Longitudinal Employer - Household Dynamics

Data codebook DATA-OTM-2.0.3

LEHD OnTheMap Technical documentation

Contents

Chapter 1

LEHD OnTheMap Technical documentation

Contents

1.1 Overview

OnTheMap has been a very popular web-based, interactive mapping application since its formal launch

under the Local Employment Dynamics partnership in February 2006[1]. Its objective is to show where people work and where workers live on maps with companion reports on their age, earnings, and industry distributions. The underlying data (OTM data) are available for access and download on the Cornell VirtualRDC, an internet-accessible computing environment dedicated to the exploration and development of synthetic data[2].

1.1.1 Background on synthetic data

OnTheMap is the first synthetic data product released by the Census Bureau. Ten (10) distinct copies of the synthetic data, known as implicates, are produced; the first implicate is used for the current implementation of OnTheMap. LEHD Program research has found that three (3) implicates are usually sufficient to determine the extent to which the confidentiality protections affect the statistical results, and these three implicates are available on the Cornell VirtualRDC.[3] Background on the development and use of synthetic data may be found in, for example, Abowd and Lane (2003), Abowd (2004), and Abowd and Lane (2004a) (also available as Abowd and Lane; 2004b). For further information on how to properly analyze multiply synthesized or imputed datasets, see Raghunathan et al. (2003); Reiter (2004b) and Reiter (2004a), or consult Sessions 8a and 8b of the online INFO 747 class at Cornell University's Cornell Institute for Social and Economic Research (CISER) at http://vrdc.ciser.cornell.edu/info747/.

1.1.2 Purpose of this document

This document describes the four types of datasets that compose the OTM data:

- Origin-Destination matrix
- Residence area characteristics
- Workplace area characteristics
- Workforce indicators

1.1.3 Changes in OnTheMap Version 2 data

While the underlying methodology for how the OTM data are derived has not changed between Version 1 and Version 2, there are a number of changes related to what data are provided in the public-use version of data. The new features of Version 2 include:

- Annual data for the period 2002-2004 (Version 1: 2002-2003)
- Cross-state patterns for all participating states, i.e. breakout of detailed location of out-of-state residents (Version 1: no breakdown of location beyond out-of-state designation)
- All files that are part of the OTM provide block-level estimates (Version 1: origin-destination matrix at the block-level, area characteristics files at the blockgroup-level)
- Estimates of all 4 job types ("all jobs", "jobs in the private sector", "all dominant jobs", "dominant jobs in the private sector") in the area characteristics files as well as in the origin-destination matrix (Version 1: only the origin-destination matrix provided estimates for all 4 job types; the residence area characteristics file provided estimates for "dominant jobs in the private sector only", and the workplace area characteristics file provided estimates for "private jobs")
- The Residence and Work Area Characteristics files now contain counts instead of proportions of workers across the various characteristics.

---

[1] http://lehd.dsd.census.gov/led/datatools/onthemap.html
[2] Census and Cornell announced the availability on October 19, 2006.
[3] Users who wish to explore the OnTheMap (OTM) data with additional implicates, please contact LEHD directly.

1.1.4   How to access the data

1.1.4.1   The process of getting an account and access

In order to access the OTM data, interested data users need to apply for a download account on the VirtualRDC, by contacting the VirtualRDC administrators (virtualrdc@cornell.edu). There is no project approval process. Optionally, users can also obtain an account on one of our compute servers, more information is provided in the online VirtualRDC guide (http://vrdc.ciser.cornell.edu/news/?p=13). Access is open to any user wishing to use the data for research purposes. We only ask that you provide comments, analysis, feedback and/or published papers. That information can be provided through the following list servers:

- led-qwi@lists.census.gov (QWI user community)

- lehd-ltd@lists.census.gov (Local Transportation Dynamics community under development)

- ctpp_news@chrispy.net (Census Transportation Planning Package community)

The research and evaluation results will be used to enhance understanding and developmental efforts for future versions of On The Map and synthetic data in general.

1.1.4.2   Technical requirements

Downloading data and analyzing it on own computer   In order to analyze the data on their own computers, users need to bring their own statistical software, and depending on the analysis, significant memory. Access is through a regular Web browser in the
OnTheMap Download Area (http://vrdc.ciser.cornell.edu/onthemap/data/).

Accessing data and analyzing it on the VirtualRDC   In order to login to the VirtualRDC compute nodes, users will need

- an internet-connected computer

- SSH client software

- (for graphical interface, optional) VNC or NX client software

All required software is open-source and free of charge, see the online VirtualRDC guide[4] for download links and detailed instructions. Windows and Linux clients are supported, but all software is known to work on Mac OS X as well. Statistical software is provided free-of-charge for use on the VirtualRDC compute nodes for research purposes, see the Installed software page[5] for a partial list. Once logged on to the VirtualRDC compute nodes, you will find the data and some useful SAS and Stata programs under

```
/mixed/onthemap/
```

1.2   General description

1.2.1   Universe

The basis of the OTM are beginning-of-quarter 2 jobs contained in Longitudinal Employer-Household Dynamics (LEHD)'s infrastructure files. To be included in the infrastructure, an individual's job must be covered by the reporting requirements of the state's unemployment insurance system. The prime exclusions are agriculture and to some parts of the public sector, particularly federal, military, and postal works. Coverage varies across states and time, although on average, 96% of all private-sector jobs are covered. The

---

[4]http://vrdc.ciser.cornell.edu/news/?p=13
[5]http://vrdc.ciser.cornell.edu/news/?page_id=83

BLS Handbook of Methods (Bureau of Labor Statistics; 1997, p. 42) describes UI coverage as "broad and basically comparable from state to state," and claims "over 96 percent of total wage and salary civilian jobs" were covered in 1994. Stevens (2002) provides a survey of coverage for a subset of the current participant states in the LEHD system.

The data integrate the predecessor-successor relationships established through observed worker flows, which suppress spurious accesssions and separations. Finally, the data are weighted to match published U.S. Bureau of Labor Statistics (BLS)'s Quarterly Census of Employment and Wages (QCEW) totals.

Please verify Section 1.6 for any specific coverage issues.

1.2.2 Sources

1.2.2.1 The place of residence

The place of residence comes from the Census Bureau's Statistical Administrative Records System (StARS) database. StARS encompasses several administrative data sets including federal tax forms, Medicare rolls, and other files, and it uses Topologically Integrated Geographic Encoding and Referencing (TIGER) to geocode a person's residence. The OTM covers any worker working in a state, including those living in another state. Workers not geocoded to a Census block of residence (about 10%) are imputed to a block by the disclosure-proofing algorithm.

1.2.2.2 The place of work

The place of work derives from the quarterly ES-202 data supplied by partner states to the Census Bureau. The data are collected as part of the QCEW program, also known as the ES-202 program, which is jointly administered by the BLS and the Employment Security Agencies in a federal-state partnership. This cooperative program between the states and the federal government collects employment, payroll, and economic activity, and physical location information from employers covered by state unemployment insurance programs and from employers subject to the reporting requirements of the ES-202 system.[6]. The universe for these data is a 'reporting unit,' which is the QCEW establishment–the place where the employees actually perform their work. Most employers have one establishment ('single-units'), but most employment is with employers who have multiple establishments ('multi-units'). One report per establishment per quarter is filed.

The ES-202 data contain characteristics of establishments by quarter, including address information. The Census Bureau geocodes the physical address. If the physical address is blank, then the mailing address substitutes. Longitudinal edits in Quarterly Workforce Indicators (QWI) processing fill in some missing data (Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer and Woodcock; 2006, see), but a small amount (about 5%) of establishments are imputed to county centroids.

1.2.2.3 The worker-employer link

The worker-employer link derives from the UI wage records supplied by the partner states. In Minnesota, these records identify the establishment within the firm employing each worker. In all other states, a statistical model developed for the QWI generates likely establishments for each worker employed by a multi-unit firm. The model uses 5 types of information to associate workers with likely establishments: 1) the distance from the worker's home to the establishment; 2) the number of employees at the establishment; 3) the number of employees at the firm; 4) the work history of the employee; and 5) the period of the establishment's existence. For further details, consult Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer and Woodcock (2006). The Origin-Destination (OD) links a worker to all implicated units within the firm, up to 10. A weight equal to the proportion of times the model implicates each establishment applies to each worker-establishment link.

---

[6]The employer and work place reports from this system are the same as the data reported to the BLS as part of the QCEW, but are referred to in the LEHD system by their old acronym "ES-202." These data are also used to compile the Covered Employment and Wages (CEW) and and Business Employment Dynamics (BED) data at the BLS.

1.2.3 Using the data

Analysts should aggregate the block-level data to larger geographic entities such as Traffic Analysis Zone (TAZ) or Census tract. The residence and workplace geocoding systems produce blockcoding errors mostly small in distance but large in frequency. Note also that misreporting by employers in the ES-202 causes spikes of trips to some blocks. For example, a large multi-establishment employer may report as a single establishment at a single address. Moreover, noise infused into the data by the disclosure-proofing method substantially alters the data, especially in areas with small populations.

This version of the data provides 3 *implicates* (independent draws in the synthesizing algorithm) for the OD matrix and the Residence Area Characteristics (RAC) files.[7] This is reflected in the filenames, see Sections 1.4.3 and 1.4.4. For further information on how to properly analyze multiply synthesized or imputed datasets, see Raghunathan et al. (2003); Reiter (2004b) and Reiter (2004a), or consult Sessions 8a and 8b of the online INFO 747 class at Cornell University's CISER at http://vrdc.ciser.cornell.edu/info747/. A note of warning is in order, though: It is statistically incorrect to use the average of the 3 implicates unless the aggregator function is strictly linear. Adding geographic areas is linear, and forming ratios from two linearly aggregated quantities (earnings over employment, for example) can be done correctly as long as the numerator and the denominator are averaged separately.

1.2.4 Confidentiality Protection of the OD Matrix and Residence Area Characteristics

The first version of the OD matrix, produced in 2003, used cell suppression to protect confidentiality. If the number of workers residing in a block was less than 5, or if the number of different blocks in which they were employed was less than 3, then the observation was suppressed. The rules seriously limited the resulting data. For instance, counts of total travelers, travelers from a given area, travelers to a given area, and travelers from one area to another were all biased estimates of the true values to the extent that they were aggregated over areas where observations were suppressed.

Instead of cell suppression, this version of the data set has been disclosure-proofed by "synthesizing" it from the true data, a method similar to data swapping in some respects. The key statistical property to preserve in the synthetic data is the joint distribution of workers across home and work areas. We use Bayesian techniques to synthesize workers' place of residence conditional on disclosable counts of workers by place of work, industry, age, and earnings categories. The resulting algorithm offers superior confidentiality protection while preserving the analytical validity of the data.

The basic idea is to replace the true home blocks by drawing synthetic home blocks for the workers with each set of characteristics on each work block. The draws are from the distribution of actual home blocks among workers with similar age, earnings, and industry characteristics, combined in a certain measure with another distribution of home blocks known as the "prior" distribution. The actual home blocks of workers employed on a certain block are defined according to QWI disclosure rules governing counts, with one modification. In the QWI, there must be at least three entities on which to base a releaseable statistic; here, a draw of one or two substitutes for any true count of workers less than three.

In this application, the prior is simply the distribution of home blocks among a larger worker group encompassing the target group. We first construct the prior for each set of worker characteristics in the Census tract containing the target work block. In cases where this doesn't provide population over a sufficient number of residence blocks, we aggregate worker types. If this is not enough, we aggregate to the county level with distinctions between all worker types. If this is insufficient, we aggregate to county level without reference to worker types.

The prior assures that some trips in the synthetic data are absent from the true data, by allowing draws of home blocks from which none of the workers on this block actually originate. It also assures that the

---

[7]LEHD Program research has found that three (3) implicates are usually sufficient to determine the extent to which the confidentiality protections affect the statistical results. Users who wish to explore the OTM data with additional implicates, please contact LEHD directly.

home block of a unique worker can be synthesized. A technical discussion of the synthetic data model is in preparation and will be available on the Local Employment Dynamics (LED) web site.

1.2.5 Confidentiality Protection of the Destination Data

The destination data are derived from the core QWI files. Definitions and disclosure-proofing methods are the same, and described in Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer and Woodcock (2006). Disclosure status flags are as follows:

Value	Explanation
-1	unable to computed this indicator because the data are unavailable
1	OK to release fuzzed value; not significantly distorted
5	indicator does not meet publication standards
9	data significantly distorted to protect confidentiality; OK to release fuzzed value

1.3 Contact information

For questions regarding the Cornell VirtualRDC, please consult their website at

<div align="center">

http://vrdc.ciser.cornell.edu

</div>

Further information regarding the OTM data can be obtained by contacting the Longitudinal Employer-Household Dynamics (LEHD) program:

Phone +1.301.763.8303

Email dsd.lehd.production.list@census.gov

1.4 Data description and dictionary

In this section, each file and its contents are described. Note that the description applies to SAS datasets created from the distributed CSV files. The programs used to create these data are described in Section 1.5.

1.4.1 Definitions

Jobs are defined as jobs for a specific individual that meet the definition of beginning-of-quarter employment (Emp, B) (Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer and Woodcock; 2006).

Primary job is the single job with the highest pay for a specific individual that meets the definition of beginning-of-quarter employment (Emp, B). Note that the number of primary jobs is the same as the number of workers, i.e., by definition each worker holds only one primary job. (WITHIN A STATE?)

Ownership is defined by the classification of the firm in QCEW records as 'private' (private ownership). OTM only distinguishes 'private' and the more inclusive 'all' categories.

1.4.2 Filenaming conventions

Filenames are generically defined. `[ST]` means the two-character postal abbreviation for a state, `[YEAR]` is a four-digit calendar year, `_[N]` is a numerical suffix, running from 1 to 3, indicating (for some files) the implicate number. Jobs are categorized by type of ownership (`[O]` = `all` or `private`) and whether the job type is based on the primary job (`[TYPE]` = `prim`) or all jobs (`[TYPE]` = `total`).

For each variable, if applicable, the algorithm creating the variable is briefly described. Also, example summary statistics or a value table with example frequencies are provided. These are for an arbitrary state and implicate file, and do not accurately describe the universe summary statistics - they are provided to describe the scope of the variables only.

1.4.3 OD matrix

The Origin-Destination (OD) Matrix links residence to work blocks. Each observation is a unique combination of a Census block of residence and a Census block of work, with the count of all jobs and primary jobs, for the entire covered economy and for private sector only. Jobs at employers with known or imputed geography are included. The data are disclosure-proofed by the synthetic data model.

Filename: `od_[ST]_[YEAR]_[N]`

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the home	H_GEOCODEFULL	00035	15	A/N
Full geocode of the workplace	W_GEOCODEFULL	00050	15	A/N
Number of jobs to which workers made this trip	JOBS_ALL	00000	8	N
Number of primary jobs to which workers made this trip	JOBS_PRIM	00008	8	N
Number of private-sector jobs to which workers made this trip	JOBS_PRIV_ALL	00016	8	N
Number of private-sector primary jobs to which workers made this trip	JOBS_PRIV_PRIM	00024	8	N
Year YYYY	YEAR	00032	3	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the home	H_GEOCODEFULL	00035	15	A/N

CHARACTERISTICS

UNITS	Geocode
ALGORITHM	see below
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

NOTE: The code is constructed as follows: FIPS State(2)||FIPS County(3)||Census Tract(6)||Census Block(4)

NOTE: The home geocode is zero-filled if the worker originates from another state.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the workplace	W_GEOCODEFULL	00050	15	A/N

CHARACTERISTICS

UNITS	Geocode
ALGORITHM	see below
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

NOTE: The code is constructed as follows: FIPS State(2)||FIPS County(3)||Census Tract(6)||Census Block(4)

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of jobs to which workers made this trip	JOBS_ALL	00000	8	N

CHARACTERISTICS

UNITS	Count
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	1.58
Std	4.59
P5	1.00
P25	1.00
P50	1.00
P75	2.00
P95	3.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of primary jobs to which workers made this trip	JOBS_PRIM	00008	8	N

CHARACTERISTICS

UNITS	Count
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	1.49
Std	4.48
P5	1.00
P25	1.00
P50	1.00
P75	2.00
P95	3.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of private-sector jobs to which workers made this trip	JOBS_PRIV_ALL	00016	8	N

CHARACTERISTICS

UNITS	Count
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	1.31
Std	4.31
P5	0.00
P25	1.00
P50	1.00
P75	1.00
P95	3.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of private-sector primary jobs to which workers made this trip	JOBS_PRIV_PRIM	00024	8	N

CHARACTERISTICS

UNITS	Count
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	1.23
Std	4.20
P5	0.00
P25	1.00
P50	1.00
P75	1.00
P95	3.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Year YYYY	YEAR	00032	3	N

CHARACTERISTICS

UNITS	Date
ALGORITHM	read-in
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

VALUE TABLE

2002   100.0

1.4.4 Residence area characteristics

The Residence Area Characteristics (RAC) file describes the residence blocks. Each observation represents a unique Census block group. The basis of the characteristics can be one of the following four job definitions:

- primary jobs in the private sector [T] = prim, [O] = private

- all jobs in the private sector [T] = all, [O] = private

- primary jobs, all sectors [T] = prim, [O] = all

- all jobs, all sectors [T] = all, [O] = all

(see Section 1.4.1 for more details. The home geocode is zero-filled if the worker originates from another state. Jobs at employers with known or imputed geography are included. The data are disclosure-proofed by the synthetic data model (REF).

Filename: rac_[ST]_[YEAR]_[O]_[T]_[N]

where [O] = all or private and [T] = prim or total

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the home	H_GEOCODE	00003	15	A/N
Number of beginning-of-Q2 [O] [T] jobs in this Block group	JOBS_[O]_[T]	00226	14	N
Count of workers in Construction	NAICS_23	00081	8	A/N
Count of workers 30 years or younger	AGE_YNG	00036	7	A/N
Count of workers 31 to 54 years	AGE_MID	00043	7	A/N
Count of workers 55 years or older	AGE_OLD	00050	7	A/N
Count of workers in Accommodation and Food Services	NAICS_72	00202	8	A/N
Count of workers in Administ, Support, Waste Mngt, Remediation Svcs	NAICS_56	00170	8	A/N
Count of workers in Agriculture, Forestry, Fishing and Hunting	NAICS_11	00057	8	A/N
Count of workers in Arts, Entertainment and Recreation	NAICS_71	00194	8	A/N
Count of workers in Educational Services	NAICS_61	00178	8	A/N
Count of workers in Finance and Insurance	NAICS_52	00138	8	A/N
Count of workers in Health Care and Social Assistance	NAICS_62	00186	8	A/N
Count of workers in Information	NAICS_51	00130	8	A/N
Count of workers in Management of Companies and Enterprises	NAICS_55	00162	8	A/N
Count of workers in Manufacturing	NAICS_31_33	00089	11	A/N
Count of workers in Mining	NAICS_21	00065	8	A/N
Count of workers in Other Services (except Public Administration)	NAICS_81	00210	8	A/N
Count of workers in Professional, Scientific and Techical Services	NAICS_54	00154	8	A/N
Count of workers in Public Administration	NAICS_92	00218	8	A/N
Count of workers in Real Estate and Rental and Leasing	NAICS_53	00146	8	A/N
Count of workers in Retail Trade	NAICS_44_45	00108	11	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Transportation and Warehousing	NAICS_48_49	00119	11	A/N
Count of workers in Utilities	NAICS_22	00073	8	A/N
Count of workers in Wholesale Trade	NAICS_42	00100	8	A/N
Count of workers with $1.2K < average Q2 monthly earnings	EARN_LO	00015	7	A/N
Count of workers with $1.2K < average Q2 monthly earnings <=$3.4K	EARN_MD	00022	7	A/N
Count of workers with $3.4K < average Q2 monthly earnings	EARN_HI	00029	7	A/N
Year YYYY	YEAR	00000	3	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the home	H_GEOCODE	00003	15	A/N

CHARACTERISTICS

UNITS	Geocode
ALGORITHM	see below
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

NOTE: The code is constructed as follows: FIPS State(2)||FIPS County(3)||Census Tract(6)||Census Block(4)

NOTE: The home geocode is zero-filled if the worker originates from another state.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of beginning-of-Q2 [O] [T] jobs in this Block group	JOBS_[O]_[T]	00226	14	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Construction	NAICS_23	00081	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers 30 years or younger	AGE_YNG	00036	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers 31 to 54 years	AGE_MID	00043	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers 55 years or older	AGE_OLD	00050	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Accommodation and Food Services	naics_72	00202	8	A/N

CHARACTERISTICS

Units	Percent
Algorithm	computed
Sourcefile	
Alternate documentation	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Administ, Support, Waste Mngt, Remediation Svcs	NAICS_56	00170	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Agriculture, Forestry, Fishing and Hunting	NAICS_11	00057	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Arts, Entertainment and Recreation	NAICS_71	00194	8	A/N

CHARACTERISTICS

Units	Percent
Algorithm	computed
Sourcefile	
Alternate documentation	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Educational Services	naics_61	00178	8	A/N

CHARACTERISTICS

Units	Percent
Algorithm	computed
Sourcefile	
Alternate documentation	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Finance and Insurance	NAICS_52	00138	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Health Care and Social Assistance	NAICS_62	00186	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Information	NAICS_51	00130	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Management of Companies and Enterprises	NAICS_55	00162	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Manufacturing	NAICS_31_33	00089	11	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Mining	NAICS_21	00065	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Other Services (except Public Administration)	NAICS_81	00210	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Professional, Scientific and Techical Services	NAICS_54	00154	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Public Administration	NAICS_92	00218	8	A/N

CHARACTERISTICS

Units	Percent
Algorithm	computed
Sourcefile	
Alternate documentation	n.a.

VALUE TABLE

0.000	99.92
0.002	0.04
0.011	0.04

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Real Estate and Rental and Leasing	NAICS_53	00146	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Retail Trade	NAICS_44_45	00108	11	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Transportation and Warehousing	naics_48_49	00119	11	A/N

CHARACTERISTICS

Units	Percent
Algorithm	computed
Sourcefile	
Alternate documentation	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Utilities	NAICS_22	00073	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers in Wholesale Trade	NAICS_42	00100	8	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers with $1.2K < average quarter2 monthly earnings	EARN_LO	00015	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers with \$1.2K < average quarter2 monthly earnings <=\$3.4K	EARN_MD	00022	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of workers with $3.4K < average quarter2 monthly earnings	EARN_HI	00029	7	A/N

CHARACTERISTICS

UNITS	Percent
ALGORITHM	computed
SOURCEFILE	
ALTERNATE DOCUMENTATION	n.a.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Year YYYY	YEAR	00000	3	N

CHARACTERISTICS

Units	Date
Algorithm	read-in
Sourcefile	
Alternate documentation	n.a.

1.4.5 Workplace area characteristics

The Workplace Area Characteristics (WAC) file is similar to the Residence Area Characteristics file, with the block group identified by *w_geocode* instead of *h_geocode*. The basis of the characteristics can be one of the following four job definitions:

- primary jobs in the private sector [T] = prim, [O] = private

- all jobs in the private sector [T] = all, [O] = private

- primary jobs, all sectors [T] = prim, [O] = all

- all jobs, all sectors [T] = all, [O] = all

(see Section 1.4.1 for more details. Jobs at employers with known or imputed geography are included. The data are disclosure-proofed by the QWI method (Abowd, Stephens and Vilhuber; 2006).

Filename: wac_[ST]_[YEAR]_[O]_[T]

Field name	Data dictionary reference name	Starting position	Field size	Data type
FIPS State(2)——FIPS County(3)——Census Tract(6)——Census Block Group(1)	W_GEOCODE	00003	15	A/N
Number of [O] [T] jobs	JOBS_[O]_[T]	00226	14	N
Count of jobs in Construction	NAICS_23	00081	8	A/N
Count of jobs held by workers 30 years or younger	AGE_YNG	00036	7	A/N
Count of jobs held by workers 31 to 54 years	AGE_MID	00043	7	A/N
Count of jobs held by workers 55 years or older	AGE_OLD	00050	7	A/N
Count of jobs in Accommodation and Food Services	NAICS_72	00202	8	A/N
Count of jobs in Administ, Support, Waste Mngt, Remediation Svcs	NAICS_56	00170	8	A/N
Count of jobs in Agriculture, Forestry, Fishing and Hunting	NAICS_11	00057	8	A/N
Count of jobs in Arts, Entertainment and Recreation	NAICS_71	00194	8	A/N
Count of jobs in Educational Services	NAICS_61	00178	8	A/N
Count of jobs in Finance and Insurance	NAICS_52	00138	8	A/N
Count of jobs in Health Care and Social Assistance	NAICS_62	00186	8	A/N
Count of jobs in Information	NAICS_51	00130	8	A/N
Count of jobs in Management of Companies and Enterprises	NAICS_55	00162	8	A/N
Count of jobs in Manufacturing	NAICS_31_33	00089	11	A/N
Count of jobs in Mining	NAICS_21	00065	8	A/N
Count of jobs in Other Services (except Public Administration)	NAICS_81	00210	8	A/N
Count of jobs in Professional, Scientific and Techical Services	NAICS_54	00154	8	A/N
Count of jobs in Public Administration	NAICS_92	00218	8	A/N
Count of jobs in Real Estate and Rental and Leasing	NAICS_53	00146	8	A/N
Count of jobs in Retail Trade	NAICS_44_45	00108	11	A/N
Count of jobs in Transportation and Warehousing	NAICS_48_49	00119	11	A/N
Count of jobs in Utilities	NAICS_22	00073	8	A/N
Count of jobs in Wholesale Trade	NAICS_42	00100	8	A/N
Count of jobs with $1.2K < average Q2 monthly earnings	EARN_LO	00015	7	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of jobs with $1.2K < average Q2 monthly earnings <=$3.4K	EARN_MD	00022	7	A/N
Count of jobs with $3.4K < average Q2 monthly earnings	EARN_HI	00029	7	A/N
Year YYYY	YEAR	00000	3	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full geocode of the workplace	W_GEOCODE	00003	15	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
Number of [O] [T] jobs	JOBS_[O]_[T]	00226	14	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
Count of jobs in Construction	NAICS_23	00081	8	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
Count of jobs held by workers 30 years or younger	AGE_YNG	00036	7	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
Count of jobs held by workers 31 to 54 years	AGE_MID	00043	7	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
Count of jobs held by workers 55 years or older	AGE_OLD	00050	7	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
Count of jobs in Accommodation and Food Services	NAICS_72	00202	8	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
Count of jobs in Administ, Support, Waste Mngt, Remediation Svcs	NAICS_56	00170	8	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
Count of jobs in Agriculture, Forestry, Fishing and Hunting	NAICS_11	00057	8	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
Count of jobs in Arts, Entertainment and Recreation	NAICS_71	00194	8	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
Count of jobs in Educational Services	NAICS_61	00178	8	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
Count of jobs in Finance and Insurance	NAICS_52	00138	8	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
Count of jobs in Health Care and Social Assistance	NAICS_62	00186	8	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
Count of jobs in Information	NAICS_51	00130	8	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
Count of jobs in Management of Companies and Enterprises	naics_55	00162	8	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
Count of jobs in Manufacturing	NAICS_31_33	00089	11	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
Count of jobs in Mining	NAICS_21	00065	8	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
Count of jobs in Other Services (except Public Administration)	NAICS_81	00210	8	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
Count of jobs in Professional, Scientific and Techical Services	NAICS_54	00154	8	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
Count of jobs in Public Administration	NAICS_92	00218	8	A/N

CHARACTERISTICS

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

VALUE TABLE

0.000   100.0

Field name	Data dictionary reference name	Starting position	Field size	Data type
Count of jobs in Real Estate and Rental and Leasing	NAICS_53	00146	8	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
Count of jobs in Retail Trade	naics_44_45	00108	11	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
Count of jobs in Transportation and Warehousing	naics_48_49	00119	11	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
Count of jobs in Utilities	naics_22	00073	8	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
Count of jobs in Wholesale Trade	NAICS_42	00100	8	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
Count of jobs with $1.2K < average Q2 monthly earnings	EARN_LO	00015	7	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
Count of jobs with $1.2K < average Q2 monthly earnings <=$3.4K	EARN_MD	00022	7	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
Count of jobs with \$3.4K < average Q2 monthly earnings	EARN_HI	00029	7	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
Year YYYY	YEAR	00000	3	N

CHARACTERISTICS

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

VALUE TABLE

2002   100.0

1.4.6 Workforce indicators

The Quarterly Workforce Indicators (QWI) file contains fields describing other characteristics of a block. Each observation corresponds to a unique Census block group. The basis of the characteristics are jobs in the private sector in quarter 2. Jobs at employers with known or imputed geography are included. The data are disclosure-proofed by the QWI method (Abowd, Stephens and Vilhuber; 2006). Detailed definitions of variables are available in Appendix A of Abowd, Stephens, Vilhuber, Andersson, McKinney, Roemer and Woodcock (2006).

Filename: `qwi_[ST]_[YEAR]`

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average earnings of full-quarter employees	Z_W3	00064	8	N
Average earnings of full-quarter new hires	Z_WH3	00080	8	N
Average earnings of separations from full-quarter status	Z_WFS	00072	8	N
Beginning-of-period employment	B	00008	8	N
Flow Employment	M	00048	8	N
Full-quarter employment	F	00016	8	N
Full-quarter new hires	H3	00040	8	N
Indicates the calendar date when data set was created (SAS date format)	CREATION_DATE	00000	8	N
Job creation	JC	00024	8	N
Job destruction	JD	00032	8	N
Number of Establishments	N_M	00088	8	N
Quarter	QUARTER	00099	3	N
Separations from full-quarter status	FS	00056	8	N
State——County——Tract——Blockgroup	W_GEOCODE	00132	15	A/N
Status: average earnings of full-quarter employees	STATUS_Z_W3	00120	3	N
Status: average earnings of full-quarter new hires	STATUS_WH3	00126	3	N
Status: average earnings of separations from full-quarter status	STATUS_Z_WFS	00123	3	N
Status: beginning-of-period employment	STATUS_B	00102	3	N
Status: flow employment	STATUS_M	00117	3	N
Status: full-quarter employement	STATUS_F	00105	3	N
Status: full-quarter new hires	STATUS_H3	00114	3	N
Status: job creation	STATUS_JC	00108	3	N
Status: job destruction	STATUS_JD	00111	3	N
Status: separations from full-quarter status	STATUS_FS	00129	3	N
Year YYYY	YEAR	00096	3	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average earnings of full-quarter employees	Z_W3	00064	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2179.84
Std	738.17
P5	1235.00
P25	1697.50
P50	2045.50
P75	2522.00
P95	3547.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average earnings of full-quarter new hires	Z_WH3	00080	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	1638.64
Std	802.01
P5	693.00
P25	1153.00
P50	1499.00
P75	1941.00
P95	3051.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average earnings of separations from full-quarter status	Z_WFS	00072	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	995.82
Std	637.77
P5	374.00
P25	644.00
P50	868.00
P75	1164.00
P95	2025.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Beginning-of-period employment	B	00008	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	470.38
Std	917.17
P5	19.00
P25	72.00
P50	181.00
P75	443.00
P95	1892.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow Employment	M	00048	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	578.54
Std	1049.85
P5	25.00
P25	99.00
P50	241.00
P75	574.00
P95	2255.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter employment	F	00016	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	416.64
Std	866.86
P5	16.00
P25	60.00
P50	153.00
P75	379.00
P95	1644.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter new hires	H3	00040	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	22.86
Std	21.42
P5	3.00
P25	8.00
P50	15.00
P75	31.00
P95	72.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Indicates the calendar date when data set was created (SAS date format)	CREATION_DATE	00000	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	16679.00
Std	0.00
P5	16679.00
P25	16679.00
P50	16679.00
P75	16679.00
P95	16679.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Job creation	JC	00024	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	21.87
Std	20.87
P5	1.00
P25	7.00
P50	15.00
P75	29.00
P95	69.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Job destruction	JD	00032	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	15.20
Std	16.84
P5	0.00
P25	4.00
P50	10.00
P75	20.00
P95	50.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of Establishments	N_M	00088	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	40.20
Std	59.17
P5	6.00
P25	14.00
P50	24.00
P75	43.00
P95	124.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
Quarter	QUARTER	00099	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.00
Std	0.00
P5	2.00
P25	2.00
P50	2.00
P75	2.00
P95	2.00

VALUE TABLE

2	100.0

Field name	Data dictionary reference name	Starting position	Field size	Data type
Separations from full-quarter status	FS	00056	8	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	24.41
Std	21.53
P5	3.00
P25	9.00
P50	17.00
P75	34.00
P95	72.00

Field name	Data dictionary reference name	Starting position	Field size	Data type
State——County——Tract——Blockgroup	GEOCODE	00132	15	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
Status: average earnings of full-quarter employees	STATUS_Z_W3	00120	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.95
Std	3.43
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: average earnings of full-quarter new hires	STATUS_WH3	00126	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	4.47
Std	3.89
P5	1.00
P25	1.00
P50	1.00
P75	9.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: average earnings of separations from full-quarter status	STATUS_Z_WFS	00123	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	4.10
Std	3.84
P5	1.00
P25	1.00
P50	1.00
P75	9.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: beginning-of-period employment	status_B	00102	3	N

CHARACTERISTICS

Units	Persons/Percent
Algorithm	read-in
Sourcefile	
Alternate documentation	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	2.49
Std	3.10
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: flow employment	STATUS_M	00117	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.49
Std	3.10
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: full-quarter employement	STATUS_F	00105	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.54
Std	3.14
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: full-quarter new hires	status_H3	00114	3	N

CHARACTERISTICS

Units	Persons/Percent
Algorithm	read-in
Sourcefile	
Alternate documentation	n.a.

EXAMPLE SUMMARY STATISTICS

Mean	3.37
Std	3.45
P5	1.00
P25	1.00
P50	1.00
P75	9.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: job creation	STATUS_JC	00108	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.42
Std	3.04
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: job destruction	STATUS_JD	00111	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	2.34
Std	2.97
P5	1.00
P25	1.00
P50	1.00
P75	1.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Status: separations from full-quarter status	STATUS_FS	00129	3	N

CHARACTERISTICS

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

EXAMPLE SUMMARY STATISTICS

Mean	3.16
Std	3.37
P5	1.00
P25	1.00
P50	1.00
P75	5.00
P95	9.00

VALUE TABLE

unable to computed this indicator because the data are unavailable	-1
OK to release fuzzed value; not significantly distorted	1
indicator does not meet publication standards	5
data significantly distorted to protect confidentiality	9

Field name	Data dictionary reference name	Starting position	Field size	Data type
Year YYYY	YEAR	00096	3	N

CHARACTERISTICS

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

VALUE TABLE

2002   100.0

1.5 Useful programs

This section describes two SAS programs that can be used for reading in the published OTM data. They were tested with Version 1 data, but need minor edits in order to work with Version 2 data. The most current versions of these programs, as well as programs for other software packages, can be found at

<div align="center">

http://vrdc.ciser.cornell.edu/onthemap/

</div>

1.5.1 readin.sas

This program is the master program. It calls a macro which does the actual work.

```
/*========================================================================*/
/* Wrapper program to read in multiply-imputed OD data                    */
/* $Id: readin.tex 2183 2008-02-22 00:21:51Z vilhuber $                   */
/* Original author: Lars Vilhuber <lars.vilhuber@cornell.edu>             */
/*========================================================================*/
```

Instructions: Copy this program to your project directory, adjust the parameters to the macro call, and run. If using this program outside of the VirtualRDC, you will need to adjust the paths accordingly.

```
/* Adjust this to your system - where are the macro definitions */
*let macropath=/mixed/onthemap/programs;
%let macropath=.;
options sasautos=("!SASROOT/sasautos","&macropath");
```

Example 1. Call without arguments to get options (output is in the log).

```
%readin_otm();
```

Example 2. Readin a single implicate of the OD file for Oregon. Leave all other values at their default.

```
* options mprint symbolgen mlogic;
%readin_otm(state=or,year=2003,type=od,implicate=02);
%readin_otm(state=or,year=2003,type=rac,implicate=02,owner=private,jobs=prim);
/* there are no implicates for these files */
%readin_otm(state=or,year=2003,type=wac,owner=all,jobs=total);
%readin_otm(state=or,year=2003,type=qwi);
```

1.5.2 readin_otm.sas

This program is a macro that does the actual readin work.

```
/*========================================================================*/
/* Program to read in multiply-imputed OD data                            */
/* $Id: readin_otm.tex 2183 2008-02-22 00:21:51Z vilhuber $               */
/* Original author: Fredrik Andersson <fredrik.andersson@census.gov>      */
/*========================================================================*/
%macro
readin_otm(state=,year=,type=od,owner=,jobs=,implicate=all,dir=/mixed/onthemap,o
utlib=work);
%let help=no;
```

Configure some parameters.

```
      %let yy=%eval(&year.-2000);
      %if ( &yy. < 10 ) %then %let yy=0&yy.;
      %if ( &yy. < 0 ) %then %let help=yes;
      %let filebase=&type._&state._&year.;
      %let fileext=csv;
      /* there are no implicates for the WAC or QWI files */
      %if ( "&type." = "wac" ) %then %let implicate=1;
      %if ( "&type." = "qwi" ) %then %let implicate=1;
      %if ( &implicate. = all ) %then %do;
          %let implicate_start=1;
          %let implicate_end=3;
      %end;
      %else %do;
          %let implicate=%eval(&implicate.*1);
          %let implicate_start=&implicate.;
          %let implicate_end=&implicate.;
      %end;
      /* check for inputs - exit if empty */
      %if ( "&state." = "" ) %then %let help=yes;
      %if ( "&year." = "" ) %then %let help=yes;
```

Some auxiliary macros

```
      %macro _print_help;
          %put
      -----------------------------------------------------------------;
          %put  READIN OF ON-THE-MAP MULTIPLY-IMPUTED DATA;
          %put  ;
          %put  readin_otm(state=,year=,type=od,;
          %put                owner=,jobs=,;
          %put                 implicate=all,dir=/mixed/onthemap,outlib=work);
          %put  ;
          %put  STATE    =&state.;
          %put  YEAR     =&year.     (format=YYYY);
          %put  TYPE     =&type.     (type=od,rac,wac,qwi);
          %put  OWNER    =&owner.    (owner=all,private);
          %put  JOBS     =&jobs.     (jobs=total,prim);
          %put  IMPLICATE=&implicate (which implicate load, number or 'all');
          %put  DIR      =&dir.      (location of data);
          %put
      -----------------------------------------------------------------;
      %mend;
```

Type-specific labelling.

```
      %macro label_type(type=,owner=,jobs=);
      %if ( &owner = private ) %then %let _owner=priv;
      %else %let _owner=all;
```

Common labels

```
      label
        year='Year YYYY'
```

Specific to the OD files.

```
      %if ( "&type." = "od"  ) %then %do;
      h_geocodefull='FIPS State(2)||FIPS County(3)||Census
      Tract(6)||Census Block(4), home'
      w_geocodefull='FIPS State(2)||FIPS County(3)||Census
      Tract(6)||Census Block(4), work'
      jobs_all_total='Number of jobs to which workers made this trip'
      jobs_all_prim='Number of primary jobs to which workers made
      this trip'
      jobs_priv_total='Number of private-sector jobs to which workers
      made this trip'
      jobs_priv_prim='Number of private-sector primary jobs to which
      workers made this trip'
      %end;
```

Specific to the RAC files.

```
%if ( "&type." = "rac" ) %then %do;
h_geocode='FIPS State(2)||FIPS County(3)||Census Tract(6)||Census Block
Group(1)'
earn_lo='Count of workers with $1.2K < average quarter2 monthly
earnings'
earn_md='Count of workers with $1.2K < average quarter2 monthly
earnings <=$3.4K'
earn_hi='Count of workers with $3.4K < average quarter2 monthly
earnings'
age_yng='Count of workers 30 years or younger'
age_mid='Count of workers 31 to 54 years'
age_old='Count of workers 55 years or older'
naics_11='Count of workers in Agriculture, Forestry, Fishing and
Hunting'
naics_21='Count of workers in Mining'
naics_22='Count of workers in Utilities'
naics_23='Count of workers  in Construction'
naics_31_33='Count of workers in Manufacturing'
naics_42='Count of workers in Wholesale Trade'
naics_44_45='Count of workers in Retail Trade'
naics_48_49='Count of workers in Transportation and Warehousing'
  naics_51='Count of workers in Information'
naics_52='Count of workers in Finance and Insurance'
naics_53='Count of workers in Real Estate and Rental and Leasing'
naics_54='Count of workers in Professional, Scientific and Techical
Services'
naics_55='Count of workers in Management of Companies and Enterprises'
naics_56='Count of workers in Administ, Support, Waste Mngt,
Remediation Svcs'
naics_61='Count of workers in Educational Services'
naics_62='Count of workers in Health Care and Social Assistance'
naics_71='Count of workers in Arts, Entertainment and Recreation'
naics_72='Count of workers in Accommodation and Food Services'
naics_81='Count of workers in Other Services (except Public
Administration)'
naics_92='Count of workers in Public Administration'
jobs_&_owner._&jobs.='Number of beginning-of-Q2 &owner.-sector jobs in
this Block group'
%end;
```

Specific to the WAC files.

```
%if ( "&type." = "wac" ) %then %do;
w_geocode='FIPS State(2)||FIPS County(3)||Census Tract(6)||Census Block
Group(1)'
earn_lo='Count of jobs with $1.2K < average quarter2 monthly earnings'
earn_md='Count of jobs with $1.2K < average quarter2 monthly earnings
<=$3.4K'
earn_hi='Count of jobs with $3.4K < average quarter2 monthly earnings'
age_yng='Count of jobs held by workers 30 years or younger'
age_mid='Count of jobs held by workers 31 to 54 years'
age_old='Count of jobs held by workers 55 years or older'
naics_11='Count of jobs in Agriculture, Forestry, Fishing and Hunting'
naics_21='Count of jobs in Mining'
naics_22='Count of jobs in Utilities'
naics_23='Count of jobs  in Construction'
naics_31_33='Count of jobs in Manufacturing'
naics_42='Count of jobs in Wholesale Trade'
naics_44_45='Count of jobs in Retail Trade'
naics_48_49='Count of jobs in Transportation and Warehousing'
  naics_51='Count of jobs in Information'
naics_52='Count of jobs in Finance and Insurance'
naics_53='Count of jobs in Real Estate and Rental and Leasing'
naics_54='Count of jobs in Professional, Scientific and Techical
Services'
naics_55='Count of jobs in Management of Companies and Enterprises'
naics_56='Count of jobs in Administ, Support, Waste Mngt, Remediation
Svcs'
naics_61='Count of jobs in Educational Services'
naics_62='Count of jobs in Health Care and Social Assistance'
naics_71='Count of jobs in Arts, Entertainment and Recreation'
naics_72='Count of jobs in Accommodation and Food Services'
naics_81='Count of jobs in Other Services (except Public
Administration)'
naics_92='Count of jobs in Public Administration'
jobs_&_owner._&jobs.='Number of beginning-of-Q2 &owner.-sector jobs in
this Block group'
%end;
```

QWI-specific labels

```
        %if ( "&type." = "qwi" ) %then %do;
        w_geocode='State||County||Tract||Blockgroup'
        B='Beginning-of-period employment'
        F='Full-quarter employment'
        FS='Separations from full-quarter status'
        H3='Full-quarter new hires'
        JC='Job creation'
        JD='Job destruction'
        M='Flow Employment'
        Z_W3='Average earnings of full-quarter employees'
        Z_WFS='Average earnings of separations from full-quarter status'
        Z_WH3='Average earnings of full-quarter new hires'
        creation_date='Indicates the calendar date when data set was created
        (SAS date format)'
        n_M='Number of Establishments'
        quarter='Quarter'
        status_B='Status: beginning-of-period employment'
        status_F='Status: full-quarter employement'
        status_FS='Status: separations from full-quarter status'
        status_H3='Status: full-quarter new hires'
        status_JC='Status: job creation'
        status_JD='Status: job destruction'
        status_M='Status: flow employment'
        status_WH3='Status: average earnings of full-quarter new hires'
        status_Z_W3='Status: average earnings of full-quarter employees'
        status_Z_WFS='Status: average earnings of separations from full-quarter
        status'
        %end;
        ;
        %mend;
        /*==============================================================================
        */
```

Now start processing.

```
        %if ( &help = yes ) %then %do;
            %if ( "&state." = "" ) %then %let state=Two-character state abbreviation;
            %if ( "&year." = "" )  %then %let year=YYYY;
            %_print_help;
            %put Please fix your macro call to specify the missing parameters;
            %put
        ---------------------------------------------------------------------;
            %return;
        %end;
        %_print_help;
```

Read-in the CSV data. Loop through the implicates, if not implicate=all.

```
        %let cur_implicate=&implicate_start.;
        %do %while ( &cur_implicate. <= &implicate_end. );
        /* there are no implicates for the WAC or QWI files */
        %if ( "&type." = "wac" ) %then %let infile=&dir./&filebase._&owner._&jobs.;
        %else %if ( "&type." = "qwi" ) %then %let infile=&dir./&filebase.;
        %else %if ( "&type." = "rac" ) %then %let
            infile=&dir./&filebase._&owner._&jobs._&cur_implicate.;
        %else %let infile=&dir./&filebase._&cur_implicate.;
        %let outfile=&infile.;
        %put =====================================================================;
        %put Starting readin of &infile..&fileext.;
        %put Writing to &outfile.;
        %put ---------------------------------------------------------------------;
        proc import datafile="&infile..&fileext."
                out=temp dbms=csv replace;
        getnames=no;
        run;
        data &outlib..&outfile. ;
        set temp (firstobs=2);
        drop var:;
```

Process OD files.

```
%if ( "&type." = "od" ) %then %do;
length year 3
              h_geocodefull w_geocodefull $15
              jobs_total jobs_prim jobs_priv_total jobs_priv_prim 8;
year=var1;
h_geocodefull=var2;
w_geocodefull=var3;
jobs_all_total=var4;
jobs_all_prim=var5;
jobs_priv_total=var6;
jobs_priv_prim=var7;
%end; /* end of od processing */
```

Process RAC files.

```
%if ( "&type." = "rac" ) %then %do;
length year 3 h_geocode $15   jobs_&_owner._&jobs. 8;
year=var1;
earn_lo=var2;
earn_md=var3;
earn_hi=var4;
age_yng=var5;
age_mid=var6;
age_old=var7;
naics_11=var8;
naics_21=var9;
naics_22=var10;
naics_23=var11;
naics_31_33=var12;
naics_42=var13;
naics_44_45=var14;
naics_48_49=var15;
  naics_51=var16;
naics_52=var17;
naics_53=var18;
naics_54=var19;
naics_55=var20;
naics_56=var21;
naics_61=var22;
naics_62=var23;
naics_71=var24;
naics_72=var25;
naics_81=var26;
naics_92=var27;
h_geocode=var28;
jobs_&_owner._&jobs.=var29;
%end; /* end rac processing */
```

Process WAC files.

```
%if ( "&type." = "wac" ) %then %do;
length year 3 w_geocode $15 jobs_&_owner._&jobs. 8;;
year=var1;
earn_lo=var2;
earn_md=var3;
earn_hi=var4;
age_yng=var5;
age_mid=var6;
age_old=var7;
naics_11=var8;
naics_21=var9;
naics_22=var10;
naics_23=var11;
naics_31_33=var12;
naics_42=var13;
naics_44_45=var14;
naics_48_49=var15;
  naics_51=var16;
naics_52=var17;
naics_53=var18;
naics_54=var19;
naics_55=var20;
naics_56=var21;
naics_61=var22;
naics_62=var23;
naics_71=var24;
naics_72=var25;
naics_81=var26;
naics_92=var27;
w_geocode=var28;
jobs_&_owner._&jobs.=var29;
%end; /* end wac processing */
```

Process QWI files.

```
%if ( "&type." = "qwi" ) %then %do;
length creation_date 8
       year  quarter 3
       w_geocode $15
       B F JC JD H3 M FS 8
       Z_W3 Z_WFS Z_WH3 8
       n_M 8
       status_B status_F status_JC status_JD status_H3
       status_M status_Z_W3 status_Z_WFS status_WH3 status_FS 3
    ;
format creation_date yymmdd10.;
creation_date=inputn(var1,'date7.');
year=var2;
quarter=var3;
w_geocode=var4;
B=var5;
F=var6;
JC=var7;
JD=var8;
H3=var9;
Z_W3=var10;
Z_WFS=var11;
Z_WH3=var12;
status_B=var13;
status_F=var14;
status_JC=var15;
status_JD=var16;
status_H3=var17;
status_M=var18;
status_Z_W3=var19;
status_Z_WFS=var20;
status_WH3=var21;
M=var22;
FS=var23;
n_M=var24;
status_FS=var25;
%end; /* end qwi processing */
        %label_type(type=&type.);
run;
```

Read obs.

```
%let dsid = %sysfunc(open(&outlib..&outfile.));
%let nobs=%sysfunc(attrn(&dsid,nobs));
%let rc = %sysfunc(close(&dsid));
%put ---------------------------------------------------------------------;
%put Finished reading in &outfile.. &nobs. observations read.;
%put =================================================================;
%let cur_implicate=%eval(&cur_implicate.+1);
%end; /* end of implicate loop */
%mend;
```

1.6   Notes

1.6.1   Unavailable data

- 2002 data for AR (Arkansas) do not exist

- 2002 and 2003 data for MS (Mississippi) do not exist

Bibliography

Abowd, J. M. (2004). Synthetic data: A new future for public use micro-data?, *Presentation PN-2004-01*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/presentations/Synthetic-Data-Census-20041207.pdf* 1.1.1

Abowd, J. M. and Lane, J. (2003). Synthetic data and confidentiality protection, *Technical paper TP-2003-10*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/techpapers/tp-2003-10.pdf* 1.1.1

Abowd, J. M. and Lane, J. (2004a). New approaches to confidentiality protection: Synthetic data, remote access and Research Data Centers, Springer, New York. 1.1.1

Abowd, J. M. and Lane, J. I. (2004b). New approaches to confidentiality protection synthetic data, remote access and research data centers, *Technical paper TP-2004-03*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/techpapers/tp-2004-03.pdf* 1.1.1

Abowd, J. M., Stephens, B. E. and Vilhuber, L. (2006). Confidentiality protection in the Census Bureau's Quarterly Workforce Indicators, *Technical Paper TP-2006-02*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/techpapers/tp-2006-02.pdf* 1.4.5, 1.4.6

Abowd, J. M., Stephens, B. E., Vilhuber, L., Andersson, F., McKinney, K. L., Roemer, M. and Woodcock, S. D. (2006). The LEHD infrastructure files and the creation of the Quarterly Workforce Indicators, *Technical Paper TP-2006-01*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/techpapers/tp-2006-01.pdf* 1.2.2.2, 1.2.2.3, 1.2.5, 1.4.1, 1.4.6

Bureau of Labor Statistics (1997). *BLS Handbook of Methods*, U.S. Bureau of Labor Statistics, Division of Information Services, Washington DC. http://www.bls.gov/opub/hom/. 1.2.1

Raghunathan, T., Reiter, J. and Rubin, D. (2003). Multiple imputation for statistical disclosure limitation, *Journal of Official Statistics* **19**(1): 1–16. 1.1.1, 1.2.3

Reiter, J. (2004a). Simultaneous Use of Multiple Imputation for Missing Data and Disclosure Limitation, *Technical report*, none. 1.1.1, 1.2.3

Reiter, J. P. (2004b). New approaches to data dissemination: A glimpse into the future (?), *Chance* **17**: 12–16. 1.1.1, 1.2.3

Stevens, D. W. (2002). Employment that is not covered by state unemployment insurance laws, *Technical paper TP-2002-16*, LEHD, U.S. Census Bureau.
URL: *http://lehd.dsd.census.gov/led/library/techpapers/tp-2002-16.pdf* 1.2.1

Appendix A

Abbreviations

BED Business Employment Dynamics

BLS U.S. Bureau of Labor Statistics

CEW Covered Employment and Wages

CISER Cornell Institute for Social and Economic Research

CRADC Cornell Restricted Access Data Center

DER Detailed Earnings Record

IRS Internal Revenue Service

KDE kernel density estimator

LBD Longitudinal Business Database

LED Local Employment Dynamics

LEHD Longitudinal Employer-Household Dynamics

MBR Master Beneficiary Record

MEF Master Earnings File

MER Master Earnings Record

NIA National Institute on Aging

NSF National Science Foundation

OD Origin-Destination

ODM Origin-Destination Matrix

OTM OnTheMap data

QCEW BLS's Quarterly Census of Employment and Wages

QWI Quarterly Workforce Indicators

RAC Residence Area Characteristics

SIPP Survey of Income and Program Participation

SRMI Sequential Regression Multiple Imputation

SSA Social Security Administration

SSN Social Security Number

StARS Statistical Administrative Records System

TAZ Traffic Analysis Zone

TIGER Topologically Integrated Geographic Encoding and Referencing

WAC Workplace Area Characteristics