

Entry and Exit Documentation

June 16, 2003

Contact:

Nicole Nestoriak nicole.nestoriak@census.gov or (301) 763-5289

1. Datasets

File Names:

Long panel: 4 state pooled files

entry_long.sas7bdat – contains entry rates

wentry_long.sas7bdat – contains employment-weighted entry rates

exit_long.sas7bdat – contains exit rates

wexit_long.sas7bdat – contains employment-weighted exit rates

Not yet completed...

Short panel: 8 state pooled files

entry_short.sas7bdat – contains entry rates

wentry_short.sas7bdat – contains employment-weighted entry rates

exit_short.sas7bdat – contains exit rates

wexit_short.sas7bdat – contains employment-weighted exit rates

Years of Coverage:

Both the long and the short panel contain statistics from 1992 through 2001 (generated from data available from 1991 through 2001). The two separate panels are maintained for easy comparison with other LEHD statistics.

Unit of Observation:

Each file contains counts of entry or exit for every year and quarter for various combinations of “by-groups” defined by the variable `_type_`. The groups include major SIC division, sector, location, firm size class and firm age.

2. Variables on Files

The following terms are used in the definitions below for employment weighted variables:

$$e_{iqt} = .5(emp_{iqt} + emp_{iqt-1}), \quad e_{qt} = \sum_i e_{iqt}, \quad \text{and} \quad w_{iqt} = e_{iqt} / e_{qt}$$

where emp_{iqt} = month 3 employment at establishment i in quarter q of year t ,

e_{iqt} = the average of this employment in quarter q of year $t-1$ and quarter q of year t

e_{qt} = the sum of e_{iqt} across all establishments in a given cell (all in semiconductors, for example), and

w_{iqt} = employment share, or the ratio of the average employment of the establishment in year t and t-1 to the average employment in year t and t-1 of the cell.

Average Number of Establishments – unweighted and weighted (avg_employers)

This variable indicates the average of the number of active establishments in the cell in year t and year t-1. More specifically,

$active_{iqt} = 1$ if establishment i has positive employment in month 3 of quarter q of year t; and

$$avg_employers_{iqt} = .5(active_{iqt-1} + active_{iqt}), \text{ and}$$

$$avg_employers_{qt} = \sum_i avg_employers_{iqt} \text{ on unweighted files, and}$$

$$avg_employers_{qt} = \sum_i w_{iqt} * avg_employers_{iqt} \text{ on employment weighted files.}$$

In both cases, summing occurs across all establishments in a given cell. This variable serves as the denominator in calculation of both entry and exit rates. A zero indicates that there are no employers in this cell in either year. Currently, only unweighted average employer counts have been provided. Weighted average employer counts have been used only to construct rates.

Entry Count of Establishments – unweighted and weighted (entry)

This variable is the count of entering establishments in the cell. This is an annual entry count. For example, an entering establishment in year t and quarter q is defined as one that is not active in year t-1 quarter q but is active in year t quarter q. Thus, an annual entry count is provided for each quarter. On unweighted files, each entering establishment contributes 1 to this count. On employment weighted files, each entering establishment contributes its share of cell employment for that year and quarter, or w_{iqt} . More formally,

$$entry_{iqt} = 1 \text{ if } active_{iqt-1} = 0 \text{ and } active_{iqt} = 1 \text{ and}$$

$$entry_{qt} = \sum_i entry_{iqt} \text{ on unweighted files, and}$$

$$entry_{qt} = \sum_i w_{iqt} * entry_{iqt} \text{ on employment weighted files.}$$

In both cases, summing occurs across all establishments in a given cell. This variable serves as the numerator in calculation of entry rates. A data value of zero indicates that there are employers in this cell in either year t-1 or year t but no entrants in year t. Currently, only unweighted entry counts have been provided. Weighted entry counts have been used only to construct rates.

Exit Count of Establishments – unweighted and weighted (exit)

This variable is the count of exiting businesses in the cell. This is an annual exit count. For example, an exiting establishment in year t and quarter q is defined as one that is active in year t-1 quarter q but is no longer active in year t quarter q. Thus, an annual exit count is provided for each quarter. On unweighted files, each exiting establishment

contributes 1 to this count. On employment weighted files, each exiting establishment contributes its share of cell employment for that year and quarter, or w_{iqt} .

More formally,

$$exit_{iqt} = 1 \text{ if } active_{iqt-1} = 1 \text{ and } active_{iqt} = 0 \text{ and}$$

$$exit_{qt} = \sum_i exit_{iqt} \text{ on unweighted files, and}$$

$$exit_{qt} = \sum_i w_{iqt} * exit_{iqt} \text{ on employment weighted files.}$$

In both cases, summing occurs across all establishments in a given cell. This variable serves as the numerator in calculation of exit rates. A data value of zero indicates that there are employers in this cell in either year t-1 or year t but no exiters in year t. Currently, only unweighted exit counts have been provided. Weighted exit counts have been used only to construct rates.

Entry Rates of Establishments - unweighted and weighted (rentry)

Formally, on both unweighted and weighted files, the entry rate is calculated as

$$entry_{qt} / avg_employers_{qt}.$$

On weighted files, both numerator and denominator are weighted counts of establishments (where the weight is the employment share w_{iqt} as noted above).

A data value of zero indicates that there are employers in this cell in either year t-1 or year t but no entrants in year t.

Exit Rates of Establishments - unweighted and weighted (rexit)

Formally, on both unweighted and weighted files, the entry rate is calculated as

$$exit_{qt} / avg_employers_{qt}.$$

On weighted files, both numerator and denominator are weighted counts of establishments (where the weight is the employment share w_{iqt} as noted above). A data value of zero indicates that there are employers in this cell in either year t-1 or year t but no exiters in year t.

Establishment age (in years) categories (firm_age_cat)

As noted in section 3 below, there are some observations for which we do not have enough information to assign establishment age. Observations for which we could not define an age are dropped from the firm_age_cat cells.

This variable categorizes establishment age in years

1 = establishment age ≥ 0 and establishment age less than 5

2 = establishment age ≥ 5 and establishment age less than 10

3 = establishment age ≥ 10

Labor Force Density (location)

This variable categorizes workers per square mile in the county in which the establishment is located and serves as a proxy for urban/rural status. Observations for which county information was unavailable are dropped from the location cells.

- 1 = workers per square mile less than 25
- 2 = workers per square mile \geq 25 and workers per square mile less than 500
- 3 = workers per square mile \geq 500

Quarter

The files contain counts separately for each of the four quarters in each year.

- 1 = first quarter (January-March)
- 2 = second quarter (April-June)
- 3 = third quarter (July-September)
- 4 = fourth quarter (October-December)

Sector

This variable identifies Sloan industries using 1987 SIC

Semiconductors

SIC='3674' or SIC='3559'

Software

SIC \geq '7371' and SIC \leq '7373'

Retail Food

SIC='5399' or 2-digit SIC='54' or SIC='5541'

Trucking

SIC='4212' or SIC='4213' or SIC='4214'

Financial Services

SIC='6011' or SIC='6021' or SIC='6022' or SIC='6029'
or SIC='6081' or SIC='6035' or SIC='6036' or SIC='6061'
or SIC='6062' or SIC='6141' or SIC='6153' or SIC='6159'
or SIC='6111' or SIC='6162' or SIC='6082' or SIC='5932'
or SIC='6019' or SIC='6163' or SIC='6099' or SIC='7389'
or SIC='6211' or SIC='6799' or SIC='6221' or SIC='6231'
or SIC='6792' or SIC='6282' or SIC='6371' or SIC='6733'
or SIC='6091' or SIC='6289' or SIC='6311' or SIC='6321'
or SIC='6324' or SIC='6331' or SIC='6351' or SIC='6361'
or SIC='6399' or SIC='6411' or SIC='6722' or SIC='6798'
or SIC='6726'

All Other

if sector="" then sector="Other";

Major industry Division (sic_div)

This variable identifies the major industry division

Sizeclass

This variable categorizes number of workers at a business unit. Both “number of workers” and “business unit” are defined more precisely below.

1 = employment < 20

2 = 20 <= employment < 100

3 = employment >= 100

Year

Counts are provided for each year and quarter from 1991 through 2001.
“yyyy”

TYPE

Cell type defined by combinations of class variables as follows:

- 1 year quarter
- 2 year quarter sizeclass
- 3 year quarter firm_age_cat
- 4 year quarter location
- 5 year quarter sic_div
- 6 year quarter sector
- 7 year quarter sector size
- 8 year quarter sector firm_age_cat
- 9 year quarter sector location

3. Data Concepts

Establishment

An establishment is an SEIN (state employer identification number) SEINunit combination. Note that an SEIN is different from a federal EIN (the business identifier found on many administrative business files).

Employment

We use the "best" measure of employment in the third month of each quarter obtained either directly from monthly ES-202 data (whenever available) or imputed from quarterly UI data when monthly data is not available. This imputation method is summarized below:

Step 1: If the business unit is not found in the ES-202 but is found on Unemployment Insurance (UI) wage record files, employment in month 3 is defined as end of quarter employment (e – see workforce indicators documentation).

Step 2: If this employment count is not greater than zero, we define month 3 employment as beginning of quarter employment b (approximates a point-in-time measure but for earlier in the quarter).

Step 3: If beginning of quarter employment is also not greater than zero, we define month 3 employment as being equal to total UI employment at any point during the quarter (m from workforce indicators documentation).

Active Establishments

The active status definition varies depending on the quarter we use to define business traits (we do the analysis separately for each of the 4 quarters).

For example, when using quarter 1, an establishment is defined as "active" for a given year if the establishment has positive employment in the first quarter of that year. We use the employment concept defined above to determine activity status.

Entry and Exit

The term "entry" encompasses both entry and re-entry. For example, if an establishment is "active" in 1992 quarter 1, is not active in 1993 quarter 1, and is active in 1994 quarter 1, this establishment is defined as an "exit" in 1993 quarter 1 and as an "entry" in 1994 quarter 1.

Firm Age

For each state, we use first year available ES-202 data for that state to define initial age for establishments active in this first year. For example, for Illinois, we have data available beginning in 1990. For each establishment, ES-202 data provides an initial date of UI liability. If an establishment is active in the first year, this variable is used to define age for all years of continuous employment from this year forward. Specifically, firm age is defined as:

$(\text{current year}+1) - (\text{date of initial liability as reported in our first year of data})$.

Thus, one year is the smallest age ever assigned to an establishment. Also, in some cases, establishments that are present in the initial year have missing values for date of initial liability or have a date that is inconsistent (greater than current year, for example). To "fill in" values for these observations, we look forward in the ES-202 data to find the first reported date of initial UI liability (or the first date that appears to be consistent). If none is found, we drop these establishments from firm age tabulations.

For businesses that are not present in the data in this first year or for businesses that exit and the re-enter (are inactive for at least one quarter), then the age of the establishment is re-set to one and accumulates annually upon re-entry. Thus, once an exit occurs (and for all new entrants), no further information from the ES-202 initial date of liability is used to determine the age of an establishment.

Lastly, note that the "age" for an exiting business (for purposes of tabulation) is defined as the age of the establishment in the year prior to exit. For example, using the quarter 1 definition, if a business is active in first quarter 1994 and inactive in first quarter 1995, the business is defined as an exit in 1995, and the

age reported for this exiting business in 1995 is actually age as calculated in 1994. In 1995, this establishment's age is defined as missing.

4. Using the Data to Create Annual Entry and Exit Rates

The entry and exit counts are divided by the average number of employers for a given cell to create annual entry and exit rates for that year. In a few cases, this division will generate a rate that is greater than one. This occurs because the denominator is an average of number of employers in year t and year $t-1$, so it is possible for an entry or exit count to exceed this total, and the numbers should not be regarded as inaccurate. However, this is rare and occurs only in cells with extremely low numbers of employers.