

**Baldwin-Green Study
Canada-US Census of Industry
1867 - 1940**

See Documentation held in Social Science Data Centre
Documents Unit, Queen's University
Kingston, Ontario
K7L 5C4

Documentation Call No.: CA2 ON UQ 90C17 Vols 1 & 2

COMPARISON OF CANADIAN AND U.S. INDUSTRY: 2-DIGIT SIC CODES, 1867-1940
-- CANADIAN DATA --

FILE NAME: MERGCA.txt

CONTENTS: Documentation for the file "MERGCA.xls" (Microsoft Excel 2000)
File resaved in MS-EXCEL 2000 format in April 2002

This file documents the column headings found in the spreadsheet files listed above. All files listed here are identical -- they have just been saved in different spreadsheet formats.

OVERVIEW:

"MERGCA.xls" is the final merged file for the CANADIAN data from the Green-Baldwin SSHRC project conducted in the 1980's. These data represent the final merged file for the CANADA-US MANUFACTURING COMPARISON at the 2-digit SIC code level. Data span from 1870 to 1940.

A series of historical price indices for various industries have been included, covering the period 1867 to 1940. They originated from Urquhart and Buckley(ed), HISTORICAL STATISTICS OF CANADA, VOL. I, series J34-44. This data is located in the last few columns.

FILE CREATION:

Data were transferred from a SAS 82.3 print output dated 1987. The work of transferring these data was done by Mark Smith, a graduate student in Economics, in January 1994.

ACKNOWLEDGEMENTS:

The Baldwin-Green project was graciously supported by SSHRC Grant # 410-83-0439 R-1. For more complete information on matchings, sources and derivation of the numbers, see the 'Background' binders available from Queen's Professor Alan G. Green or the Queen's Documents Library. The corresponding paper written by Baldwin and Green, for the Belagio Conference in 1986, exists only in hard copy form, and may be available from Professor Green. A copy of this paper has been added to the data file documentation in the Documents Library.

HEADINGS:

Each industry has a similar set of headings, with different individual industry codes. Replace the asterisk with the appropriate industry code (see next section).

- | | |
|-----------------------------|-------|
| 1) Number of establishments | NEST* |
| 2) Number of employees | NEMP* |

- | | |
|--------------------------------------|--------------------------|
| 3) Salary and wages | SAW* |
| 4) Gross Value of Production | GRPRO* |
| 5) Cost of materials used | CMAT* |
| 6) Value added | VA* |
| 7) Average value added | AVA* = [VA*/NEST*] |
| 8) Average value of gross production | AGRPRO* = [GRPRO*/NEST*] |

INDUSTRY CODES:

Each Canadian variable has an extra letter added to the end of the variable name, which relates it to its industry.

The following letter codes have been used:

Code	Example	Industry
----	-----	-----
S	VAS	TRANSPORTATION EQUIPMENT
W	VAW	WOOD
T	VAT	TOBACCO PRODUCTS
F	VAF	FOOD AND BEVERAGE
K	VAK	KNITTING MILLS (Note: Doesn't generate AVA or AGRPRO) (note: does not include VAK or AVAK)
I	VAI	IRON AND STEEL PRODUCTS
L	VAL	LEATHER
Z	VAZ	MISCELLANEOUS MAUFACTURING
M	VAM	NON-METALLIC MINERALS
N	VAN	NON-FERROUS METALS
P	VAP	PAPER AND ALLIED INDUSTRIES
O	VAO	PETROLEUM AND COAL PRODUCTS
R	VAR	RUBBER AND PLASTICS
X	VAX	TEXTILES
B	VAB	PRINTING AND PUBLISHING
H	VAH	CHEMICAL PRODUCTS
C	VAC	CLOTHING INDUSTRY
E	VAE	ELECTRICAL PRODUCTS

PRICE INDICES:

The following price indices are included in the file:

Industry	Index label
<hr/>	
General index for Canada, including gold.....	GENC
Vegetable products.....	VEGC
Animal products.....	ANIMC
Chemical industries.....	CHEMC
Non-ferrous metals.....	NFERC
Non-metalic minerals	NMETC
Wood and allied industries	WOODC
Textiles and Allied Industies	TEXTC
Vegetable products	VEGC
Iron and steel Industry	IRONC
General Manufacturing	GENC

The following variables are the numbers from above, divided by 100.

NEWANIMC
 NEWCHEMC
 NEWNFERC
 NEWNMETC
 NEWWOODC
 NEWTEXTC
 NEWVEGC
 NEWIRONC
 NEWGENC

=====
July 1994
updated April 2002

**Baldwin-Green Study
Canada-US Census of Industry
1867 - 1940**

See Documentation held in Social Science Data Centre
Documents Unit, Queen's University
Kingston, Ontario
K7L 5C4

Documentation Call No.: **CA2 ON UQ 90C17** Vols 1 & 2

COMPARISON OF CANADIAN AND U.S. INDUSTRY: 2-DIGIT SIC CODES, 1867-1940
-- US DATA --

FILE NAME: MERGUS.txt

CONTENTS: Documentation for the file "MERGUS.xls" (Microsoft Excel 2000)
File resaved in MS-EXCEL 2000 format in April 2002

This file documents the column headings found in the spreadsheet
files listed above. All files listed here are identical --
they have just been saved in different spreadsheet formats.

OVERVIEW:

"MERGUS.xls" is the final merged file for the UNITED STATES data
from the Green-Baldwin SSHRC project conducted in the 1980's.
These data represent the final merged file for the CANADA-US
MANUFACTURING COMPARISON at the 2-digit SIC code level.
Data span from 1900 to 1940.

A series of historical price indices for various industries have
been included, covering the period 1890 to 1940. These data were
taken from THE HISTORICAL STATISTICS OF THE UNITED STATES (HSUS),
VOLUME 1. The series used were E13-24. The 'NEW' series, as they
are labeled, are the prices from the HSUS re-calibrated to
have 1935-39=1.0. This was done to match the Canadian data.
These price indices are found in the last few columns of the
spreadsheet.

FILE CREATION:

Data were transferred from a SAS 82.3 print output dated 1987.
The work of transferring these data was done by Mark Smith,
a graduate student in Economics, in July 1994.

ACKNOWLEDGEMENTS:

The Baldwin-Green project was graciously supported by
SSHRC Grant # 410-83-0439 R-1. For more complete information on
matchings, sources and derivation of the numbers, see the
'Background' binders available from Queen's Professor
Alan G. Green or the Queen's Documents Library. The corresponding
paper written by Baldwin and Green, for the Belagio Conference
in 1986, exists only in hard copy form, and may be available from
Professor Green. A copy of this paper has been added to the
data file documentation in the Documents Library.

HEADINGS:

Each industry has a similar set of headings, with different

individual industry codes. Replace the asterisk with the appropriate industry code (see next section).

- 1) Number of establishments..... NEST*U
- 2) Number of employees..... NEMP*U
- 3) Salary and wages..... SAW*U
- 4) Gross Value of Production..... GRPRO*U
- 5) Cost of materials used..... CMAT*U
- 6) Value added..... VA*U
- 7) Average value added..... AVA*U = [VA*U/NEST*U]
- 8) Average value of gross production.. AGRPRO*U = [GRPRO*U/NEST*U]

INDUSTRY CODES:

Each US variable has an extra letter added to the end of the variable name, which relates it to its industry.

A "U" has been added to the end of each variable, indicating "US", to differentiate these variables from those in the Canadian file.

The following letter codes have been used:

Industry	Character	example
Wood	W	VAWU
Tobacco products	T	NESTTU
Food and Beverage	F	AVAFU
Leather	L	GRPROLU
Paper and allied industries	P	VAPU
Petroleum and Coal Products	O	VAOU
Rubber and plastics	R	AVARU
Textiles	X	VAXU
Printing and publishing	B	VABU
Chemical products	H	VAHU
Apparell and clothing	C	VACU
Electrical	E	VAEU
Instruments	V	VAVU
Stone, Minerals and Glass	G	VAGU
Furniture	U	VAJU

{note: the following have only 2 years are entered, 1938 & 1940}

Fabricated metal products	D	VADU
Primary metal industries	J	VAJU
Machinery (except elec.)	Q	VAQU
Transportation Equipment	S	VASU

COMMENTS:

Matching records indicate that, at one time, census records matched for all the Canadian data. It is strange, therefore, that some of the files are incomplete, and major industries such as IRON AND STEEL have been excluded. Where do 'instruments' or 'glass' fit? Canadian entries for NON-METALLIC and NON-FERROUS seem to have no like entries here. Should 'KNITTING' be included with TEXTILES? Should FURNITURE be included with WOOD? These questions may be never answered.

PRICE INDICES:

The following price indices are included in the file:

All comodities	GENU
Food products	FOODU
Farm products	FARMU
Hides and leather products	LEATU
Building materials	BUILU

fuel and lighting	FUELU
Chemical and allied products	CHEMU
Metals and metal products	METAU
House furnishing products	HOUSU
Textiles	TEXTU

These price indices are set with 1926 = 100.

The following series use 1935-39 = 1.00 as a base, and are denoted 'NEW'.

NEWGENU
NEWFUELU
NEWTEXTU
NEWCHEMU
NEWFARMU
NEWFOODU
NEWMETAU
NEWHOUSU
NEWLEATU
NEWBUILU

5

5

5

BALDWIN - GREEN STUDY
CANADA - U.S. CENSUS OF INDUSTRY, 1867 - 1940
BACKGROUND REPORT

VOLUME 1

①

GUIDE TO THE BACKGROUND REPORT ON CROSS COUNTRY AND LONGITUDINAL MATCHING:

The report is contained in 2 Volumes, which together comprise 9 sectors of manufacturing activities. These are:

	<u>Vol. No.</u>	
Vegetable Products	1	KR
Animal Products	1	KR
Textiles	1	KR
Wood and Paper	2	RB
Iron and Steel	2	RB
Non-Ferrous Metals	2	RB
Non-Metallic Minerals	2	RB
Chemicals	2	KR
Other	2	KR/RB

Volume 1 was primarily prepared by Kevin Reed and Volume 2 primarily by Ruth Beck.

The report is organized by sector. The format for each sector is:

1. - Title Page
2. - List of Sector Subcategories
3. - Canadian Industry Lists 1890 - 1929.
4. - U.S. Industry Lists 1890 - 1929.
5. - Matched and unmatched industries in each Subcategory:
 - a) Cross country
 - b) Longitudinal (between adjacent periods)
 - c) Summary of possible multi-period longitudinal matches

NOTES:

2. - The sectors are divided into subcategories which correspond to logical groupings of industries. Occasionally, matched and unmatched industries are included in

subcategories (or sectors even) which do not appear () to be logical. In these cases, the subcategory or sector is always referenced on the Industry Lists (see below)

3,4. The industry lists also provide a directory to information in this report. For example:

1930 IRON AND STEEL:

- C1 Bolts, nuts, rivets and washers (6) [7]
- C2 Lightning rods (see Non-ferrous metals) [6]
- C3 Rolling mills (11)

U

① The number in round brackets () refers to the location (subcategory) in which information on the cross country matching of this industry is to be found.

② The number in square brackets [] refers to the location of information on longitudinal matching of this industry. The square bracket is included only if the location of longitudinal matching information is different from the location of cross country matching information.

I'm not sure whether this has been done correctly.

③ The letter U in the left hand margin means that the industry is unmatched for cross country purposes.)

④ One further example of directory information is:

C4 Dishwashers (5) [5, 11]

In this case, 2 numbers are given in square brackets. The first refers to the backwards longitudinal match and the second, to the forwards match.

⑤ In the US 1920 industry lists, one more variation is possible. Since some information is duplicated in the data base, some industries are unmatched due to the fact that they provide duplicate information. These are noted by U-DUPL in the left-hand margin.

GUIDE TO REPORT. contd.

- 5. a) The cross country matches are presented in order of year - from 1890 - 1929. The numbers 0, 1, 2, 3 refer to 1890, 1900, 1910, 1920 and 1929 respectively and are noted in the left-hand margin of the text for easy reference.
- b) For longitudinal matches 0/0, 0/1, 1/2 and 2/3, referring to the 1890/1900, 1900/1910, 1910/1920 and 1920/1929 matches respectively, are placed in the left-hand margin as a reference to longitudinal matches.
- c) Both matched and unmatched industries are included for each year, for each subcategory. This has been done to isolate similar industries and to provide a focus for matching decisions.
- d) The notation of possible multi-period matches is a preliminary assessment of the potential for extending longitudinal matchings over a longer time period. The information is presented in such a manner as to facilitate the creation of a table whereby one could identify the total number of matches available for a particular string of years. The table could take the form:

BOLTS, NUTS	DIES AND TAPS	ENGINES	MACHINERY	OR	SEWING MACHINES	WASHING MACHINES	ROLLING STOCK	SAWS
0	0	0			0	0	0	0
1	1	1	1	OR	1	1	1	1
2		2	2		2	2	2	
			3		3	3	3	

- numbers could be marked in pencil where the quality of match is $\leq E$ due to goodness of fit, industry structure or heterogeneous products etc
 - alternative choices could be noted such as in the case of machinery.

- this type of information would allow one to see if there is a large enough data base of multi-period matches (say 0, 1, 2 - or 1900 to 1920) to enable one to perform various econometric techniques etc.