

---

CORNELL

UNIVERSITY

VIRTUALRDC

---

DOCUMENT NO. 2007-01

LBD Synthetic Data

---

Version : 2007-05-01  
Author : Lars Vilhuber  
Date : May 7, 2007

---

---

This document is part of the VirtualRDC documentation. The VirtualRDC (<http://vrdc.ciser.cornell.edu>) is partially supported by the National Science Foundation Grants #0427889 #0339191 #9978093 to Cornell University (Cornell Institute for Social and Economic Research).

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Dataset descriptions</b>	<b>3</b>
2.1	Variable availability . . . . .	3
2.2	LBD zero-obs datasets . . . . .	4
2.3	LBD synthetic datasets . . . . .	5
2.3.1	Details . . . . .	6

# Chapter 1

## Introduction

This document describes the first release of the Longitudinal Business Database (LBD) synthetic data.

The first release is a limited-functionality proof-of-concept. It only contains synthetic data for a single industry (SIC 57300), and only a subset of the variables. For reference, this document also reports the structure of the internal-use LBD, zero-obs versions of which are available on the VirtualRDC. The complete data for the internal-use LBD can only be accessed through the Census Research Data Center network, consult <http://www.ces.census.gov> for more details.

The synthetic LBD data can be accessed when logged on to the VirtualRDC under

```
/economic/lbd/synthetic
```

To obtain an account on the VirtualRDC, please contact [virtualrdc@cornell.edu](mailto:virtualrdc@cornell.edu).

## Chapter 2

# Dataset descriptions

### 2.1 Variable availability

Variable name	Variable label	Available on	
		internal dataset	synthetic dataset
act	Activity Code	1	0
cbp		1	0
cfm	Census File Number	1	0
county	County FIPS codes	1	0
emp	March 12 Employment	1	1
firstflag	First Link Flag	1	0
firstyear	First Year Establishment is Observed	1	1
flaga	Type of Link Flag	1	0
flagb	Birth-Death-Continuer Link Flag	1	0
imputation	Imputation Number	0	1
lastflag	Last Link Flag	1	0
lastyear	Last Year Establishment is Observed	1	1
lbdnum	LBD number	1	1
lfo	Legal Form of Organization	1	0
mfsic1	Most Frequent SIC 1	1	0
mu	Single-Multi Identifier	1	0
pay	Reported Annual Payroll (in \$1000)	1	1
pdiv	Processing (Economic) Division Code	1	0
rapf	Reported Annual Payroll Flag	1	0
recnum	SSEL Record Number	1	0
sic	Standard Industrial Classification Code	1	1
state	State FIPS codes	1	0
toc	Type of Operation Code	1	0
yr	Year	1	1

## 2.2 LBD zero-obs datasets

Field name	Data dictionary reference name	Starting position	Field size	Data type
	CBP	00064	8	N
Activity Code	ACT	00116	1	A/N
Birth-Death-Continuer Link Flag	FLAGB	00095	9	A/N
Census File Number	CFN	00084	10	A/N
County FIPS codes	COUNTY	00120	3	A/N
First Link Flag	FIRSTFLAG	00123	1	A/N
First Year Establishment is Observed	FIRSTYEAR	00040	8	N
LBD Number	LBDNUM	00072	12	A/N
Last Link Flag	LASTFLAG	00124	1	A/N
Last Year Establishment is Observed	LASTYEAR	00048	8	N
Legal Form of Organization	LFO	00114	2	A/N
March 12 Employment	EMP	00008	8	N
Most Frequent SIC 1	MFSIC1	00125	2	A/N
Processing (Economic) Division Code	PDIV	00110	1	A/N
Reported Annual Payroll (in \$1,000)	PAY	00016	8	N
Reported Annual Payroll Flag	RAPF	00117	1	A/N
SSEL Record Number	RECNUM	00024	8	N
Single-Multi Identifier	MU	00032	8	N
Standard Industrial Classification Code	SIC	00104	6	A/N
State FIPS codes	STATE	00118	2	A/N
Type of Link Flag	FLAGA	00094	1	A/N
Type of Operation Code	TOC	00111	2	A/N
Year	YR	00000	8	N

## 2.3 LBD synthetic datasets

Field name	Data dictionary reference name	Starting position	Field size	Data type
First Year Establishment is Observed	FIRSTYEAR	00008	8	N
Imputation Number	IMPUTATION	00000	8	N
LBD number	LBDNUM	00057	12	A/N
Last Year Establishment is Observed	LASTYEAR	00016	8	N
March 12 Employment	EMP	00032	8	N
Reported Annual Payroll (in \$1000)	PAY	00024	8	N
Standard Industrial Classification Code	SIC	00048	9	A/N
Year	YR	00040	8	N

### 2.3.1 Details

The statistics refer to the 1993 file.

Field name	Data dictionary reference name	Starting position	Field size	Data type
First Year Establishment is Observed	FIRSTYEAR	00008	8	N

#### CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

#### EXAMPLE SUMMARY STATISTICS

Mean	85.76
Std	6.21
P5	75.00
P25	81.00
P50	87.00
P75	91.00
P95	93.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Imputation Number	IMPUTATION	00000	8	N

### CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

### EXAMPLE SUMMARY STATISTICS

Mean	1.00
Std	0.00
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	1.00



Field name	Data dictionary reference name	Starting position	Field size	Data type
LBD number	LBDNUM	00057	12	A/N

CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Last Year Establishment is Observed	LASTYEAR	00016	8	N

### CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

### EXAMPLE SUMMARY STATISTICS

Mean	41.95
Std	47.15
P5	1.00
P25	1.00
P50	1.00
P75	96.00
P95	98.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
March 12 Employment	EMP	00032	8	N

CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	6.56
Std	8.49
P5	0.00
P25	2.00
P50	4.00
P75	8.00
P95	22.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Reported Annual Payroll (in \$1000)	PAY	00024	8	N

### CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

### EXAMPLE SUMMARY STATISTICS

Mean	149.53
Std	355.94
P5	3.00
P25	16.00
P50	49.00
P75	142.00
P95	602.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Standard Industrial Classification Code	SIC	00048	9	A/N

CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

VALUE TABLE

573000	100.0
--------	-------

Field name	Data dictionary reference name	Starting position	Field size	Data type
Year	YR	00040	8	N

### CHARACTERISTICS

UNITS	Persons/Percent
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

### EXAMPLE SUMMARY STATISTICS

Mean	93.00
Std	0.00
P5	93.00
P25	93.00
P50	93.00
P75	93.00
P95	93.00