

LBD Codebook

Author: Javier Miranda

Date: April 2011

Variable name	Label	Type	Length
act	activity code	text	1
bestnaics	best NAICS	text	10
bestsic	best SIC	text	6
cbp		num	8
cbp_bestnaics		num	8
cbp_bestsic		num	8
cbp_naics		num	8
cbp_sic		num	8
cenflag	census in-scope flag	num	8
cfn	CFN if yr<2002, survuid if yr>=2001	text	10
county	county FIPS code	text	3
emp	March 12 employment	num	8
firmid		text	10
firstflag	first link flag	text	1
firstyear	first year establishment is observed	num	8
flag_i		text	18
flaga	type of link flag	text	1
flagb	birth-death-continuer link flag	text	1
lastflag	last link flag	text	1
last year	last year establishment is observed	num	8
lbdnum	LBD number	text	12
lfo	legal form of organization	text	2
mfsic1	modal 1 digit SIC	text	2
mfsic4	modal 4 digit SIC	text	4
mnaics2	modal 2 digit NAICS	text	2
mnaics4	modal 4 digit NAICS	text	4
mu	single-multi identifier	num	8
naics		text	10
pay	reported annual payroll (in \$1000)	num	8
pdiv	processing (economic) division code	text	1
rapf	reported annual payroll flag	text	1
recnum	SSEL record number	num	8
sic	Standard Industrial Classification Code	text	6
state	state FIPS code	text	2
toc	type of operation code	text	2
yr	year	num	8
zip		text	9

ACT

Label: activity code
Values: have varied over time
Description: Indicates whether the record represents an active establishment or is some form of placeholder record.

BESTNAICS

Label: best NAICS
Values:
Description: This is the most recent census year NAICS code reported for the establishment. Census years are those which end in 2 or 5 (ie 1982, 1995)

BESTSIC

Label: best SIC
Values:
Description: This is the most recent census year SIC code reported for the establishment. Census years are those which end in 2 or 5 (ie 1982, 1995)

CBP

Label:
Values: 0,1
Description: indicates if the establishment was considered in-scope for County Business Patterns. This is currently set to the same value as CBP_BESTSIC for 1976-2001 and CBP_BESTNAICS for 2002-present. A value of 1 indicates in-scope.

CBP_BESTNAICS

Label:
Values: 0,1
Description: This flag relies on BESTNAICS to determine if the establishment is County Business Patterns in-scope. A value of 1 indicates in-scope.

CBP_BESTSIC

Label:
Values: 0,1
Description: This flag relies on BESTSIC to determine if the establishment is County Business Patterns in-scope. A value of 1 indicates in-scope.

CBP_NAICS

Label:
Values: 0,1
Description: This flag relies on NAICS to determine if the establishment is County Business Patterns in-scope. A value of 1 indicates in-scope.

CBP_SIC

Label:
Values: 0,1
Description: This flag relies on SIC to determine if the establishment is County Business Patterns in-scope. A value of 1 indicates in-scope.

CENFLAG

Label: census in-scope flag
Values: 0,1
Description

CFN

Label: CFN if yr<2002, SURVUID if yr>=2002
Values:
Description: Census File Number. A CFN beginning with a zero is a single unit firm, while all others are multi-unit. The first six digits identify the firm, and the last four identify the plant. Each change in MU status will force a change in CFN.

COUNTY

Label: county FIPS codes
Values:
Description: Three digit county FIPS code. Uniquely identifies the county within the state. Must be combined with the state FIPS code to uniquely identify the county within the United States.

EMP

Label: March 12 employment
Values:
Description: Number of employees, both full and part time, working at the establishment as of March 12, reported to the Internal Revenue Service. In some cases this value is imputed due to missing or invalid data.

FIRMID

Label:
Values:

Description: For single unit firms, FIRMID=CFN. For multi-unit firms, FIRMID is the first six digits of CFN, followed by four zeroes.

FIRSTFLAG

Label: first link flag
Values: 3,5,A,B,I,J,M,R,Y,Z
Description: This is the establishment's FLAGA value on the first year it is observed.

FIRSTYEAR

Label: first year establishment is observed
Values: 1975-present
Description: In conjunction with LASTYEAR, allows users to quickly determine the tenure of an establishment from any point in the data series.

FLAG_I

Label:
Values: existing BR record, imputed record, prior yr fix
Description:

FLAGA

Label: type of link flag
Values: 1-8,A-E,I,J,M,P,R,V,W,Y,Z
Description: Refer to table 1 at end of document

FLAGB

Label: birth-death-continuer link flag
Values: Birth, Death, Continuer
Description: Birth flag coincides with the first year that the establishment appears in the Business Register. Death flag is recorded in the year following the last year the establishment appears in the Business Register.

LASTFLAG

Label: last link flag
Values: 6,A,C,D,E,P
Description: This is the establishment's FLAGA value on the last year it is observed.

LASTYEAR

Label: last year establishment is observed
Values: 1975-present
Description: In conjunction with FIRSTYEAR, allows users to quickly determine the tenure of an establishment from any point in the data series.

LBDNUM

Label: LBD number
Values:
Description: Uniquely identifies the establishment longitudinally through all years of data.

LFO

Label: legal form of organization
Values: varies over time
Description: This variable attempts to describe the structure of a firm with regard to ownership, tax obligations, and potential liabilities. Structures include corporations, s-corporations, sole proprietorships, tax-exempt, etc.

MFSIC1

Label: modal 1 digit SIC
Values:
Description:

MFSIC4

Label: modal 4 digit SIC
Values:
Description: The establishment's most commonly reported 4 digit SIC.

MNAICS2

Label: modal 2 digit NAICS
Values:
Description: The establishment's most commonly reported 2 digit NAICS.

MNAICS4

Label: modal 4 digit NAICS
Values:
Description:

MU

Label: single-multi identifier
Values: 0,1

Description: A value of 1 indicates the establishment is a member of a multi-unit firm. A value of 0 indicates a firm with only one establishment. A change in MU status will force a change in CFN.

NAICS

Label:

Values:

Description: North American Industry Classification System. There are three distinct versions of NAICS: 1997, 2002, and 2007.

PAY

Label: reported annual payroll (in \$1000)

Values:

Description:

PDIV

Label: processing (economic) division code

Values: have varied over time

Description: Represents the economic division or an entity based on the industrial classification codes and various other characteristics from survey results or administrative records. Refer to Business Register documentation for year specific code definitions.

RAPF

Label: reported annual payroll flag

Values: have varied over time

Description: This flag designates the source of annual payroll value. Refer to Business register documentation for year specific code definitions.

RECNUM

Label: SSEL record number

Values:

Description:

SIC

Label: standard industrial classification code
Values:
Description:

STATE

Label: state FIPS codes
Values:
Description: Two digit state level FIPS code

TOC

Label: type of operation code
Values: have varied over time
Description: Used in conjunction with SIC/NAICS to further classify the primary activity of the establishment. Refer to Business Register Documentation for year specific code definitions.

YR

Label: year
Values: 1975-present
Description:

ZIP

Label:
Values:
Description: Nine digit Zip+4.

Table 1: Definition of the FLAGA flags

Description	Code	Status	Detailed Definition	Notes
Missing Continuer: Year 1	1	Continuer	These only apply to a few years and represent continuing estabs that are missing in the source data for the current year. They were, therefore incorrectly identified as births in the following year. For year t, these are identified by matching year t+1 births to year t-1 deaths. We used only numeric matching methods. The matches are removed from year t-1 deaths, added to year t continuers and reallocated from year t+1 births to year t+1 continuers. The "year1" flag fills the hole in continuers caused by missing data for that year.	Assigned during post-matching fixes.
Missing Continuer: Year 2	2	Continuer	These are the year t+1 births reallocated to year t+1 continuers as explained directly above. The "year2" flag reallocates bogus births as continuers	Assigned during post-matching fixes.
Missing Birth: Year 1	3	Birth	As with the year to year fixes that repair the linkages for continuers that are broken due to missing source data, the birth fixes identify births that are missed due to missing source data and are therefore, incorrectly allocated to the following year. That is, due to missing records in the SSEL for year t, some year t births are picked up in the LBD as year t+1 births. To fix this, we take year t+1 births and see if they have positive prior year (year t) payroll and/or employment. If so, those estabs are added to year t births (although flagged separately) and are reallocated from births to continuers in year t+1. The "reallocated to" picks up the births in the correct year.	Assigned during post-matching fixes.
Missing Birth: Year 2	4	Continuer	These are the year t+1 births reallocated to year t+1 continuers as explained directly above.	Assigned during post-matching fixes.
Repaired Initial Flag	5	Birth	These are establishments that had misassigned initial flag. Often these were coded as Y's	Assigned during post-matching fixes.
Repaired Terminal Flag	6	Death	These are establishments that had misassigned terminal flags. These were often B's for single year cases or Z's	Assigned during post-matching fixes.
Repaired Bogus Birth	7	Continuer	These are establishments that had an incorrectly assigned birth flag when it had activity in both adjacent years. Flaga was changed from a "B" to a "7".	Assigned during post-matching fixes.
Repaired Bogus Death	8	Continuer	These are establishment that had an incorrectly assigned death flag. Most died in the subsequent year.	Assigned during post-matching fixes.
Automatched	A	Continuer	These are records that link to the prior year via name and address matching.	Assigned during match.
Birth	B	Birth	These are active records that do not link to the prior year.	Assigned during match.
CFN, PPN Continuer	C	Continuer	These are active records that match to prior year records by by CFN and PPN.	Assigned during match.
Death	D	Death	These are inactive records that match to the prior year. That is the estab was active in the prior, and is inactive in the current and subsequent years.	Assigned during match.
Other Numeric Continuer	E	Continuer	These are records that link to the prior year by other numeric identifiers such as OLDID and NEWID.	Assigned during match.
Beginning of SSEL Inactive Spell	i	Psuedo Continuer	These are records that were active in the prior year and are not in the current. However, they will become active again. These were flagged as deaths in the first version of the LBD. See Jarmin and Miranda (2002a) for details	Assigned during post-matching fixes for version 2 of the 75-98 LBD
End of SSEL Inactive Spell	j	Psuedo Continuer	These are records that were inactive in the prior year and are active in the current year. However, they will become active again. These were flagged as births in the first version of the LBD. See Jarmin and Miranda	Assigned during post-matching fixes for version 2 of the 75-98 LBD

(2002a) for details.

Mid-Year Reorganization	M	Continuer	These are records that link to the previous year under one set of numeric identifiers (e.g, CFN, PPN) and to the subsequent year on a different set. These matches are based on name and address information.	Assigned during match.
PPN Continuer	P	Continuer	These are records that link to the prior year by PPN and not CFN.	Assigned during match.
Birth Reorganization	R	Birth	These are births under one set of numeric identifiers that link to the subsequent year under a different set. These links are made via name and address matching.	Assigned during match.
Interim Splitter	V	Continuer	These are splitter records created to fill the gap between the original birth record and the retimed birth.	Assigned during splitter processing
Retimed Splitter Birth	W	Continuer	These are records originally classified as births, but where the birth was retimed to an earlier year as a result of splitter processing	Assigned during splitter processing
EIN Continuer	Y	Continuer	These are active records that link to the prior year by EIN by not CFN or PPN.	Assigned during match.
CFN Continuer	Z	Continuer	These are records that link to the prior year by CFN only. These are used in the pre-PPN era.	Assigned during match.