Introduction:
The Use of Census Manuscript Data for Historical Research

KRIS INWOOD
RICHARD REID*

CENSUS AND CENSUS-LIKE enumerations have been undertaken for many centuries and have been used by historians for almost as long. The censuses of early imperial China continue to support new historical research. Annual censuses taken by the Inca in Peru are lost to us, but they may have inspired the Spanish to collect data which has become enormously important in the writing of early Latin American history. Demographic historians of the Mediterranean basin have used population listings dating from the tenth century, and earlier censuses are known to have existed. Indeed, the Norman enquiry that resulted in the cadastral-like Domesday Book apparently was an administrative form traceable to Arab Sicily and Byzantium. The eleventh-century Domesday and the fourteenth-century Poll Tax were followed by intermittent British enumerations at the local and regional levels, long before the national census of 1801. The gradual emergence of the modern national census in a wide range of countries

* Kris Inwood is a professor of economics and Richard Reid an associate professor of history at the University of Guelph.
during the nineteenth and twentieth centuries is well known, as is the use of individual census manuscripts for historical research. In the United States, historians such as James C. Malin in the 1930s and Oscar Handlin and Frank L. Owsley in the 1940s produced major innovative works based on manuscript census research.\(^6\)

The extensive use to which manuscript censuses have been put over the past few decades, by a wide range of researchers, is a reflection of the increasing diversification of academic inquiries. New questions that deal with issues of class, wealth, gender, occupation, political behaviour, and social structures have led researchers to exploit more fully manuscript records formerly neglected because of problems of size and complexity. This research has been facilitated by rapidly accelerating advances in computer technology, which have allowed the gathering and analysis of very large amounts of data in ways that would have been impossible just a short time ago. Indeed, the changes have come so rapidly that long-term projects set up in the 1970s now seem simplistic or restricted when compared to the more sophisticated and technologically driven undertakings of the 1990s. Much more can be expected in the near future.

In the United States, for instance, plans are underway to utilize the capacity of supercomputers, which can handle such extremely large sets of information that it will soon be possible to store and analyze the information recorded on all of the country’s national manuscript censuses. Organized out of the University of Illinois's National Center for Supercomputing Applications, the project intends to use the next generation of optical character scanners to read all manuscript records that exist for the United States censuses from 1790 to 1920. Once stored in the supercomputer, the data, which would include personal census information for approximately 500 million persons, would provide a basis for large-file data manipulation of issues involving the whole population. Researchers using personal computers anywhere could access the supercomputer’s data base, retrieve results from the programmes, and analyze or manipulate smaller data sets. Increasingly sophisticated and task-specific software designed to accommodate researchers with only moderate computer skills will make utilization of the supercomputer’s information even easier.

Other countries will, no doubt, be engaged in similar projects. Researchers will then be able to use the whole population of a country, as evidenced in the manuscript census, for their analyses. In addition, collaborative ventures among research groups, including graduate students, sharing large sets of information, will employ new ways to analyze the data. The potential is enormous, but some caution is necessary. As technology provides the capacity to handle and massage larger amounts of data in shorter time spans,

---

researchers must match that sophistication with an increased awareness of the constraints of their primary source.

As it becomes easier and easier to have access to more and more information, users must become even more alert to just how the information was generated and the ways in which the original collection of the data has created certain limitations and restrictions in what that data can be used to say. The ways in which nineteenth-century censuses were conceived, their intended purpose, how officials structured and changed their categories, and some of the implicit social assumptions deeply buried within the categories must be understood by researchers who would use them. The specific mechanisms involved in the collection of a national census, the length of time taken, the selection of the enumerators, their explicit instructions, and their implicit attitudes will all affect the form and content of the information collected. For these reasons and more, it is critical that researchers become as fully informed about the restrictions as they are about the potentials of census manuscript information. Indeed, greater knowledge about the social constructions imbedded in the manuscript censuses opens the way for wholly new questions to be answered and for an important source to become even more influential.

One way to appreciate the significance of census manuscripts for historical research in Canada is to observe the number and diversity of data sets derived from the census and currently used by historians. In Appendix A we summarize an informal and admittedly partial inventory of Canadian historical micro data bases. We examine only data that describe a national or provincial population and that have figured in recent or current historical research. The list excludes many important data sets describing individual cities, communities, and sub-provincial regions only because they are too numerous for a satisfactory inventory.

It is interesting to observe that the census provides the foundation for all of the Canadian and most of the Ontario micro data bases identified in Appendix A. Even the Ontario probate data being collected by Livio Di Mateo have been linked to the census. The census has been rather less important in Quebec, in part because of the richness of her notarial records and registration of vital events. The inventory for Atlantic Canada encompasses both census and non-census sources, although even non-census sources such as the Gwyn-Siddiq probate records from Nova Scotia and Gail Campbell’s polling, school, and marriage data in New Brunswick have been linked to the census. The wedding of rich detail from other sources to census data, with their great strength of representativeness, is a fertile source of new historical insight.

Our imprecise survey of national and provincial data bases currently or recently used by Canadian historians illustrates the importance of census

---

7 Note, however, the manuscript census component of the important IREP data base.
manuscript data and the diversity of uses to which they are being put. One problem arising with their use is the astonishing lack of information about the workings of the census itself. In Canada, as elsewhere, the Census Bureau tells us very little about census administration in general or, more specifically, enumeration and the editing of manuscript data. A few important studies have begun to fill in the details, but already it is clear that census micro data require careful interpretation just like any other source.

The article by Bruce Curtis provides a useful example of the willingness of historians to examine the context in which a census is taken, the details of the process by which information is collected, and the uses made of the data. Curtis begins his story with the controversy surrounding the Montreal population returns in the 1871 national census. The disappointment and distrust felt by many Montrealers prompted the city to initiate its own check census. Subsequent events provide a fascinating glimpse into the workings of a census and the thinking of contemporaries who clearly cared a great deal about the population count. The story of the Montreal check census also stimulates Curtis to useful epistemological reflections about the historical understanding of census materials.

In one way or another, all of the articles in this issue illustrate the interpretative challenge posed by census manuscripts and the value of that information to the judicious investigator. The earliest census examined here is the mid-seventeenth-century Bohemian population listing discussed by Sheilagh Ogilvie and Markus Cerman. The interpretative difficulty with this source is not so much its incompleteness and unevenness as knowing exactly who the *hausgenossen* were and what their place was in Bohemian household organization and social structure. The Bohemian census illustrates

---

8 Some concern about the "editing" of manuscript information prior to tabulation by historical census authorities is raised by recent research comparing U.S. micro data and published tables; see Susan Carter and Richard Sutch, "Fixing the Facts: Editing of the 1880 United States Census of Occupations with Implications for Long-Term Labor-Force Trends and the Sociology of Official Statistics", *Historical Methods*, vol. 29, no. 1 (Winter 1996), pp. 5-24. Difference between published tables and modern retabulations is evident as well in the Canadian industrial data examined in this issue by Kris Inwood.

a familiar problem: we would like the source to tell us about social patterns, but we need to know something about social patterns (in this case inheritance) to make sense of the source.

Several of the articles concern the ways in which the definition of occupations and activities articulated by the designers of the census, or present in the minds of the enumerators, affect how the researcher can interpret the data. Like Ogilvie and Cerman, Kris Inwood identifies ambiguity in the basic unit of observation described by the census. He examines the definition of what constituted an industrial establishment in the Canadian censuses from 1871 to 1891, the inconsistencies this created in how enumerators recorded production figures, and the difficulties they encountered in identifying multi-product operations.

Harvey Smith looks at a very different issue. In discussing the population records for nineteenth-century rural France, he indicates the ways in which census listings can be seen as reflections of a broader state process. The censuses revealed a shift from a clear language of social rank in 1836 to a complex one of occupations by 1911. It was a transition from censuses prepared under the influence of a local landowning elite to ones designed by state bureaucrats.

Eilidh Garrett is also concerned with the changes in enumeration conventions and in the classification of occupations in published reports. Her study of women's involvement in formal labour markets in Britain from 1891 to 1921 suggests that a changing female occupational structure was masked by the official census publications.

The article by Mike Wayne shows how traditional literature on an issue can be revised by a careful reading of the manuscript census. The belief that large numbers of African-Americans fled the United States to Canada after the Fugitive Slave Law took effect and that they resided in large communities until after the Civil War, when the vast majority returned to the U.S., is contradicted by the 1861 manuscript census. Indeed, even the published census figures, Wayne illustrates, contain a high degree of error.

A different problem of inaccuracy within census returns is raised in Richard Reid's paper, which focuses on the issue of under-representation in census returns. While any census will miss some individuals, nineteenth-century censuses often did not pick up significant numbers of people, and certain groups were more poorly recorded than others. Reid's article looks at one such group — southern African-Americans in the United States census of 1870. Although the under-reporting for this subpopulation was exceptionally large, the reasons why many were missed by enumerators may resonate in other national censuses.

The article by Tracy McKenzie deals, in part, with issues of structural change in agriculture in the post-bellum South and the extent to which the American agricultural schedules are capable of resolving the question. He suggests that the agricultural schedules may better describe changing economic behaviour than structural alterations within southern agriculture.
Several of the articles included in this collection stem from a 1993 conference at the University of Guelph that focused on the use of manuscript data for historical research. In total, two dozen conference papers examined the source’s potential, its limitations, and the methodologies available to extract maximum information from the data. Revised versions of several other papers from the conference appear in the September 1995 issue of *Historical Methods*.

One theme running through all the articles is a caution to interpret census data, whether aggregate or manuscript, with a sensitivity to the biases and idiosyncrasies of the enumeration and tabulation process. At the same time, it is clear that a census, and especially the manuscript record of enumeration, constitutes an important source for historians who respect its limitations and exercise ingenuity to overcome them.

**APPENDIX A**

**Select Canadian and Provincial Micro Data Bases**

**Canada**

*1871 personal census*

Personal information on 60,000 people in 10,000 households (modified random stratified sample of households). Created by Gordon Darroch and Michael Orenstein, Department of Sociology, York University.


*1871 industrial census*

Production records of all industrial establishments (45,000 observations). Created by Kris Inwood (Department of Economics) and Gerald and Elizabeth Bloomfield (Department of Geography), University of Guelph.


*1891 personal census*

Personal characteristics of all unemployed people in seven cities (40,000 observations) and personal characteristics of people in 10 per cent of households in seven cities (35,000 observations). Created by Peter Baskerville and Eric Sager, Department of History, University of Victoria.

1901 personal census
Personal characteristics of people in 10 per cent of households in seven cities (42,000 observations). Created by Peter Baskerville and Eric Sager, Department of History, University of Victoria.


1901 personal census
National sample of schedules 1 and 2 of the 1901 Census, being organized by Project Director Eric Sager.

Québec
1608–1850 vital events
All baptisms, marriages, and deaths for about 1.4 million people in what is now Quebec (pre-1800 data are complete and work continues on the nineteenth century). Created by the Programme de recherche en démographie historique, Université de Montréal.


1792–1835 estate inventories
About 1,000 inventories with compositional detail. Created by Gilles Paquet, Department of Economics, University of Ottawa, and Jean-Pierre Wallot, Archivist and Director, National Archives of Canada.


1851 industrial enumerations
3,456 industrial "units of production" enumerations. Created by Serge Courville, Jean-Claude Robert, and Normand Séguin.

Sample publications: Serge Courville, Jean-Claude Robert, and Normand

1842–1971 vital events and personal census
660,000 vital events 1842–1871, personal census information 1851–1871, and related data describing individuals in the Saguenay region and adjacent parts of eastern Quebec. Created by the Centre interuniversitaire de recherches sur les populations (SOREP), Director Gérard Bouchard.


1800–1971 marriage records
All marriage records for Quebec during the nineteenth and twentieth centuries (about half of the 4.3 million marriages have been entered). Created by the Centre interuniversitaire de recherches sur les populations (SOREP), Director Gérard Bouchard.

Atlantic Canada
1783–1810 New Brunswick land grants and census schedules
9,300 land grants and 400 households in three censuses. Created by The Loyalist Project, Department of History, University of New Brunswick.

1820–1937 British registry of ships and shipowners
Description of 65,000 ships and 408,000 owners, with personal and share information on the latter. Created by The Atlantic Canada Shipping Project.


1820–1914 crew lists and voyages
Wage agreements for 30,000 voyages from four Atlantic Canadian ports involving 250,000 sailors and 95,000 ports of call; the data include personal information about the crew members (such as birthplace, previous experience, reason for discharge). There is also a one-per-cent sample of all crew lists for the British Empire apart from Canada. Created by The Atlantic Canada Shipping Project.

**1851–1871 New Brunswick census schedules, 1845–1870 marriage records, 1845–1857 poll books**

Nominal census records increase from 31,000 people in 1851 to 38,000 in 1871, 6,000 poll book entries, 5,000 marriages, school attendance. Created by Gail Campbell, Department of History, University of New Brunswick.


**1851/1871 Nova Scotia personal census and 1852/1872 probates**

Wealth of 240/471 decedents with personal characteristics from the census of the preceding year. Created by Julian Gwyn, Department of History, University of Ottawa, and Fazley Siddiq, Department of Economics, Dalhousie University.


**1871–1921 Nova Scotia personal census**

Personal and household data for household heads and anyone returning an occupation in three towns, 40,000 observations. Created by Del Muise, Department of History, Carleton University.


**1901–1914 strikes**

411 strikes in the Maritimes recorded by the Canadian Department of Labour. Created by Ian McKay, Department of History, Dalhousie University.

Ontario
1852 household and farm census
2,600 farms with production data and personal data of farm family members (20 per cent random sample of each county). Created by Bill Marr, Department of Economics, Wilfrid Laurier University.

1861 household and farm census
1,100 farms with production data and personal data of farm family members (random within settlement strata), now being expanded to 1,600 observations. Created by Marvin McInnis, Department of Economics, Queen’s University.

1861 and 1871 personal and household census
34,000 and 40,000 individuals plus other household members with personal farm, industrial, and other enumerated information (random sample by surname clusters in central Ontario); 16,000 individuals from the two years have been linked and different schedules within each of the censuses are linked. Created by Gordon Darroch and Michael Orenstein, Department of Sociology, York University.

1871 personal census
5,600 individuals with personal characteristics and assets. Created by Gordon Darroch, Department of Sociology, York University, and Lee Soltow, Department of Economics, Ohio University.
Sample publications: Gordon Darroch and Lee Soltow, Poverty and

1871 personal census
Personal data for 417,000 family heads and household members with other surnames (100 per cent or all records). Created by Bruce Elliott, Department of History, Carleton University, and 400 members of the Ontario Genealogical Society.


1871 personal, farm and industrial census
Personal data for members of households in which at least one person is named as a proprietor in the industrial schedule; all female proprietors and a random 10 per cent of male proprietors in the province (5,000 households total). Created by Kris Inwood, Department of Economics, University of Guelph.

1891 personal census and 1892 probates
3,515 probates from 1892 deaths linked to 1891 census. Created by Livio Di Matteo, Department of Economics, Lakehead University.


Saskatchewan
1881 and 1891 personal census
Personal characteristics of all reported residents (10,000 observations entered to date; in process). Created by Dave DeBrou, Department of History, University of Saskatoon.