctions), leading the authors to recommend tailoring data management services using discipline-specific approaches (Akers & Doty, 2013). Another study attempted to examine differences in research data practices by discipline and by methodology but the findings were of limited generalizability due to the low response rate and a survey instrument which confounded research strategies (e.g., qualitative, quantitative, and mixed methods) with research designs (e.g., experimental, survey, field work), and data collection methods (e.g., oral history, textual analysis) (Weller & Monroe-Gulick, 2014).

Motivated by research funding agency requirements for data management plans, most studies have focused on the data curation behaviors and attitudes (such as data preservation and data sharing) of science researchers (e.g., Scaramozzino, Ramirez, & McGaughey, 2012), or if institution-wide, have grouped business scholars with social science or professional schools (e.g., Akers & Doty, 2013), thus providing little, if any, insight into variations within the business discipline. In the next section, I discuss how deeper insight into a researcher's choice of data collection methods and patterns of secondary data use within a discipline can be acquired by conducting a content analysis of scholarly research publications such as journal articles, theses, and dissertations.

### **Content Analysis**

Content analysis, which is a nonreactive or unobtrusive data collection method, enables researchers to overcome some of the weaknesses of survey research, such as low response rates, sampling errors, or unclear question wording (Neuman, 2003). Several studies provided insight into business research designs

at the disciplinary level of analysis. Researchers investigating the prevalence of mixed methods research designs in business and management dissertations conducted a systematic content analysis of 186 Doctor of Business Administration theses and confirmed the use of a diverse set research designs and data collection methods (Miller & Cameron, 2011). In a similar study, McLennan, Moyle, and Weiler (2013) explored the role of economics in tourism postgraduate research by conducting a content analysis of 118 doctoral dissertations completed in the United States, Canada, Australia, and New Zealand between 2000 and 2010. Their examination of the frequency of use of specific research approaches methodologies found that 60% of tourism economics theses used quantitative approaches, 21% used qualitative approaches, and 9% used mixed methods approaches, while their analysis of the data collection methods employed identified a diverse range of techniques including interviews, surveys, case studies, econometric forecasting, observation, and econometric modeling (McLennan, Moyle, & Weiler, 2013, p. 186).

Other content analysis studies explored research trends and practices within specific business subfields (e.g., accounting, logistics and supply chain management), thus providing insight into primary and secondary data use at the sub disciplinary level of analysis. An examination of trends in accounting research over 50 year period found that archival research, defined as 'papers using data from historical market information [such as] stock prices' (Oler, Oler, & Skousen, 2010, p. 668), has been the dominant research methodology in published accounting papers since the 1980s and comprised more than 60% of all papers published between 2000 and 2007. A review of articles published over a two year span in the *Journal of Business Logistics* found that 62% of empirically-based studies used primary data from surveys or case studies, while 21% of studies used secondary data methods (Rabinovich & Cheon, 2011). While logistics and supply chain researchers appear to rely less heavily on secondary sources than do accounting researchers, a broader review of recent research in the logistics and supply chain field identified extensive use of secondary data sources for archival data collection, simulation, content analysis, event studies, and meta-analysis, leading Rabinovich and Cheon to advocate for the extension of traditional secondary data methods to include logistics research.

Several studies conducted by librarians also illustrate the value of the content analysis method in uncovering discipline-specific data practices. Nicholson and Bennett (2009) explored the nature of primary and secondary data use and availability within business ethics research through a content analysis of 48 doctoral dissertations. Their analysis revealed that 51% of the dissertations contained only primary data, 12% relied exclusively on secondary data, and 32% collected both primary and secondary data. A review of primary data collection methods identified four main categories: observations, surveys, experiments, and structured interviews, while a review of the secondary data collected identified a range of data types including numeric datasets, corporate annual reports, government filings and regulatory cases (Nicholson & Bennett, 2009). More recently, Williams (2013) analysed the content of 124 journal articles published by 64 faculty members in crop sciences for evidence of data usage and data sharing, in order to identify faculty candidates for data services. An advantage of the bibliographic study (sic) approach was that it revealed a diversity of discipline-specific data practices, but it was time consuming to conduct, because if data sets were used, they

were typically not cited in the bibliography, but within the text of the article (Williams, 2013, p. 207).

In summary, content analysis is an unobtrusive discovery method which can provide insight into the prevalence of various research designs and data collection methods in order to determine patterns of primary and secondary data use within specific disciplines, but few studies have examined variations between or within business subfields. This study attempts to fill that gap by reporting on the findings of an exploratory content analysis study of business master's theses.

### **Purpose**

The purpose of this study was to investigate the research designs and data collection methods of students in a research-based Master of Science in Management (MSCM) program in order to better understand the extent to which these researchers collected and analyzed primary or secondary data. A content analysis of a corpus of 32 master's theses explored differences between and within subfields of business with respect to research designs and data collection methods. In cases where a thesis used secondary data, attempts were made to identify whether the data sources could be considered open data, or commercial data. This study explored the following research questions:

- 1 What is the distribution of theses by area of specialization and how does it compare to the distribution of core (supervisory) faculty?
- 2 What is the overall distribution of theses by research design and by data collection method? What are the patterns of data collection method use within each type of research design?
- 3 What is the distribution of research designs and data collection methods by area of specialization?
- 4 What is the overall nature of primary and secondary data collection and use (across all specializations)?
- 5 What types of secondary sources are used in business research? Do these researchers use open data sources, proprietary/commercial data sources, or both?

### **Procedures**

#### **Setting**

Brock University is a large comprehensive university located in Canada which offers a wide variety of undergraduate and graduate programs across seven faculties. Brock University's Goodman School of Business (GSB) is accredited by AACSB International and has undergraduate and graduate degree programs in accounting and business administration, an enrollment of 2500 FTE students, and a faculty complement of 95 (Brock University Institutional Analysis & Planning, 2014). The GSB launched a research-based Master of Science in Management program during the 2007/2008 academic year with two goals in mind: first, to prepare students to conduct research in industry and government settings, and second, to prepare students for doctoral level studies in business (Brock University, 2007). The MSCM is a two year program which culminates in a thesis based on independent and original research, and currently offers specializations in accounting, finance, operations and information systems management (O & ISM), (formerly known as management science), marketing, and organization studies. The organization studies stream was first offered during the 2010/2011 academic year (Brock University, 2010). Although Brock University does not currently offer a doctoral level degree in Business, at one point in time the GSB's medium to long term plan included the development of a

research-based doctoral degree in Business, perhaps jointly with another university (Brock University Faculty of Business, 2005).

**Method**

In order to better understand the extent to which business student researchers collect and analyze primary or secondary data, I conducted a systematic content analysis of a corpus of 32 Master of Science in Management theses which were deposited in Brock University Library's Digital Repository<sup>2</sup>. Master's theses must be published in the digital repository as a graduation requirement, so this sample represented 100% of the MSCM degrees awarded since the inception of the program. Each thesis was hand coded using a hybrid approach of manifest and latent coding, similar to the approach taken by Nicholson and Bennett (2009), in order to identify the business subfield, research design, and data collection method employed, and the extent and nature of secondary data use (see Appendix A). The full text of each thesis was reviewed, with particular attention paid to the title page, acknowledgements, abstract, table of contents, methods, and data sections.

Brock University's MSCM program offers five subfields (referred to as areas of specialization): which are (a) accounting, (b) finance, (c) O & ISM, (d) marketing, and (e) organization studies. If the area of specialization was not specifically stated on the title page, a code was assigned based on the topic of the theses, and the home subject area of the student's thesis advisor (who was often cited in the acknowledgements).

Each thesis was coded according to the choice of research design and data collection method and the coding form allowed for the possible use of more than one research design and data collection method, as might be the case in a mixed method research strategy. A core list of research designs and data collection methods was compiled after a review of accounting, business, and social science research methods textbooks (see Appendices B and C).

Finally, each thesis was analyzed for evidence of secondary data use. Each secondary data source was identified by name and by type (i.e., open or commercial / proprietary). Further investigation was required in some cases to determine if a secondary source was a commercial or an open source.

**Findings**

**Distribution of Theses by Area of Specialization**

Table 1 presents a comparison of the distribution of theses and core faculty by area of specialization. The largest proportion of

Table 1 Distribution of Theses and Core Faculty by Area of Specialization			
Area of Specialization	Theses (%)	Core Faculty (%)	Over or Under-Represented
Accounting	5 (16%)	13 (25%)	Under
Finance	15 (47%)	8 (15%)	Over
O & ISM	3 (9%)	8 (15%)	Under
Marketing	7 (22%)	7 (13%)	Over
Organization Studies	2 (6%)	16 (30%)	Under
Total	32 (100%)	52 (100%)	

theses came from the finance area, followed by the marketing area. The proportions for these two areas were larger than one might expect, based on the distribution of core faculty by area of specialization as currently listed on the program's website (Brock University Goodman School of Business, 2015). Three of the five subject area specializations were under-represented (when compared to the distribution of core faculty) including: accounting, O & ISM, and organization studies. The differences in proportions might be a result of several factors such as the relative newness of the MSCM program, the growth of the program over time, and variations in student interest in each of the specialized streams. According to the Appraisal Brief for the MSCM program (Brock University Faculty of Business, 2005), at the time the degree program was proposed there were 21 core faculty distributed across four areas of specialization: (a) eight faculty in accounting, (b); five faculty in finance, (c); five faculty in management science, and (d) three faculty in marketing. Given the two year length of the program, and the fact that the organization studies specialization was not added until the 2010-2011 academic year, it is not as surprising to have just two theses completed in organization studies. The 2014-2015 Graduate Calendar notes that the specialized streams may not be offered every year if there is insufficient student interest (Brock University, 2014).

**Distribution of Theses by Research Design and Data Collection Method**

Three types of research designs were employed in MSCM theses: archival / secondary analysis, survey, and experimental (see Figure 1). There were no examples of case study designs, and none of the theses employed more than one research design. The analysis of data collection method use, as shown in Figure 2, noted three different types of data gathering methods: archival-empirical/quantitative, questionnaires, and archival-content analysis. Patterns of data collection method use within each type of research design appear in Table 2. Both examples of theses with experimental designs used questionnaires for data collection, as did all seven of the theses with survey research designs. Of the 23 theses which employed the archival / secondary analysis research design, only one engaged in a qualitative content analysis, while the other 22 engaged in the empirical analysis of quantitative data. None of the theses employed more than one type of method for the collection of data.

Figure 1 Research Design Use

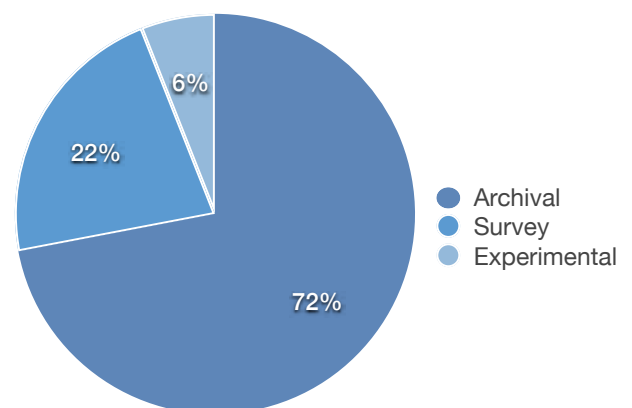


Figure 1 Overall patterns of research design use in MSCM theses (N=32).

Figure 2 Data Collection Method Use

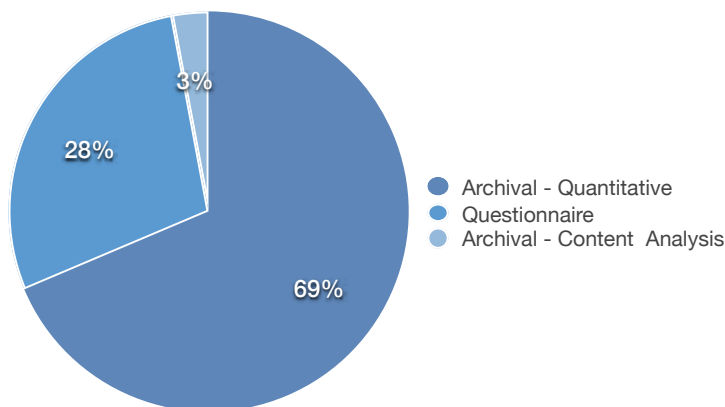


Figure 2 Overall patterns of data collection method use in MSCM theses (N=32).

Research Design	Questionnaire	Archival – Content Analysis	Archival – Quantitative	Total:
Experimental	2 (100%)	0 (0%)	Under	2 (100%)
Survey	7 (100%)	0 (0%)	Over	7 (100%)
Archival	0 (0%)	1 (4.3%)	Under	23 (100%)
Total	9 (28%)	1 (3.15)	22 (68.7%)	32 (100%)

**Distributions of Research Designs and Data Collection Methods by Area of Specialization**

The distribution of research designs and data collection methods by area of specialization are presented in Table 3 and Table 4. The archival /secondary analysis research design was employed at least once within each area of specialization, but was most heavily used within the finance and accounting specializations. Survey designs were employed in four of the five areas of specialization, while experimental designs were used in two of the marketing theses. The marketing area exhibited the widest variety of research designs, while finance used only one type of research design. An analysis of data collection method use by area of specialization revealed widespread use of the archival – empirical/quantitative method, with evidence of use within four of the five areas of specialization.

In order to make sense of the patterns of research design and data collection method use by area of specialization, I also examined the course descriptions for each area of specialization in the MSCM program. Students in all specializations except finance take a two course research methodology sequence which covers topics such as: multivariate statistical techniques, advanced regression analysis, measurement and scaling, survey research and questionnaire design, sampling methods, qualitative research, and structural equation modeling (Brock University, 2014). Students in the finance specialization take a two course sequence in empirical finance which covers empirical research methods and econometric techniques in investment finance (Brock University, 2014). Students in the

accounting stream also take additional courses which cover accounting theory and research methods in behavioural accounting research and market-based research, while the O & ISM specialization includes courses on modeling, data mining, mathematical programming, simulation, and forecasting. Looking again at Table 3 and Table 4, patterns of use begin to emerge, with finance, accounting, and O & ISM theses favouring archival designs and quantitative analysis methods, while marketing and organization studies theses used a variety of research designs and data collection methods. Insights from a case study of an MSc program in finance in the United Kingdom confirmed that students were exposed to secondary data and regression analysis as the model to follow in their own research (Belghitar & Belghitar, 2010, p.578).

**Primary and secondary data collection**

This study also explored the nature of primary and secondary data collection and use across all areas of specialization, and within each specialization. Table 5 presents the patterns of primary and secondary data collection across all areas of specialization. MSCM theses showed a greater reliance on secondary data sources, less reliance on primary data collection, and no evidence of combining primary and secondary data collection, when compared to the Nicholson and Bennett (2009) analysis of business ethics dissertations. Primary data collection methods were used in four of the five areas of specialization, and secondary data sources were used in all five areas of specialization (see Table 6). The Finance area relied exclusively on secondary data sources, as did the majority of

Specialization	Experimental	Survey	Case	Archival	Total
Accounting	0 (0%)	1 (20%)	0 (0%)	4 (80%)	5 (100%)
Finance	0 (0%)	0 (0%)	0 (0%)	15 (100%)	15 (100%)
O & ISM	0 (0%)	1 (33.3%)	0 (0%)	2 (66.6%)	3 (100%)
Marketing	2 (28.5%)	4 (57.1%)	0 (0%)	1 (14.2%)	7 (100%)
Organ. Studies	0 (0%)	1 (50%)	0 (0%)	1 (50%)	2 (100%)
Total	2 (6.2%)	7 (21.8%)	0 (0%)	23 (71.8%)	32 (100%)

Specialization	Questionnaire	Archival – Content Analysis	Archival – Quantitative	Total
Accounting	1(20%)	0 (0%)	4 (80%)	5 (100%)
Finance	0 (0%)	0 (0%)	15 (100%)	15 (100%)
O & ISM	1 (33.3%)	0 (0%)	2 (66.6%)	3 (100%)
Marketing	6 (85.7%)	0 (0%)	1 (14.2%)	7 (100%)
Organ. Studies	1 (50%)	1 (50%)	0 (0%)	2 (100%)
Total	9 (28%)	1 (3.1%)	22 (68.7%)	32 (100%)

**Table 5**  
**Primary and Secondary Data Collection Across All Areas of Specialization (N = 32)**

Description	Number ( % )
Theses with primary data collection	9 (28%)
Theses with secondary data collection	23 (72%)
• Collected only open data sources	2 (6%)
• Collected only commercial data sources	9 (28%)
• Collected both open and commercial sources	12 (37%)
Theses with both primary and secondary data collection	0 (0%)
<b>Number of secondary data sources used</b>	
1 Source	3 (13%)
2 Sources	6 (26%)
3 Sources	3 (13%)
4 Sources	4 (17%)
5 Sources or more	4 (17%)
<b>Total</b>	<b>23 (100%)</b>

**Table 6**  
**Primary and Secondary Data Collection by Area of Specialization, Percentage of Row Totals**

Specialization	Primary	Secondary	Total
Accounting	1 (20%)	4 (80%)	5 (100%)
Finance	0 (0%)	15 (100%)	15 (100%)
O & ISM	1 (25%)	3 (75%)	4 (100%)
Marketing	6 (85%)	1(15%)	7 (100%)
Organ. Studies	1 (50%)	1(50%)	2 (100%)
<b>Total</b>	<b>9 (28%)</b>	<b>23 (72%)</b>	<b>32 (100%)</b>

accounting and operations and information systems management theses. All the theses which collected primary data relied on some form of a questionnaire for data collection (see Table 4). Although not a focus of this study, a latent analysis revealed that a variety of methods were used to collect questionnaire data including printed questionnaires and online software packages (e.g., MediaLab, SurveyMonkey, and Qualtrics). However, in some cases it was not possible to determine if the questionnaires were administered using paper or online instruments.

#### **Secondary sources types used in business research**

This study investigated both the nature and types of secondary sources used and revealed that very few theses relied exclusively on open data sources. 28% of theses used commercial data sources exclusively, and 37% of theses used a combination of open and commercial data sources (see Table 5). Many theses used multiple secondary data sources, including four theses which used five or more secondary data sources (Table 5). A detailed listing of these open and commercial secondary data sources,

including source type, example names, and frequency of use, is presented in Appendix D. The secondary sources included many of the data types identified in the Nicholson and Bennett (2009) study, but given this study's Canadian location, it was not surprising to find some Canadian equivalents, such as SEDAR (for corporate financial reports and filings) and CANSIM (for socioeconomic data). Some, but not all, of the theses that used secondary sources relied on commercial numeric datasets hosted by the Library or Business School (e.g., Bloomberg Professional, CFMRC, Compustat, CRSP, Datastream). Other theses used datasets such as the Internet Retailer's Top 500 Guide, which may have been purchased directly the student or provided by the student's faculty advisor.

## **Discussion**

### **Implications for research data management**

Libraries planning research data services may use a life cycle model to describe the real-world activities of their researchers (Carlson, 2014). The research data lifecycle consists of the following data related activities: discovery and planning, data collection, data processing and analysis, publishing and sharing, long term management, and reusing data (Corti et al., 2014 p. 17). Viewed through the lens of the research data lifecycle, this study's analysis of business master's theses identified variations between and within business subfields with respect to research design and data collection method use which have implications for the development of research data management services for business researchers.

Over 70% of graduate student business researchers could be categorized as data consumers, while less than 30% could be considered data producers. Archival / secondary analysis research designs were employed at least once within each business subfield, and constituted the majority of theses in the accounting, finance, and operations and information system management subfields. The discovery and acquisition of secondary data sources are crucial real-world activities for these business data consumers. A minority of theses employed survey or experimental research designs and

collected research data via questionnaires. Due to the small sample size, clear sub disciplinary patterns could not be found, but the researchers collecting primary data in this study were more likely to be from the marketing subfield. The real-world data activities of these business data producers, such as planning data collection protocols and obtaining informed consent, closely align with social science researchers employing survey or experimental research designs. This segment of researchers could be targeted for future research data management services.

Unlike our counterparts in the United States and the United Kingdom, there is less urgency with respect to developing research data management plans in Canada due to a lack of public funding agency data sharing mandates. The Tri-Agency Open Access Policy on Publication, which applies to peer-reviewed journal publications, was announced on February 27, 2015, but the sharing of publication-related research data is only required by Canadian Institutes of Health Research funding recipients (Tri-Agency, 2015). Barriers to publishing and sharing business data exist due to the

proprietary nature of many of the data sets used in the accounting and financial research which rely on econometric methods. While some top economic journals have mandatory data availability policies which require authors to submit the datasets used in their research, most journals do allow exemptions for research based on proprietary or confidential sources such as Thomson Reuters Datastream (Vlaeminck, 2012, 2013).

#### **Implications for reference support and collection development**

Liaison librarians and data specialists often provide data reference support to novice researchers such as undergraduate students (Partlo, 2010). According to the 'Seven Ages of Research' model, master's degree students (who are in the first age of the researcher's lifecycle) are engaged in research for a limited period of time and may not be seeking a career in academia (Bent, Gannon-Leary, & Webb, 2007, p. 85). Researchers in this early career stage often turn to thesis supervisors for guidance on the research process, academic writing, and information retrieval, to varying degrees of satisfaction (Bent, Gannon-Leary, & Webb, 2007, p. 89). Librarians serving business schools with research-based master's degree programs with an accounting or finance emphasis, where students are expected to engage in archival quantitative research, may want to direct their energy toward providing higher levels of support for secondary data discovery and extraction, and then promoting this service to faculty and graduate students (e.g., Kellam, 2011).

Liaison librarians and data specialists who develop data collections to support academic programs in their institutions will need to work closely with disciplinary faculty in accounting, finance, or other subfields that rely heavily on secondary data, to identify key data sources. Thesis and dissertation content analysis can provide much needed evidence of student usage, and strengthen the argument that these databases support the curriculum (e.g., when writing the library or data services support portions of program appraisal briefs or reaccreditation reviews). A content analysis of faculty journal publications can uncover hidden data sources (e.g., owned or licensed by faculty members for use by their own research groups, but not listed on the Library website or a Departmental resources web page), which may be missed by traditional database benchmarking efforts that examine such listings. One must also keep in mind that traditional citation analysis studies which only look at bibliographies (for a review, see Hoffman & Doucette, 2012), will drastically undercount the use of secondary data sources which are often only cited within the methods sections of dissertations and journal articles. In times of Library materials budget cuts, expensive subscription-based financial datasets with high cost per use may be prime targets for cancellation, but are fundamental to the research process for financial scholars.

#### **Limitations**

There were three aspects of this study which limited the generalization of the findings to the broader population of academic business researchers. The first limitation was due to the small sample size which, while representative of 100% of the population of Brock University's MSCM students, did not reflect the sub disciplinary distribution of core faculty supervisors in the MSCM program. The second limitation was due to the focus of the study on only one career stage of researchers. While a graduate student's choice of research designs and data collection methods may mirror the standard protocols within a discipline, the short duration of master's degree programs may limit his or

her choices (see Belghitar & Belghitar, 2010, p. 579), and therefore may not be representative of research designs employed within doctoral theses or studies published by faculty in peer-reviewed journal articles. The third limitation was due to the use of a single coder, rather than multiple coders. However, the results were strengthened by the nature of the coding scheme, which relied primarily on manifest coding, and focused on measuring the frequency of occurrence of each variable (Neuman, 2003). This exploratory study served as a pilot project to test the thesis coding scheme, which could easily be adapted for broader use at other institutions by using a broader list of business subfields, such as the list of 16 subfields employed by the Council of Canadian Academies in their bibliometric analysis (CCA, 2009b, p. 11).

#### **Conclusion**

This study used the content analysis method to provide insight into primary and secondary data use by master's level business students at one institution. The content analysis found variations in the choice of research designs and data collection methods use across the business discipline, as well as variations within business subfields. The results of this exploratory study can serve as a benchmark for future discipline-specific content analysis studies, and the content analysis method can be used to examine variations in data practices within other social science disciplines. The study could be extended to examine business research from multiple academic institutions, multiple career stages (e.g., doctoral students, faculty, and postdoctoral researchers), or a variety of types of research output including doctoral dissertations and peer-reviewed journal articles. The use of multiple coders would allow for the examination of a larger and more representative sample of the broader business research landscape. Given the Canadian Association of Research Libraries' interest in developing a collaborative, networked approach to building a research data management infrastructure (for a discussion see Canadian Association of Research Libraries, 2013), a collaborative approach to conducting future discipline-specific content analysis studies is recommended.

#### **References**

- Akers, K.G. & Doty, J. (2013). Disciplinary differences in faculty research data management practices and perspectives. *International Journal of Digital Curation*, 8(2), p. 5-26.
- Belghitar, Y. & Belghitar, G.S. (2010). The role of critical evaluation in finance education: Insights from an MSc Programme. *Accounting Education: An International Journal*, 19(6), p. 569-586.
- Bent, M., Gannon-Leary, & Webb, J. (2007). Information literacy in a researcher's learning life: The seven ages of research. *New Review of Information Networking*, 13(2), p. 84-99.
- Brock University (2014). Master of Science in Management Program Description, Brock University Graduate Calendar, 2014-2015. [Online]. Available from: <http://www.brocku.ca/webca/2014/graduate/mgmt.html>
- Brock University (2010). Master of Science in Management Program Description, Brock University Graduate Calendar, 2010-2011. [Online]. Available from: <http://www.brocku.ca/webcal/2010/graduate/mgmt.html>
- Brock University (2007). Master of Science in Management Program Description, Brock University Graduate Calendar, 2007-2008. [Online]. Available from: <http://brocku.ca/webcal/2007/graduate/mgmt.html>
- Brock University Faculty of Business (2005). Appraisal Brief Master of Science (M.Sc.) in Management. [Online]. Available from: [http://www.brocku.ca/webfm\\_send/1289](http://www.brocku.ca/webfm_send/1289)

- Brock University Goodman School of Business (2015). Msc in Management: Core Faculty/Supervisors. [Online]. Available from <http://www.brocku.ca/business/future/graduate/researchdegrees/msc/core-faculty>
- Brock University Institutional Analysis & Planning (2014). Brock Facts and Supplemental Dynamic Reports. [Online]. Available from: <http://brocku.ca/institutional-analysis>
- Brock University Library (2012). Brock University Library Strategic Plan. [Online]. Available from: [https://www.brocku.ca/webfm\\_send/23579](https://www.brocku.ca/webfm_send/23579)
- Bryman, A., Bell, E., Mills, A.J. & Yue, A.R. (2011). *Business research methods*, Canadian Edition. Don Mills, ON: Oxford University Press.
- Canadian Association of Research Libraries (2013). Facilitation, collaboration, and cooperation: A Canadian research data management network. [Online]. Available from: [http://www.carl-abrc.ca/uploads/SCC/Canadian\\_RDMN-Dec-2-2013-summary.pdf](http://www.carl-abrc.ca/uploads/SCC/Canadian_RDMN-Dec-2-2013-summary.pdf)
- Carlson, J. (2012). Demystifying the data interview: developing a foundation for reference librarians to talk with researchers about their data. *Reference Services Review*, 40(1), p. 7-23.
- Carlson, J. (2014). The use of life cycle models in developing and supporting data services. In Ray, J.M (ed.). *Research data management: practical strategies for information professionals*. West Lafayette: Purdue University Press.
- Corti, L., Van den Eynden, V., Bishop, L. & Woollard, M. (2014). *Managing and sharing research data: A guide to good practice*. Los Angeles: Sage.
- Council of Canadian Academies. Expert Panel on Management, Business, and Finance Research (2009a). *Better research for better business*. Ottawa: Council of Canadian Academies. [Online]. Available from: <http://www.scienceadvice.ca/en/assessments/completed/research-business.aspx>
- Council of Canadian Academies. Expert Panel on Management, Business, and Finance Research (2009b). *Better research for better business*. Report Appendices. [Online]. Ottawa: Council of Canadian Academies. Available from: <http://www.scienceadvice.ca/en/assessments/completed/research-business.aspx>
- Creswell, J.W. (2003) *Research design: Qualitative, quantitative, and mixed methods*, Second Edition. Thousand Oaks, CA: Sage.
- Hoffman, K. & Doucette, L. (2012). A Review of Citation Analysis Methodologies for Collection Management. *College & Research Libraries*, 73(4), p. 321-335.
- Hong, E. & Lowry, L. (2007). Business data: issues and challenges from the Canadian perspective. *IASSIST Quarterly*, (Spring), p. 9-13.
- Kellam, L.M. (2011). *Numeric data services and sources for the general reference librarian*. Oxford: Chandos Publishing.
- Key Perspectives (2010). *Data dimensions: disciplinary differences in research data sharing, reuse, and long term viability*. SCARP Synthesis study. [Online]. Edinburgh: Digital Curation Centre. Available from: <http://www.dcc.ac.uk>
- McClure, M., Level, A.V., Cranston, C.L., Oehlerts, B. & Culbertson, M. (2014). *Data curation: a study of researcher practices and needs*. Portal: Libraries and the Academy, 14(2), p. 139-164.
- McLennan, C.J., Moyle, B.D. & Weiler, B.V. (2013). The role of economics in tourism postgraduate research: An analysis of doctoral dissertations completed between 2000 - 2010. *Journal of Applied Economics and Business Research*, 3(4), p. 181-191.
- Miller, P.J. & Cameron, R. (2011). Mixed method research designs: a case study of their adoption in a doctor of business administration program. *International Journal of Multiple Research Approaches*, 5(3), p. 382-402.
- Neuman, W.L. (2003). *Social research methods: qualitative and quantitative approaches*, Fifth Edition. Boston: Allyn and Bacon.
- Nicholson, S.W. & Bennett, T.B. (2009). Transparent practices: primary and secondary data in business ethics dissertations. *Journal of Business Ethics*, 84(3), p. 417-425.
- Oler, D.K, Oler, M.J. & Skousen, C.J. (2010). Characterizing accounting research. *Accounting Horizons*, 24(4), p. 635-670.
- Parham, S.W., Bodnar, J., & Fuchs, S. (2012). Supporting tomorrow's research: assessing faculty data curation needs at Georgia Tech. *College & Research Libraries News*, 73(1), p. 10-13.
- Partlo, K. (2010). The pedagogical data interview. *IASSIST Quarterly* (Winter/Spring), p. 6-10.
- Rabinovich, E. & Cheon, S. (2011). Expanding horizons and deepening understanding via the use of secondary data sources. *Journal of Business Logistics*, 32(4), p. 303-316.
- Scaramozzino, J.M., Ramirez, M.L. & McGaughey, K.J. (2012). A study of faculty data curation behaviors and attitudes at a teaching-centered university. *College & Research Libraries*, 73(4), p. 349-365.
- Smith, M. (2011). *Research methods in accounting*, 2nd edition. Los Angeles, Sage.
- Steeleworthy, M. (2014). *Research data management and the Canadian academic library: An organizational consideration of data management and data stewardship*. [Online]. Partnership: The Canadian Journal of Library and Information Practice and Research, 9(1). Available from: <https://journal.lib.uoguelph.ca/index.php/perj/article/view/2990/3278>
- Tri-Agency (CIRH, NSERC, & SSHRC). (2015). *Tri-Agency Open Access Policy on Publications*. [Online]. 27th February. Available from: <http://www.science.gc.ca/default.asp?lang=En&n=F6765465-1>
- Vlaeminck, S. (2012). *Data policies of economics journals: research data management in economic journals*. [Online]. 10th December. Available from <http://openeconomics.net/resources/data-policies-of-economics-journals/>
- Vlaeminck, S. (2013). *Data management in scholarly journals and possible roles for libraries – some insights from EDaWaX*. *Liber Quarterly*, 23(1), p. 48-79.
- Weller, T. & Monroe-Gulick, A. (2014). *Understanding methodological and disciplinary differences in the data practices of academic researcher*. *Library Hi Tech*, 32(3), p. 467-482.
- Whyte, A. (2014). *A pathway to sustainable research data services: from scoping to sustainability*. In Pryor, G., Jones, S. & Whyte, A. (eds.) *Delivering research data management services: Fundamentals of good practice*. London: Facet Publishing.
- Williams, S.C. (2013). *Using a bibliographic study to identify faculty candidates for data services*. *Science & Technology Libraries*. 32 (2), p. 202-209.

## Notes

- 1 Linda D. Lowry is the business and economics liaison librarian at Brock University in St. Catharines, Ontario, Canada. She can be reached by email: [lloewry@brocku.ca](mailto:lloewry@brocku.ca)
- 2 <https://dr.library.brocku.ca/>



## Appendix A

Content Analysis Coding Form	
Variable Name	Codes
Area of Specialization (circle one)	Accounting (1); Finance (2); Operations and Information Systems Management (3); Marketing (4); Organization Studies (5)
Research Design (circle all that apply)	Experimental (1); Survey (2); Case study/Field Research (3); Archival / Secondary Analysis (5); Other (specify) (6)
Data Collection Method (circle all that apply)	Questionnaire (1); Structured interview (2); Focus group (3); Ethnography/observation (4); Archival-content analysis (5); Archival-empirical/quantitative (6); Other (specify) (7)
Type of Secondary Data Collected	Open only (specify sources or names of data sets) (1); Proprietary only (specify sources or names of data sets) (2); Both (3) (specify sources or names of data sets); Not applicable (primary data only) (4)

## Appendix B

Descriptions of Research Designs		
Research Design	Description	Examples from textbooks
Experimental	Includes classical experimental designs (random assignment, pretest, post-test, experimental group, control group) or quasi-experimental designs (Neuman, 2003, p. 247).	Experimental (Bryman et al., 2011; Smith, 2011; Neuman, 2003).
Survey	“Quantitative social research in which one systematically asks many people the same questions, then records and analyzes their answers” (Neuman, 2003, p. 546).	Cross-sectional or Social Survey (Bryman et al., 2011); Survey (Smith, 2011; Neuman, 2003).
Case Study / Field Research	Case study “entails the detailed and intensive analysis of a single case” (Bryman et al., 2011, p. 571); Field research is “a type of qualitative research in which a researcher directly observes the people being studied in a natural setting for an extended period” (Neuman, 2003, p. 535).	Case Study (Bryman et al., 2011); Fieldwork (Smith, 2011); Field Research (Neuman, 2003).
Archival / Secondary Analysis	Archival: research using secondary sources such as historical documents, texts, journal articles, corporate annual reports, and company disclosures to conduct time-series, cross-section data analysis, content analysis or critical analysis (Smith, 2011); Secondary Analysis: “research in which one does not gather data oneself, but re-examines data previously gathered by someone else and asks new questions” (Neuman, 2003, p.544).	Archival (Smith, 2011); Secondary Analysis (Neuman, 2003).

## Appendix C

Descriptions of Data Collection Methods		
Data Collection Method	Description	Examples from textbooks
Questionnaire	“A collection of questions administered to respondents” (Bryman et al., 2011, p. 579).	Self-completion questionnaires (Bryman et al., 2011); Mail and online surveys (Smith 2011); Mail and self-administered questionnaires (Neuman, 2003).
Structured Interview	“A research interview in which all respondents are asked exactly the same questions in the same order with the aid of a formal interview schedule” (Bryman et al., 2011, p. 581).	Structured interviewing (Bryman et al., 2011); Interviews (Smith, 2011); Telephone and face-to-face interviews (Neuman, 2003)
Focus Group	“A form of group interview in which: there are several participants; there is an emphasis in the questioning on a particular fairly tightly defined topic; and the emphasis is upon interaction with the group and the joint construction of meaning” (Bryman et al., 2011, p. 575).	Focus groups (Bryman et al., 2011; Neuman, 2003).
Ethnography/ Observation	A composite category which includes ethnography (immersion in a social setting) and structured observation (observing and recording behaviour) (Bryman, et al., p. 574, 581).	Structured observation, ethnography & participant observation (Bryman et al., 2011); Complete participant, complete observer, participant-observer (Smith, 2011); Nonreactive / unobtrusive observation; Ethnography (Neuman, 2003).
Archival – Content Analysis	“a systematic analysis of texts (which may be printed or visual) to determine the presence, association, and meaning of images, words, phrases, concepts, and/or themes (Bryman et al., p. 375)	Content analysis (Bryman et al., 2011; Smith, 2011; Neuman, 2003)
Archival – Empirical / Quantitative	The empirical analysis of secondary data sources (such as cross-sectional or time-series data).	Secondary analysis (Bryman et al., 2011; Neuman, 2003); Archival (econometric analysis of cross-sectional or time series data) (Smith, 2011).

## Appendix D

### Open and Commercial Secondary Data Sources Used in MSCM Theses

Classification	Data Type	Examples	Number
Open Data Sources	Canadian Public Company & Mutual Fund filings	SEDAR	5
	Socioeconomic Data	CANSIM; FRED	3
	Stock Exchange websites	FTSE; ME; NASDAQ	3
	Canadian regulatory filings	IIROC	1
	US Government websites	EPA	1
	Bankruptcy Cases	UCLA-LoPucki Bankruptcy Research Database	1
	Other academic datasets	Hasbrouck's Liquidity Estimates	1
	Organizational websites	Ontario Winery websites	1
Commercial / Proprietary Sources			
Library subscriptions	Numeric Databases	Compustat,	8
		Datstream,	7
		CFMRC,	5
		Bloomberg,	5
		CRSP,	4
		Execucomp,	2
		Fundata,	2
		Risk Metrics,	2
		TRACE,	1
		IBES,	1
Carbon Disclosure Project,	1		
Library subscriptions	Bibliographic Databases and Full Text Publications	Lexis-Nexus	4
		TSX E-Review	2
		CBCA	1
		Hoover's,	1
		Law Source,	1
Other (not Library-hosted)	Ecommerce data; web search analytics, financial trading and ownership data	Internet Retailer Top 500 Guide	3
		SpyFu	2
		Keyword Spy	1
		Econoday,	1
		Espeed,	1
		Global Hysales,	1
		13f Spectrum	1