cirlin 21

CENSUS OF MANUFACTURING STATISTICS, 1958-1976

BSTRACT: This census dataset contains establishment statistics and industry structure data for 343 4-digit CENSUS of MANUFACTURING Industries. All are SIC industries in the 2000 to 4000 range (i.e. manufacturing). This dataset was created by using the 1972-1976 census dataset ('PRFL7276') on the tape 'DATABK', the 1958-1971 census dataset ('CENSUS') on the tape 'DATABK', and information provided in the 1972 SIC classification manual. The manual enabled us to adjust many industries that had been split up, redefined, or renumbered during this 19 yr. period so as to give us a valid time-series dataset for as many industries (343) as it was possible to maintain complete and accurate data on over time. A fuller discussion of the adjustment procedures used, and copies of the intermediate programs used to create this dataset can be found in the files of the Research Program in Competition and Business Policy.

For a full description of the 1958-1971 and 1972-1976 datasets that were used, see their respective documentations on file at the RPCBP.

CL: The dataset is stored in fixed format and can be read using the following fortran program and JCL.

```
      INTEGER SIC(343),NAME(343,19),YR(343,19),FLAG(343)
      REAL VAR(19,19,343)
      DO 1 I=1,343
      READ(8,101) SIC(I),(NAME(I,II),II=1,19),
     *(YR(I,J),(VAR(I,J,K),K=1,19),J=1,19),FLAG(I)
  101 FORMAT(I4,19A4,19(I4,19F9.1),I2)
    1 CONTINUE
C
C be clear that 'k' indexes the variable, and 'j' the yr.
C
      STOP
      END
//GO.FT08F001 DD DSN=cen5876,DISP=OLD,UNIT=TAPE6250,
// DCB=(RECFM=FB,LRECL=3407,BLKSIZE=10221),LABEL=(21,SL,,IN),
// VOL=SER=~~DATABK~~ DTA063
```

57

ITATION: RESEARCH PROGRAM IN COMPETITION AND BUSINESS POLICY, GSM, UCLA

ATA SET ORGANIZATION: There are 343 SICs, one record for each, and hence 343 records in total. For each record the dataset provides the SIC, the name of the industry, the year, data on 19 variables for that year, and a flag. The SIC and the name are given once per record. The years (and corresponding data) are given in ascending order, the flag follows all the other data for the SIC.

For example, a typical record would look as follows:

2015 Poultry Dressing Plants 1958 yar1---yar19 1959

VARIABLES:

There are 19 variables included on the dataset.
They are written in the following order:

1.	number of employees	(thousands)
2.	payroll	($millions)
3.	number of production workers	(thousands)
4.	man-hours	(millions)
5.	wages	($millions)
6.	value added	($millions)
7.	cost of materials	($millions)
8.	value of shipments	($millions)
9.	capital expenditures	($millions)
10.	end of year inventories	($millions)
11.	gross book value of fixed assets	($millions)
12.	4-firm value of shipments CR	(%)
13.	8-firm value of shipments CR	(%)
14.	20-firm value of shipments CR	(%)
15.	50-firm value of shipments CR	(%)
16.	number of companies	exact figure
17.	primary product specialization ratio	(%)
18.	coverage ratio	(%)
19.	number of establishments (for years 1963 and 1967) 4-firm product group CR (for year 1958) code (for years 1972-1976)	

The code variable will nearly always be zero. If it is equal to '9.0', this indicates that value added (rather than value of shipments) was used to calculate the concentration ratios (CR). The code will equal (-1.0) whenever an adjustment was made to the 1972-1976 industries due to an SIC definition change.

Flag-this entry will have one of 4 values.

0-indicates that no SIC definition change was made in 1972, and hence no adjustment was required for this industry.
1-indicates that several 1958-1971 SICs were aggregated.
2-indicates that several 1972-1976 SICs were aggregated.
3-indicates that the 1972 SIC is different from the earlier SIC, but the industry is exactly the same. (i.e. simply a number change).

SOURCES:

1. The 1972-1976 census data, found on the tape 'DATABK', dsn=PRFL7276.
2. The 1958-1971 census data, found on the tape 'DATABK', dsn=CENSUS.
3. The 1972 Standard Industrial Classification Manual-1972, prepared by the Office of Management and Budget.
4. 'Lincards'-part of the documentation on file with the RPCBP, it details all the adjustments that were made.

NOTES:

to be added together in order to get comparable figures for the pre (or post) definition change SIC, a '-1.0' is inserted for the variables that could not be added together. These variables were the four concentration ratios, the primary product specialization ratio, and the coverage ratio. For those wishing to use "weighted average" concentration ratios for these cases (a procedure that we feel is invalid), the original (pre-adjustment) data resides on 'DATABK', and can be used for your own purposes.

2. A '-1.0' (or a '0.0') will in all other contexts indicate that the data was "withheld to avoid disclosing figures for individual companies". When and if a '-3.0' occurs, it indicates the same thing. A '-4.0' signifies that the value was "withheld because the estimate did not meet publication standards."

3. The documentation for Census and PRFL7276 discuss data availability for the years in question at some length.

4. On file as well is data that includes the number of establishments for the 1972 SICs. Perhaps this can soon be included in this dataset.