Canadian Century Research Infrastructure Project
Occupation and Industry Background

Root variable: Occupation and Industry

Synopsis of code set construction:

In coding occupations and industry (place of employment) CCRI has used the code sets used by IPUMS: OCC50 and IND50. These were based on the industry and occupation code sets created by the US Census Bureau for the 1950 US census. Since the two sets of codes are interdependent (the correct code for occupation often depends on knowing the industry entry), we coded both fields simultaneously: coders had before them all unique combinations of enumerator entries in both fields. To code each field separately would have resulted in an unacceptable degree of variation and imprecision.

It is important to note that when analyzing occupations users must use both fields: occupation and industry. The codes entered are not fully detailed occupation codes (a unique 6-digit code – 3 under occupation and 3 under industry – for every unique combination of enumerator entries in the two fields). As an example, several different occupation titles are coded as 690 (operative and kindred workers n.e.c.) under occupation. We have tried, however, to avoid excessive aggregation or placing workers into the 690 category where the information in both fields allows a more precise coding.

There are some limits to the OCC50 and IND50 codes that we borrow from IPUMS. Users have the option of applying their own code sets, such as HISCO, if they wish.

We have made small adjustments to the OCC50 and IND50 codes taken from IPUMS. These changes are noted in bold in the code lists in part 2. Our changes were usually made to respect particular Canadian realities. Thus we have added a code under industry for Hunting/trapping, in order to avoid placing workers in this sector with agriculture.

There were instances when IPUMS would code industries or occupations in different ways when the occupations entered were not different. We decided to create more strict guidelines for ourselves in some of these circumstances. There were, therefore, times when we differed in how we would code an occupation versus how IPUMS would code an occupation. Following are a few examples:

A. OCC Blacksmith IND Blacksmith Shop
1. IPUMS Response: OCC 501 "Blacksmith" IND varied between 817 "Misc. repair services", and multiple manufacturing industries including 399, 499 and 347.
2. CCRI Response: OCC 501 "Blacksmith" IND 399 "Misc. manuf industries"
3. DETAILS: We felt consistency in how the industry was labeled would help the end user. If a Blacksmith worked for a specified industry, such as, railroad repair shop, we would code it as such, 379 "Railroad and misc. transportation equipment", otherwise we would use 399 "Misc. Manuf industries."

Canadian Century Research Infrastructure Project
Occupation and Industry Background

B. OCC Warden IND forestry department
1. IPUMS Response: OCC 53 "Forester" IND 116 "Forestry"
2. CCRI Response: OCC 53 "Forester" IND 916 "Federal Govt"
3. DETAILS: This is the only gov't department that IPUMS would code as such, for example the Ministry of Agriculture would be coded as a Gov't Level not as 105 Agriculture, therefore, to be consistent we coded it as a level of govt.

Understanding the numeric code:

This code is six digits long. The first three numbers indicate the occupation. The last three numbers indicate the industry. There are no parent codes.

Code set application notes:

Codes added by CCRI

Code	Ind/Occ	Title	Details
106	IND	Own Farm	Was used specifically if "own farm" was written in the industry column to differentiate it from other farmers.
127	IND	Hunting/Trapping	Was added as hunters/trappers were coded by IPUMS as "agriculture/farmers". The IND 127 code was used with OCC 970 "labourer". We thought this highlighted a difference between IPUMS and the Canadian context of CCRI.
247	IND	Bridges and Roads	Was added as bridges/roads to differentiate it from construction (where IPUMS would put all bridge/road employees)
569	IND	Wharf, dock, harbour	IPUMS would split these employees between working for 546 Water Transportation or 568 Services Incidental to Transportation but we decided a specific code would cover their specific location better.
628	IND	HBC	Added for employees working for the Hudson's Bay Company as in many cases, such as "clerks" or "laborers", we did not know if they worked in retail, manufacturing, or transportation. This highlighted a difference between IPUMS and the Canadian context of CCRI.
981	IND	At home (not housework)	Was added as there were numerous individuals who worked in a variety of occupations ranging from "basket weaver" to "carpenter" to "seamstress" to "novelist" who listed "at home" as their industry. The "at home" code in IPUMS was 982 "Housework at

			home" which we did not cover the fact that these individuals were engaged in an occupation at home.
979	IND	With family member	As with 981 "at home" we added this industry code as there were numerous individuals who worked "with mother/father/brother" at an occupation that would be hard to code without it, such as a "labourer" who worked "with father".
988	ND	Jail	Added as IPUMS would code all Jail employees as working for different levels of Government. In this way a "cook" for a jail" was coded as IND 946 "Gov't" and OCC 754 "cook" which would lose the specific work done by the individual.
769	OCC	Caretaker	There was no code that sufficiently worked for a "caretaker" of a building. IPUMS would switch between using 764 "Housekeeper" and 230 "Manager, building"
997	OCC	Well/oil digging/drilling	IPUMS coded these individuals as "labourers" which was decided did not suffice for covering their specific work.
996	OCC	Assistant	Added because there were numerous individuals who it would give an Industry (ranging from manufacturing/retail/wholesale/transport) and then say "assistant" under occupation, so we would not even know what sort of occupation they were assisting. Instead of coding these individuals as "Uncodeable" we added this code.
947	IND	Indian Reserve	IPUMS would either code people who wrote "Indian Reserve" under Industry as "Uncodeable" or as "Government" and we decided neither sufficed and thought it highlighted a difference between IPUMS and the Canadian context of CCRI.
900	IND	Library, museum and other related institutions	Added because IPUMS would code individuals working at a library or museum as working for "Government" or "Educational Services".
900	OCC	Sealer (as in fishing)	We decided calling a "sealer" a "fisherman" would lose the significant difference between the occupations and highlighted a difference between IPUMS and the Canadian context of CCRI.
980	IND	Housewife	We added this code because IPUMS had the code within the program but it was not available for us to

			select.
992	OCC	Income	We used this code when the enumerator wrote "income" in the occupation column but we were not clear on whether the individual was "retired" or not.

Note:

Users may notice inconsistencies in the coding of occupations. In most cases the inconsistencies may not be the result of errors in the coding process. Coding is an interpretative process, and in any specific case there may be more than one acceptable interpretation, especially when the original enumeration entry is unclear or ambiguous.

A good example is a case where the entry under Place of Employment reads simply "shop" and the entry under Occupation is "tailor." The Occupation code should be 590 but the Industry (Place) code could be either 448 (Manufacturing - apparel) or 656 (Retail - apparel). With such cases we tried to ensure consistency especially within each census year.

Nevertheless, analysts must use appropriate caution when aggregating occupations into categories, either by occupation or sector, and it is essential to use both Occupation and Industry (Place) fields in aggregation and analysis.

Other inconsistencies may result from our efforts to apply the code sets developed by the IPUMS project. We have tried to use the Occupation code 690 (operative and kindred workers) sparingly, and to apply a more precise code when the specific occupation is known. We have also tried to adapt the codes to the Canadian context where possible.

It is inevitable, however, that there will be some inconsistencies when two fields are being coded simultaneously, and by reference to precedents for specific occupation-place code combinations applied by IPUMS. The ultimate assurance of consistency rests with the user, who must use both fields in any analysis relating to occupation.

In future a user may wish to apply HISCO codes (Historical International Standard Classification of Occupations).

For those with an occupation about 40 percent of the Place of Employment variable is coded to 99999001 "Blank" or to 99999006 "Missing -- Mandatory Field" For those with rural based occupations 67 percent of the Place of Employment variable is either blank or Missing--Mandatory Field and for those with an urban based occupation 21 percent of the Place of Employment variable is either blank or Missing--Mandatory Field.

These are not data entry errors. Enumerators often left the Place of Employment variable blank when they assumed that the occupation title was a sufficient indicator of Place of Employment. Thus the occupation farmer often has no Place of Employment filled in. So too the occupation

clergyman often has no Place of Employment data.