LEHD INFRASTRUCTURE FILES IN THE CENSUS RDC – OVERIEW

by

Lars Vilhuber
U.S. Census Bureau

Kevin McKinney
U.S. Census Bureau

Abstract

The Longitudinal Employer-Household Dynamics (LEHD) Program at the U.S. Census Bureau, with the support of several national research agencies, maintains a set of infrastructure files using administrative data provided by state agencies, enhanced with information from other administrative data sources, demographic and economic (business) surveys and censuses. The LEHD Infrastructure Files provide a detailed and comprehensive picture of workers, employers, and their interaction in the U.S. economy. This document describes the structure and content of the 2011 Snapshot of the LEHD Infrastructure files as they are made available in the Census Bureaus secure and restricted-access Research Data Center network. The document attempts to provide a comprehensive description of all researcher-accessible files, of their creation, and of any modifcations made to the files to facilitate researcher access.

Contents

CONTENTS

List of Tables

List of Figures

Chapter 1.
Overview of LEHD Infrastructure

The Longitudinal Employer-Household Dynamics (LEHD) Infrastructure files available in the Research Data Center (RDC) are structured as individual components. A big-picture overview of it can be found in Abowd et al. 2006a, which was published as Abowd et al. 2009. Figure 1.1 provides an overview of the flow of data elements through the LEHD data creation process.

Currently, the core outputs of the data creation process are the Quarterly Workforce Indicators (QWI), shown in Figure 1.1, and the OnTheMap (OTM) data. The LEHD Infrastructure files in the RDC environment do not contain any information related to the disclosure limitation measures used in the QWI (for more information on the disclosure limitation techniques, see Abowd et al. 2006a and Abowd, Stephens, and Vilhuber 2006 for a discussion). Public-Use QWI (QWIPU) are available for the first time, see Chapter 12. Note that use of the QWIPU data precludes access to the confidential files, but has certain other advantages (see Chapter 12 for more details).

After pulling the files from LEHD production archives, several research-related improvements are made to the files, fixing minor data inconsistencies or updating documentation. Since the S008 Snapshot, the SAS header of the files contains an identifier tag that allows to uniquely track (most) files. A SAS "proc contents" can show that information.

1.1   UPDATES: APRIL 2013: S2011 RELEASE

This is the third release of the LEHD Infrastructure files. It contains data for the period through the end of 2011, and includes Q1 of 2012. We refer to it as the 'S2011' snapshot of the LEHD Infrastructure files. The data was pulled from LEHD archives as a coherent ensemble in 2012Q4 and 2013Q1. The LEHD Snapshot S2011 covers 49 states and the District of Columbia. Massachusetts, the Virgin Islands, and Puerto Rico have not yet had infrastructure files produced.

We should highlight the fact that not all states have full-quality data through Q1 of 2012. Problematic interior quarters or lower-quality variables will generally be included in the Snapshot and are highlighted in their respective sections (in particular EHF and ECF) and through appropriate data quality flags. States with recent data delivery or quality issues may have shorter time series overall (data may end earlier than 2012Q1). Table XX shows the available time periods by state and product.

Information on previous updates can be found in Section 1.2.

1.2   UPDATE HISTORY

Figure 1.1: Data flow view of LEHD Infrastructure
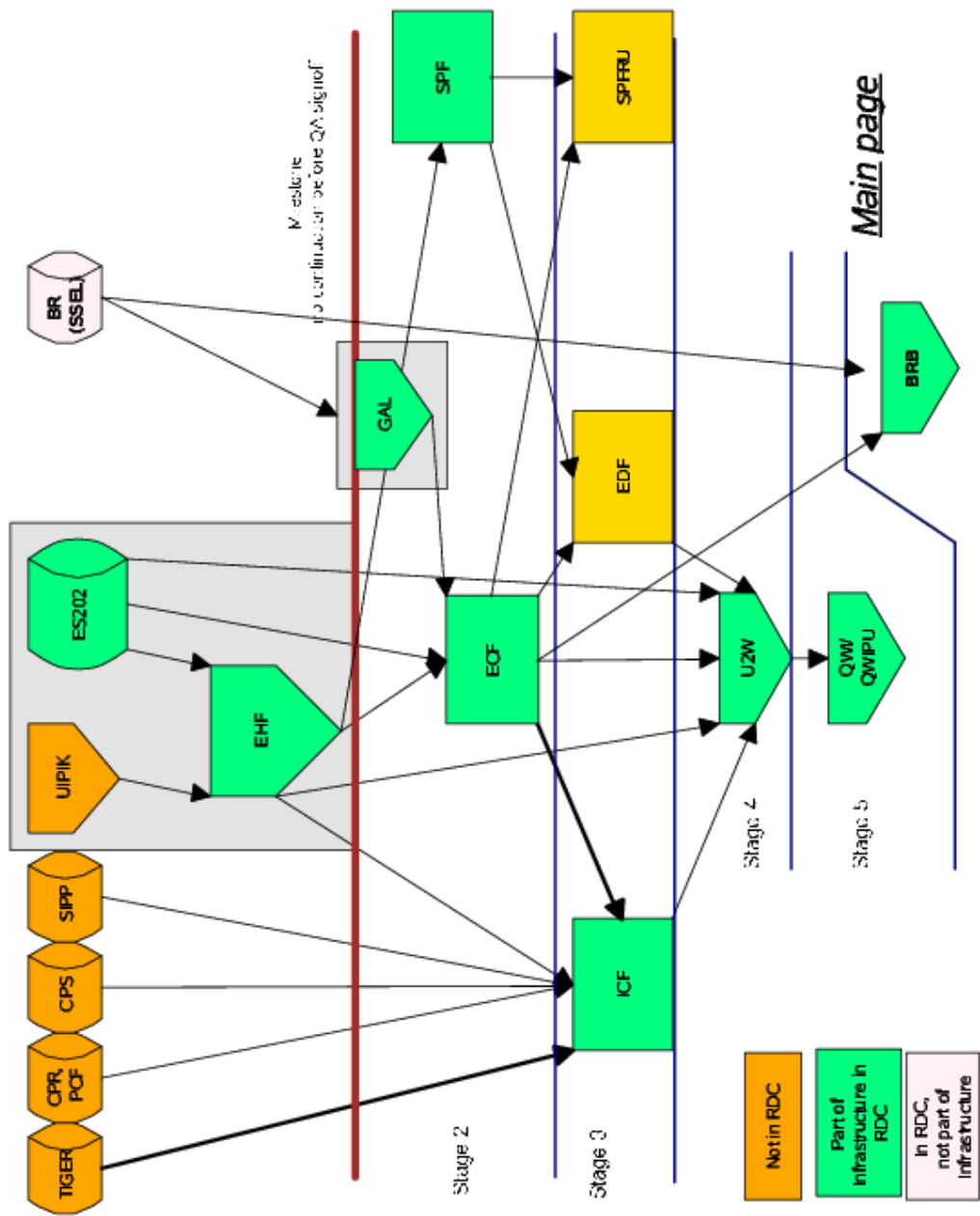
1.2.1   October 2010: S2008 release

The S2008 release is the second release of the LEHD Infrastructure files. It contains data that covers the years up to and including 2008Q1. The data was pulled from LEHD archives as a coherent ensemble in October 2009. For detailed information, see McKinney and Vilhuber (2011a).

```
Process ID      Latest creation date
-------------------------------
brb                2005-05-21
ecf                2009-08-12
edf                2009-08-12
ehf                2009-08-07
es202              2009-08-05
gal                2009-08-05
icf                2009-08-12
qwi                2009-08-25
spf                2009-08-12
u2w                2009-08-18
```

After pulling the files from LEHD production archives, several research-related improvements are made to the files, fixing minor data inconsistencies or updating documentation. In the S008 Snapshot, the SAS header of the files contains an identifier tag that allows to uniquely track (most) files. A "proc contents" can show that information.

1.2.2 August 2008: S2004 release

The S2004 snapshot is the first release of the LEHD Infrastructure files. It contains data that covers the years up to and including 2004Q1. The data was pulled from LEHD archives as a coherent ensemble over the course of 2005 and 2006. For detailed information, see McKinney and Vilhuber (2011b).

Improvements are made to the files, fixing minor data inconsistencies or updating documentation. To identify the version of the files in the data archive, a file called *version.txt* is at the root of each data directory, e.g., *u2w/version.txt*. The file will contain the name of the data, the snapshot number, and the date stamp of the most recent file within the data. As of the writing of this document,

```
./brb/version.txt: BRB S2004 2005-06-23
./ecf/version.txt: ECF S2004 2007-05-17
./ehf/version.txt: EHF S2004 2006-03-29
./gal/version.txt: GAL S2004 2008-03-27
./icf/version.txt: ICF S2004 2007-06-01
./u2w/version.txt: U2W S2004 2008-03-27
./qwi/version.txt: QWI S2004 2007-03-30
./spf/version.txt: SPF S2004 2006-06-28
./es202/version.txt: ES202 S2004 2007-02-09
./ecft26/version.txt: ECFT26 S2004 2007-05-17
./galt26/version.txt: GALT26 S2004 2008-03-07
./icft26/version.txt: ICFT26 S2004 2007-06-03
```

Table 1.1: LEHD components

Name and abbreviation	CES abbr. if different	Name of FTI version	CES abbreviation of FTI version
Business Register Bridge (BRB)		(all)	
Employer Characteristics File (ECF)		ECFT26	ect
Employment History Files (EHF)			
ES-202 (ES-202)	es2	ECFT26	ect
Individual Characteristics File (ICF)		ICFT26	ict
Geocoded Address List (GAL)		GALT26	gat
Quarterly Workforce Indicators (QWI) (establishment level)			
Successor-Predecessor File (SPF)			
Unit-to-Worker Impute (U2W)			

1.3 TREATMENT OF FEDERAL TAX INFORMATION

Some components of the LEHD Infrastructure include Title-26 protected variables. In the Snapshot, these are stored as separate datasets for tracking and monitoring purposes, but are not documented separately. Such T26 components need to be requested separately, and as of the writing of this documentation, will trigger additional proposal review. Table 1.1 shows the nine components and their Federal Tax Information (FTI) counterparts, if present, as they are available in the RDC.

1.4 IDENTIFIERS

In general, linkages between the different files are created using deterministic match-merge techniques. Person, firm, and establishment identifiers allow users to link all LEHD Infrastructure files. Throughout, all Social Security Numbers (SSNs) have been replaced by Protected Identity Keys (PIKs) - no SSNs are available anywhere in these data. Linkage to other person-level data products at the Census Bureau require crosswalks keyed to the PIK, which are not available as part of the LEHD Snapshot and must be requested separately.[1]

Firm identifiers are called State employer identification numbers (SEINs). The identifiers are constructed internally by LEHD, and generally, but not always, reflect an entity reporting unemployment insurance (UI) taxes to state authorities. "Establishments" (more precisely: reporting units) are identified by SEIN reporting unit (SEINUNIT). Establishments and firms are structured as one would expect with establishments listed hierarchically within each firm. Therefore to uniquely identify an establishment both the SEIN and SEINUNIT must be used. The firm and establishment identifiers are state and firm-structure-specific - within the LEHD Infrastructure files, there is no straighforward method of linking units of a firm with multiple tax reporting entities (SEINs). Although the vast majority of firms have only one SEIN, a firm, depending on its structure may have multiple SEINs operating both within and across state boundaries. Although the federal Employer Identification Number (EIN) is available and can be used to link SEINs within and across states, the EIN suffers from similar problems as the SEIN. The identifier is not necessarily unique within a firm, is designed for tax reporting, and the structure of EINs within a firm is arbitrary. The Census Bureau recognizes the limitations of administrative identifiers and has addressed this problem on the Business Register (BR) and the Longitudinal Business Database (LBD). The BRB files as well as the EIN stored on the ECF are used to link to the Business Register (BR), Longitudinal Business Database (LBD) and other Census economic data. Note that

---

1. Previous versions of the ICF provided additional person identifiers linking to Census survey data (Current Population Survey (CPS), and Survey of Income and Program Participation (SIPP)). Starting with the S2011 Snapshot, these are no longer maintained as part of the LEHD Snapshot.

the BRB is in general a many-to-many link file. The BRB does permit assigning all SEINs and SEINUNITs to a common *alpha* (the overall firm identifier inthe BR). However, exact identifier-based establishment-to-establishment matches between BR/LBD and LEHD data are generally not possible for establishments part of multi-establishment firms.

For any further information, refer to the component-specific documentation.

1.5 AVAILABILITY OF DATA

Availability of LEHD Infrastructure files is conditional on (i) the data files having been processed in the LEHD Production system, and subsequently integrated into the LEHD Infrastructure and (ii) permission for use in research having been granted by LEHD's state partner. The standard Memorandum of Understanding (MOU) between the Census Bureau and its state partners precludes access to person and firm names and physical addresses as provided in the ES-202 data. As described below, there are geographic identifiers that are derived in the GAL that can be used for analysis and integrating data for appropriate and approved purposes. In addition to data provided by the states, and processed through the LEHD Production system, data provided by Office of Personnel Management (OPM) are also available (in experimental mode).

As of June 20, 2014, 50 states (including the District of Columbia) have been processed for the complete set of LEHD data files and integrated. In general, LEHD Infrastructure files are available from 2000 onwards. However, the availability of historical data prior to 2000 varies significantly across states. Table 1.2 tabulates the availability data source (state UI or OPM) in the S2011snapshot (Figure 1.2 graphically depicts availability for UI/EHF data). A full list of files for each type of file is provided in each detailed section. Note that for certain states, availability of UI files (as captured by the EHF) differs from historical availability of Quarterly Census of Employment and Wages (QCEW) files (as captured by the ECF). Finally, a shorter time-series for the QWI indicates certain serious data issues interrupting the data series, sufficient to block publication of the official QWI, but possibly without consequences for certain research uses. Data sources not currently available for the entire time period may become available in the next update to the LEHD Infrastructure, or as a revision to the current snapshot.

Table 1.2: Availability by data source

| Data source | Start of data series | | | End |
	EHF	ECF	QWI	quarter
OPM	2000Q1	2000Q1	2000Q1	2011Q4
Alaska	1990Q1	1990Q1	2000Q1	2012Q1
Alabama	2001Q1	2001Q1	2001Q1	2012Q1
Arkansas	2002Q3	2002Q3	2002Q3	2012Q1
Arizona	1992Q1	1992Q1	2004Q1	2012Q1
California	1991Q3	1991Q1	1991Q3	2012Q1
Colorado	1990Q1	1990Q1	1993Q2	2012Q1
Connecticut	1996Q1	1996Q1	1996Q1	2012Q1
District of Columbia	2002Q2	2000Q4	2005Q2	2012Q1
Delaware	1998Q3	1997Q1	1998Q3	2012Q1
Florida	1992Q4	1989Q1	1992Q4	2012Q1
Georgia	1994Q1	1994Q1	1998Q1	2012Q1
Hawaii	1995Q4	1995Q4	1995Q4	2012Q1
Iowa	1998Q4	1990Q1	1998Q4	2012Q1
Idaho	1990Q1	1990Q1	1991Q1	2012Q1
Illinois	1990Q1	1990Q1	1990Q1	2012Q1
Indiana	1990Q1	1990Q1	1998Q1	2012Q1

(continued on next page)

Table 1.2 – Continued

	Start of data series			End
Data source	EHF	ECF	QWI	quarter
Kansas	1990Q1	1990Q1	1993Q1	2012Q1
Kentucky	1996Q4	1996Q4	2001Q1	2012Q1
Louisiana	1990Q1	1990Q1	1995Q1	2012Q1
Maryland	1985Q2	1985Q2	1990Q1	2012Q1
Maine	1996Q1	1996Q1	1996Q2	2012Q1
Michigan	1998Q1	1998Q1	2000Q3	2012Q1
Minnesota	1994Q3	1994Q3	1994Q3	2012Q1
Missouri	1990Q1	1990Q1	1995Q1	2012Q1
Mississippi	2003Q3	2003Q3	2003Q3	2012Q1
Montana	1993Q1	1993Q1	1993Q1	2012Q1
North Carolina	1991Q1	1990Q1	1992Q4	2011Q4
North Dakota	1998Q1	1998Q1	1998Q1	2012Q1
Nebraska	1999Q1	1999Q1	1999Q1	2012Q1
New Hampshire	2003Q1	2003Q1	2003Q1	2012Q1
New Jersey	1996Q1	1995Q1	1996Q1	2012Q1
New Mexico	1995Q3	1990Q1	1995Q3	2012Q1
Nevada	1998Q1	1998Q1	1998Q1	2012Q1
New York	1995Q1	1990Q1	2000Q1	2012Q1
Ohio	2000Q1	2000Q1	2000Q1	2012Q1
Oklahoma	2000Q1	1999Q1	2000Q1	2012Q1
Oregon	1991Q1	1990Q1	1991Q1	2012Q1
Pennsylvania	1991Q1	1991Q1	1997Q1	2012Q1
Rhode Island	1995Q1	1990Q1	1995Q1	2012Q1
South Carolina	1998Q1	1998Q1	1998Q1	2012Q1
South Dakota	1994Q1	1994Q1	1998Q1	2012Q1
Tennessee	1998Q1	1998Q1	1998Q1	2012Q1
Texas	1995Q1	1990Q1	1995Q1	2012Q1
Utah	1999Q1	1990Q1	1999Q3	2012Q1
Virginia	1998Q1	1995Q3	1998Q1	2012Q1
Vermont	2000Q1	2000Q1	2000Q1	2012Q1
Washington	1990Q1	1990Q1	1990Q1	2012Q1
Wisconsin	1990Q1	1990Q1	1990Q1	2012Q1
West Virginia	1997Q1	1990Q1	1997Q1	2012Q1
Wyoming	1992Q1	1992Q1	2001Q1	2012Q1

The data underlying this table is attached to this document as CSV.

Availablility of core Infrastructure files for research is dependent on a state's participation in the Local Employment Dynamics (LED) program, and on permission having been given to make the files accessible in the RDC.

1.6   PROCESSING FILES

LEHD Infrastructure files are significantly larger than even traditionally large research files such as the decennial census. In the current version, in all available states and years combined, wage, job, and other information is presented for

- 1,579,392,898 jobs (from EHF_PHF) held by

Figure 1.2: Data availability (UI/EHF) by data source
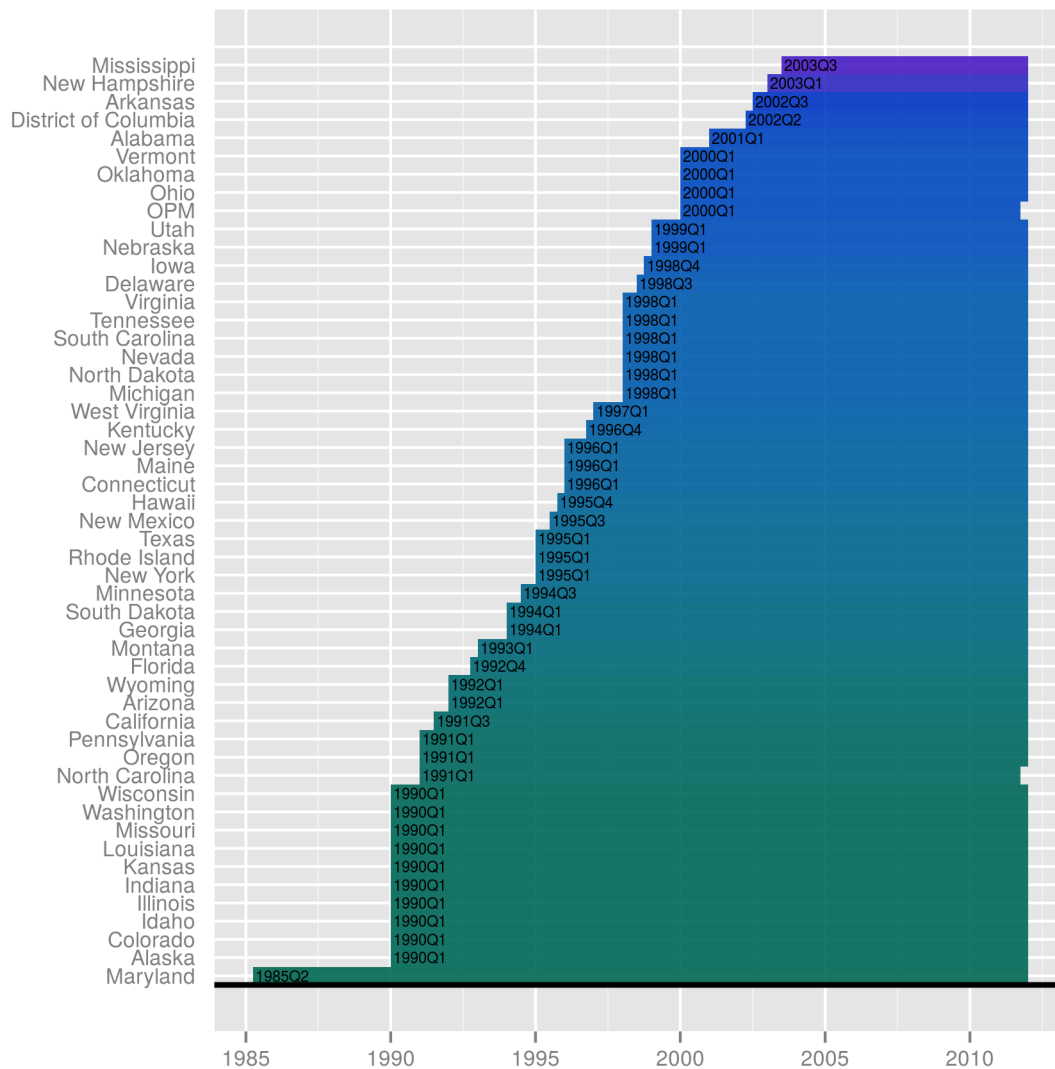
- 262,106,337 people (from ICF_US) working for

- 21,794,809 firms (from EHF_SHF)

Careful planning is required to ensure that adequate resources are available. To facilitate researchers in this endeavor, the research versions of the LEHD Infrastructure files in the RDC environment have additional random variables that allow for the selection of uniform random subsamples of firms (SEIN), establishments (SEINUNIT), and individuals (PIK). No such random variable is available on the EHF, since there is no single good strategy for selecting jobs. Tables in the documentation for individual components also contain information about the size on-disk of each file.

1.7 DISCLOSURE LIMITATION

Special disclosure and data use rules apply to analyses based on the micro-data from the LEHD Infrastructure file system. These data underlie the QWI, and research results are therefore subject to restrictions that ensure the QWI disclosure limitation mechanism is not compromised. Disclosure limitation for the QWI uses noise infusion of the micro-data. The Disclosure Review Board (DRB) does not allow the release of any tabulations for sub-state geography that do not use the QWI noise infusion process. In addition, the required noise factors have not been placed on the RDC snapshot files as part of the DRB's normal rules limiting access to the specific parameters of its approved disclosure limitation methods. Only the DRB may approve the release of tabular output from the LEHD infrastructure file system. Sub-state geography tables will not be approved. National or multi-state tables may be approved provided they do not compromise the protection system. Model-based output is normally allowed. The chief disclosure officer for the RDC network will coordinate the reviews.

The underlying micro-data in the LEHD infrastructure file system were provided to the Census Bureau by states' Labor Market Information (LMI) offices under Memoranda of Understanding (also called Data Use Agreements) negotiated with each state. This process is part of the LED federal/state partnership, and places additional restrictions on the results that may be published. Current members of the LED partnership are shown on the LEHD main web page.

Publicly disclosing a single state's data, or any sub-state information such as Metropolitan Statistical Area (MSA) or Core-Based Statistical Area (CBSA), in identifiable form requires the permission of the state's LMI officer. When reporting results from studies that include multiple states, the results should be pooled across the states. State-specific controls can be included, but no coefficients therefrom reported. The identity of the LED member states is obviously not confidential. You may say which states were used in your analysis, and that you controlled for state-specific factors. The chief disclosure officer for the RDC network will review compliance with this requirement in consultation with the Assistant Division Chief for LEHD.

Additional rules may apply to the use of the ICF (Chapter 9). Please see Section 9.1.3 for more information.

1.8 CITING THE DATA AND SPONSORS

Sponsors

The LEHD Snapshot draws on a data infrastructure that received substantial funding from a number of funding agencies and foundations. We strongly encourage researchers to acknowledge that funding in their paper's "Acknowledgements" or data appendix. The following statement can be used:

Data access

In addition, as more and more journals and funding agencies have stringent data availability requirements (National Science Foundation 2011; American Economic Association 2014; Review of Economics and Statistics 2014; Journal of Labor Economics 2009), researchers will need to work with the Census Bureau to ensure availability of their programs and research extracts. The following statement has been successfully used for accepted papers (provided by John M. Abowd, Cornell University):

```
The data used for this paper were prepared in the U.S. Census Bureau's secure computing
facilities under an authorized project using the Research Data Center network.  The
exact analysis files have been fully archived so that the programming sequence submitted
in compliance with the [JOURNAL]'s editorial policy can be run in its entirety, except
for the component that extracts the analysis sample from the underlying confidential
databases.  I grant any researchers with appropriate Census-approved project permission
to use my exact research files provided that those files were among the ones that they
requested when the approval was obtained (a Census Bureau requirement).  In compliance
with the [JOURNAL]'s editorial policy, I am submitting the list of those files, and
the last known location of the archive on the Census Bureau's RDC network as of [date].
I authorize the editorial staff of the [JOURNAL] to release this list and my statement
of cooperation to any researcher who requests it, as well as to the U.S. Census Bureau
or any agency cooperating with the Census Bureau in supervising research that uses the
restricted-access data that I have used.
```

Data citation

A suggested data citation for each component of the LEHD Snapshot is provided in each chapter, and can be used in the bibliography of researchers' articles (see https://www.icpsr.umich.edu/icpsrweb/ICPSR/curation/citations.jsp for more details on data citations), for instance:

> U.S. Census Bureau. 2014. *Individual Characteristics Files (ICF) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

The full Bibtex file underlying the data citations is attached to this document. LATEX users can simply add the bibliography file to their sources, and cite the data in the text, as they would regular articles:

```
...
I am using the S2011 ICF \citep{S2011:icf}.
...
\bibliographystyle{chicago}
\bibliography{myfile.bib,data.bib}
...
```

which would yield

```
...
I am using the S2011 ICF (U.S.  Census Bureau 2014).
...
Bibliography
```

U.S. Census Bureau. 2014. *Individual Characteristics Files (ICF) in
   LEHD Infrastructure, S2011 Version.* [Computer file]. Wash-
   ington,DC: U.S. Census Bureau, Center for Economic Studies,
   Research Data Centers [distributor].

Users of other bibliographical software can generally import Bibtex files, and should refer to their user manual.

Provenance

Finally, each file that is part of the LEHD Snapshot is tagged with metadata indicating its provenance. We provide a listing of these in each chapter, and they are also encoded into the SAS dataset metadata (obtainable by `proc contents`). While not yet providing a full Handle or Digital Object Identifier (DOI), interested users should be able to leverage this information. The full provenance code ("SnapshotID") is composed of several components:

| snapshot : s2011 : 1 : 421726 |

Fixed name, always equal to `snapshot`
Version of snapshot (used in S2008, S2011)
Revision of snapshot
Identifier derived from LEHD `unique_table_id`

Note that the "SnapshotID" is derived from the LEHD `unique_table_id`, but the tables themselves have been modified, sometimes extensively, to be useful to researchers. Furthermore, in some cases, multile Snapshot files are derived from the same LEHD file, yielding the same provenance code. As such, the "SnapshotID" is not a unique identifier for SAS files in the Snapshot.

   The full provenance code is the entire string "`snapshot:s2011:1:421726`" (in this case for the file `ehf_ak.sas7bdat`), and can be traced back to the LEHD file identified by `unique_table_id = 421726`. For brevity, the tables in each chapter will only list the last two components ("ShortID"), except where this would lead to confusion.

   The exception to the provenance description above are the OPM files, which stem from an experi-
   mental pre-production process, and had not been assigned unique LEHD identifiers at the time of
   S2011 data preparation.

Chapter 2.
Changes to Snapshot S2011

2.1 PREVIOUS VERSIONS

This document updates, but does not replace McKinney and Vilhuber 2011b. Each Snapshot is immutable. Although users are encouraged to use the latest available snapshot, for a variety of reasons, this is not always feasible or desirable. Users who require access to the previous snapshots (S2004, S2008) should contact their RDC administrator for further details.

2.2 MAJOR CHANGES RELATIVE TO PREVIOUS SNAPSHOTS

2.2.1 Scope

The S2011snapshot covers all the states with the exception of Massachusetts, for which data was not available at the time that the snapshot was created. The snapshot may be updated at a later time to include Massachusetts.

This snapshot extends the available time series through 2012Q1, where possible. For state-specific exceptions, please see Table 1.2.

2.2.2 Changes on the ICF

Completely new structure Since the last snapshot (S2008), the ICF has been completely restructured. There now is a single national ICF, rather than state-level ICFs, and missing data is imputed (multiply) only once for any individual, then stored until observed data becomes available (in a later production cycle).

Users wishing to subset by person can condition on selected two-digit (numeric) PIK substrings (`substr(PIK,1,2)`). A separate file contains the longitudinal address information.

Access rules and conditions The National ICF is constructed based on data from the Census Numident (derived from Social Security Administration (SSA) data), Decennial Census 2000 (100 Percent Census Edited File (HCEF) for race/ethnicity, and Sample Census Edited File (SCEF) for education), as well as imputation models which leverage all of the above, plus information on coworkers and neighbors, where the links are inferred from the LEHD Infrastructure and the Composite Person Record (CPR) respectively. The longitudinal address information is derived from CPR information, and is subject to Title 26 restrictions. Address information is completed from 1999 to the most current CPR date, using longitudinal edits and imputation models that condition on contemporaneous coworker information.

Use of the National ICF is thus

- subject to approval by SSA

- subject to approval by Internal Revenue Service (IRS) when using longitudinal address information

- incompatible with simultaneous access to swapped Decennial (100 Percent Detail File (HDF) and Sample Edited Detail File (SEDF))

- subject to additional conditions for the (planned) release of results, above and beyond general RDC and LEHD conditions.

The most recent version of these restrictions and rules are available from the RDC administrators or in the CES Researcher Handbook. We discuss the release restrictions in the next paragraph.

Disclosure avoidance rules for ICF Special rules apply for Census 2000 and ACS tabulations in general, and transfer to the ICF. Note that the National ICF (S2011) itself does not contain or use ACS information. The following is an extract from a memo to LEHD staff by LEHD Senior Management, which was first issued in 2003, and is continuously updated. The text below is from a draft 2013 version, and provided here for reference only. The latest memo always applies, and can be obtained through the RDC Administrator or the LEHD Research Branch Chief.

a. A research project is deemed to use Census 2000 data if any variable used in the production of the tables or research results comes from the HCEF/SCEF Decennial Census file system in use at LEHD.

[...]

c. A research project uses a "special tabulation" from Census 2000 or the ACS if it produces a table of results using input files that contain a variable from Census 2000 (definition 3.a) or ACS (definition 3.b). All special tabulations from Census 2000 or ACS must be directly reviewed by the Disclosure Review Board, except as noted below. See the attached memos for guidelines in preparing such tables. Note, in particular, the population definition rules, the rounding rules, and the required methodology for computing percentiles.

d. The finest level of detail that may be shown for Group Quarters data is Institutional/ Noninstitutional. There are no exceptions to this rule.

e. Special tabulations with geographic detail that is national or state-level may be released without prior DRB approval. LEHD disclosure review is still required.

f. Model-based statistical results (coefficients, standard errors) that were prepared from national or state-level geography may be released without prior DRB approval. If the model includes geographic controls at the sub-state level, the coefficients on these controls may not be released without DRB approval. It is OK to note on the table of coefficients: includes controls for [insert geography]."

The gist is that if researchers do state or national tabulations, they are OK, anything else will require DRB review. Researchers do not need approval by individual states, but the use of the ICF is subject to approval by SSA.

Dropping link variables to SIPP and CPS Furthermore, the ICF's function as a crosswalk to SIPP and CPS was no longer being actively maintained, and has been dropped - no crosswalked identifiers are stored on the ICF anymore, and must be obtained separately by researchers.

2.2.3 Changes to EHF

New Job History File Researchers often combine the EHF with the U2W, in order to obtain establishment-level information on jobs. Both of the inputs have been in previous snapshots. The resulting file, internally called Person History File (PHF)_b, has been available to internal researchers, but not external researchers. The file, with a researcher-friendly name of "Job History File" (JHF), is available in this snapshot. Note that whereas the LEHD production system constructs this variable in the QWI sequence, it is available in the Snapshot as part of the EHF files.

2.2.4 Changes on the ECF

New firm characteristics and link variables on ECF New variables on the ECF provide firm-level age and size data, where a "firm is defined as the economic entity at the national level (across state boundaries). Improved cleaning and coding on the EIN is also incorporated. The new variable FIRMID allows to link to business files such as the LBD or the BR, and from there to many of the economic datasets in the RDC. These variables are labelled "beta" and should be used with caution. More information on their construction is available in Haltiwanger et al. (2014). These data are in active use in public-use QWI, see for instance "Quarterly Workforce Indicators: New Jobs Data by Firm Age and Firm Size" and "Quarterly Workforce Indicators 101." However, because these variables are derived from the BR and LBD, they are subject to Title 26 restrictions (see Section 5.5.7).

New sort order The default sort order of ECF files has been modified to be more convenient for typical researcher use. Researchers are advised that re-sorting files is time-consuming (their problem) and resource intensive (in SAS, with negative externalities for all researchers).

2.2.5 Changes to QWI establishment files

The QWI_SEINUNIT files (internally known as UFF_B) have been expanded. Each file contains the statistics known from the public-use QWI, for each interaction of demographic characteristics. Prior to S2011, only the "WIA" tabulations were available, and the files were simply called "QWI_SEINUNIT". With the release of race, ethnicity, and education tabulations, two additional files have been created, and one file modified:

- QWI_SEINUNIT_WIA is the new name of the previously available file for *age x sex* statistics

- QWI_SEINUNIT_RH contains the same statistics for *race x ethnicity* groups

- QWI_SEINUNIT_SE contains the same statistics for *sex x education* groups

In addition, for the convenience of researchers, a smaller file containing only the marginal categories (i.e., no breakouts by specific groups) was created, as QWI_SEINUNIT_estabtots.

Note that the use of the QWI_SEINUNIT files is incompatible with the use of the QWI public-use files also now part of the S2011 snapshot. Researchers must choose one or the other.

Further note that since release R2013Q2 of the public-use QWI, the shorthand for demographic characteristics "sex-age" has changed from WIA to SA. This is *not* reflected in the S2011 snapshot, which is based on earlier data.

2.2.6 Availability of Successor-Predecessor File

The SPF, which computes worker-flows between firms, and tracks administratively recorded successor-predecessor relationships, is available in this release.

2.2.7 Addition of OPM data on Federal workers

LEHD has been working on integrating OPM data on Federal workers. The current efforts have been contributed to the Snapshot. The value-added to these data are labelled "beta". Data available will complement the EHF, ECF, ICF, U2W, the new Job History File (JHF), and the QWI SEINUNIT-level file, in direct analogies of the existing file structures. RDC users should be able to access these files by requesting a "OPM" dataset. Access to the OPM data do not require state permissions.

2.2.8 Dropping of BRB/LBDB

The BRB and the related LBD Bridge (LBDB) are being dropped as part of the LEHD Snapshot. They are not actively maintained as part of the LEHD statistical production system, and have been used exclusively as a research file. This does not mean that the BRB and LBDB are being dropped from the set of research files available to researchers at the Census Bureau and in the RDC system, only that they won't be refreshed as part of the LEHD Snapshot. Note that an alternate link variable is now available as part of the ECF.

2.2.9 Dropping of GAL crosswalks to AHS, BR, ACS-POW

We are dropping the GAL crosswalks to American Housing Survey (AHS), BR, American Community Survey Place of Work file (ACS-POW), because either the related files are not useful in the RDC (ACS-POW), or because the relevant crosswalks have not been updated in the LEHD production system for over a decade, and are thus of doubtful utility (AHS, BR). We note that this does not affect in any way the availability of the AHS, BR, or American Community Survey (ACS) in the RDC - this only affects the crosswalk created as part of GAL at LEHD to a particular version of those files.

2.2.10 Addition of public-use QWI

The most frequently used files outside of the RDC are the QWIPU tabulations by North American Industry Coding System (NAICS) sub-sector (NAICS3) and county, by the "classic" age-by-sex ("WIA"), sex-by-education (SE), and race-by-ethnicity (RH) tabulations, as well as the beta-release of firm-age and firm-size tabulations by those same demographic classifications. The files are consistent with the overall snapshot (R2012Q4). The total size is approximately 1TB.

Note that the use of the QWI_SEINUNIT files is incompatible with the use of the QWI public-use files. Researchers must choose one or the other. However, use of the QWI public-use files is not subject to any approvals.

2.3 MINOR CHANGES

2.3.1 Geocode

The reference geography for the S2011 has changed to the 2010 (Decennial) geography.
See http://www.census.gov/geo/maps-data/data/tiger.html.

Chapter 3.
Business Register Bridge (BRB) and LBD Bridge (LBDB)

3.1 OVERVIEW

The Business Register Bridge (BRB) is no longer maintained, and has been excluded from the current snapshot. Users should reference the S2008 snapshot (McKinney and Vilhuber 2011a) for the last version.

The LBD Bridge (LBDB) will be updated shortly, and documentation will be made available either in a subsequent release of this document, or as a separate technical paper.

Researchers wishing to link to the LBD should also consider the use of the EIN on the ECF (Chapter 5).

3.2 DATA CITATION

> U.S. Census Bureau. 2014. *Business Register Bridge (BRB) in LEHD Infrastructure, S2008 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

Chapter 4.
Composite Person Record (CPR)

The Composite Person Record (CPR) is a legacy file that, until 2011, was used by LEHD to attach residence information to the infrastructure files. It is generally not available for external projects, and is not documented in the public-use version of this document. The residence information is available, subject to relevant approvals, in the **??**cha:icf).

Chapter 5.
Employer Characteristics File (ECF)

5.1 OVERVIEW

The Employer Characteristics File (ECF) consolidates LEHD employer microdata information on size, location, industry, etc., into two easily accessible files. For each firm identified by SEIN, establishment-level data, identified by SEIN-SEINUNIT, is stored in the "SEINUNIT file." Some information is aggregated to the SEIN level, and stored in the "SEIN file." The SEIN file contains no new information, and should be viewed merely as an easier and/or more efficient way of accessing data aggregated to the firm level. Each file contains one record for every YEAR QUARTER a firm and/or establishment is present in either the ES-202 or the UI. All information is subject to extensive data edits and imputation, and the final files contain no missing information. The files can be linked to other Census data through the use of the LEHD SEIN as well as the EIN.

5.1.1 Changes in Snapshot S2011

New firm characteristics and link variables on ECF New variables on the ECF provide firm-level age and size data, where a "firm is defined as the economic entity at the national level (across state boundaries). Improved cleaning and coding on the EIN is also incorporated. The new variable FIRMID allows to link to business files such as the LBD or the BR, and from there to many of the economic datasets in the RDC. These variables are labelled "beta" and should be used with caution. More information on their construction is available in Haltiwanger et al. (2014). These data are in active use in public-use QWI, see for instance "Quarterly Workforce Indicators: New Jobs Data by Firm Age and Firm Size" and "Quarterly Workforce Indicators 101." However, because these variables are derived from the BR and LBD, they are subject to Title 26 restrictions (see Section 5.5.7).

New sort order The default sort order of ECF files has been modified to be more convenient for typical researcher use. Researchers are advised that re-sorting files is time-consuming (their problem) and resource intensive (in SAS, with negative externalities for all researchers).

5.2 DATA CITATION

> U.S. Census Bureau. 2014. *Composite Person Record (CPR) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

5.3 DETAILED DESCRIPTION

5.3.1 Input Files

- The ES202 (also called Quarterly Census of Employment and Wages (QCEW)) data from the states is the primary input to the ECF file creation process.

- UI data is used to supplement information on the ES202. As part of the creation of the EHF, `ehf_sein_employment` is created. This file contains E (end of period employment), B (beginning of period employment), M (employed anytime in the quarter), and W1 (total wages) calculated similarly to the same measures on the QWI (see Abowd et al. 2006a, 2009). For more details on this file, see Chapter 6.

- GAL data containing lat/long coordinates of the establishments, plus county, Workforce Investment Board (WIB) areas, and CBSA geography. For more details, see Chapter 8.

- Existing files with permanent distortion ("fuzz") factors must be available if data for the state has been officially released (these files are not available to external researchers.)

- SIC and NAICS impute datasets, used for probabilistic SIC-NAICS crosswalks and to impute partially missing industry coding (may be available upon demand).

- BLS-derived control totals, produced by the EHF, see Chapter 6.

5.3.2 Processing Overview

1. First data is read in from the yearly ES202 files and stacked one on top of the other. General and state specific consistency checks are then performed. The COUNTY, NAICS, and EIN data are checked for invalid values. The SIC invalid check is a little more sophisticated. If a 4 digit SIC code is present, but is not valid, then the SIC code undergoes a conditional impute based on the first 2 or 3 digits. If the first 2 or 3 digits are not valid either, then SIC is set to missing (this value will eventually be filled).

   The ES202 data contains a "master" record for multi-unit firms that must be removed. Information in the master record is preserved if data is not available in the establishment records (data is initially allocated equally to each establishment). Various inconsistencies in the record structure are also dealt with, such as 2 records (master and establishment) appearing for a single-unit.

2. The UI data is then integrated with the ES202 data and totals are calculated at the SEIN YEAR QUARTER level.

3. Using both UI and ES202 data a "best" series of variables for payroll and employment is created (these variables are available on auxiliary datasets, see Section 5.5.4 and Section 5.5.6).

4. The allocation process implemented above (master to establishments) does not incorporate any information on the structure of the firm. A flat prior is used in the allocation process (each establishment is assumed to have equal employment and payroll). We improve on this by examining firms with allocated data that previously reported as a multi-unit. The structure of their reports from a previous quarter is then used to allocate payroll and employment. The new records are integrated back into the data, hopefully improving longitudinal consistency at the establishment level.

At this point, the SEIN YEAR QUARTER SEINUNIT dataset record structure is finalized.

5. The GAL is brought into the ECF.

6. The COUNTY, SIC, NAICS, and EIN data are transformed from long to wide format for each SEINUNIT. This dataset is used to fill missing values in these variables with information from other periods for the same establishment.

7. The modal COUNTY, SIC, NAICS, OWNER_CODE, and EIN are calculated (both establishment and employment weighted) for each SEIN in a given YEAR and QUARTER.

8. The SEIN level mode variables (SIC, NAICS, etc) are then transformed from long to wide and the missing values are filled with data from the closest YEAR and QUARTER, if available.

At this point, if an SEIN mode variable has a missing value, then that missing value must be present for every YEAR and QUARTER. The distribution of employment across 4 digit SIC in 1997 is calculated and is used to impute the industry code for each SEIN with missing SIC. These SIC codes are also assigned to the SEINUNIT level data. (Similar processing happens for NAICS)

9. The weights are calculated, based on the expanded BLS controltotals acquired from the EHF.

10. The final step is to apply fuzz (noise distortion) factors to each dataset. The fuzz factor process is done separately for the SEIN and the SEINUNIT data see Abowd et al. 2006b, for more details. Once this is completed the datasets are written to their final location and the master fuzz files are updated.

5.3.3 A note on NAICS codes on the ECF

Enhanced NAICS variables are available on all ECF since February 2003. The variable lists (Section 5.5.3 and Section 5.5.4) show that there are 75 new variables for NAICS alone. The variables can be differentiated mainly by the source(s) and coding system used in their creation. There are two sources of data; the ES202 and the Longitudinal Data Base (LDB) from the Bureau of Labor Statistics (BLS): and three coding systems; NAICS1997, NAICS2002, and NAICS2007 (see the Census web site for more info.). Every NAICS variable uses at least one source and one coding system.

The ESO and FNL variables are of primary importance to the user community. The ESO variables use ONLY information from the ES202 and ignore any information that may be available on the LDB (see Section5.3.5 for some analysis on why this may be preferred). The FNL variables incorporate information from both the ES202 and the LDB, with the LDB being the dominant source. The ES_NAICS_FNL1997 and ES_NAICS_FNL2002 should be used to create the QWI estimates. Neither the ESO and the FNL variables contain missing values.

5.3.4 A note on naming conventions

The variable naming conventions used for internal LEHD files, from which the RDC version of the ECF is derived, stems from the early days of the LEHD program in 1999, and the ES-202 file layout at the time. Since then, the BLS and its partners have implemented a name change for NAICS-related variables (see *ES-202 Technical Memorandum No. S-02-01*):

- NAICS → NSTA (NAICS-SIC Treatment of Auxiliaries)

- AUXNAICS → NAICS (official NAICS coding)

At LEHD, the internal ES202 variable naming scheme for NAICS/NAICS_AUX remains unchanged for compatibility reasons, and this naming scheme carries through into the ECF. Please keep this in mind while reading this document, and while using the ECF.

5.3.5 LDB versus LEHD NAICS backcoding

The Longitudinal Data Base (LDB) algorithm is to some extent a black box and testing has shown that it does a relatively poor job of capturing firm industry changes that occurred during the 1990's. In fact, the LDB appears to be a simple backfill that does not take into account a firm's entire Standard Industry Classification (SIC) history.

Although some of the SIC changes over time may be spurious, a firm's SIC code history contains valuable information that we have attempted to preserve in our imputation algorithm. Overall, the effect of the different approaches is relatively small, since very few firms change industry, in particular relative to the proportion of firms that change geography. In the following, we present a summary of research done at LEHD in 2004 on the ESO vs. FNL NAICS codes. (This research was first completed for the S2004 snapshot, and has not been updated for S2011)

The NAICS_LDB variable is used for about 85% of the records for Illinois, the rest are filled with information from the ES202 (not sure why only 85% of the records on our ES202 files are in the LDB. The results weighted by employment are about the same suggesting that activity was not a criterion for being included on the LDB). First and not surprisingly, in later years and quarters (1999+) when NAICS is actively coded by the states, the codes look almost identical when available.

Second, there is little variation in the LDB NAICS codes over time compared with SIC. Among all of the active SEIN SEINUNITs over the period, a little over 8% experience at least one SIC change compared with about 1.5% on the LDB (almost all of these are 1999+). While this is not entirely unexpected, it is something to keep in mind when comparing NAICS_FNL versus SIC or NAICS_ESO employment totals. Many of these changes in industry appear to be real and are not captured on the LDB.

One effect of this is that as we go back in time a larger portion of employment can be found in NAICS_FNL codes that are different than one would expect given the SIC code on the ECF. For example, in 1990 about 13% of employment is in a NAICS_FNL code that is different than what we would expect based on the SIC. By 2001 this number falls to 3%. The ES202 based NAICS variable does a better job tracking SIC, since more SIC information is used in putting it together (about 3% consistently over the period).

The main source of the discrepancy is due to entities that experience a change in their SIC code prior to 2000. The LDB appears to ignore this change, while the ESO NAICS variable uses an SIC based impute for these SEINUNITS. The result is a series that exhibits similar patterns of change over time as SIC, while still preserving the value added in the NAICS codes for entities that did not experience a change. Also, users should keep in mind that for early years (<1997) some of the NAICS industries have yet to come into existence. We have no estimates on the prevalence of this problem.

5.3.6 Coding of MISS and SRC variables

Each new NAICS variable has several associated variables of which the `miss` and `src` variable are the most important.

5.3.6.1 MISS Variable Codes

If information from another period is used, the flag variable reports how many quarters away the NAICS value was found. Values greater than six should only appear in SEINUNIT level variables. If NAICS is missing for all quarters, then the SEINUNIT value has been filled with the SEIN value. The SEINUNIT codes represent the SEIN value +5.

5.3.6.2 SRC Variable Codes

The ESO and FNL variables use the following source codes. If more detail is desired about the source of the NAICS code, the user must look to the SRC code for that source. For example, if the ESO source code for ES_NAICS_ESO1997 says NCS, then the actual SRC information will be found in ES_NAICS1997_SRC.

The AUX, LDB and standard NAICS codes have the following source variables.

5.3.7 NAICS algorithm precedence ordering

Four basic sources of industry information are available on the ECF; NAICS, NAICS_AUX, SIC, and the NAICS_LDB. The NAICS, NAICS_AUX, and NAICS_LDB missing values were filled using the following pref-

Table 5.1: MISS Variable Codes

0	=	Valid value available in that period
1	=	Missing
1.5	=	(1999 and earlier only) Filled using impute based on SIC due to an SIC change over the period.
2	=	Filled using own code from another period
3	=	Filled from another source contemporaneously
5	=	Filled using the non-employ weight mode (SEIN mode var only)
6	=	Unconditionally imputed (SEIN mode var only)
6	=	NAICS imputed using SIC unconditional impute (SEIN mode var only)
7	=	Filled using the SEIN mode from another period (sic, fnl and eso vars only)
11	=	Filled using unconditional impute of SEIN value (sic, fnl and eso vars only)

Table 5.2: SRC Variable: ESO, FNL

AUX	=	Source is the ES202 NAICS AUX variable
LDB	=	Source is the LDB NAICS variable
NCS	=	Source is the ES202 NAICS variable
SIC	=	Source is the ES202 SIC code

Table 5.3: SRC Variable: AUX, LDB, NAICS

SIC	=	Source is the ES202 SIC code
NO7	=	Source is a NAICS 2007 Code
NO2	=	Source is a NAICS 2002 Code
N97	=	Source is a NAICS 1997 Code

erence ordering. SIC is filled similarly, except miss=1.5 is not used and NAICS, not SIC, would be the basis for the impute when miss=3.

1. Valid 6 digit industry code (miss=0)

2. Imputed code based on first 3,4, or 5 digits when no valid six digit code is available in another period (miss=0)

3. Imputed code based on contemporaneous SIC if SIC changed prior to 2000 (miss=1.5)

4. Valid 6 digit code from another period (miss=2)

5. Valid code from another source (for example if NAICS1997 is missing, NAICS2002 or SIC may be available) (miss=3)

6. Use SEIN mode value (miss=5,7)

7. Unconditional impute (miss=6,11)

5.3.8   ESO and FNL variables

The ESO and FNL variables are made up of combinations of the various sources of industry information. The ESO variable uses the NAICS and NAICS_AUX variables as input. Information from the variable with the lowest MISS value is preferred although in case of a tie the NAICS_AUX value is used.

The FNL variable uses the ESO and LDB variables. Information from the variable with the lowest MISS value is preferred although in case of a tie the NAICS_LDB value is used. Keep in mind that although the source of an ESO or FNL variable may be equal to NCS, the actual source can only be ascertained by going back to the original.

5.3.9   Employment Flag Variable Codes

All current uses of the ECF have been forced to assume that employment and payroll information has been reported by the firm, although under certain conditions the ES202 processing specifications require imputation of missing values. The flag values below allow the user to determine when imputation has occurred.

The master record contains valuable information that has been preserved in the master_empl_month1_flg –master_total_wages_flg variables. For example, one should theoretically be able to distinguish 0 prorated codes from 0 unknowns by looking at multi units with masters that reported (code=1) and subunits with a zero.

The following information stems from an email exchange between Kevin McKinney (U.S. Census Bureau) and George Putnam (Illinois) on 12/15/2003.

Employment Flag Variable Codes Prior to late 1995:

0	*=*	*unknown*
1	*=*	*not imputed*
2	*=*	*imputed (including prorated multiple worksite data)*

Late 1995 or early 1996:

0	*=*	*prorated data (multiple worksites)*
1	*=*	*actual or not imputed data*
2	*=*	*estimated data*

1997 first quarter forward (ES202 processing manual, Appendix B):

Blank	=	*reported data*
R	=	*reported data*
A	=	*estimated from CES report*
C	=	*changed (re-reported)*
D	=	*reported from missing data notice*
E	=	*imputed single unit employment or imputed worksite employment*
		prorated from imputed parent record
H	=	*hand-imputed (not system generated)*
L	=	*late reported (overrides prior imputation)*
M	=	*missing data*
N	=	*zero-filled pending resolution of long-term delinquent reporter*
P	=	*prorated from reported master to worksite*
S	=	*aggregated master from reported MWR or EDI data*
W	=	*estimated from wage record employment*
X	=	*non-numeric employment zero-filled pending further action*

5.3.10 Multi-Unit Code or MEEI

The MULTI_UNIT variable on the ECF is determined by counting the number of SEINUNIT records for a given SEIN once the master records have been removed. However, some multiunit firms refuse to report detailed information for their sub-units and appear as single units on the ECF. The table below provides an estimate of the magnitude of multiunit firms refusing to report detailed unit information using data from Illinois.

	MULTI_UNIT	
MULTI_UNIT_CODE	0	1
1	1,485,000	0
2	0	0
3	> 0	155,000
4	5,000	0
5	0	> 0
6	15,000	0

Prior to 1997 (ES202 processing manual sent from George Putnam):

1	=	Single establishment unit
2	=	Multi-unit master record
3	=	Subunit establishment level record for a multi-unit employer
4	=	Multi-establishment employer reporting as a single unit due to unavailability of data, including refusals
5	=	A subunit record that actually represents a combination of establishments; finer level breakouts are not yet available
6	=	Known multi establishment employer reporting as a single unit and not solicited for disaggregation because of small employment (< 10) in all secondary establishments combined

1997 first quarter forward (ES202 processing manual, Appendix B):

1	=	Single establishment unit
2	=	Multi-unit master record
3	=	Subunit establishment level record for a multi-unit employer
4	=	Multi-establishment employer reporting as a single unit due to unavailability of data, including refusals
5	=	A subunit record that actually represents a combination of establishments; finer level breakouts are not yet available
6	=	Known multi establishment employer reporting as a single unit and not solicited for disaggregation because of small employment (< 10) in all secondary establishments combined

5.3.11   Auxiliary Code

This variable gives detailed information about firm locations that do not directly engage in production related activities.

Prior to 1997 (ES202 processing manual sent from George Putnam):

0	=	Unknown
1	=	Central administrative office
2	=	Performs research, development or testing services
3	=	Provides storage or warehouse services
5	=	Does not provide auxiliary services, it is an operating establishment
9	=	Performs auxiliary services that are not described above

1997 first quarter forward (ES202 processing manual, Appendix B):

0	=	Auxiliary status not known
1	=	Central administrative office
2	=	Performs research, development or testing services
3	=	Provides storage or warehouse services
5	=	Does not provide auxiliary services, it is an operating establishment
6	=	Headquarters
7	=	Administrative, Other than Headquarters
9	=	Performs auxiliary services that are not described above

5.4   ECF RESEARCH VERSION, TITLE 26, AND THE STRUCTURE OF FILES IN THE CENSUS RESEARCH ENVIRONMENT

Because some data elements on the internal-use ECF are considered Title 26-protected, the structure of the files has been slightly modified for the RDC environment to facilitate and streamline project proposals by clearly identifying files without any Title 26-protected data.

All Title 26-protected information has been removed from the main ECF files, and stored in separate files with the same record count. These files have the necessary unique record identifiers to be merged back to the main ECF files, if the researcher has acquired the appropriate permissions.

Furthermore, since the S2008 snapshot, the ECF files have been restructured (relative to the LEHD Production source files and earlier snapshots) to be more user-friendly. By construction, the ECF files produced by LEHD have much auxiliary diagnostic information, helpful for identifying sources of imputations and edits, and allowing for the use of multiple sources of industry coding, geography, and other establishment and firm characteristics. However, most researchers do not need that kind of detail. We separated the auxiliary output from the core ECF content, creating a leaner, easier-to-use ECF. The auxiliary output remains available in a separate file (suffixed with _aux).

We also made a cleaner distinction between SEIN (firm) level variables and SEINUNIT (establishment) variables. Whereas (currently) the LEHD Production ECF has some SEIN-level variables on the SEINUNIT file, this is not the case for the research files.

Finally, names on research files have been made more consistent, highlighting data sources more consistently, and thus may differ from names on internal files. Table 5.4 on page 5-30 shows the correspondence between internal and research names. (Note: this document has not been fully revised, and in some descriptive sections, the internal names are still used.)

We note that no new variables were created, and no variables were removed. The information is simply organized in a (hopefully) simpler way. Users who previously used the S2004 snapshot may need to restructure their programs. See Section 5.6 for one possible program to assist in that endeavor. There is no straightforward way to reconstruct the previous data structure from the current data files.

Users will find all Title 13-protected content in the directories

```
ecf/ecf_XX_sein.sas7bdat
ecf/ecf_XX_sein_aux.sas7bdat
ecf/ecf_XX_seinunit.sas7bdat
ecf/ecf_XX_seinunit_aux.sas7bdat
```

Title 26-protected content can be found in

```
ecft26/ecf_XX_t26.sas7bdat
```

Additional Census-confidential information (in directory `ecfcc`) is not available to researchers.

5.5   DATA SET DESCRIPTIONS

5.5.1   Naming scheme

There are five files in the ECF/ECFT26 group, with an additional three files in the CC group: SAS datasets with zero observations are attached to this document:[1]

- ecf/ecf_zz_sein.sas7bdat

- ecf/ecf_zz_sein_aux.sas7bdat

- ecf/ecf_zz_seinunit.sas7bdat

- ecf/ecf_zz_seinunit_aux.sas7bdat

- ecfcc/ecf_zz_cc.sas7bdat

- ecfcc/ecf_zz_leg.sas7bdat

- ecfcc/ecf_zz_sein_fuzz.sas7bdat

- ecfcc/ecf_zz_seinunit_fuzz.sas7bdat

- ecft26/ecf_zz_t26.sas7bdat

`ZZ` stands for the state postal abbreviation. Files with `_t26` contain FTI, are stored in separate subdirectories and require a separate set of permissions. Files with `_cc` or `_fuzz` contain Census-confidential information and are generally not available to external researchers. Either set of files are of little use without the regular ECF group data.

5.5.2   Data location

The files are stored in three main directories, with state-specific subdirectories:

```
ecf/ZZ/        for most files
ecft26/ZZ      for files with Title 26 protected content
ecfcc/ZZ       for files with Census-confidential content
```

On the RDC network, all directories can be found under

```
/mixed/lehd/s2011
```

---

1. Also visible on the attachment tab - Adobe Reader may be required.

5.5.3 Main SEINUNIT dataset: ecf_zz_seinunit

SEINUNIT-level file, research variables only.

Record identifier: SEIN SEINUNIT YEAR QUARTER

Sort order: SEIN YEAR QUARTER SEINUNIT

File indexes: none

Entity "establishment" or State Employment Security Agency (SESA)

Unique Entity Key SEIN SEINUNIT

Note that SEINUNIT is only unique within any given time period within SEIN.

Field name	Data dictionary reference name	Starting position	Field size	Data type
0=ok,1=not found,2+found off qtr	ES_COUNTY_MISS	00079	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_FNL1997_MISS	00085	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_FNL2002_MISS	00088	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_FNL2007_MISS	00091	3	N
0=ok,1=not found,2+found off qtr	ES_OWNER_CODE_MISS	00082	3	N
0=ok,1=not found,2+found off qtr	ES_SIC_MISS	00076	3	N
1=UI only,2=202 only,3=both	SOURCE	00059	3	N
5-digit Core-Based Statistical Area	LEG_CBSA	00135	5	A/N
Best UI/202 Employment Month 1	BEST_EMP1	00044	4	N
Best UI/202 Employment Month 2	BEST_EMP2	00048	4	N
Best UI/202 Employment Month 3	BEST_EMP3	00052	4	N
Best UI/202 Wages	BEST_WAGES	00062	5	N
CBSA Type 1=Metro, 2=Micro, Else=9	LEG_CBSA_MEMI	00140	1	A/N
Census Block suffix 1	LEG_BLOCK_SUF1	00133	1	A/N
Census Block suffix 2	LEG_BLOCK_SUF2	00134	1	A/N
Census block within tract	LEG_BLOCK	00129	4	A/N
Cleaned ES202 FIPS County CCC	ES_COUNTY	00229	3	A/N
Cleaned GEO FIPS County CCC	LEG_COUNTY	00190	3	A/N
Cleaned GEO FIPS County CCC, pre-longitudinal impute	LEG_COUNTY_ORIG	00222	3	A/N
Cleaned GEO State SS	LEG_STATE	00186	2	A/N
Cleaned OWNER_CODE O	ES_OWNER_CODE	00232	1	A/N
Cleaned SIC Code IIII	ES_SIC	00225	4	A/N
Cleaned SIC Division I	ES_SIC_DIV	00263	1	A/N
Continuous Time YEAR QUARTER	YR_QTR	00106	6	A/N
ES202 FIPS State SS	ES_STATE	00188	2	A/N
FIPS state——FIPS county——Census tract	LEG_GEOCODE	00118	11	A/N
Final 1997 NAICS Code NNNNNN	ES_NAICS_FNL1997	00233	6	A/N
Final 2002 NAICS Code NNNNNN	ES_NAICS_FNL2002	00239	6	A/N
Final 2007 NAICS Code NNNNNN	ES_NAICS_FNL2007	00245	6	A/N
Final GALID	LEG_GALID	00157	29	A/N
Firm engaged (not) engaged in production	QCEW_AUXILIARY_CODE	00117	1	A/N
GALID, pre-longitudinal impute	LEG_GALID_ORIG	00193	29	A/N
Latitude, 6 implied decimal places	LEG_LATITUDE	00000	8	N
Longitude, 6 implied decimal places	LEG_LONGITUDE	00008	8	N
Number of Establishments	NUM_ESTABS	00032	4	N
Quality of final geography	LEG_GEO_QUAL	00070	3	N
Quality of geography, pre-longitudinal impute	LEG_GEO_QUAL_ORIG	00073	3	N
Quarter QQ	QUARTER	00040	4	N
Random sample selector for SEIN	SAMPLE_SEIN	00016	8	N
Random sample selector for SEINUNIT	SAMPLE_SEINUNIT	00024	8	N
SEIN w/2+ records on 202	MULTI_UNIT	00056	3	N
Source of Ind Code	ES_NAICS_FNL1997_SRC	00254	3	A/N
Source of Ind Code	ES_NAICS_FNL2002_SRC	00257	3	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Source of Ind Code	ES_NAICS_FNL2007_SRC	00260	3	A/N
Source of Ind Code	ES_SIC_SRC	00251	3	A/N
Source of best_ data	BEST_FLAG	00067	3	N
State Employer Identification Number	SEIN	00094	12	A/N
State UI Reporting Unit Number	SEINUNIT	00112	5	A/N
Sub-county Geography from the LEG	LEG_SUBCTYGEO	00147	10	A/N
Workforce Investment Board area	LEG_WIB	00141	6	A/N
Year YYYY	YEAR	00036	4	N

5.5.4 Auxiliary SEINUNIT dataset: ecf_zz_seinunit_aux

SEINUNIT-level file, auxiliary and diagnostic variables only.

Record identifier: SEIN SEINUNIT YEAR QUARTER

Sort order: SEIN YEAR QUARTER SEINUNIT

File indexes: none

Entity "establishment" or SESA

Unique Entity Key SEIN SEINUNIT

This file can be merged onto the main SEINUNIT file using the specified identifiers in sort order. It is generally not needed by researchers, but made available for those requiring more detailed longitudinal information on imputes and edits.

Field name	Data dictionary reference name	Starting position	Field size	Data type
0 if seinunit=00000	SEINUNIT_TYPE	00035	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS1997_MISS	00067	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS2002_MISS	00070	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS2007_MISS	00073	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_AUX1997_MISS	00076	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_AUX2002_MISS	00079	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_AUX2007_MISS	00082	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_ESO1997_MISS	00166	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_ESO2002_MISS	00169	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_ESO2007_MISS	00172	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_IMP1997_MISS	00000	8	N
0=ok,1=not found,2+found off qtr	ES_NAICS_IMP2002_MISS	00008	8	N
0=ok,1=not found,2+found off qtr	ES_NAICS_IMP2007_MISS	00016	8	N
0=ok,1=not found,2+found off qtr	ES_NAICS_LDB1997_MISS	00085	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_LDB2002_MISS	00088	3	N
0=ok,1=not found,2+found off qtr	ES_NAICS_LDB2007_MISS	00091	3	N
0=ok,1=not found,2+found off qtr label	GEO_VARS_MISS	00094	3	N
1=county 2=es_county (long edit) 3=mode_leg_county_emp 4=mode_leg_county 5=leg_c	COUNTY_IMPUTE_SOURCE	00175	3	N
Cleaned 1997 NAICS Code NNNNNN	ES_NAICS1997	00300	6	A/N
Cleaned 1997 NAICS Code NNNNNN	ES_NAICS_AUX1997	00318	6	A/N
Cleaned 1997 NAICS Code NNNNNN	ES_NAICS_LDB1997	00336	6	A/N
Cleaned 2002 NAICS Code NNNNNN	ES_NAICS2002	00306	6	A/N
Cleaned 2002 NAICS Code NNNNNN	ES_NAICS_AUX2002	00324	6	A/N
Cleaned 2002 NAICS Code NNNNNN	ES_NAICS_LDB2002	00342	6	A/N
Cleaned 2007 NAICS Code NNNNNN	ES_NAICS2007	00312	6	A/N
Cleaned 2007 NAICS Code NNNNNN	ES_NAICS_AUX2007	00330	6	A/N
Cleaned 2007 NAICS Code NNNNNN	ES_NAICS_LDB2007	00348	6	A/N
ES202 ONLY 1997 NAICS Code NNNNNN	ES_NAICS_ESO1997	00354	6	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
ES202 ONLY 2002 NAICS Code NNNNNN	ES_NAICS_ESO2002	00360	6	A/N
ES202 ONLY 2007 NAICS Code NNNNNN	ES_NAICS_ESO2007	00366	6	A/N
Flag, number of quarters to find geocodes	LEG_FLAG_GEO	00064	3	N
GALID of address on es202	ES_GALID	00271	29	A/N
Multiunit Imputed Record Structure	STRUCTURE_FIX	00061	3	N
NAICS Code not Valid	NAICS_1997_INVALID	00236	1	A/N
NAICS Code not Valid	NAICS_2002_INVALID	00237	1	A/N
NAICS Code not Valid	NAICS_2007_INVALID	00238	1	A/N
NAICS Code not Valid	NAICS_AUX_1997_INVALID	00239	1	A/N
NAICS Code not Valid	NAICS_AUX_2002_INVALID	00240	1	A/N
NAICS Code not Valid	NAICS_AUX_2007_INVALID	00241	1	A/N
NAICS Code not Valid	NAICS_LDB_1997_INVALID	00250	1	A/N
NAICS Code not Valid	NAICS_LDB_2002_INVALID	00251	1	A/N
NAICS Code not Valid	NAICS_LDB_2007_INVALID	00252	1	A/N
Original ES202 County	QCEW_COUNTY	00247	3	A/N
Original ES202 Employment Month 1	QCEW_EMPL_MONTH1	00038	5	N
Original ES202 Employment Month 2	QCEW_EMPL_MONTH2	00043	5	N
Original ES202 Employment Month 3	QCEW_EMPL_MONTH3	00048	5	N
Original ES202 SIC	QCEW_SIC	00243	4	A/N
Original ES202 wages	QCEW_TOTAL_WAGES	00053	5	N
Original NAICS 1997 Code	QCEW_NAICS1997	00199	6	A/N
Original NAICS 2002 Code	QCEW_NAICS2002	00205	6	A/N
Original NAICS 2007 Code	QCEW_NAICS2007	00211	6	A/N
Original NAICS AUX 1997 Code	QCEW_NAICS_AUX1997	00217	6	A/N
Original NAICS AUX 2002 Code	QCEW_NAICS_AUX2002	00223	6	A/N
Original NAICS AUX 2007 Code	QCEW_NAICS_AUX2007	00229	6	A/N
Original NAICS LDB 1997 Code	QCEW_NAICS_LDB1997	00253	6	A/N
Original NAICS LDB 2002 Code	QCEW_NAICS_LDB2002	00259	6	A/N
Original NAICS LDB 2007 Code	QCEW_NAICS_LDB2007	00265	6	A/N
Original Owner Code	QCEW_OWNER_CODE	00235	1	A/N
Quarter QQ	QUARTER	00028	4	N
Quarters Away County data found	ES_COUNTY_FLAG	00127	3	N
Quarters Away LEG variables found	GEO_VARS_FLAG	00133	3	N
Quarters Away NAICS data found	ES_NAICS1997_FLAG	00100	3	N
Quarters Away NAICS data found	ES_NAICS2002_FLAG	00103	3	N
Quarters Away NAICS data found	ES_NAICS2007_FLAG	00106	3	N
Quarters Away NAICS data found	ES_NAICS_AUX1997_FLAG	00109	3	N
Quarters Away NAICS data found	ES_NAICS_AUX2002_FLAG	00112	3	N
Quarters Away NAICS data found	ES_NAICS_AUX2007_FLAG	00115	3	N
Quarters Away NAICS data found	ES_NAICS_LDB1997_FLAG	00118	3	N
Quarters Away NAICS data found	ES_NAICS_LDB2002_FLAG	00121	3	N
Quarters Away NAICS data found	ES_NAICS_LDB2007_FLAG	00124	3	N
Quarters Away OWNER_CODE data found	ES_OWNER_CODE_FLAG	00130	3	N
Quarters Away SIC data found	ES_SIC_FLAG	00097	3	N
Reported or imputed Month 1 Employment	QCEW_EMPL_MONTH1_FLG	00195	1	A/N
Reported or imputed Month 2 Employment	QCEW_EMPL_MONTH2_FLG	00196	1	A/N
Reported or imputed Month 3 Employment	QCEW_EMPL_MONTH3_FLG	00197	1	A/N
Reported or imputed Total Wages	QCEW_TOTAL_WAGES_FLG	00198	1	A/N
SEINUNIT data non-numeric	SEINUNIT_BAD	00032	3	N
SIC Code not Valid	QCEW_SIC_INVALID	00242	1	A/N
SIC IMP 1997 NAICS Code NNNNNN	ES_NAICS_IMP1997	00372	6	A/N
SIC IMP 2002 NAICS Code NNNNNN	ES_NAICS_IMP2002	00378	6	A/N
SIC IMP 2007 NAICS Code NNNNNN	ES_NAICS_IMP2007	00384	6	A/N
Seinunit has some NAICS info	ES_NAICS1997_VALID	00139	3	N
Seinunit has some NAICS info	ES_NAICS2002_VALID	00142	3	N
Seinunit has some NAICS info	ES_NAICS2007_VALID	00145	3	N
Seinunit has some NAICS info	ES_NAICS_AUX1997_VALID	00148	3	N
Seinunit has some NAICS info	ES_NAICS_AUX2002_VALID	00151	3	N
Seinunit has some NAICS info	ES_NAICS_AUX2007_VALID	00154	3	N
Seinunit has some NAICS info	ES_NAICS_LDB1997_VALID	00157	3	N
Seinunit has some NAICS info	ES_NAICS_LDB2002_VALID	00160	3	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Seinunit has some NAICS info	ES_NAICS_LDB2007_VALID	00163	3	N
Seinunit has some SIC info	ES_SIC_VALID	00136	3	N
Source of Ind Code	ES_NAICS1997_SRC	00390	3	A/N
Source of Ind Code	ES_NAICS2002_SRC	00393	3	A/N
Source of Ind Code	ES_NAICS2007_SRC	00396	3	A/N
Source of Ind Code	ES_NAICS_AUX1997_SRC	00399	3	A/N
Source of Ind Code	ES_NAICS_AUX2002_SRC	00402	3	A/N
Source of Ind Code	ES_NAICS_AUX2007_SRC	00405	3	A/N
Source of Ind Code	ES_NAICS_ESO1997_SRC	00426	3	A/N
Source of Ind Code	ES_NAICS_ESO2002_SRC	00429	3	A/N
Source of Ind Code	ES_NAICS_ESO2007_SRC	00432	3	A/N
Source of Ind Code	ES_NAICS_IMP1997_SRC	00417	3	A/N
Source of Ind Code	ES_NAICS_IMP2002_SRC	00420	3	A/N
Source of Ind Code	ES_NAICS_IMP2007_SRC	00423	3	A/N
Source of Ind Code	ES_NAICS_LDB1997_SRC	00408	3	A/N
Source of Ind Code	ES_NAICS_LDB2002_SRC	00411	3	A/N
Source of Ind Code	ES_NAICS_LDB2007_SRC	00414	3	A/N
State Employer Identification Number	SEIN	00178	12	A/N
State UI Reporting Unit Number	SEINUNIT	00190	5	A/N
Year YYYY	YEAR	00024	4	N
candidate for structure fix	SPECIAL_HANDLE	00058	3	N

5.5.5 Main SEIN dataset: ecf_zz_sein

SEIN-level file, with variables aggregated from the establishment level.

Record identifier: SEIN YEAR QUARTER

Sort order: SEIN YEAR QUARTER

File indexes: none

Entity "firm"

Unique Entity Key SEIN

Note that SEIN is unique within any given time period across all states, but may not be uniquely identify an entity over time within a state, as the underlying UI account numbers can and do get re-used.

Field name	Data dictionary reference name	Starting position	Field size	Data type
1=UI only,2=202 only,3=both	SOURCE	00065	3	N
ES202 multi-unit (non) reporter	MULTI_UNIT_CODE	00151	1	A/N
Emp Mode Cleaned County	MODE_ES_COUNTY_EMP	00182	3	A/N
Emp Mode Cleaned GEO CBSA	MODE_LEG_CBSA_EMP	00253	5	A/N
Emp Mode Cleaned GEO CBSA type	MODE_LEG_CBSA_MEMI_EMP	00252	1	A/N
Emp Mode Cleaned GEO COUNTY	MODE_LEG_COUNTY_EMP	00223	3	A/N
Emp Mode Cleaned GEO COUNTY	MODE_LEG_SUBCTYGEO_EMP	00236	10	A/N
Emp Mode Cleaned GEO COUNTY, pre-longitudinal impute	MODE_LEG_COUNTY_ORIG_EMP	00249	3	A/N
Emp Mode Cleaned GEO STATE	MODE_LEG_STATE_EMP	00218	2	A/N
Emp Mode Cleaned GEO WIB	MODE_LEG_WIB_EMP	00210	6	A/N
Emp Mode Cleaned NAICS 1997	MODE_ES_NAICS_FNL1997_EMP	00185	6	A/N
Emp Mode Cleaned NAICS 2002	MODE_ES_NAICS_FNL2002_EMP	00191	6	A/N
Emp Mode Cleaned NAICS 2007	MODE_ES_NAICS_FNL2007_EMP	00197	6	A/N
Emp Mode Cleaned OWNER_CODE	MODE_ES_OWNER_CODE_EMP	00203	1	A/N
Emp Mode Cleaned SIC	MODE_ES_SIC_EMP	00178	4	A/N
First Quarter SEIN on 202	MULTI_FIRST_QUARTER	00062	3	N
First Year SEIN on 202	MULTI_FIRST_YEAR	00059	3	N
Missing Value	MODE_ES_COUNTY_EMP_MISS	00118	3	N
Missing Value	MODE_ES_COUNTY_MISS	00085	3	N
Missing Value	MODE_ES_NAICS_FNL1997_EMP_MISS	00109	3	N
Missing Value	MODE_ES_NAICS_FNL1997_MISS	00076	3	N
Missing Value	MODE_ES_NAICS_FNL2002_EMP_MISS	00112	3	N
Missing Value	MODE_ES_NAICS_FNL2002_MISS	00079	3	N
Missing Value	MODE_ES_NAICS_FNL2007_EMP_MISS	00115	3	N
Missing Value	MODE_ES_NAICS_FNL2007_MISS	00082	3	N
Missing Value	MODE_ES_OWNER_CODE_EMP_MISS	00121	3	N
Missing Value	MODE_ES_OWNER_CODE_MISS	00088	3	N
Missing Value	MODE_ES_SIC_EMP_MISS	00106	3	N
Missing Value	MODE_ES_SIC_MISS	00073	3	N
Missing Value	MODE_LEG_CBSA_EMP_MISS	00127	3	N
Missing Value	MODE_LEG_CBSA_MISS	00094	3	N
Missing Value	MODE_LEG_COUNTY_EMP_MISS	00133	3	N
Missing Value	MODE_LEG_COUNTY_MISS	00100	3	N
Missing Value	MODE_LEG_COUNTY_ORIG_EMP_MISS	00008	8	N
Missing Value	MODE_LEG_COUNTY_ORIG_MISS	00000	8	N
Missing Value	MODE_LEG_STATE_EMP_MISS	00130	3	N
Missing Value	MODE_LEG_STATE_MISS	00097	3	N
Missing Value	MODE_LEG_SUBCTYGEO_EMP_MISS	00136	3	N
Missing Value	MODE_LEG_SUBCTYGEO_MISS	00103	3	N
Missing Value	MODE_LEG_WIB_EMP_MISS	00124	3	N
Missing Value	MODE_LEG_WIB_MISS	00091	3	N
Number of Establishments	NUM_ESTABS	00032	4	N

Chapter 5: Employer Characteristics File (ECF)

Field name	Data dictionary reference name	Starting position	Field size	Data type
Quarter QQ	QUARTER	00040	4	N
Random sample selector for SEIN	SAMPLE_SEIN	00024	8	N
SEIN Best UI/202 Month 1, Employment	SEIN_BEST_EMP1	00044	4	N
SEIN Best UI/202 Month 2, Employment	SEIN_BEST_EMP2	00048	4	N
SEIN Best UI/202 Month 3, Employment	SEIN_BEST_EMP3	00052	4	N
SEIN Best UI/202 Payroll	SEIN_BEST_WAGES	00068	5	N
SEIN w/2+ records on 202	MULTI_UNIT	00056	3	N
State Employer Identification Number	SEIN	00139	12	A/N
Unit Mode Cleaned County	MODE_ES_COUNTY	00156	3	A/N
Unit Mode Cleaned GEO CBSA	MODE_LEG_CBSA	00258	5	A/N
Unit Mode Cleaned GEO CBSA type	MODE_LEG_CBSA_MEMI	00263	1	A/N
Unit Mode Cleaned GEO COUNTY	MODE_LEG_COUNTY	00220	3	A/N
Unit Mode Cleaned GEO COUNTY	MODE_LEG_SUBCTYGEO	00226	10	A/N
Unit Mode Cleaned GEO COUNTY, pre-longitudinal impute	MODE_LEG_COUNTY_ORIG	00246	3	A/N
Unit Mode Cleaned GEO STATE	MODE_LEG_STATE	00216	2	A/N
Unit Mode Cleaned GEO WIB	MODE_LEG_WIB	00204	6	A/N
Unit Mode Cleaned NAICS 1997	MODE_ES_NAICS_FNL1997	00159	6	A/N
Unit Mode Cleaned NAICS 2002	MODE_ES_NAICS_FNL2002	00165	6	A/N
Unit Mode Cleaned NAICS 2007	MODE_ES_NAICS_FNL2007	00171	6	A/N
Unit Mode Cleaned OWNER_CODE	MODE_ES_OWNER_CODE	00177	1	A/N
Unit Mode Cleaned SIC	MODE_ES_SIC	00152	4	A/N
Weight sum(B_UI)=sum(month1_BLS)	QWI_UNIT_WEIGHT	00016	8	N
Year YYYY	YEAR	00036	4	N

5.5.6 Auxiliary SEIN dataset: ecf_zz_sein_aux

SEIN-level file, auxiliary and diagnostic variables only.

Record identifier: SEIN YEAR QUARTER

Sort order: SEIN YEAR QUARTER

File indexes: none

Entity "firm"

Unique Entity Key SEIN

This file can be merged onto the main SEIN file using the specified identifiers in sort order. It is generally not needed by researchers, but made available for those requiring more detailed longitudinal information on imputes and edits.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Best SEIN UI Employment	EMP1_UI	00048	4	N
Best SEIN UI Employment	EMP2_UI	00052	4	N
Best SEIN UI Employment	EMP3_UI	00056	4	N
Emp Mode Cleaned NAICS 1997	MODE_ES_NAICS_ESO1997_EMP	00239	6	A/N
Emp Mode Cleaned NAICS 2002	MODE_ES_NAICS_ESO2002_EMP	00245	6	A/N
Emp Mode Cleaned NAICS 2007	MODE_ES_NAICS_ESO2007_EMP	00251	6	A/N
MULTI ever ES202 wages	EVER_WAGES	00073	3	N
MULTI ever has ES202 month 1 employment	EVER_EMP1	00076	3	N
MULTI ever has ES202 month 2 employment	EVER_EMP2	00079	3	N
MULTI ever has ES202 month 3 employment	EVER_EMP3	00082	3	N
Missing Value	MODE_ES_NAICS_ESO1997_EMP_MISS	00153	3	N
Missing Value	MODE_ES_NAICS_ESO1997_MISS	00102	3	N
Missing Value	MODE_ES_NAICS_ESO2002_EMP_MISS	00156	3	N
Missing Value	MODE_ES_NAICS_ESO2002_MISS	00105	3	N
Missing Value	MODE_ES_NAICS_ESO2007_EMP_MISS	00159	3	N
Missing Value	MODE_ES_NAICS_ESO2007_MISS	00108	3	N
Original UI Payroll Info W1	UI_PAYROLL	00060	5	N
Quarter QQ	QUARTER	00020	4	N
Quarters Away Data Found	MODE_ES_COUNTY_EMP_FLAG	00183	3	N
Quarters Away Data Found	MODE_ES_COUNTY_FLAG	00132	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO1997_EMP_FLAG	00165	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO1997_FLAG	00114	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO2002_EMP_FLAG	00168	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO2002_FLAG	00117	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO2007_EMP_FLAG	00171	3	N
Quarters Away Data Found	MODE_ES_NAICS_ESO2007_FLAG	00120	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL1997_EMP_FLAG	00174	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL1997_FLAG	00123	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL2002_EMP_FLAG	00177	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL2002_FLAG	00126	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL2007_EMP_FLAG	00180	3	N
Quarters Away Data Found	MODE_ES_NAICS_FNL2007_FLAG	00129	3	N
Quarters Away Data Found	MODE_ES_OWNER_CODE_EMP_FLAG	00186	3	N
Quarters Away Data Found	MODE_ES_OWNER_CODE_FLAG	00135	3	N
Quarters Away Data Found	MODE_ES_SIC_EMP_FLAG	00162	3	N
Quarters Away Data Found	MODE_ES_SIC_FLAG	00111	3	N
Quarters Away Data Found	MODE_LEG_CBSA_EMP_FLAG	00192	3	N
Quarters Away Data Found	MODE_LEG_CBSA_FLAG	00141	3	N
Quarters Away Data Found	MODE_LEG_COUNTY_EMP_FLAG	00198	3	N
Quarters Away Data Found	MODE_LEG_COUNTY_FLAG	00147	3	N
Quarters Away Data Found	MODE_LEG_COUNTY_ORIG_EMP_FLAG	00008	8	N
Quarters Away Data Found	MODE_LEG_COUNTY_ORIG_FLAG	00000	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Quarters Away Data Found	MODE_LEG_STATE_EMP_FLAG	00195	3	N
Quarters Away Data Found	MODE_LEG_STATE_FLAG	00144	3	N
Quarters Away Data Found	MODE_LEG_SUBCTYGEO_EMP_FLAG	00201	3	N
Quarters Away Data Found	MODE_LEG_SUBCTYGEO_FLAG	00150	3	N
Quarters Away Data Found	MODE_LEG_WIB_EMP_FLAG	00189	3	N
Quarters Away Data Found	MODE_LEG_WIB_FLAG	00138	3	N
SEIN 202 Employment Month 1	QCEW_SEIN_EMP1	00036	4	N
SEIN 202 Employment Month 2	QCEW_SEIN_EMP2	00040	4	N
SEIN 202 Employment Month 3	QCEW_SEIN_EMP3	00044	4	N
SEIN 202 Wages	QCEW_SEIN_WAGES	00065	5	N
SEIN UI Wages	UI_WAGES	00097	5	N
SEIN ever multi unit	EVER_MULTI	00070	3	N
SEIN ever on 202	EVER_202	00085	3	N
SEIN ever on UI	EVER_UI	00088	3	N
SEIN in ES202	IN_202	00094	3	N
SEIN in UI	IN_UI	00091	3	N
State Employer Identification Number	SEIN	00204	12	A/N
Stored Master Multi Code	MASTER_MULTI_UNIT_CODE	00220	1	A/N
Stored Master Record Flag	MASTER_EMPL_MONTH1_FLG	00216	1	A/N
Stored Master Record Flag	MASTER_EMPL_MONTH2_FLG	00217	1	A/N
Stored Master Record Flag	MASTER_EMPL_MONTH3_FLG	00218	1	A/N
Stored Master Record Flag	MASTER_TOTAL_WAGES_FLG	00219	1	A/N
UI Employment B	UI_SEINSIZE_B	00032	4	N
UI Employment E	UI_SEINSIZE_E	00028	4	N
UI Employment M	UI_SEINSIZE_M	00024	4	N
Unit Mode Cleaned NAICS 1997	MODE_ES_NAICS_ESO1997	00221	6	A/N
Unit Mode Cleaned NAICS 2002	MODE_ES_NAICS_ESO2002	00227	6	A/N
Unit Mode Cleaned NAICS 2007	MODE_ES_NAICS_ESO2007	00233	6	A/N
Year YYYY	YEAR	00016	4	N

5.5.7 Auxiliary T26 dataset: ecf_zz_t26

T26 variables associated with both the SEIN and the SEINUNIT-level file. For California, this includes the EIN. For all states, this includes any variables derived from T26 datasets, primarily the BR. National firm-age and firm-size variables are on this file.

Record identifier: SEIN SEINUNIT YEAR QUARTER

Sort order: SEIN SEINUNIT YEAR QUARTER

File indexes: none

Entity "establishment" or SESA

Unique Entity Key SEIN SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
1 if EIN does not match LBD in this year or +/- 1 year (0=this year, 2=next year, 3=previous year)	FAS_EIN_MATCH_LBD	00000	8	N
1=LBD, 2=ECF	FAS_SOURCE_SIZE	00101	1	A/N
1=LBD, 2=ECF, 3=imputed	FAS_SOURCE_AGE	00100	1	A/N
Best firm age	FAS_FIRM_AGE	00008	8	N
Best initial firm size (size March 12 of last year, current size if new)	FAS_FIRM_SIZE	00016	8	N
EIN cleaned for Firm Age	FAS_EIN	A/N	00076	ECF_ZZ_T26
FAS: Firm is a multi-unit on BDS	FAS_MULTI_UNIT_BDS	00024	8	N
FAS: Imputation flag for firm_age	FAS_FIRM_AGE_FLAG	00032	8	N
FAS: Imputation flag for firm_size	FAS_FIRM_SIZE_FLAG	00040	8	N
Firm Alpha from LBD	FAS_FIRM_ID	00090	10	A/N
Info about fas_ein variable	FAS_EIN_FLAG	00056	3	N
Quarter Q	QUARTER	00052	4	N
State Employer Identification Number	SEIN	00059	12	A/N
State UI Reporting Unit Number	SEINUNIT	00071	5	A/N
Year YYYY	YEAR	00048	4	N

5.5.8 Auxiliary SEINUNIT T26 dataset: ecf_zz_seinunit_t26

(S2011) The information in this file has been subsumed into the `ecf_zz_t26` file. For earlier snapshots, consult the relevant documentation.

5.5.9 Auxiliary SEIN T26 dataset: ecf_zz_sein_t26

(S2011) The information in this file has been subsumed into the `ecf_zz_t26` file. For earlier snapshots, consult the relevant documentation.

5.5.10   Summary information on datasets

Table 5.10: Number of observations for ECF

Group	Number of datafiles	Records (1000s)	Filesize (GB)
ECF	200	1,312,055	379

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 5.11: List of data files for ECF, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Alaska (ak)					
ecf_ak_sein	1990Q1	2012Q1	64	< 1	1:423575
ecf_ak_sein_aux	1990Q1	2012Q1	64	< 1	1:423575
ecf_ak_seinunit	1990Q1	2012Q1	1,615	< 1	1:423575
ecf_ak_seinunit_aux	1990Q1	2012Q1	1,615	< 1	1:423575
Alabama (al)					
ecf_al_sein	2001Q1	2012Q1	221	< 1	1:423197
ecf_al_sein_aux	2001Q1	2012Q1	221	< 1	1:423197
ecf_al_seinunit	2001Q1	2012Q1	6,310	2	1:423197
ecf_al_seinunit_aux	2001Q1	2012Q1	6,310	2	1:423197
Arkansas (ar)					
ecf_ar_sein	2002Q3	2012Q1	143	< 1	1:423239
ecf_ar_sein_aux	2002Q3	2012Q1	143	< 1	1:423239
ecf_ar_seinunit	2002Q3	2012Q1	3,211	1	1:423239
ecf_ar_seinunit_aux	2002Q3	2012Q1	3,211	1	1:423239
Arizona (az)					
ecf_az_sein	1992Q1	2012Q1	388	< 1	1:421590
ecf_az_sein_aux	1992Q1	2012Q1	388	< 1	1:421590
ecf_az_seinunit	1992Q1	2012Q1	9,164	2	1:421590
ecf_az_seinunit_aux	1992Q1	2012Q1	9,164	3	1:421590
California (ca)					
ecf_ca_sein	1991Q1	2012Q1	3,506	1	1:421707
ecf_ca_sein_aux	1991Q1	2012Q1	3,506	< 1	1:421707
ecf_ca_seinunit	1991Q1	2012Q1	96,870	24	1:421707
ecf_ca_seinunit_aux	1991Q1	2012Q1	96,870	31	1:421707
Colorado (co)					
ecf_co_sein	1990Q1	2012Q1	554	< 1	1:444396
ecf_co_sein_aux	1990Q1	2012Q1	554	< 1	1:444396
ecf_co_seinunit	1990Q1	2012Q1	13,078	3	1:444396
ecf_co_seinunit_aux	1990Q1	2012Q1	13,078	4	1:444396
Connecticut (ct)					

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 5.11 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ecf_ct_sein	1996Q1	2012Q1	264	< 1	1:424112
ecf_ct_sein_aux	1996Q1	2012Q1	264	< 1	1:424112
ecf_ct_seinunit	1996Q1	2012Q1	7,131	2	1:424112
ecf_ct_seinunit_aux	1996Q1	2012Q1	7,131	2	1:424112
District of Columbia (dc)					
ecf_dc_sein	2000Q4	2012Q1	70	< 1	1:423351
ecf_dc_sein_aux	2000Q4	2012Q1	70	< 1	1:423351
ecf_dc_seinunit	2000Q4	2012Q1	1,474	< 1	1:423351
ecf_dc_seinunit_aux	2000Q4	2012Q1	1,474	< 1	1:423351
Delaware (de)					
ecf_de_sein	1997Q1	2012Q1	60	< 1	1:430016
ecf_de_sein_aux	1997Q1	2012Q1	60	< 1	1:430016
ecf_de_seinunit	1997Q1	2012Q1	1,674	< 1	1:430016
ecf_de_seinunit_aux	1997Q1	2012Q1	1,674	1	1:430016
Florida (fl)					
ecf_fl_sein	1989Q1	2012Q1	1,734	< 1	1:424190
ecf_fl_sein_aux	1989Q1	2012Q1	1,734	< 1	1:424190
ecf_fl_seinunit	1989Q1	2012Q1	43,968	11	1:424190
ecf_fl_seinunit_aux	1989Q1	2012Q1	43,968	15	1:424190
Georgia (ga)					
ecf_ga_sein	1994Q1	2012Q1	629	< 1	1:423302
ecf_ga_sein_aux	1994Q1	2012Q1	629	< 1	1:423302
ecf_ga_seinunit	1994Q1	2012Q1	16,763	4	1:423302
ecf_ga_seinunit_aux	1994Q1	2012Q1	16,763	5	1:423302
Hawaii (hi)					
ecf_hi_sein	1995Q4	2012Q1	82	< 1	1:427211
ecf_hi_sein_aux	1995Q4	2012Q1	82	< 1	1:427211
ecf_hi_seinunit	1995Q4	2012Q1	2,375	1	1:427211
ecf_hi_seinunit_aux	1995Q4	2012Q1	2,375	1	1:427211
Iowa (ia)					
ecf_ia_sein	1990Q1	2012Q1	226	< 1	1:424309
ecf_ia_sein_aux	1990Q1	2012Q1	226	< 1	1:424309
ecf_ia_seinunit	1990Q1	2012Q1	8,111	2	1:424309
ecf_ia_seinunit_aux	1990Q1	2012Q1	8,111	3	1:424309
Idaho (id)					
ecf_id_sein	1990Q1	2012Q1	164	< 1	1:426050
ecf_id_sein_aux	1990Q1	2012Q1	164	< 1	1:426050
ecf_id_seinunit	1990Q1	2012Q1	4,088	1	1:426050
ecf_id_seinunit_aux	1990Q1	2012Q1	4,088	1	1:426050
Illinois (il)					
ecf_il_sein	1990Q1	2012Q1	986	< 1	1:442790
ecf_il_sein_aux	1990Q1	2012Q1	986	< 1	1:442790
ecf_il_seinunit	1990Q1	2012Q1	31,166	8	1:442790
ecf_il_seinunit_aux	1990Q1	2012Q1	31,166	10	1:442790

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 5.11 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Indiana (in)					
ecf_in_sein	1990Q1	2012Q1	415	< 1	1:427805
ecf_in_sein_aux	1990Q1	2012Q1	415	< 1	1:427805
ecf_in_seinunit	1990Q1	2012Q1	12,269	3	1:427805
ecf_in_seinunit_aux	1990Q1	2012Q1	12,269	4	1:427805
Kansas (ks)					
ecf_ks_sein	1990Q1	2012Q1	245	< 1	1:427812
ecf_ks_sein_aux	1990Q1	2012Q1	245	< 1	1:427812
ecf_ks_seinunit	1990Q1	2012Q1	7,270	2	1:427812
ecf_ks_seinunit_aux	1990Q1	2012Q1	7,270	2	1:427812
Kentucky (ky)					
ecf_ky_sein	1996Q4	2012Q1	205	< 1	1:423782
ecf_ky_sein_aux	1996Q4	2012Q1	205	< 1	1:423782
ecf_ky_seinunit	1996Q4	2012Q1	6,188	2	1:423782
ecf_ky_seinunit_aux	1996Q4	2012Q1	6,188	2	1:423782
Louisiana (la)					
ecf_la_sein	1990Q1	2012Q1	322	< 1	1:443294
ecf_la_sein_aux	1990Q1	2012Q1	322	< 1	1:443294
ecf_la_seinunit	1990Q1	2012Q1	10,165	3	1:443294
ecf_la_seinunit_aux	1990Q1	2012Q1	10,165	3	1:443294
Maryland (md)					
ecf_md_sein	1985Q2	2012Q1	533	< 1	1:427239
ecf_md_sein_aux	1985Q2	2012Q1	533	< 1	1:427239
ecf_md_seinunit	1985Q2	2012Q1	14,865	4	1:427239
ecf_md_seinunit_aux	1985Q2	2012Q1	14,865	5	1:427239
Maine (me)					
ecf_me_sein	1996Q1	2012Q1	120	< 1	1:433221
ecf_me_sein_aux	1996Q1	2012Q1	120	< 1	1:433221
ecf_me_seinunit	1996Q1	2012Q1	3,323	1	1:433221
ecf_me_seinunit_aux	1996Q1	2012Q1	3,323	1	1:433221
Michigan (mi)					
ecf_mi_sein	1998Q1	2012Q1	545	< 1	1:433241
ecf_mi_sein_aux	1998Q1	2012Q1	545	< 1	1:433241
ecf_mi_seinunit	1998Q1	2012Q1	15,359	4	1:433241
ecf_mi_seinunit_aux	1998Q1	2012Q1	15,359	5	1:433241
Minnesota (mn)					
ecf_mn_sein	1994Q3	2012Q1	341	< 1	1:421880
ecf_mn_sein_aux	1994Q3	2012Q1	341	< 1	1:421880
ecf_mn_seinunit	1994Q3	2012Q1	12,103	3	1:421880
ecf_mn_seinunit_aux	1994Q3	2012Q1	12,103	4	1:421880
Missouri (mo)					
ecf_mo_sein	1990Q1	2012Q1	560	< 1	1:426027
ecf_mo_sein_aux	1990Q1	2012Q1	560	< 1	1:426027
ecf_mo_seinunit	1990Q1	2012Q1	15,803	4	1:426027

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 5.11 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ecf_mo_seinunit_aux	1990Q1	2012Q1	15,803	5	1:426027
Mississippi (ms)					
ecf_ms_sein	2003Q3	2012Q1	111	< 1	1:426057
ecf_ms_sein_aux	2003Q3	2012Q1	111	< 1	1:426057
ecf_ms_seinunit	2003Q3	2012Q1	2,440	1	1:426057
ecf_ms_seinunit_aux	2003Q3	2012Q1	2,440	1	1:426057
Montana (mt)					
ecf_mt_sein	1993Q1	2012Q1	107	< 1	1:430023
ecf_mt_sein_aux	1993Q1	2012Q1	107	< 1	1:430023
ecf_mt_seinunit	1993Q1	2012Q1	3,007	1	1:430023
ecf_mt_seinunit_aux	1993Q1	2012Q1	3,007	1	1:430023
North Carolina (nc)					
ecf_nc_sein	1990Q1	2011Q4	649	< 1	1:445006
ecf_nc_sein_aux	1990Q1	2011Q4	649	< 1	1:445006
ecf_nc_seinunit	1990Q1	2011Q4	21,401	6	1:445006
ecf_nc_seinunit_aux	1990Q1	2011Q4	21,401	7	1:445006
North Dakota (nd)					
ecf_nd_sein	1998Q1	2012Q1	55	< 1	1:430029
ecf_nd_sein_aux	1998Q1	2012Q1	55	< 1	1:430029
ecf_nd_seinunit	1998Q1	2012Q1	1,437	< 1	1:430029
ecf_nd_seinunit_aux	1998Q1	2012Q1	1,437	< 1	1:430029
Nebraska (ne)					
ecf_ne_sein	1999Q1	2012Q1	115	< 1	1:430044
ecf_ne_sein_aux	1999Q1	2012Q1	115	< 1	1:430044
ecf_ne_seinunit	1999Q1	2012Q1	3,009	1	1:430044
ecf_ne_seinunit_aux	1999Q1	2012Q1	3,009	1	1:430044
New Hampshire (nh)					
ecf_nh_sein	2003Q1	2012Q1	84	< 1	1:432237
ecf_nh_sein_aux	2003Q1	2012Q1	84	< 1	1:432237
ecf_nh_seinunit	2003Q1	2012Q1	1,789	< 1	1:432237
ecf_nh_seinunit_aux	2003Q1	2012Q1	1,789	1	1:432237
New Jersey (nj)					
ecf_nj_sein	1995Q1	2012Q1	695	< 1	1:421873
ecf_nj_sein_aux	1995Q1	2012Q1	695	< 1	1:421873
ecf_nj_seinunit	1995Q1	2012Q1	19,462	5	1:421873
ecf_nj_seinunit_aux	1995Q1	2012Q1	19,462	6	1:421873
New Mexico (nm)					
ecf_nm_sein	1990Q1	2012Q1	155	< 1	1:431614
ecf_nm_sein_aux	1990Q1	2012Q1	155	< 1	1:431614
ecf_nm_seinunit	1990Q1	2012Q1	4,337	1	1:431614
ecf_nm_seinunit_aux	1990Q1	2012Q1	4,337	1	1:431614
Nevada (nv)					
ecf_nv_sein	1998Q1	2012Q1	168	< 1	1:443454
ecf_nv_sein_aux	1998Q1	2012Q1	168	< 1	1:443454

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 5.11 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ecf_nv_seinunit	1998Q1	2012Q1	3,638	1	1:443454
ecf_nv_seinunit_aux	1998Q1	2012Q1	3,638	1	1:443454
New York (ny)					
ecf_ny_sein	1990Q1	2012Q1	1,935	< 1	1:443256
ecf_ny_sein_aux	1990Q1	2012Q1	1,935	< 1	1:443256
ecf_ny_seinunit	1990Q1	2012Q1	48,106	13	1:443256
ecf_ny_seinunit_aux	1990Q1	2012Q1	48,106	16	1:443256
Ohio (oh)					
ecf_oh_sein	2000Q1	2012Q1	488	< 1	1:443461
ecf_oh_sein_aux	2000Q1	2012Q1	488	< 1	1:443461
ecf_oh_seinunit	2000Q1	2012Q1	14,256	4	1:443461
ecf_oh_seinunit_aux	2000Q1	2012Q1	14,256	5	1:443461
Oklahoma (ok)					
ecf_ok_sein	1999Q1	2012Q1	201	< 1	1:421894
ecf_ok_sein_aux	1999Q1	2012Q1	201	< 1	1:421894
ecf_ok_seinunit	1999Q1	2012Q1	5,274	1	1:421894
ecf_ok_seinunit_aux	1999Q1	2012Q1	5,274	2	1:421894
Oregon (or)					
ecf_or_sein	1990Q1	2012Q1	383	< 1	1:421904
ecf_or_sein_aux	1990Q1	2012Q1	383	< 1	1:421904
ecf_or_seinunit	1990Q1	2012Q1	10,035	3	1:421904
ecf_or_seinunit_aux	1990Q1	2012Q1	10,035	3	1:421904
Pennsylvania (pa)					
ecf_pa_sein	1991Q1	2012Q1	881	< 1	1:424197
ecf_pa_sein_aux	1991Q1	2012Q1	881	< 1	1:424197
ecf_pa_seinunit	1991Q1	2012Q1	27,164	7	1:424197
ecf_pa_seinunit_aux	1991Q1	2012Q1	27,164	9	1:424197
Rhode Island (ri)					
ecf_ri_sein	1990Q1	2012Q1	116	< 1	1:434249
ecf_ri_sein_aux	1990Q1	2012Q1	116	< 1	1:434249
ecf_ri_seinunit	1990Q1	2012Q1	3,050	1	1:434249
ecf_ri_seinunit_aux	1990Q1	2012Q1	3,050	1	1:434249
South Carolina (sc)					
ecf_sc_sein	1998Q1	2012Q1	265	< 1	1:437461
ecf_sc_sein_aux	1998Q1	2012Q1	265	< 1	1:437461
ecf_sc_seinunit	1998Q1	2012Q1	6,585	2	1:437461
ecf_sc_seinunit_aux	1998Q1	2012Q1	6,585	2	1:437461
South Dakota (sd)					
ecf_sd_sein	1994Q1	2012Q1	68	< 1	1:430057
ecf_sd_sein_aux	1994Q1	2012Q1	68	< 1	1:430057
ecf_sd_seinunit	1994Q1	2012Q1	1,973	1	1:430057
ecf_sd_seinunit_aux	1994Q1	2012Q1	1,973	1	1:430057
Tennessee (tn)					
ecf_tn_sein	1998Q1	2012Q1	343	< 1	1:426214

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 5.11 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ecf_tn_sein_aux	1998Q1	2012Q1	343	< 1	1:426214
ecf_tn_seinunit	1998Q1	2012Q1	8,295	2	1:426214
ecf_tn_seinunit_aux	1998Q1	2012Q1	8,295	3	1:426214
Texas (tx)					
ecf_tx_sein	1990Q1	2012Q1	1,525	< 1	1:423789
ecf_tx_sein_aux	1990Q1	2012Q1	1,525	< 1	1:423789
ecf_tx_seinunit	1990Q1	2012Q1	45,644	12	1:423789
ecf_tx_seinunit_aux	1990Q1	2012Q1	45,644	15	1:423789
Utah (ut)					
ecf_ut_sein	1990Q1	2012Q1	242	< 1	1:433251
ecf_ut_sein_aux	1990Q1	2012Q1	242	< 1	1:433251
ecf_ut_seinunit	1990Q1	2012Q1	5,995	2	1:433251
ecf_ut_seinunit_aux	1990Q1	2012Q1	5,995	2	1:433251
Virginia (va)					
ecf_va_sein	1995Q3	2012Q1	515	< 1	1:437563
ecf_va_sein_aux	1995Q3	2012Q1	515	< 1	1:437563
ecf_va_seinunit	1995Q3	2012Q1	13,608	4	1:437563
ecf_va_seinunit_aux	1995Q3	2012Q1	13,608	4	1:437563
Vermont (vt)					
ecf_vt_sein	2000Q1	2012Q1	48	< 1	1:430037
ecf_vt_sein_aux	2000Q1	2012Q1	48	< 1	1:430037
ecf_vt_seinunit	2000Q1	2012Q1	1,203	< 1	1:430037
ecf_vt_seinunit_aux	2000Q1	2012Q1	1,203	< 1	1:430037
Washington (wa)					
ecf_wa_sein	1990Q1	2012Q1	764	< 1	1:426034
ecf_wa_sein_aux	1990Q1	2012Q1	764	< 1	1:426034
ecf_wa_seinunit	1990Q1	2012Q1	17,977	5	1:426034
ecf_wa_seinunit_aux	1990Q1	2012Q1	17,977	6	1:426034
Wisconsin (wi)					
ecf_wi_sein	1990Q1	2012Q1	410	< 1	1:421887
ecf_wi_sein_aux	1990Q1	2012Q1	410	< 1	1:421887
ecf_wi_seinunit	1990Q1	2012Q1	13,202	3	1:421887
ecf_wi_seinunit_aux	1990Q1	2012Q1	13,202	4	1:421887
West Virginia (wv)					
ecf_wv_sein	1990Q1	2012Q1	87	< 1	1:441338
ecf_wv_sein_aux	1990Q1	2012Q1	87	< 1	1:441338
ecf_wv_seinunit	1990Q1	2012Q1	4,311	1	1:441338
ecf_wv_seinunit_aux	1990Q1	2012Q1	4,311	1	1:441338
Wyoming (wy)					
ecf_wy_sein	1992Q1	2012Q1	69	< 1	1:430051
ecf_wy_sein_aux	1992Q1	2012Q1	69	< 1	1:430051
ecf_wy_seinunit	1992Q1	2012Q1	1,626	< 1	1:430051
ecf_wy_seinunit_aux	1992Q1	2012Q1	1,626	1	1:430051

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

5.6 HELPFUL PROGRAMS

The following programs might be found to be useful when using the data.

5.6.1 Renaming from internal to research ECF names

The program ecf_rename.sas will rename internal ECF files to research (RDC) versions of ECF.

```
/* $Id: ecf_rename.sas 11171 2014-03-28 21:36:22Z vilhuber $ */
/* $URL: https://trac.vilhuber.org/svn/LEHD/branches/ticket983-snapshot2012/05_documentation/ecf_rename.sas $ */
/* renames internal ECF variable names
   to research ECF variable names used in the S2008 and forward.

   Usage: include this SAS program in any SAS programming sequence
          (either %include it or put it into a SAS macro autocall
           location), and then insert '%ecf_rename;' into
           a data step.
*/

%macro ecf_rename;

rename=
        auxiliary_code=qcew_auxiliary_code
        county         =qcew_county
        ecf_only       =fas_ein_match_lbd
        ein            =qcew_ein
        ein_bad        =qcew_ein_bad
        ein_defect     =qcew_ein_defect
        empl_month1    =qcew_empl_month1
        empl_month2    =qcew_empl_month2
        empl_month3    =qcew_empl_month3
        empl_month1_flg=qcew_empl_month1_flg
        empl_month2_flg=qcew_empl_month2_flg
        empl_month3_flg=qcew_empl_month3_flg
        firm_age_flag  =fas_firm_age_flag
        firm_size_flag =fas_firm_size_flag
        firmage        =fas_firm_age
        firmid         =fas_firm_id
        firmsize       =fas_firm_size
        firmsize_fuzz  =fas_firm_size_fuzz
        multi_unit_bds=fas_multi_unit_bds
        naics1997      =qcew_naics1997
        naics2002      =qcew_naics2002
        naics2007      =qcew_naics2007
        naics_aux1997  =qcew_naics_aux1997
        naics_aux2002  =qcew_naics_aux2002
        naics_aux2007  =qcew_naics_aux2007
        naics_ldb1997  =qcew_naics_ldb1997
        naics_ldb2002  =qcew_naics_ldb2002
        naics_ldb2007  =qcew_naics_ldb2007
        owner_code     =qcew_owner_code
        payroll=ui_payroll
        sein_emp1=qcew_sein_emp1
        sein_emp2=qcew_sein_emp2
        sein_emp3=qcew_sein_emp3
        sein_wages=qcew_sein_wages
        seinsize_b=ui_seinsize_b
        seinsize_e=ui_seinsize_e
        seinsize_m=ui_seinsize_m
        sic            =qcew_sic
        sic_invalid    =qcew_sic_invalid
        source_age     =fas_source_age
```

```
        source_size   =fas_source_size
        total_wages   =qcew_total_wages
        total_wages_flg=qcew_total_wages_flg
        valid_ein=qcew_valid_ein
        wages_UI=ui_wages
;
%mend;
```

5.6.2   Selecting a random sample of establishments

The ECF files can be large, and researchers may wish to analyze only a random subsample of firms. The variables SAMPLE_SEIN and SAMPLE_SEINUNIT can be used to select a random sample of the ECF. To do this in a space-efficient way, the following code can be used as a template.

```
%let state=tx;
libname INLIB "/mixed/lehd/s2011/ecf/&state./";

data mydata/view=mydata;
   set INLIB.ecf_&state._seinunit
       (where=(sample_seinunit <= 0.05));
run;

proc reg data=mydata;
model y= x w z;
run;
```

The code above uses a VIEW, which means the dataset is constructed on the fly as it is used in the analysis procedure. Although overall disk usage is not necessarily smaller when using random access (as the SAS regression procedure apparently does), it is still faster. For other processes using sequential access only, in particular simple DATA steps, a view will be space-efficient because only the relevant observations are streamed into any intermediate data files.

5.7   NOTES

Table 5.4: Renaming of ECF variables in the RDC

Internal name → External name
auxiliary_code → qcew_auxiliary_code
county → qcew_county
ecf_only → fas_ein_match_lbd
ein → qcew_ein
ein_bad → qcew_ein_bad
ein_defect → qcew_ein_defect
empl_month1 → qcew_empl_month1
empl_month2 → qcew_empl_month2
empl_month3 → qcew_empl_month3
empl_month1_flg → qcew_empl_month1_flg
empl_month2_flg → qcew_empl_month2_flg
empl_month3_flg → qcew_empl_month3_flg
firm_age_flag → fas_firm_age_flag
firm_size_flag → fas_firm_size_flag
firmage → fas_firm_age
firmid → fas_firm_id
firmsize → fas_firm_size
firmsize_fuzz → fas_firm_size_fuzz
multi_unit_bds → fas_multi_unit_bds
naics1997 → qcew_naics1997
naics2002 → qcew_naics2002
naics2007 → qcew_naics2007
naics_aux1997 → qcew_naics_aux1997
naics_aux2002 → qcew_naics_aux2002
naics_aux2007 → qcew_naics_aux2007
naics_ldb1997 → qcew_naics_ldb1997
naics_ldb2002 → qcew_naics_ldb2002
naics_ldb2007 → qcew_naics_ldb2007
owner_code → qcew_owner_code
payroll → ui_payroll
sein_emp1 → qcew_sein_emp1
sein_emp2 → qcew_sein_emp2
sein_emp3 → qcew_sein_emp3
sein_wages → qcew_sein_wages
seinsize_b → ui_seinsize_b
seinsize_e → ui_seinsize_e
seinsize_m → ui_seinsize_m
sic → qcew_sic
sic_invalid → qcew_sic_invalid
source_age → fas_source_age
source_size → fas_source_size
total_wages → qcew_total_wages
total_wages_flg → qcew_total_wages_flg
valid_ein → qcew_valid_ein
wages_UI → ui_wages

Chapter 6.
Employment History Files (EHF)

6.1 OVERVIEW

The Employment History Files (EHF) are designed to store the complete in-state history of employment, for each individual that appears in the UI wage records employed at some firm, and for each firm and establishment that appear in the QCEW records with positive employment at some time.

The core EHF for each state contains one record for each employee-employer combination–a job–in that state in each year. Both annual and quarterly earnings variables are available in the EHF. Individuals who are employed, but never have strictly positive earnings at their employing SEIN (a theoretical possibility) in a given year do not have a record in the EHF for that year.

To facilitate analysis, the EHF data are restructured into "wide" file containing one observation per job (PIK-SEIN combination), with all quarterly earnings and activity information available on that record. The restructured file is called the PHF. It should be noted that the actual file structure is at the PIK-SEIN-SEINUNIT-YEAR level for the EHF, and at the PIK-SEIN-SEINUNIT level for the PHF. Although only one state (Minnesota) has non-zero values for SEINUNIT, this allows the file structure to be homogeneous across states. An active job within a quarter, the primary job-level economic activity measure, is defined as having strictly positive quarterly earnings for the individual-employer pair that define the job.

Researchers often combine the EHF with the U2W, in order to obtain establishment-level information on jobs. Both of the inputs have been in previous snapshots. The resulting file of such a merge, internally called PHF_b, has been available to internal researchers, but not external researchers. Starting with S2011, the merged file, with a researcher-friendly name of Job History File (JHF), is available to all researchers. Note that whereas the LEHD production system constructs this variable in the QWI sequence, it is available in the Snapshot as part of the EHF files.

A history of observed activity (positive employment) in the QCEW records, is available and computed at the SEINUNIT level (Unit History File, UHF) and the SEIN level (SEIN History File, SHF).

6.1.1 Changes in Snapshot S2011

New Job History File   Researchers often combine the EHF with the U2W, in order to obtain establishment-level information on jobs. Both of the inputs have been in previous snapshots. The resulting file, internally called PHF_b, has been available to internal researchers, but not external researchers. The file, with a researcher-friendly name of "Job History File" (JHF), is available in this snapshot. Note that whereas the LEHD production system constructs this variable in the QWI sequence, it is available in the Snapshot as part of the EHF files.

6.2 DATA CITATION

> U.S. Census Bureau. 2014. *Employer Characteristics Files (ECF) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

6.3   INPUT FILES

6.3.1   Wage records: UI

Wage records correspond to the report of an individual's UI-covered earnings by an employing entity, identified by a state UI account number (called the SEIN in the LEHD system). An individual's UI wage record is retained in the processing if at least one employer reports earnings of at least one dollar for that individual during the quarter. Thus, an in-scope job must produce at least one dollar of UI-covered earnings during a given quarter in the LEHD universe. Maximum earnings reported are defined in a specific state's unemployment insurance system, and observed top-coding varies across states and over time.

A record is completed with information on the individual's Social Security Number (later replaced with the PIK within the LEHD system), first name, last name, and middle initial. A few states include additional information: the firm's reporting unit or establishment (SEINUNIT), available for Minnesota, and a crucial component to the Unit-to-Worker impute described later; weeks worked, available for some years in Florida; hours worked, available for Washington and Minnesota state.

Current UI wage records are reported for the quarter that ended approximately six months prior to the reporting date at Census (the first day of the calendar quarter). Wage records are also reported for the quarter that the state considers "final" in the sense that revisions to its administrative UI wage record data base after that date are relatively rare. This quarter typically ends nine months prior to the reporting date. Historical UI wage records were assembled by the partner states from their administrative record backup systems.

6.3.2   Employer reports: QCEW - ES-202

The employer reports are based on information from each state's Department of Employment Security. The data are collected as part of the Covered Employment and Wages (CEW) 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. 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." 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. These data are also used to compile the QCEW and the Business Employment Dynamics (BED) data at the BLS.

The information contained in the ES-202 reports has increased substantially over the years. Employers report wages subject to statutory payroll taxes on this form, together with some other information. Common to all years, and critical to LEHD processing, are information on the employer's identity (the SEIN), the reporting unit's identify (SEINUNIT), ownership information, employment on the 12th of each month covered by the quarter, and total wages paid over the course of the quarter. Additional information pertains to industry classifications (initially SIC, and later NAICS). Other information include the federal EIN, geography both at a high level (county or MSA) and low level (physical location street address and mailing address). A recent expansion of the standard report's record layout has increased the informational content substantially. The LEHD Infrastructure File system is, fundamentally, a job-based frame designed to be represent the universe of individual-employer pairs covered by state unemployment insurance system reporting requirements. Thus, the underlying data are wage records extracted from Unemployment Insurance (UI) administrative files from each LED partner state. In addition to the UI wage records, LED partner states also deliver an extract of the file reported to the Bureau of Labor Statistic's Quarterly Census of Employment and Wages (QCEW, formerly known as ES-202). These data are received by LEHD on a quarterly basis, with historical time series extending back to the early 1990s for many states.

6.4 DATA SET DESCRIPTIONS

6.4.1 Naming scheme

All files start with `ehf`.

SAS datasets with zero observations are attached to this document:[1]

- ehf/ehf_zz.sas7bdat

- ehf/ehf_zz_controltotals.sas7bdat

- ehf/ehf_zz_phf.sas7bdat

- ehf/ehf_zz_sein_employment.sas7bdat

- ehf/ehf_zz_shf.sas7bdat

- ehf/ehf_zz_uhf.sas7bdat

- ehf/ehf_zz_uniqpik.sas7bdat

- ehf/jhf_zz.sas7bdat

`ZZ` stands for the state postal abbreviation. The main EHF file has no suffix, other files have a suffix.

6.4.2 Data location

The files are stored in two main directories, with state-specific subdirectories:

```
ehf/ZZ/ for most files
```

No files in the EHF process contain Title 26 data. On the RDC network, the directory can be found under

```
/mixed/lehd/s2011
```

---

1. Also visible on the attachment tab - Adobe Reader may be required.

6.4.3   UI-based Output Files

6.4.3.1   EHF

The EHF is designed to store the complete in-state work history for each individual that appears in the UI wage records. The EHF for each state contains one record for each employee-employer combination in that state in each year. Every individual who is employed during a given year will then have one observation per employer for that year. Annual earnings and quarterly earnings variables are present on the file. The presence of positive quarterly earnings is used in the job flow analysis not only to compute earnings and payroll statistics but also to determine an individual's employment status each quarter.

The EHF (`ehf_&state.`) is organized by PIK-SEIN-SEINUNIT-YEAR. Note that all states except Minnesota (MN) have SEINUNIT='00000', so this reverts back to PIK-SEIN-YEAR for all states except MN.

Record identifier PIK-SEIN-SEINUNIT-YEAR

Sort order PIK-SEIN-SEINUNIT-YEAR

Entity Job

Unique Entity Key PIK-SEIN-SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
Annual earnings	EARN_ANN	00003	5	N
Calendar year	YEAR	00000	3	N
Protected Identification Key	PIK	00028	9	A/N
Qtr 1 earnings	EARN1	00008	5	N
Qtr 2 earnings	EARN2	00013	5	N
Qtr 3 earnings	EARN3	00018	5	N
Qtr 4 earnings	EARN4	00023	5	N
Source of data (FIPS state code/0=Fed)	SOURCE	00037	2	A/N
State Employer Identification Number	SEIN	00041	12	A/N
State UI Reporting Unit Number	SEINUNIT	00053	5	A/N
Type of source	SOURCETP	00039	2	A/N

6.4.3.2 (proto-)PHF

The proto PHF is a reformatted version of the EHF. Rather than having one record per year, the PHF is organized by "job", or unique employee-employer combination, identified by PIK-SEIN(-SEINUNIT), with complete historical arrays for earnings and employment status, but where only *observed* SEINUNIT are used. It is not to be confused with the PHF_B of the QWI sequence (called JHF in the Snapshot, see Section 6.4.3.3), which is augmented with information from the U2W process for non-MN states.

The PHF (`ehf_&state._phf`) is organized by PIK-SEIN-SEINUNIT. Note that all states except MN have SEINUNIT='00000', so this reverts back to PIK-SEIN for all states except MN.

Record identifier PIK-SEIN-SEINUNIT

Sort order PIK-SEIN-SEINUNIT

Entity Job

Unique Entity Key PIK-SEIN-SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
Binary workhistory ...00111000... 1=employed	WORK	00264	80	A/N
Employment in QTIME=33	E33	00000	5	N
Employment in QTIME=34	E34	00005	5	N
Employment in QTIME=35	E35	00010	5	N
Employment in QTIME=36	E36	00015	5	N
Employment in QTIME=37	E37	00020	5	N
Employment in QTIME=38	E38	00025	5	N
Employment in QTIME=39	E39	00030	5	N
Employment in QTIME=40	E40	00035	5	N
Employment in QTIME=41	E41	00040	5	N
Employment in QTIME=42	E42	00045	5	N
Employment in QTIME=43	E43	00050	5	N
Employment in QTIME=44	E44	00055	5	N
Employment in QTIME=45	E45	00060	5	N
Employment in QTIME=46	E46	00065	5	N
Employment in QTIME=47	E47	00070	5	N
Employment in QTIME=48	E48	00075	5	N
Employment in QTIME=49	E49	00080	5	N
Employment in QTIME=50	E50	00085	5	N
Employment in QTIME=51	E51	00090	5	N
Employment in QTIME=52	E52	00095	5	N
Employment in QTIME=53	E53	00100	5	N
Employment in QTIME=54	E54	00105	5	N
Employment in QTIME=55	E55	00110	5	N
Employment in QTIME=56	E56	00115	5	N
Employment in QTIME=57	E57	00120	5	N
Employment in QTIME=58	E58	00125	5	N
Employment in QTIME=59	E59	00130	5	N
Employment in QTIME=60	E60	00135	5	N
Employment in QTIME=61	E61	00140	5	N
Employment in QTIME=62	E62	00145	5	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Employment in QTIME=63	E63	00150	5	N
Employment in QTIME=64	E64	00155	5	N
Employment in QTIME=65	E65	00160	5	N
Employment in QTIME=66	E66	00165	5	N
Employment in QTIME=67	E67	00170	5	N
Employment in QTIME=68	E68	00175	5	N
Employment in QTIME=69	E69	00180	5	N
Employment in QTIME=70	E70	00185	5	N
Employment in QTIME=71	E71	00190	5	N
Employment in QTIME=72	E72	00195	5	N
Employment in QTIME=73	E73	00200	5	N
Employment in QTIME=74	E74	00205	5	N
Employment in QTIME=75	E75	00210	5	N
Employment in QTIME=76	E76	00215	5	N
Employment in QTIME=77	E77	00220	5	N
Employment in QTIME=78	E78	00225	5	N
Employment in QTIME=79	E79	00230	5	N
Employment in QTIME=80	E80	00235	5	N
Protected Identification Key	PIK	00243	9	A/N
SEINUNIT imputed (never true, compatibility)	FLAG_SEINUNIT_IMPUTED	00240	3	N
State Employer Identification Number	SEIN	00252	12	A/N
State UI Reporting Unit Number	SEINUNIT	00344	5	A/N

6.4.3.3   JHF

The JHF (jhf_&state.) is created by combining the U2W with the EHF_PHF. This creates a file with multiple imputed establishment assignments for each job, where establishment assignments are missing (multi-units in states other than Minnesota). Internally, this file is called PHF_B, produced by the QWI process. For observed establishments, flag_seinunit_imputed=0 and only one SEINUNIT will be observed. Otherwise, ten implicates seinunit1-seinunit10 are kept on the file.

Record identifier PIK-SEIN

Sort order PIK-SEIN

Entity Job

Unique Entity Key PIK-SEIN

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date of Record Creation	LAST_UPDATE	00000	8	N
Date of Record Creation	SUCC_PRED_LAST_UPDATE	00008	8	N
Employment in QTIME=100	E100	00246	5	N
Employment in QTIME=101	E101	00251	5	N
Employment in QTIME=102	E102	00256	5	N
Employment in QTIME=103	E103	00261	5	N
Employment in QTIME=104	E104	00266	5	N
Employment in QTIME=105	E105	00271	5	N
Employment in QTIME=106	E106	00276	5	N
Employment in QTIME=107	E107	00281	5	N
Employment in QTIME=108	E108	00286	5	N
Employment in QTIME=109	E109	00291	5	N
Employment in QTIME=110	E110	00296	5	N
Employment in QTIME=111	E111	00301	5	N
Employment in QTIME=112	E112	00306	5	N
Employment in QTIME=61	E61	00051	5	N
Employment in QTIME=62	E62	00056	5	N
Employment in QTIME=63	E63	00061	5	N
Employment in QTIME=64	E64	00066	5	N
Employment in QTIME=65	E65	00071	5	N
Employment in QTIME=66	E66	00076	5	N
Employment in QTIME=67	E67	00081	5	N
Employment in QTIME=68	E68	00086	5	N
Employment in QTIME=69	E69	00091	5	N
Employment in QTIME=70	E70	00096	5	N
Employment in QTIME=71	E71	00101	5	N
Employment in QTIME=72	E72	00106	5	N
Employment in QTIME=73	E73	00111	5	N
Employment in QTIME=74	E74	00116	5	N
Employment in QTIME=75	E75	00121	5	N
Employment in QTIME=76	E76	00126	5	N
Employment in QTIME=77	E77	00131	5	N
Employment in QTIME=78	E78	00136	5	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Employment in QTIME=79	E79	00141	5	N
Employment in QTIME=80	E80	00146	5	N
Employment in QTIME=81	E81	00151	5	N
Employment in QTIME=82	E82	00156	5	N
Employment in QTIME=83	E83	00161	5	N
Employment in QTIME=84	E84	00166	5	N
Employment in QTIME=85	E85	00171	5	N
Employment in QTIME=86	E86	00176	5	N
Employment in QTIME=87	E87	00181	5	N
Employment in QTIME=88	E88	00186	5	N
Employment in QTIME=89	E89	00191	5	N
Employment in QTIME=90	E90	00196	5	N
Employment in QTIME=91	E91	00201	5	N
Employment in QTIME=92	E92	00206	5	N
Employment in QTIME=93	E93	00211	5	N
Employment in QTIME=94	E94	00216	5	N
Employment in QTIME=95	E95	00221	5	N
Employment in QTIME=96	E96	00226	5	N
Employment in QTIME=97	E97	00231	5	N
Employment in QTIME=98	E98	00236	5	N
Employment in QTIME=99	E99	00241	5	N
Employment indicators for first six quarters at successor job spell	SUCC_ACC_FIRST6	00414	6	A/N
Employment indicators for last six quarters at predecessor job spell	PRED_SEP_LAST6	00408	6	A/N
First quarter of employment at this job spell; spell has predecessor	SUPPRESS_ACCESSION	00024	8	N
Flag: SEINUNIT is imputed=1	FLAG_SEINUNIT_IMPUTED	00048	3	N
Last quarter of employment at this job spell; spell has successor	SUPPRESS_SEPARATION	00032	8	N
Protected Identification Key	PIK	00363	9	A/N
Quarter of accession at successor	SUCC_QA	00044	4	N
Quarter of separation from predecessor	PRED_QS	00040	4	N
SEIN of predecessor	SEIN_PRED	00384	12	A/N
SEIN of successor	SEIN_SUCC	00396	12	A/N
Selector based on random PIK	RANDOM_PIK_GROUP	00420	9	A/N
Spell count as per U2W	SPELL_U2W	00016	8	N
State (FIPS)	STATE	00361	2	A/N
State Employer Identification Number	SEIN	00372	12	A/N
State UI Reporting Unit Number (Impute 1)	SEINUNIT1	00311	5	A/N
State UI Reporting Unit Number (Impute 10)	SEINUNIT10	00356	5	A/N
State UI Reporting Unit Number (Impute 2)	SEINUNIT2	00316	5	A/N
State UI Reporting Unit Number (Impute 3)	SEINUNIT3	00321	5	A/N
State UI Reporting Unit Number (Impute 4)	SEINUNIT4	00326	5	A/N
State UI Reporting Unit Number (Impute 5)	SEINUNIT5	00331	5	A/N
State UI Reporting Unit Number (Impute 6)	SEINUNIT6	00336	5	A/N
State UI Reporting Unit Number (Impute 7)	SEINUNIT7	00341	5	A/N
State UI Reporting Unit Number (Impute 8)	SEINUNIT8	00346	5	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
State UI Reporting Unit Number (Impute 9)	SEINUNIT9	00351	5	A/N

6.4.3.4 UNIQPIK file

The UNIQPIK file is an input to the ICF. It also contains some diagnostic information, such as the number of records overall per PIK. It used to be produced by the (legacy) UIPIK sequence (called `ssnall` there).

The UNIQPIK file (`ehf_&state._uniqpik`) is organized by PIK.

Record identifier PIK

Sort order PIK

Entity Person

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
	SSNFLAG			
Illegal SSN Range Flag		00000	1	A/N
	PIK			
Protected Identification Key		00001	9	A/N
	CUT			
cut=substr(pik,1,2)		00010	9	A/N

6.4.3.5 SEIN_EMPLOYMENT

The SEIN_EMPLOYMENT is a SEIN-level measure of employment based on UI data.

The SEIN_EMPLOYMENT file (`ehf_&state._sein_employment`) is organized by SEIN-YEAR. No SEINUNIT version exists.

Record identifier SEIN-YEAR

Sort order SEIN-YEAR

Entity Firm

Unique Entity Key SEIN

Field name	Data dictionary reference name	Starting position	Field size	Data type
Beginning of quarter employment	B	00016	8	N
Beginning of quarter employment	E	00008	8	N
Flow employment	M	00000	8	N
Quarter	QUARTER	00035	3	N
State Employer Identification Number	SEIN	00038	12	A/N
Total earnings during the quarter	W1	00024	8	N
Year	YEAR	00032	3	N
Year-Quarter YYYY:Q	YR_QTR	00050	6	A/N

6.4.4 ES202-based Output Files

6.4.4.1 UHF

The UHF (Unit History File) used to be produced by the SPF (prior to version 3.1.12). It contains a full history of activity for each SEIN-SEINUNIT (wide file). It is still used as an input to the SPF. It replaces seinunit_history_es.sas7bdat.

The UHF file (ehf_&state._uhf) is organized by SEIN-SEINUNIT.

Record identifier SEIN-SEINUNIT

Sort order SEIN-SEINUNIT

Entity Establishment

Unique Entity Key SEIN-SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
...1... if part of multi-establishment, ...2... if master unit	MU_CODE	02034	80	A/N
=...1... if positive employment in quarter i	ACTIVE_EMPLOY_ES	01954	80	A/N
Ever had positive employment	ACTIVE_EVER_ES	01920	8	N
First QTIME with positive employment	ACTIVE_BEG_QTR_ES	01928	3	N
Last QTIME with positive employment	ACTIVE_END_QTR_ES	01931	3	N
Maximum monthly employment in QTIME=1	EMP_ES1	00640	8	N
Maximum monthly employment in QTIME=10	EMP_ES10	00712	8	N
Maximum monthly employment in QTIME=11	EMP_ES11	00720	8	N
Maximum monthly employment in QTIME=12	EMP_ES12	00728	8	N
Maximum monthly employment in QTIME=13	EMP_ES13	00736	8	N
Maximum monthly employment in QTIME=14	EMP_ES14	00744	8	N
Maximum monthly employment in QTIME=15	EMP_ES15	00752	8	N
Maximum monthly employment in QTIME=16	EMP_ES16	00760	8	N
Maximum monthly employment in QTIME=17	EMP_ES17	00768	8	N
Maximum monthly employment in QTIME=18	EMP_ES18	00776	8	N
Maximum monthly employment in QTIME=19	EMP_ES19	00784	8	N
Maximum monthly employment in QTIME=2	EMP_ES2	00648	8	N
Maximum monthly employment in QTIME=20	EMP_ES20	00792	8	N
Maximum monthly employment in QTIME=21	EMP_ES21	00800	8	N
Maximum monthly employment in QTIME=22	EMP_ES22	00808	8	N
Maximum monthly employment in QTIME=23	EMP_ES23	00816	8	N
Maximum monthly employment in QTIME=24	EMP_ES24	00824	8	N
Maximum monthly employment in QTIME=25	EMP_ES25	00832	8	N
Maximum monthly employment in QTIME=26	EMP_ES26	00840	8	N
Maximum monthly employment in QTIME=27	EMP_ES27	00848	8	N
Maximum monthly employment in QTIME=28	EMP_ES28	00856	8	N
Maximum monthly employment in QTIME=29	EMP_ES29	00864	8	N
Maximum monthly employment in QTIME=3	EMP_ES3	00656	8	N
Maximum monthly employment in QTIME=30	EMP_ES30	00872	8	N
Maximum monthly employment in QTIME=31	EMP_ES31	00880	8	N
Maximum monthly employment in QTIME=32	EMP_ES32	00888	8	N
Maximum monthly employment in QTIME=33	EMP_ES33	00896	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Maximum monthly employment in QTIME=34	EMP_ES34	00904	8	N
Maximum monthly employment in QTIME=35	EMP_ES35	00912	8	N
Maximum monthly employment in QTIME=36	EMP_ES36	00920	8	N
Maximum monthly employment in QTIME=37	EMP_ES37	00928	8	N
Maximum monthly employment in QTIME=38	EMP_ES38	00936	8	N
Maximum monthly employment in QTIME=39	EMP_ES39	00944	8	N
Maximum monthly employment in QTIME=4	EMP_ES4	00664	8	N
Maximum monthly employment in QTIME=40	EMP_ES40	00952	8	N
Maximum monthly employment in QTIME=41	EMP_ES41	00960	8	N
Maximum monthly employment in QTIME=42	EMP_ES42	00968	8	N
Maximum monthly employment in QTIME=43	EMP_ES43	00976	8	N
Maximum monthly employment in QTIME=44	EMP_ES44	00984	8	N
Maximum monthly employment in QTIME=45	EMP_ES45	00992	8	N
Maximum monthly employment in QTIME=46	EMP_ES46	01000	8	N
Maximum monthly employment in QTIME=47	EMP_ES47	01008	8	N
Maximum monthly employment in QTIME=48	EMP_ES48	01016	8	N
Maximum monthly employment in QTIME=49	EMP_ES49	01024	8	N
Maximum monthly employment in QTIME=5	EMP_ES5	00672	8	N
Maximum monthly employment in QTIME=50	EMP_ES50	01032	8	N
Maximum monthly employment in QTIME=51	EMP_ES51	01040	8	N
Maximum monthly employment in QTIME=52	EMP_ES52	01048	8	N
Maximum monthly employment in QTIME=53	EMP_ES53	01056	8	N
Maximum monthly employment in QTIME=54	EMP_ES54	01064	8	N
Maximum monthly employment in QTIME=55	EMP_ES55	01072	8	N
Maximum monthly employment in QTIME=56	EMP_ES56	01080	8	N
Maximum monthly employment in QTIME=57	EMP_ES57	01088	8	N
Maximum monthly employment in QTIME=58	EMP_ES58	01096	8	N
Maximum monthly employment in QTIME=59	EMP_ES59	01104	8	N
Maximum monthly employment in QTIME=6	EMP_ES6	00680	8	N
Maximum monthly employment in QTIME=60	EMP_ES60	01112	8	N
Maximum monthly employment in QTIME=61	EMP_ES61	01120	8	N
Maximum monthly employment in QTIME=62	EMP_ES62	01128	8	N
Maximum monthly employment in QTIME=63	EMP_ES63	01136	8	N
Maximum monthly employment in QTIME=64	EMP_ES64	01144	8	N
Maximum monthly employment in QTIME=65	EMP_ES65	01152	8	N
Maximum monthly employment in QTIME=66	EMP_ES66	01160	8	N
Maximum monthly employment in QTIME=67	EMP_ES67	01168	8	N
Maximum monthly employment in QTIME=68	EMP_ES68	01176	8	N
Maximum monthly employment in QTIME=69	EMP_ES69	01184	8	N
Maximum monthly employment in QTIME=7	EMP_ES7	00688	8	N
Maximum monthly employment in QTIME=70	EMP_ES70	01192	8	N
Maximum monthly employment in QTIME=71	EMP_ES71	01200	8	N
Maximum monthly employment in QTIME=72	EMP_ES72	01208	8	N
Maximum monthly employment in QTIME=73	EMP_ES73	01216	8	N
Maximum monthly employment in QTIME=74	EMP_ES74	01224	8	N
Maximum monthly employment in QTIME=75	EMP_ES75	01232	8	N
Maximum monthly employment in QTIME=76	EMP_ES76	01240	8	N
Maximum monthly employment in QTIME=77	EMP_ES77	01248	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Maximum monthly employment in QTIME=78	EMP_ES78	01256	8	N
Maximum monthly employment in QTIME=79	EMP_ES79	01264	8	N
Maximum monthly employment in QTIME=8	EMP_ES8	00696	8	N
Maximum monthly employment in QTIME=80	EMP_ES80	01272	8	N
Maximum monthly employment in QTIME=9	EMP_ES9	00704	8	N
Month 1 employment in QTIME=1	BPEMP_ES1	00000	8	N
Month 1 employment in QTIME=10	BPEMP_ES10	00072	8	N
Month 1 employment in QTIME=11	BPEMP_ES11	00080	8	N
Month 1 employment in QTIME=12	BPEMP_ES12	00088	8	N
Month 1 employment in QTIME=13	BPEMP_ES13	00096	8	N
Month 1 employment in QTIME=14	BPEMP_ES14	00104	8	N
Month 1 employment in QTIME=15	BPEMP_ES15	00112	8	N
Month 1 employment in QTIME=16	BPEMP_ES16	00120	8	N
Month 1 employment in QTIME=17	BPEMP_ES17	00128	8	N
Month 1 employment in QTIME=18	BPEMP_ES18	00136	8	N
Month 1 employment in QTIME=19	BPEMP_ES19	00144	8	N
Month 1 employment in QTIME=2	BPEMP_ES2	00008	8	N
Month 1 employment in QTIME=20	BPEMP_ES20	00152	8	N
Month 1 employment in QTIME=21	BPEMP_ES21	00160	8	N
Month 1 employment in QTIME=22	BPEMP_ES22	00168	8	N
Month 1 employment in QTIME=23	BPEMP_ES23	00176	8	N
Month 1 employment in QTIME=24	BPEMP_ES24	00184	8	N
Month 1 employment in QTIME=25	BPEMP_ES25	00192	8	N
Month 1 employment in QTIME=26	BPEMP_ES26	00200	8	N
Month 1 employment in QTIME=27	BPEMP_ES27	00208	8	N
Month 1 employment in QTIME=28	BPEMP_ES28	00216	8	N
Month 1 employment in QTIME=29	BPEMP_ES29	00224	8	N
Month 1 employment in QTIME=3	BPEMP_ES3	00016	8	N
Month 1 employment in QTIME=30	BPEMP_ES30	00232	8	N
Month 1 employment in QTIME=31	BPEMP_ES31	00240	8	N
Month 1 employment in QTIME=32	BPEMP_ES32	00248	8	N
Month 1 employment in QTIME=33	BPEMP_ES33	00256	8	N
Month 1 employment in QTIME=34	BPEMP_ES34	00264	8	N
Month 1 employment in QTIME=35	BPEMP_ES35	00272	8	N
Month 1 employment in QTIME=36	BPEMP_ES36	00280	8	N
Month 1 employment in QTIME=37	BPEMP_ES37	00288	8	N
Month 1 employment in QTIME=38	BPEMP_ES38	00296	8	N
Month 1 employment in QTIME=39	BPEMP_ES39	00304	8	N
Month 1 employment in QTIME=4	BPEMP_ES4	00024	8	N
Month 1 employment in QTIME=40	BPEMP_ES40	00312	8	N
Month 1 employment in QTIME=41	BPEMP_ES41	00320	8	N
Month 1 employment in QTIME=42	BPEMP_ES42	00328	8	N
Month 1 employment in QTIME=43	BPEMP_ES43	00336	8	N
Month 1 employment in QTIME=44	BPEMP_ES44	00344	8	N
Month 1 employment in QTIME=45	BPEMP_ES45	00352	8	N
Month 1 employment in QTIME=46	BPEMP_ES46	00360	8	N
Month 1 employment in QTIME=47	BPEMP_ES47	00368	8	N
Month 1 employment in QTIME=48	BPEMP_ES48	00376	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Month 1 employment in QTIME=49	BPEMP_ES49	00384	8	N
Month 1 employment in QTIME=5	BPEMP_ES5	00032	8	N
Month 1 employment in QTIME=50	BPEMP_ES50	00392	8	N
Month 1 employment in QTIME=51	BPEMP_ES51	00400	8	N
Month 1 employment in QTIME=52	BPEMP_ES52	00408	8	N
Month 1 employment in QTIME=53	BPEMP_ES53	00416	8	N
Month 1 employment in QTIME=54	BPEMP_ES54	00424	8	N
Month 1 employment in QTIME=55	BPEMP_ES55	00432	8	N
Month 1 employment in QTIME=56	BPEMP_ES56	00440	8	N
Month 1 employment in QTIME=57	BPEMP_ES57	00448	8	N
Month 1 employment in QTIME=58	BPEMP_ES58	00456	8	N
Month 1 employment in QTIME=59	BPEMP_ES59	00464	8	N
Month 1 employment in QTIME=6	BPEMP_ES6	00040	8	N
Month 1 employment in QTIME=60	BPEMP_ES60	00472	8	N
Month 1 employment in QTIME=61	BPEMP_ES61	00480	8	N
Month 1 employment in QTIME=62	BPEMP_ES62	00488	8	N
Month 1 employment in QTIME=63	BPEMP_ES63	00496	8	N
Month 1 employment in QTIME=64	BPEMP_ES64	00504	8	N
Month 1 employment in QTIME=65	BPEMP_ES65	00512	8	N
Month 1 employment in QTIME=66	BPEMP_ES66	00520	8	N
Month 1 employment in QTIME=67	BPEMP_ES67	00528	8	N
Month 1 employment in QTIME=68	BPEMP_ES68	00536	8	N
Month 1 employment in QTIME=69	BPEMP_ES69	00544	8	N
Month 1 employment in QTIME=7	BPEMP_ES7	00048	8	N
Month 1 employment in QTIME=70	BPEMP_ES70	00552	8	N
Month 1 employment in QTIME=71	BPEMP_ES71	00560	8	N
Month 1 employment in QTIME=72	BPEMP_ES72	00568	8	N
Month 1 employment in QTIME=73	BPEMP_ES73	00576	8	N
Month 1 employment in QTIME=74	BPEMP_ES74	00584	8	N
Month 1 employment in QTIME=75	BPEMP_ES75	00592	8	N
Month 1 employment in QTIME=76	BPEMP_ES76	00600	8	N
Month 1 employment in QTIME=77	BPEMP_ES77	00608	8	N
Month 1 employment in QTIME=78	BPEMP_ES78	00616	8	N
Month 1 employment in QTIME=79	BPEMP_ES79	00624	8	N
Month 1 employment in QTIME=8	BPEMP_ES8	00056	8	N
Month 1 employment in QTIME=80	BPEMP_ES80	00632	8	N
Month 1 employment in QTIME=9	BPEMP_ES9	00064	8	N
Number of establishments in QTIME=1	NUMRUNS1	01280	8	N
Number of establishments in QTIME=10	NUMRUNS10	01352	8	N
Number of establishments in QTIME=11	NUMRUNS11	01360	8	N
Number of establishments in QTIME=12	NUMRUNS12	01368	8	N
Number of establishments in QTIME=13	NUMRUNS13	01376	8	N
Number of establishments in QTIME=14	NUMRUNS14	01384	8	N
Number of establishments in QTIME=15	NUMRUNS15	01392	8	N
Number of establishments in QTIME=16	NUMRUNS16	01400	8	N
Number of establishments in QTIME=17	NUMRUNS17	01408	8	N
Number of establishments in QTIME=18	NUMRUNS18	01416	8	N
Number of establishments in QTIME=19	NUMRUNS19	01424	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of establishments in QTIME=2	NUMRUNS2	01288	8	N
Number of establishments in QTIME=20	NUMRUNS20	01432	8	N
Number of establishments in QTIME=21	NUMRUNS21	01440	8	N
Number of establishments in QTIME=22	NUMRUNS22	01448	8	N
Number of establishments in QTIME=23	NUMRUNS23	01456	8	N
Number of establishments in QTIME=24	NUMRUNS24	01464	8	N
Number of establishments in QTIME=25	NUMRUNS25	01472	8	N
Number of establishments in QTIME=26	NUMRUNS26	01480	8	N
Number of establishments in QTIME=27	NUMRUNS27	01488	8	N
Number of establishments in QTIME=28	NUMRUNS28	01496	8	N
Number of establishments in QTIME=29	NUMRUNS29	01504	8	N
Number of establishments in QTIME=3	NUMRUNS3	01296	8	N
Number of establishments in QTIME=30	NUMRUNS30	01512	8	N
Number of establishments in QTIME=31	NUMRUNS31	01520	8	N
Number of establishments in QTIME=32	NUMRUNS32	01528	8	N
Number of establishments in QTIME=33	NUMRUNS33	01536	8	N
Number of establishments in QTIME=34	NUMRUNS34	01544	8	N
Number of establishments in QTIME=35	NUMRUNS35	01552	8	N
Number of establishments in QTIME=36	NUMRUNS36	01560	8	N
Number of establishments in QTIME=37	NUMRUNS37	01568	8	N
Number of establishments in QTIME=38	NUMRUNS38	01576	8	N
Number of establishments in QTIME=39	NUMRUNS39	01584	8	N
Number of establishments in QTIME=4	NUMRUNS4	01304	8	N
Number of establishments in QTIME=40	NUMRUNS40	01592	8	N
Number of establishments in QTIME=41	NUMRUNS41	01600	8	N
Number of establishments in QTIME=42	NUMRUNS42	01608	8	N
Number of establishments in QTIME=43	NUMRUNS43	01616	8	N
Number of establishments in QTIME=44	NUMRUNS44	01624	8	N
Number of establishments in QTIME=45	NUMRUNS45	01632	8	N
Number of establishments in QTIME=46	NUMRUNS46	01640	8	N
Number of establishments in QTIME=47	NUMRUNS47	01648	8	N
Number of establishments in QTIME=48	NUMRUNS48	01656	8	N
Number of establishments in QTIME=49	NUMRUNS49	01664	8	N
Number of establishments in QTIME=5	NUMRUNS5	01312	8	N
Number of establishments in QTIME=50	NUMRUNS50	01672	8	N
Number of establishments in QTIME=51	NUMRUNS51	01680	8	N
Number of establishments in QTIME=52	NUMRUNS52	01688	8	N
Number of establishments in QTIME=53	NUMRUNS53	01696	8	N
Number of establishments in QTIME=54	NUMRUNS54	01704	8	N
Number of establishments in QTIME=55	NUMRUNS55	01712	8	N
Number of establishments in QTIME=56	NUMRUNS56	01720	8	N
Number of establishments in QTIME=57	NUMRUNS57	01728	8	N
Number of establishments in QTIME=58	NUMRUNS58	01736	8	N
Number of establishments in QTIME=59	NUMRUNS59	01744	8	N
Number of establishments in QTIME=6	NUMRUNS6	01320	8	N
Number of establishments in QTIME=60	NUMRUNS60	01752	8	N
Number of establishments in QTIME=61	NUMRUNS61	01760	8	N
Number of establishments in QTIME=62	NUMRUNS62	01768	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Number of establishments in QTIME=63	NUMRUNS63	01776	8	N
Number of establishments in QTIME=64	NUMRUNS64	01784	8	N
Number of establishments in QTIME=65	NUMRUNS65	01792	8	N
Number of establishments in QTIME=66	NUMRUNS66	01800	8	N
Number of establishments in QTIME=67	NUMRUNS67	01808	8	N
Number of establishments in QTIME=68	NUMRUNS68	01816	8	N
Number of establishments in QTIME=69	NUMRUNS69	01824	8	N
Number of establishments in QTIME=7	NUMRUNS7	01328	8	N
Number of establishments in QTIME=70	NUMRUNS70	01832	8	N
Number of establishments in QTIME=71	NUMRUNS71	01840	8	N
Number of establishments in QTIME=72	NUMRUNS72	01848	8	N
Number of establishments in QTIME=73	NUMRUNS73	01856	8	N
Number of establishments in QTIME=74	NUMRUNS74	01864	8	N
Number of establishments in QTIME=75	NUMRUNS75	01872	8	N
Number of establishments in QTIME=76	NUMRUNS76	01880	8	N
Number of establishments in QTIME=77	NUMRUNS77	01888	8	N
Number of establishments in QTIME=78	NUMRUNS78	01896	8	N
Number of establishments in QTIME=79	NUMRUNS79	01904	8	N
Number of establishments in QTIME=8	NUMRUNS8	01336	8	N
Number of establishments in QTIME=80	NUMRUNS80	01912	8	N
Number of establishments in QTIME=9	NUMRUNS9	01344	8	N
Number of quarters with positive employment	ACTIVE_QTRS_ES	01934	3	N
State Employer ID Number	SEIN	01937	12	A/N
State UI Reporting Unit Number	SEINUNIT	01949	5	A/N

6.4.4.2 SHF

The SHF (SEIN History File) used to be produced by the SPF (prior to version 3.1.12) as an internal file only. It contains a full history of activity for each SEIN (wide file). It is still used as an input to the SPF. It replaces sein_history_es.sas7bdat.

The SHF file (`ehf_&state._shf`) is organized by SEIN.

Record identifier SEIN

Sort order SEIN

Entity Firm

Unique Entity Key SEIN

Field name	Data dictionary reference name	Starting position	Field size	Data type
in QTIME=1	ESTABS_ES1	01280	8	N
in QTIME=10	ESTABS_ES10	01352	8	N
in QTIME=11	ESTABS_ES11	01360	8	N
in QTIME=12	ESTABS_ES12	01368	8	N
in QTIME=13	ESTABS_ES13	01376	8	N
in QTIME=14	ESTABS_ES14	01384	8	N
in QTIME=15	ESTABS_ES15	01392	8	N
in QTIME=16	ESTABS_ES16	01400	8	N
in QTIME=17	ESTABS_ES17	01408	8	N
in QTIME=18	ESTABS_ES18	01416	8	N
in QTIME=19	ESTABS_ES19	01424	8	N
in QTIME=2	ESTABS_ES2	01288	8	N
in QTIME=20	ESTABS_ES20	01432	8	N
in QTIME=21	ESTABS_ES21	01440	8	N
in QTIME=22	ESTABS_ES22	01448	8	N
in QTIME=23	ESTABS_ES23	01456	8	N
in QTIME=24	ESTABS_ES24	01464	8	N
in QTIME=25	ESTABS_ES25	01472	8	N
in QTIME=26	ESTABS_ES26	01480	8	N
in QTIME=27	ESTABS_ES27	01488	8	N
in QTIME=28	ESTABS_ES28	01496	8	N
in QTIME=29	ESTABS_ES29	01504	8	N
in QTIME=3	ESTABS_ES3	01296	8	N
in QTIME=30	ESTABS_ES30	01512	8	N
in QTIME=31	ESTABS_ES31	01520	8	N
in QTIME=32	ESTABS_ES32	01528	8	N
in QTIME=33	ESTABS_ES33	01536	8	N
in QTIME=34	ESTABS_ES34	01544	8	N
in QTIME=35	ESTABS_ES35	01552	8	N
in QTIME=36	ESTABS_ES36	01560	8	N
in QTIME=37	ESTABS_ES37	01568	8	N
in QTIME=38	ESTABS_ES38	01576	8	N
in QTIME=39	ESTABS_ES39	01584	8	N
in QTIME=4	ESTABS_ES4	01304	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
in QTIME=40	ESTABS_ES40	01592	8	N
in QTIME=41	ESTABS_ES41	01600	8	N
in QTIME=42	ESTABS_ES42	01608	8	N
in QTIME=43	ESTABS_ES43	01616	8	N
in QTIME=44	ESTABS_ES44	01624	8	N
in QTIME=45	ESTABS_ES45	01632	8	N
in QTIME=46	ESTABS_ES46	01640	8	N
in QTIME=47	ESTABS_ES47	01648	8	N
in QTIME=48	ESTABS_ES48	01656	8	N
in QTIME=49	ESTABS_ES49	01664	8	N
in QTIME=5	ESTABS_ES5	01312	8	N
in QTIME=50	ESTABS_ES50	01672	8	N
in QTIME=51	ESTABS_ES51	01680	8	N
in QTIME=52	ESTABS_ES52	01688	8	N
in QTIME=53	ESTABS_ES53	01696	8	N
in QTIME=54	ESTABS_ES54	01704	8	N
in QTIME=55	ESTABS_ES55	01712	8	N
in QTIME=56	ESTABS_ES56	01720	8	N
in QTIME=57	ESTABS_ES57	01728	8	N
in QTIME=58	ESTABS_ES58	01736	8	N
in QTIME=59	ESTABS_ES59	01744	8	N
in QTIME=6	ESTABS_ES6	01320	8	N
in QTIME=60	ESTABS_ES60	01752	8	N
in QTIME=61	ESTABS_ES61	01760	8	N
in QTIME=62	ESTABS_ES62	01768	8	N
in QTIME=63	ESTABS_ES63	01776	8	N
in QTIME=64	ESTABS_ES64	01784	8	N
in QTIME=65	ESTABS_ES65	01792	8	N
in QTIME=66	ESTABS_ES66	01800	8	N
in QTIME=67	ESTABS_ES67	01808	8	N
in QTIME=68	ESTABS_ES68	01816	8	N
in QTIME=69	ESTABS_ES69	01824	8	N
in QTIME=7	ESTABS_ES7	01328	8	N
in QTIME=70	ESTABS_ES70	01832	8	N
in QTIME=71	ESTABS_ES71	01840	8	N
in QTIME=72	ESTABS_ES72	01848	8	N
in QTIME=73	ESTABS_ES73	01856	8	N
in QTIME=74	ESTABS_ES74	01864	8	N
in QTIME=75	ESTABS_ES75	01872	8	N
in QTIME=76	ESTABS_ES76	01880	8	N
in QTIME=77	ESTABS_ES77	01888	8	N
in QTIME=78	ESTABS_ES78	01896	8	N
in QTIME=79	ESTABS_ES79	01904	8	N
in QTIME=8	ESTABS_ES8	01336	8	N
in QTIME=80	ESTABS_ES80	01912	8	N
in QTIME=9	ESTABS_ES9	01344	8	N
=...1... if positive employment in quarter i	ACTIVE_EMPLOY_ES	01957	80	A/N
Ever had positive employment	ACTIVE_EVER_ES	01920	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
First QTIME with positive employment	ACTIVE_BEG_QTR_ES	01936	3	N
Last QTIME with positive employment	ACTIVE_END_QTR_ES	01939	3	N
Maximum monthly employment in QTIME=1	EMP_ES1	00640	8	N
Maximum monthly employment in QTIME=10	EMP_ES10	00712	8	N
Maximum monthly employment in QTIME=11	EMP_ES11	00720	8	N
Maximum monthly employment in QTIME=12	EMP_ES12	00728	8	N
Maximum monthly employment in QTIME=13	EMP_ES13	00736	8	N
Maximum monthly employment in QTIME=14	EMP_ES14	00744	8	N
Maximum monthly employment in QTIME=15	EMP_ES15	00752	8	N
Maximum monthly employment in QTIME=16	EMP_ES16	00760	8	N
Maximum monthly employment in QTIME=17	EMP_ES17	00768	8	N
Maximum monthly employment in QTIME=18	EMP_ES18	00776	8	N
Maximum monthly employment in QTIME=19	EMP_ES19	00784	8	N
Maximum monthly employment in QTIME=2	EMP_ES2	00648	8	N
Maximum monthly employment in QTIME=20	EMP_ES20	00792	8	N
Maximum monthly employment in QTIME=21	EMP_ES21	00800	8	N
Maximum monthly employment in QTIME=22	EMP_ES22	00808	8	N
Maximum monthly employment in QTIME=23	EMP_ES23	00816	8	N
Maximum monthly employment in QTIME=24	EMP_ES24	00824	8	N
Maximum monthly employment in QTIME=25	EMP_ES25	00832	8	N
Maximum monthly employment in QTIME=26	EMP_ES26	00840	8	N
Maximum monthly employment in QTIME=27	EMP_ES27	00848	8	N
Maximum monthly employment in QTIME=28	EMP_ES28	00856	8	N
Maximum monthly employment in QTIME=29	EMP_ES29	00864	8	N
Maximum monthly employment in QTIME=3	EMP_ES3	00656	8	N
Maximum monthly employment in QTIME=30	EMP_ES30	00872	8	N
Maximum monthly employment in QTIME=31	EMP_ES31	00880	8	N
Maximum monthly employment in QTIME=32	EMP_ES32	00888	8	N
Maximum monthly employment in QTIME=33	EMP_ES33	00896	8	N
Maximum monthly employment in QTIME=34	EMP_ES34	00904	8	N
Maximum monthly employment in QTIME=35	EMP_ES35	00912	8	N
Maximum monthly employment in QTIME=36	EMP_ES36	00920	8	N
Maximum monthly employment in QTIME=37	EMP_ES37	00928	8	N
Maximum monthly employment in QTIME=38	EMP_ES38	00936	8	N
Maximum monthly employment in QTIME=39	EMP_ES39	00944	8	N
Maximum monthly employment in QTIME=4	EMP_ES4	00664	8	N
Maximum monthly employment in QTIME=40	EMP_ES40	00952	8	N
Maximum monthly employment in QTIME=41	EMP_ES41	00960	8	N
Maximum monthly employment in QTIME=42	EMP_ES42	00968	8	N
Maximum monthly employment in QTIME=43	EMP_ES43	00976	8	N
Maximum monthly employment in QTIME=44	EMP_ES44	00984	8	N
Maximum monthly employment in QTIME=45	EMP_ES45	00992	8	N
Maximum monthly employment in QTIME=46	EMP_ES46	01000	8	N
Maximum monthly employment in QTIME=47	EMP_ES47	01008	8	N
Maximum monthly employment in QTIME=48	EMP_ES48	01016	8	N
Maximum monthly employment in QTIME=49	EMP_ES49	01024	8	N
Maximum monthly employment in QTIME=5	EMP_ES5	00672	8	N
Maximum monthly employment in QTIME=50	EMP_ES50	01032	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Maximum monthly employment in QTIME=51	EMP_ES51	01040	8	N
Maximum monthly employment in QTIME=52	EMP_ES52	01048	8	N
Maximum monthly employment in QTIME=53	EMP_ES53	01056	8	N
Maximum monthly employment in QTIME=54	EMP_ES54	01064	8	N
Maximum monthly employment in QTIME=55	EMP_ES55	01072	8	N
Maximum monthly employment in QTIME=56	EMP_ES56	01080	8	N
Maximum monthly employment in QTIME=57	EMP_ES57	01088	8	N
Maximum monthly employment in QTIME=58	EMP_ES58	01096	8	N
Maximum monthly employment in QTIME=59	EMP_ES59	01104	8	N
Maximum monthly employment in QTIME=6	EMP_ES6	00680	8	N
Maximum monthly employment in QTIME=60	EMP_ES60	01112	8	N
Maximum monthly employment in QTIME=61	EMP_ES61	01120	8	N
Maximum monthly employment in QTIME=62	EMP_ES62	01128	8	N
Maximum monthly employment in QTIME=63	EMP_ES63	01136	8	N
Maximum monthly employment in QTIME=64	EMP_ES64	01144	8	N
Maximum monthly employment in QTIME=65	EMP_ES65	01152	8	N
Maximum monthly employment in QTIME=66	EMP_ES66	01160	8	N
Maximum monthly employment in QTIME=67	EMP_ES67	01168	8	N
Maximum monthly employment in QTIME=68	EMP_ES68	01176	8	N
Maximum monthly employment in QTIME=69	EMP_ES69	01184	8	N
Maximum monthly employment in QTIME=7	EMP_ES7	00688	8	N
Maximum monthly employment in QTIME=70	EMP_ES70	01192	8	N
Maximum monthly employment in QTIME=71	EMP_ES71	01200	8	N
Maximum monthly employment in QTIME=72	EMP_ES72	01208	8	N
Maximum monthly employment in QTIME=73	EMP_ES73	01216	8	N
Maximum monthly employment in QTIME=74	EMP_ES74	01224	8	N
Maximum monthly employment in QTIME=75	EMP_ES75	01232	8	N
Maximum monthly employment in QTIME=76	EMP_ES76	01240	8	N
Maximum monthly employment in QTIME=77	EMP_ES77	01248	8	N
Maximum monthly employment in QTIME=78	EMP_ES78	01256	8	N
Maximum monthly employment in QTIME=79	EMP_ES79	01264	8	N
Maximum monthly employment in QTIME=8	EMP_ES8	00696	8	N
Maximum monthly employment in QTIME=80	EMP_ES80	01272	8	N
Maximum monthly employment in QTIME=9	EMP_ES9	00704	8	N
Month 1 employment in QTIME=1	BPEMP_ES1	00000	8	N
Month 1 employment in QTIME=10	BPEMP_ES10	00072	8	N
Month 1 employment in QTIME=11	BPEMP_ES11	00080	8	N
Month 1 employment in QTIME=12	BPEMP_ES12	00088	8	N
Month 1 employment in QTIME=13	BPEMP_ES13	00096	8	N
Month 1 employment in QTIME=14	BPEMP_ES14	00104	8	N
Month 1 employment in QTIME=15	BPEMP_ES15	00112	8	N
Month 1 employment in QTIME=16	BPEMP_ES16	00120	8	N
Month 1 employment in QTIME=17	BPEMP_ES17	00128	8	N
Month 1 employment in QTIME=18	BPEMP_ES18	00136	8	N
Month 1 employment in QTIME=19	BPEMP_ES19	00144	8	N
Month 1 employment in QTIME=2	BPEMP_ES2	00008	8	N
Month 1 employment in QTIME=20	BPEMP_ES20	00152	8	N
Month 1 employment in QTIME=21	BPEMP_ES21	00160	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Month 1 employment in QTIME=22	BPEMP_ES22	00168	8	N
Month 1 employment in QTIME=23	BPEMP_ES23	00176	8	N
Month 1 employment in QTIME=24	BPEMP_ES24	00184	8	N
Month 1 employment in QTIME=25	BPEMP_ES25	00192	8	N
Month 1 employment in QTIME=26	BPEMP_ES26	00200	8	N
Month 1 employment in QTIME=27	BPEMP_ES27	00208	8	N
Month 1 employment in QTIME=28	BPEMP_ES28	00216	8	N
Month 1 employment in QTIME=29	BPEMP_ES29	00224	8	N
Month 1 employment in QTIME=3	BPEMP_ES3	00016	8	N
Month 1 employment in QTIME=30	BPEMP_ES30	00232	8	N
Month 1 employment in QTIME=31	BPEMP_ES31	00240	8	N
Month 1 employment in QTIME=32	BPEMP_ES32	00248	8	N
Month 1 employment in QTIME=33	BPEMP_ES33	00256	8	N
Month 1 employment in QTIME=34	BPEMP_ES34	00264	8	N
Month 1 employment in QTIME=35	BPEMP_ES35	00272	8	N
Month 1 employment in QTIME=36	BPEMP_ES36	00280	8	N
Month 1 employment in QTIME=37	BPEMP_ES37	00288	8	N
Month 1 employment in QTIME=38	BPEMP_ES38	00296	8	N
Month 1 employment in QTIME=39	BPEMP_ES39	00304	8	N
Month 1 employment in QTIME=4	BPEMP_ES4	00024	8	N
Month 1 employment in QTIME=40	BPEMP_ES40	00312	8	N
Month 1 employment in QTIME=41	BPEMP_ES41	00320	8	N
Month 1 employment in QTIME=42	BPEMP_ES42	00328	8	N
Month 1 employment in QTIME=43	BPEMP_ES43	00336	8	N
Month 1 employment in QTIME=44	BPEMP_ES44	00344	8	N
Month 1 employment in QTIME=45	BPEMP_ES45	00352	8	N
Month 1 employment in QTIME=46	BPEMP_ES46	00360	8	N
Month 1 employment in QTIME=47	BPEMP_ES47	00368	8	N
Month 1 employment in QTIME=48	BPEMP_ES48	00376	8	N
Month 1 employment in QTIME=49	BPEMP_ES49	00384	8	N
Month 1 employment in QTIME=5	BPEMP_ES5	00032	8	N
Month 1 employment in QTIME=50	BPEMP_ES50	00392	8	N
Month 1 employment in QTIME=51	BPEMP_ES51	00400	8	N
Month 1 employment in QTIME=52	BPEMP_ES52	00408	8	N
Month 1 employment in QTIME=53	BPEMP_ES53	00416	8	N
Month 1 employment in QTIME=54	BPEMP_ES54	00424	8	N
Month 1 employment in QTIME=55	BPEMP_ES55	00432	8	N
Month 1 employment in QTIME=56	BPEMP_ES56	00440	8	N
Month 1 employment in QTIME=57	BPEMP_ES57	00448	8	N
Month 1 employment in QTIME=58	BPEMP_ES58	00456	8	N
Month 1 employment in QTIME=59	BPEMP_ES59	00464	8	N
Month 1 employment in QTIME=6	BPEMP_ES6	00040	8	N
Month 1 employment in QTIME=60	BPEMP_ES60	00472	8	N
Month 1 employment in QTIME=61	BPEMP_ES61	00480	8	N
Month 1 employment in QTIME=62	BPEMP_ES62	00488	8	N
Month 1 employment in QTIME=63	BPEMP_ES63	00496	8	N
Month 1 employment in QTIME=64	BPEMP_ES64	00504	8	N
Month 1 employment in QTIME=65	BPEMP_ES65	00512	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Month 1 employment in QTIME=66	BPEMP_ES66	00520	8	N
Month 1 employment in QTIME=67	BPEMP_ES67	00528	8	N
Month 1 employment in QTIME=68	BPEMP_ES68	00536	8	N
Month 1 employment in QTIME=69	BPEMP_ES69	00544	8	N
Month 1 employment in QTIME=7	BPEMP_ES7	00048	8	N
Month 1 employment in QTIME=70	BPEMP_ES70	00552	8	N
Month 1 employment in QTIME=71	BPEMP_ES71	00560	8	N
Month 1 employment in QTIME=72	BPEMP_ES72	00568	8	N
Month 1 employment in QTIME=73	BPEMP_ES73	00576	8	N
Month 1 employment in QTIME=74	BPEMP_ES74	00584	8	N
Month 1 employment in QTIME=75	BPEMP_ES75	00592	8	N
Month 1 employment in QTIME=76	BPEMP_ES76	00600	8	N
Month 1 employment in QTIME=77	BPEMP_ES77	00608	8	N
Month 1 employment in QTIME=78	BPEMP_ES78	00616	8	N
Month 1 employment in QTIME=79	BPEMP_ES79	00624	8	N
Month 1 employment in QTIME=8	BPEMP_ES8	00056	8	N
Month 1 employment in QTIME=80	BPEMP_ES80	00632	8	N
Month 1 employment in QTIME=9	BPEMP_ES9	00064	8	N
Number of quarters with positive employment	ACTIVE_QTRS_ES	01942	3	N
SEIN was ever had multiple units	EVER_MU	01928	8	N
State Employer ID Number	SEIN	01945	12	A/N

6.4.5   Summary information on datasets

Table 6.8: Number of observations for EHF

Group	Number of datafiles	Records (1000s)	Filesize (GB)
EHF	400	7,986,525	1787

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 6.9: List of data files for EHF, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Alaska (ak)					
ehf_ak	1990Q1	2012Q1	11,526	1	1:421726
ehf_ak_controltotals	1990Q1	2012Q1	< 1	< 1	1:421738
ehf_ak_phf	1990Q1	2012Q1	4,970	2	1:421730
ehf_ak_sein_employment	1990Q1	2012Q1	1,261	< 1	1:421728
ehf_ak_shf	2000Q1	2012Q1	46	< 1	1:421736
ehf_ak_uhf	2000Q1	2012Q1	53	< 1	1:421734
ehf_ak_uniqpik	1990Q1	2012Q1	1,387	< 1	1:421732
jhf_ak	2000Q1	2012Q1	5,063	2	1:424475
Alabama (al)					
ehf_al	2001Q1	2012Q1	35,865	2	1:421346
ehf_al_controltotals	1990Q1	2012Q1	< 1	< 1	1:421352
ehf_al_phf	2001Q1	2012Q1	15,145	4	1:421348
ehf_al_sein_employment	2001Q1	2012Q1	3,641	< 1	1:421347
ehf_al_shf	2001Q1	2012Q1	219	1	1:421351
ehf_al_uhf	2001Q1	2012Q1	301	1	1:421350
ehf_al_uniqpik	2001Q1	2012Q1	4,471	< 1	1:421349
jhf_al	2001Q1	2012Q1	16,024	5	1:423620
Arkansas (ar)					
ehf_ar	2002Q3	2012Q1	20,504	1	1:422646
ehf_ar_controltotals	1990Q1	2012Q1	< 1	< 1	1:422658
ehf_ar_phf	2002Q3	2012Q1	8,606	2	1:422650
ehf_ar_sein_employment	2002Q3	2012Q1	2,238	< 1	1:422648
ehf_ar_shf	2002Q3	2012Q1	143	< 1	1:422656
ehf_ar_uhf	2002Q3	2012Q1	173	< 1	1:422654
ehf_ar_uniqpik	2002Q3	2012Q1	2,860	< 1	1:422652
jhf_ar	2002Q3	2012Q1	8,933	3	1:423802
Arizona (az)					
ehf_az	1992Q1	2012Q1	82,850	5	1:421571
ehf_az_controltotals	1990Q1	2012Q1	< 1	< 1	1:421577
ehf_az_phf	1992Q1	2012Q1	36,989	17	1:421573

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_az_sein_employment	1992Q1	2012Q1	7,728	< 1	1:421572
ehf_az_shf	2004Q1	2012Q1	263	1	1:421576
ehf_az_uhf	2004Q1	2012Q1	297	1	1:421575
ehf_az_uniqpik	1992Q1	2012Q1	8,248	< 1	1:421574
jhf_az	2004Q1	2012Q1	37,364	10	1:421718
California (ca)					
ehf_ca	1991Q3	2012Q1	532,389	32	1:421031
ehf_ca_controltotals	1990Q1	2012Q1	< 1	< 1	1:421037
ehf_ca_phf	1991Q3	2012Q1	212,825	97	1:421033
ehf_ca_sein_employment	1991Q3	2012Q1	69,973	4	1:421032
ehf_ca_shf	1991Q1	2012Q1	3,307	9	1:421036
ehf_ca_uhf	1991Q1	2012Q1	3,834	11	1:421035
ehf_ca_uniqpik	1991Q3	2012Q1	44,616	1	1:421034
jhf_ca	1991Q3	2012Q1	225,457	114	1:433732
Colorado (co)					
ehf_co	1990Q1	2012Q1	79,179	5	1:421354
ehf_co_controltotals	1990Q1	2012Q1	< 1	< 1	1:421360
ehf_co_phf	1990Q1	2012Q1	36,222	17	1:421356
ehf_co_sein_employment	1990Q1	2012Q1	9,422	1	1:421355
ehf_co_shf	1990Q1	2012Q1	517	1	1:421359
ehf_co_uhf	1990Q1	2012Q1	616	2	1:421358
ehf_co_uniqpik	1990Q1	2012Q1	8,639	< 1	1:421357
jhf_co	1993Q2	2012Q1	37,954	18	1:445058
Connecticut (ct)					
ehf_ct	1996Q1	2012Q1	43,294	3	1:422097
ehf_ct_controltotals	1990Q1	2012Q1	< 1	< 1	1:422103
ehf_ct_phf	1996Q1	2012Q1	15,518	6	1:422099
ehf_ct_sein_employment	1996Q1	2012Q1	5,459	< 1	1:422098
ehf_ct_shf	1996Q1	2012Q1	262	1	1:422102
ehf_ct_uhf	1996Q1	2012Q1	305	1	1:422101
ehf_ct_uniqpik	1996Q1	2012Q1	4,430	< 1	1:422100
jhf_ct	1996Q1	2012Q1	16,175	7	1:425776
District of Columbia (dc)					
ehf_dc	2002Q2	2012Q1	7,920	< 1	1:423245
ehf_dc_controltotals	1990Q1	2012Q1	< 1	< 1	1:423251
ehf_dc_phf	2002Q2	2012Q1	3,028	1	1:423247
ehf_dc_sein_employment	2002Q2	2012Q1	854	< 1	1:423246
ehf_dc_shf	2000Q4	2012Q1	69	< 1	1:423250
ehf_dc_uhf	2000Q4	2012Q1	73	< 1	1:423249
ehf_dc_uniqpik	2002Q2	2012Q1	1,702	< 1	1:423248
jhf_dc	2005Q2	2012Q1	3,046	1	1:423965
Delaware (de)					
ehf_de	1998Q3	2012Q1	9,642	1	1:428948
ehf_de_controltotals	1990Q1	2012Q1	< 1	< 1	1:428960

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_de_phf	1998Q3	2012Q1	3,782	1	1:428952
ehf_de_sein_employment	1998Q3	2012Q1	1,165	< 1	1:428950
ehf_de_shf	1997Q1	2012Q1	60	< 1	1:428958
ehf_de_uhf	1997Q1	2012Q1	65	< 1	1:428956
ehf_de_uniqpik	1998Q3	2012Q1	1,258	< 1	1:428954
jhf_de	1998Q3	2012Q1	3,873	1	1:431621
Florida (fl)					
ehf_fl	1992Q4	2012Q1	250,803	15	1:423423
ehf_fl_controltotals	1990Q1	2012Q1	< 1	< 1	1:423429
ehf_fl_phf	1992Q4	2012Q1	111,528	48	1:423425
ehf_fl_sein_employment	1992Q4	2012Q1	27,125	2	1:423424
ehf_fl_shf	1989Q1	2012Q1	1,724	5	1:423428
ehf_fl_uhf	1989Q1	2012Q1	2,237	6	1:423427
ehf_fl_uniqpik	1992Q4	2012Q1	24,835	< 1	1:423426
jhf_fl	1992Q4	2012Q1	118,793	58	1:433660
Georgia (ga)					
ehf_ga	1994Q1	2012Q1	120,892	7	1:422609
ehf_ga_controltotals	1990Q1	2012Q1	< 1	< 1	1:422615
ehf_ga_phf	1994Q1	2012Q1	51,725	21	1:422611
ehf_ga_sein_employment	1994Q1	2012Q1	12,157	1	1:422610
ehf_ga_shf	1998Q1	2012Q1	551	1	1:422614
ehf_ga_uhf	1998Q1	2012Q1	690	2	1:422613
ehf_ga_uniqpik	1994Q1	2012Q1	12,591	< 1	1:422612
jhf_ga	1998Q1	2012Q1	54,170	21	1:424783
Hawaii (hi)					
ehf_hi	1995Q4	2012Q1	13,295	1	1:426018
ehf_hi_controltotals	1990Q1	2012Q1	< 1	< 1	1:426024
ehf_hi_phf	1995Q4	2012Q1	4,802	2	1:426020
ehf_hi_sein_employment	1995Q4	2012Q1	1,697	< 1	1:426019
ehf_hi_shf	1995Q4	2012Q1	81	< 1	1:426023
ehf_hi_uhf	1995Q4	2012Q1	102	< 1	1:426022
ehf_hi_uniqpik	1995Q4	2012Q1	1,509	< 1	1:426021
jhf_hi	1995Q4	2012Q1	5,010	2	1:430609
Iowa (ia)					
ehf_ia	1998Q4	2012Q1	33,553	2	1:423591
ehf_ia_controltotals	1990Q1	2012Q1	< 1	< 1	1:423597
ehf_ia_phf	1998Q4	2012Q1	12,207	4	1:423593
ehf_ia_sein_employment	1998Q4	2012Q1	3,529	< 1	1:423592
ehf_ia_shf	1990Q1	2012Q1	225	1	1:423596
ehf_ia_uhf	1990Q1	2012Q1	305	1	1:423595
ehf_ia_uniqpik	1998Q4	2012Q1	3,626	< 1	1:423594
jhf_ia	1998Q4	2012Q1	12,790	5	1:431464
Idaho (id)					
ehf_id	1990Q1	2012Q1	22,430	1	1:425947

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_id_controltotals	1990Q1	2012Q1	< 1	< 1	1:425953
ehf_id_phf	1990Q1	2012Q1	9,621	5	1:425949
ehf_id_sein_employment	1990Q1	2012Q1	3,070	< 1	1:425948
ehf_id_shf	1991Q1	2012Q1	150	< 1	1:425952
ehf_id_uhf	1991Q1	2012Q1	170	< 1	1:425951
ehf_id_uniqpik	1990Q1	2012Q1	2,335	< 1	1:425950
jhf_id	1991Q1	2012Q1	9,954	5	1:427910
Illinois (il)					
ehf_il	1990Q1	2012Q1	209,478	13	1:442716
ehf_il_controltotals	1990Q1	2012Q1	< 1	< 1	1:442722
ehf_il_phf	1990Q1	2012Q1	78,573	38	1:442718
ehf_il_sein_employment	1990Q1	2012Q1	21,620	1	1:442717
ehf_il_shf	1990Q1	2012Q1	980	3	1:442721
ehf_il_uhf	1990Q1	2012Q1	1,099	3	1:442720
ehf_il_uniqpik	1990Q1	2012Q1	17,374	< 1	1:442719
jhf_il	1990Q1	2012Q1	82,162	44	1:442845
Indiana (in)					
ehf_in	1990Q1	2012Q1	107,583	6	1:424442
ehf_in_controltotals	1990Q1	2012Q1	< 1	< 1	1:424448
ehf_in_phf	1990Q1	2012Q1	43,536	21	1:424444
ehf_in_sein_employment	1990Q1	2012Q1	9,763	1	1:424443
ehf_in_shf	1998Q1	2012Q1	322	1	1:424447
ehf_in_uhf	1998Q1	2012Q1	406	1	1:424446
ehf_in_uniqpik	1990Q1	2012Q1	8,569	< 1	1:424445
jhf_in	1998Q1	2012Q1	44,899	17	1:430756
Kansas (ks)					
ehf_ks	1990Q1	2012Q1	48,694	3	1:424466
ehf_ks_controltotals	1990Q1	2012Q1	< 1	< 1	1:424472
ehf_ks_phf	1990Q1	2012Q1	19,907	10	1:424468
ehf_ks_sein_employment	1990Q1	2012Q1	5,512	< 1	1:424467
ehf_ks_shf	1990Q1	2012Q1	241	1	1:424471
ehf_ks_uhf	1990Q1	2012Q1	294	1	1:424470
ehf_ks_uniqpik	1990Q1	2012Q1	4,993	< 1	1:424469
jhf_ks	1993Q1	2012Q1	20,531	10	1:432209
Kentucky (ky)					
ehf_ky	1996Q4	2012Q1	47,877	3	1:423562
ehf_ky_controltotals	1990Q1	2012Q1	< 1	< 1	1:423568
ehf_ky_phf	1996Q4	2012Q1	19,407	7	1:423564
ehf_ky_sein_employment	1996Q4	2012Q1	4,525	< 1	1:423563
ehf_ky_shf	2001Q1	2012Q1	180	< 1	1:423567
ehf_ky_uhf	2001Q1	2012Q1	226	1	1:423566
ehf_ky_uniqpik	1996Q4	2012Q1	5,047	< 1	1:423565
jhf_ky	2001Q1	2012Q1	19,964	6	1:425366
Louisiana (la)					

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_la	1990Q1	2012Q1	69,545	4	1:442781
ehf_la_controltotals	1990Q1	2012Q1	< 1	< 1	1:442787
ehf_la_phf	1990Q1	2012Q1	29,373	14	1:442783
ehf_la_sein_employment	1990Q1	2012Q1	7,093	< 1	1:442782
ehf_la_shf	1990Q1	2012Q1	319	1	1:442786
ehf_la_uhf	1990Q1	2012Q1	398	1	1:442785
ehf_la_uniqpik	1990Q1	2012Q1	6,267	< 1	1:442784
jhf_la	1995Q1	2012Q1	30,572	13	1:443688
Maryland (md)					
ehf_md	1985Q2	2012Q1	101,541	6	1:425939
ehf_md_controltotals	1990Q1	2012Q1	< 1	< 1	1:425945
ehf_md_phf	1985Q2	2012Q1	40,691	23	1:425941
ehf_md_sein_employment	1985Q2	2012Q1	11,206	1	1:425940
ehf_md_shf	1990Q1	2012Q1	486	1	1:425944
ehf_md_uhf	1990Q1	2012Q1	569	2	1:425943
ehf_md_uniqpik	1985Q2	2012Q1	9,298	< 1	1:425942
jhf_md	1990Q1	2012Q1	42,406	22	1:431173
Maine (me)					
ehf_me	1996Q1	2012Q1	16,118	1	1:431320
ehf_me_controltotals	1990Q1	2012Q1	< 1	< 1	1:431326
ehf_me_phf	1996Q1	2012Q1	5,970	2	1:431322
ehf_me_sein_employment	1996Q1	2012Q1	2,246	< 1	1:431321
ehf_me_shf	1996Q1	2012Q1	119	< 1	1:431325
ehf_me_uhf	1996Q1	2012Q1	138	< 1	1:431324
ehf_me_uniqpik	1996Q1	2012Q1	1,605	< 1	1:431323
jhf_me	1996Q2	2012Q1	6,206	3	1:434447
Michigan (mi)					
ehf_mi	1998Q1	2012Q1	99,847	6	1:431414
ehf_mi_controltotals	1990Q1	2012Q1	< 1	< 1	1:431420
ehf_mi_phf	1998Q1	2012Q1	37,717	13	1:431416
ehf_mi_sein_employment	1998Q1	2012Q1	9,919	1	1:431415
ehf_mi_shf	1998Q1	2012Q1	536	1	1:431419
ehf_mi_uhf	1998Q1	2012Q1	620	2	1:431418
ehf_mi_uniqpik	1998Q1	2012Q1	9,525	< 1	1:431417
jhf_mi	2000Q3	2012Q1	39,254	13	1:437263
Minnesota (mn)					
ehf_mn	1994Q3	2012Q1	77,421	5	1:421611
ehf_mn_controltotals	1990Q1	2012Q1	< 1	< 1	1:421618
ehf_mn_phf	1994Q3	2012Q1	29,751	12	1:421614
ehf_mn_sein_employment	1994Q3	2012Q1	7,711	< 1	1:421613
ehf_mn_shf	1994Q3	2012Q1	338	1	1:421617
ehf_mn_uhf	1994Q3	2012Q1	423	1	1:421616
ehf_mn_uniqpik	1994Q3	2012Q1	6,267	< 1	1:421615
jhf_mn	1994Q3	2012Q1	29,751	13	1:422931

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Missouri (mo)					
ehf_mo	1990Q1	2012Q1	98,769	6	1:424432
ehf_mo_controltotals	1990Q1	2012Q1	< 1	< 1	1:424438
ehf_mo_phf	1990Q1	2012Q1	39,945	19	1:424434
ehf_mo_sein_employment	1990Q1	2012Q1	10,682	1	1:424433
ehf_mo_shf	1990Q1	2012Q1	540	1	1:424437
ehf_mo_uhf	1990Q1	2012Q1	672	2	1:424436
ehf_mo_uniqpik	1990Q1	2012Q1	8,580	< 1	1:424435
jhf_mo	1995Q1	2012Q1	41,592	18	1:429199
Mississippi (ms)					
ehf_ms	2003Q3	2012Q1	17,144	1	1:425955
ehf_ms_controltotals	1990Q1	2012Q1	< 1	< 1	1:425961
ehf_ms_phf	2003Q3	2012Q1	7,061	2	1:425957
ehf_ms_sein_employment	2003Q3	2012Q1	1,749	< 1	1:425956
ehf_ms_shf	2003Q3	2012Q1	109	< 1	1:425960
ehf_ms_uhf	2003Q3	2012Q1	134	< 1	1:425959
ehf_ms_uniqpik	2003Q3	2012Q1	2,623	< 1	1:425958
jhf_ms	2003Q3	2012Q1	7,296	2	1:427847
Montana (mt)					
ehf_mt	1993Q1	2012Q1	13,618	1	1:424414
ehf_mt_controltotals	1990Q1	2012Q1	< 1	< 1	1:424426
ehf_mt_phf	1993Q1	2012Q1	5,643	2	1:424418
ehf_mt_sein_employment	1993Q1	2012Q1	2,255	< 1	1:424416
ehf_mt_shf	1993Q1	2012Q1	106	< 1	1:424424
ehf_mt_uhf	1993Q1	2012Q1	120	< 1	1:424422
ehf_mt_uniqpik	1993Q1	2012Q1	1,471	< 1	1:424420
jhf_mt	1993Q1	2012Q1	5,777	3	1:431741
North Carolina (nc)					
ehf_nc	1991Q1	2011Q4	137,666	8	1:444336
ehf_nc_controltotals	1990Q1	2011Q4	< 1	< 1	1:444342
ehf_nc_phf	1991Q1	2011Q4	61,405	28	1:444338
ehf_nc_sein_employment	1991Q1	2011Q4	13,487	1	1:444337
ehf_nc_shf	1990Q1	2011Q4	640	2	1:444341
ehf_nc_uhf	1990Q1	2011Q4	821	2	1:444340
ehf_nc_uniqpik	1991Q1	2011Q4	12,822	< 1	1:444339
jhf_nc	1992Q4	2011Q4	64,774	30	1:445206
North Dakota (nd)					
ehf_nd	1998Q1	2012Q1	8,228	< 1	1:424415
ehf_nd_controltotals	1990Q1	2012Q1	< 1	< 1	1:424427
ehf_nd_phf	1998Q1	2012Q1	3,249	1	1:424419
ehf_nd_sein_employment	1998Q1	2012Q1	1,058	< 1	1:424417
ehf_nd_shf	1998Q1	2012Q1	55	< 1	1:424425
ehf_nd_uhf	1998Q1	2012Q1	65	< 1	1:424423
ehf_nd_uniqpik	1998Q1	2012Q1	989	< 1	1:424421

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
jhf_nd	1998Q1	2012Q1	3,335	1	1:431758
Nebraska (ne)					
ehf_ne	1999Q1	2012Q1	19,873	1	1:425596
ehf_ne_controltotals	1990Q1	2012Q1	< 1	< 1	1:425602
ehf_ne_phf	1999Q1	2012Q1	7,713	2	1:425598
ehf_ne_sein_employment	1999Q1	2012Q1	2,335	< 1	1:425597
ehf_ne_shf	1999Q1	2012Q1	115	< 1	1:425601
ehf_ne_uhf	1999Q1	2012Q1	141	< 1	1:425600
ehf_ne_uniqpik	1999Q1	2012Q1	2,263	< 1	1:425599
jhf_ne	1999Q1	2012Q1	8,020	3	1:431785
New Hampshire (nh)					
ehf_nh	2003Q1	2012Q1	9,809	1	1:431119
ehf_nh_controltotals	1990Q1	2012Q1	< 1	< 1	1:431125
ehf_nh_phf	2003Q1	2012Q1	3,759	1	1:431121
ehf_nh_sein_employment	2003Q1	2012Q1	1,366	< 1	1:431120
ehf_nh_shf	2003Q1	2012Q1	84	< 1	1:431124
ehf_nh_uhf	2003Q1	2012Q1	102	< 1	1:431123
ehf_nh_uniqpik	2003Q1	2012Q1	1,470	< 1	1:431122
jhf_nh	2003Q1	2012Q1	3,878	1	1:436496
New Jersey (nj)					
ehf_nj	1996Q1	2012Q1	100,548	6	1:421620
ehf_nj_controltotals	1990Q1	2012Q1	< 1	< 1	1:421626
ehf_nj_phf	1996Q1	2012Q1	37,403	14	1:421622
ehf_nj_sein_employment	1996Q1	2012Q1	12,076	1	1:421621
ehf_nj_shf	1995Q1	2012Q1	687	2	1:421625
ehf_nj_uhf	1995Q1	2012Q1	788	2	1:421624
ehf_nj_uniqpik	1996Q1	2012Q1	11,014	< 1	1:421623
jhf_nj	1996Q1	2012Q1	39,422	17	1:422387
New Mexico (nm)					
ehf_nm	1995Q3	2012Q1	23,113	1	1:431099
ehf_nm_controltotals	1990Q1	2012Q1	< 1	< 1	1:431108
ehf_nm_phf	1995Q3	2012Q1	9,981	4	1:431101
ehf_nm_sein_employment	1995Q3	2012Q1	2,523	< 1	1:431100
ehf_nm_shf	1990Q1	2012Q1	154	< 1	1:431106
ehf_nm_uhf	1990Q1	2012Q1	176	< 1	1:431104
ehf_nm_uniqpik	1995Q3	2012Q1	2,704	< 1	1:431102
jhf_nm	1995Q3	2012Q1	10,309	4	1:433074
Nevada (nv)					
ehf_nv	1998Q1	2012Q1	30,488	2	1:441348
ehf_nv_controltotals	1990Q1	2012Q1	< 1	< 1	1:441354
ehf_nv_phf	1998Q1	2012Q1	13,978	5	1:441350
ehf_nv_sein_employment	1998Q1	2012Q1	2,616	< 1	1:441349
ehf_nv_shf	1998Q1	2012Q1	166	< 1	1:441353
ehf_nv_uhf	1998Q1	2012Q1	204	1	1:441352

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_nv_uniqpik	1998Q1	2012Q1	3,940	< 1	1:441351
jhf_nv	1998Q1	2012Q1	14,488	6	1:443540
New York (ny)					
ehf_ny	1995Q1	2012Q1	246,444	15	1:442836
ehf_ny_controltotals	1990Q1	2012Q1	< 1	< 1	1:442842
ehf_ny_phf	1995Q1	2012Q1	92,068	36	1:442838
ehf_ny_sein_employment	1995Q1	2012Q1	29,056	2	1:442837
ehf_ny_shf	1990Q1	2012Q1	1,700	5	1:442841
ehf_ny_uhf	1990Q1	2012Q1	1,935	5	1:442840
ehf_ny_uniqpik	1995Q1	2012Q1	23,920	< 1	1:442839
jhf_ny	2000Q1	2012Q1	95,258	33	1:443301
Ohio (oh)					
ehf_oh	2000Q1	2012Q1	107,179	6	1:443281
ehf_oh_controltotals	1990Q1	2012Q1	< 1	< 1	1:443287
ehf_oh_phf	2000Q1	2012Q1	41,750	12	1:443283
ehf_oh_sein_employment	2000Q1	2012Q1	9,887	1	1:443282
ehf_oh_shf	2000Q1	2012Q1	482	1	1:443286
ehf_oh_uhf	2000Q1	2012Q1	637	2	1:443285
ehf_oh_uniqpik	2000Q1	2012Q1	10,993	< 1	1:443284
jhf_oh	2000Q1	2012Q1	43,958	15	1:443982
Oklahoma (ok)					
ehf_ok	2000Q1	2012Q1	31,969	2	1:421636
ehf_ok_controltotals	1990Q1	2012Q1	< 1	< 1	1:421642
ehf_ok_phf	2000Q1	2012Q1	14,091	4	1:421638
ehf_ok_sein_employment	2000Q1	2012Q1	3,496	< 1	1:421637
ehf_ok_shf	1999Q1	2012Q1	200	1	1:421641
ehf_ok_uhf	1999Q1	2012Q1	245	1	1:421640
ehf_ok_uniqpik	2000Q1	2012Q1	3,775	< 1	1:421639
jhf_ok	2000Q1	2012Q1	14,592	5	1:422240
Oregon (or)					
ehf_or	1991Q1	2012Q1	58,304	4	1:421628
ehf_or_controltotals	1990Q1	2012Q1	< 1	< 1	1:421634
ehf_or_phf	1991Q1	2012Q1	23,678	11	1:421630
ehf_or_sein_employment	1991Q1	2012Q1	7,525	< 1	1:421629
ehf_or_shf	1990Q1	2012Q1	380	1	1:421633
ehf_or_uhf	1990Q1	2012Q1	444	1	1:421632
ehf_or_uniqpik	1991Q1	2012Q1	5,709	< 1	1:421631
jhf_or	1991Q1	2012Q1	24,779	13	1:422671
Pennsylvania (pa)					
ehf_pa	1991Q1	2012Q1	186,270	11	1:423290
ehf_pa_controltotals	1990Q1	2012Q1	< 1	< 1	1:423296
ehf_pa_phf	1991Q1	2012Q1	67,918	32	1:423292
ehf_pa_sein_employment	1991Q1	2012Q1	20,072	1	1:423291
ehf_pa_shf	1991Q1	2012Q1	855	2	1:423295

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_pa_uhf	1991Q1	2012Q1	1,066	3	1:423294
ehf_pa_uniqpik	1991Q1	2012Q1	15,172	< 1	1:423293
jhf_pa	1997Q1	2012Q1	71,240	29	1:426064
Rhode Island (ri)					
ehf_ri	1995Q1	2012Q1	13,272	1	1:434183
ehf_ri_controltotals	1990Q1	2012Q1	< 1	< 1	1:434189
ehf_ri_phf	1995Q1	2012Q1	4,941	2	1:434185
ehf_ri_sein_employment	1995Q1	2012Q1	1,866	< 1	1:434184
ehf_ri_shf	1990Q1	2012Q1	115	< 1	1:434188
ehf_ri_uhf	1990Q1	2012Q1	124	< 1	1:434187
ehf_ri_uniqpik	1995Q1	2012Q1	1,524	< 1	1:434186
jhf_ri	1995Q1	2012Q1	5,103	2	1:436845
South Carolina (sc)					
ehf_sc	1998Q1	2012Q1	45,108	3	1:436992
ehf_sc_controltotals	1990Q1	2012Q1	< 1	< 1	1:436998
ehf_sc_phf	1998Q1	2012Q1	19,081	6	1:436994
ehf_sc_sein_employment	1998Q1	2012Q1	4,802	< 1	1:436993
ehf_sc_shf	1998Q1	2012Q1	263	1	1:436997
ehf_sc_uhf	1998Q1	2012Q1	303	1	1:436996
ehf_sc_uniqpik	1998Q1	2012Q1	5,492	< 1	1:436995
jhf_sc	1998Q1	2012Q1	19,890	8	1:439769
South Dakota (sd)					
ehf_sd	1994Q1	2012Q1	11,555	1	1:425626
ehf_sd_controltotals	1990Q1	2012Q1	< 1	< 1	1:425632
ehf_sd_phf	1994Q1	2012Q1	4,558	2	1:425628
ehf_sd_sein_employment	1994Q1	2012Q1	1,566	< 1	1:425627
ehf_sd_shf	1998Q1	2012Q1	59	< 1	1:425631
ehf_sd_uhf	1998Q1	2012Q1	72	< 1	1:425630
ehf_sd_uniqpik	1994Q1	2012Q1	1,185	< 1	1:425629
jhf_sd	1998Q1	2012Q1	4,672	2	1:432366
Tennessee (tn)					
ehf_tn	1998Q1	2012Q1	66,027	4	1:425931
ehf_tn_controltotals	1990Q1	2012Q1	< 1	< 1	1:425937
ehf_tn_phf	1998Q1	2012Q1	28,604	10	1:425933
ehf_tn_sein_employment	1998Q1	2012Q1	6,002	< 1	1:425932
ehf_tn_shf	1998Q1	2012Q1	342	1	1:425936
ehf_tn_uhf	1998Q1	2012Q1	408	1	1:425935
ehf_tn_uniqpik	1998Q1	2012Q1	7,462	< 1	1:425934
jhf_tn	1998Q1	2012Q1	29,624	11	1:428165
Texas (tx)					
ehf_tx	1995Q1	2012Q1	291,509	18	1:423572
ehf_tx_controltotals	1990Q1	2012Q1	< 1	< 1	1:423584
ehf_tx_phf	1995Q1	2012Q1	126,632	50	1:423576
ehf_tx_sein_employment	1995Q1	2012Q1	26,334	2	1:423574

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_tx_shf	1990Q1	2012Q1	1,505	4	1:423582
ehf_tx_uhf	1990Q1	2012Q1	1,992	6	1:423580
ehf_tx_uniqpik	1995Q1	2012Q1	26,559	< 1	1:423578
jhf_tx	1995Q1	2012Q1	136,074	61	1:427050
Utah (ut)					
ehf_ut	1999Q1	2012Q1	26,526	2	1:433065
ehf_ut_controltotals	1990Q1	2012Q1	< 1	< 1	1:433071
ehf_ut_phf	1999Q1	2012Q1	11,444	4	1:433067
ehf_ut_sein_employment	1999Q1	2012Q1	2,894	< 1	1:433066
ehf_ut_shf	1990Q1	2012Q1	241	1	1:433070
ehf_ut_uhf	1990Q1	2012Q1	310	1	1:433069
ehf_ut_uniqpik	1999Q1	2012Q1	3,032	< 1	1:433068
jhf_ut	1999Q3	2012Q1	12,076	4	1:434310
Virginia (va)					
ehf_va	1998Q1	2012Q1	81,415	5	1:437050
ehf_va_controltotals	1990Q1	2012Q1	< 1	< 1	1:437056
ehf_va_phf	1998Q1	2012Q1	32,207	11	1:437052
ehf_va_sein_employment	1998Q1	2012Q1	8,628	1	1:437051
ehf_va_shf	1995Q3	2012Q1	510	1	1:437055
ehf_va_uhf	1995Q3	2012Q1	640	2	1:437054
ehf_va_uniqpik	1998Q1	2012Q1	9,729	< 1	1:437053
jhf_va	1998Q1	2012Q1	34,156	13	1:439620
Vermont (vt)					
ehf_vt	2000Q1	2012Q1	6,383	< 1	1:425604
ehf_vt_controltotals	1990Q1	2012Q1	< 1	< 1	1:425610
ehf_vt_phf	2000Q1	2012Q1	2,377	1	1:425606
ehf_vt_sein_employment	2000Q1	2012Q1	978	< 1	1:425605
ehf_vt_shf	2000Q1	2012Q1	48	< 1	1:425609
ehf_vt_uhf	2000Q1	2012Q1	54	< 1	1:425608
ehf_vt_uniqpik	2000Q1	2012Q1	808	< 1	1:425607
jhf_vt	2000Q1	2012Q1	2,463	1	1:432369
Washington (wa)					
ehf_wa	1990Q1	2012Q1	103,579	6	1:424458
ehf_wa_controltotals	1990Q1	2012Q1	< 1	< 1	1:424464
ehf_wa_phf	1990Q1	2012Q1	43,182	21	1:424460
ehf_wa_sein_employment	1990Q1	2012Q1	13,688	1	1:424459
ehf_wa_shf	1990Q1	2012Q1	761	2	1:424463
ehf_wa_uhf	1990Q1	2012Q1	840	2	1:424462
ehf_wa_uniqpik	1990Q1	2012Q1	9,789	< 1	1:424461
jhf_wa	1990Q1	2012Q1	45,058	24	1:428141
Wisconsin (wi)					
ehf_wi	1990Q1	2012Q1	95,117	6	1:421273
ehf_wi_controltotals	1990Q1	2012Q1	< 1	< 1	1:421279
ehf_wi_phf	1990Q1	2012Q1	33,666	16	1:421275

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 6.9 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
ehf_wi_sein_employment	1990Q1	2012Q1	9,995	1	1:421274
ehf_wi_shf	1990Q1	2012Q1	402	1	1:421278
ehf_wi_uhf	1990Q1	2012Q1	486	1	1:421277
ehf_wi_uniqpik	1990Q1	2012Q1	6,892	< 1	1:421276
jhf_wi	1990Q1	2012Q1	36,013	19	1:423006
West Virginia (wv)					
ehf_wv	1997Q1	2012Q1	17,704	1	1:437467
ehf_wv_controltotals	1990Q1	2012Q1	< 1	< 1	1:437473
ehf_wv_phf	1997Q1	2012Q1	6,968	2	1:437469
ehf_wv_sein_employment	1997Q1	2012Q1	2,088	< 1	1:437468
ehf_wv_shf	1990Q1	2012Q1	87	< 1	1:437472
ehf_wv_uhf	1990Q1	2012Q1	125	< 1	1:437471
ehf_wv_uniqpik	1997Q1	2012Q1	2,030	< 1	1:437470
jhf_wv	1997Q1	2012Q1	7,301	3	1:441379
Wyoming (wy)					
ehf_wy	1992Q1	2012Q1	9,327	1	1:425612
ehf_wy_controltotals	1990Q1	2012Q1	< 1	< 1	1:425618
ehf_wy_phf	1992Q1	2012Q1	4,198	2	1:425614
ehf_wy_sein_employment	1992Q1	2012Q1	1,386	< 1	1:425613
ehf_wy_shf	2001Q1	2012Q1	49	< 1	1:425617
ehf_wy_uhf	2001Q1	2012Q1	57	< 1	1:425616
ehf_wy_uniqpik	1992Q1	2012Q1	1,304	< 1	1:425615
jhf_wy	2001Q1	2012Q1	4,254	1	1:432376

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

6.5 NOTES

Table 6.10: UI/EHF Summary of Information and Known Issues
with Data Coverage and Quality

State	Known Data Quality Issues (UI/EHF)	Recommendation to Researchers
CA	None	
CO	60-70% hole in UI data in 1993:3. 20% unresolved identifier mismatch on UI in [90:1-90:3]	Researchers should generally avoid use of pre-1994 EHF data in CO.
FL	(1) There appear to be changes being made in the firm identifiers on the ES202 and UI data in the mid-to-late 1990s. Specifically it looks as though some changes are made on the identifiers in the ES202 in 1996 and in 1997 the UI data is corrected in kind. In the ES202 data, 14% of firms die in 1995:4 and are born in 1996:1, indicating a shift in some firm identifiers. A similar change in magnitude occurs in the UI data between 1997:1 and 1997:4. Between these years, the rate of match between the UI and ES202 SEINs is somewhat poor (10% of UI SEINs do not appear on the ES202 between 1996:1 and 1997:3), although it is quite good both before and after. (2) The match between the ES202 and UI data is not good in 2002:4-2003:3, with 13-20% of UI SEINs not appearing in the ES-202 data.	While not a big enough problem to recommend avoiding use of these date ranges in FL, be aware that changes in firm identifiers in the mid-1990s will bias worker flow measures during this period.
IA	None	
ID	1990 UI data has firm identifier problems on approximately 40% of the records. Because of these problems, this year is not included in the EHF.	Researchers should generally avoid use of 1990 ID EHF data, which should not be too much of an issue as ES202 information is missing for this year in ID.
IL	Small hole in UI data in 1990:1 (10% missing). 1992:1 and 1993:1 are also missing UI wage records.	Note to researchers: These problems bias worker flows in those quarters, also full quarter employment in early years of IL data.
IN	None	
KS	Large holes in KS UI data at 1990:1 (¿50% missing) and 1992:4 (25% missing)	Researchers should generally avoid use of 1990 and 1992 KS EHF data; this problem will also bias full quarter employment and flows in 1993.
KY	UI identifier problem in 2000:3-2001:2 likely, due to 10%, 15% death rates in 2000:3, 2000:4, followed by 11%, 14% birth rates in 2001:1 and 2001:2. (Normal is 3-7% births/deaths in a particular quarter)	Note to researchers: These problems bias worker flows in those quarters, also full quarter employment during 2000-2001 KY data.
MD	None	
ME	None	
MN	None	
MO	1994:4 UI data is small (approximately 70% sample).	Researchers should generally avoid use of 1994 MO EHF data; this problem will also bias some full quarter employment and flows measures in 1995.

(cont)

Table 6.10 – Continued

State	Known Data Quality Issues (UI/EHF)	Recommendation to Researchers
MT		
NC	* ES202 show persistently lower employment than UI, by about 14%, except for 1991:1-1992:3 (around 0%) and 2002:1-2002:4 (5-8%). Warnings are generated when it goes above 15%. * Payroll is typically 6-8% higher on ES202 compared to UI except for 1991:1-1992:3, where it is 20-30% higher. There are also significant, but not as large deviations in 2002:1-2003:1. * Based on the BLS PU records, the ES202 data series looks fine: ES202 sums rarely go above 1% (Test 13-1 and 13-2) Conclusion: we are still missing wage records in the early periods, and some in later periods as well. The most recent wage records actually look coherent with the longest time series, but 2002 is a small problem.	Note to Researchers: Similar to problems in early years of IL, these issues bias worker flows in those quarters, also full quarter employment.
NJ	Small holes in NJ UI data at 1998:3 (5%) and 1999:1 (8-10%) and 2003:1 (10%)	Note to Researchers: Problem probably small enough to ignore for most research purposes.
NM	None	
OK	None	
OR	1994:1 is small, but not terribly so.	Note to Researchers: Problem probably small enough to ignore for most research purposes.
PA	UI wage records are 1% sample for 1996:4	Note to Researchers: Generally avoid use of 1996 PA annual earnings (particularly earnings changes between 1995-1996, 1996-1997, which will be biased), this problem will also bias some flows and full quarter employment measures in 1996 and 1997.
TX	None	
VA	1998:1 is small, and 1998:2 also looks on the small side.	Note to Researchers: Problems probably small enough to ignore for most research purposes.
WA	None	
WI	None	
WV	None	

Chapter 7.
ES-202 files (ES202)

The ES-202 program, also known as the Covered Employment and Wages program, is a joint BLS-state program that collects data on employment and wages from firms.

The BLS summarizes and publishes the data as the Quarterly Census of Employment and Wages (QCEW) at the industry level (to the 4-digit SIC from 1975 to 2000; to the 6-digit NAICS industry since 1990), by county and MSA.

The ES-202 files are inputs to the ECF, but are not separately available to external researchers.

Chapter 8.
Geo-coded Address List (GAL)

8.1 OVERVIEW

The Geocoded Address List (GAL) is a data set containing unique commercial and residential addresses in a state geocoded to the Census Block and latitude/longitude coordinates. It consists of the address list (GAL) and a crosswalk for each processed file-year. The GAL contains each unique address, identified by a GAL identifier called `galid`, its geocodes, a flag for each file-year in which it appears, data quality indicators, and data processing information, including the release date of the Geographic Reference File (GRF). The GAL Crosswalk contains the ID of each input entity and the ID of its address (`galid`).

This is the last version of the GAL in this format (internally version 3). LEHD introduced a new GAL version 4 in 2013, with expanded functionality and growing coverage.

8.2 DATA CITATION

> U.S. Census Bureau. 2014. *Employment History Files (EHF) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

8.2.1 Changes in this Snapshot

We are dropping the GAL crosswalks to AHS, BR, ACS-POW, because either the related files are not useful in the RDC (ACS-POW), or because the relevant crosswalks have not been updated in the LEHD production system for over a decade, and are thus of doubtful utility (AHS, BR). We note that this does not affect in any way the availability of the AHS, BR, or ACS in the RDC - this only affects the crosswalk created as part of GAL at LEHD to a particular version of those files.

8.3   DETAILED DESCRIPTION

8.3.1   Input Data

The input data consists of addresses, geocodes, and coordinates. As of early 2013, the source files providing addresses consisted of the following:

ACS-POW	*	American Community Survey Place of Work (2001 and later)
AHS	*	American Housing Survey (2002)
ES202		QCEW (all available years 1990 and later)
BR	*	Business Register (Standard Statistical Establishment List 1990 and later), through Business Register Bridge
MAF		Master Address File (the year following the year of the desired geographic vintage)

The linkages to data sources denoted by "*" are no longer made available to RDC users, as the linked files have not been maintained for several years at LEHD.

8.3.2   Geocodes

The source files providing geocodes and coordinates are the following:

GCP	the databases of Group1's Geographic Coding Plus software
MAF	Master Address File
GRF-C	Geographic Reference File, Codes (encompassed in the BMF)
WIB-C	Workforce Investment Board, Codes (encompassed in the BMF)
BMF	Block Map File

8.3.3   Update frequency

The internal use GAL is produced quarterly. The RDC version is produced occassionally, at the same time as the other LEHD-provided RDC files.

8.3.4   Processing description

All internal processing variables (parsed addresses in particular) are available on the RDC. All crosswalks to input files (for instance, the BR) are available as well. Note that a researcher needs to request the input files separately, and not all input files may be available in the RDC environment.

The Census-internal GAL is considered commingled data, *i.e.*, it contains information protected both under Title 13 and Title 26. Because projects requesting Title 26 data are handled differently from projects requesting only Title 13 information, the GAL is split. Before transferring the GAL to the RDC environment, all variables that refer to Title 26 data are split off, and stored in a separate file (`gal_ZZ_2010_t26flags.sas7bdat`). Furthermore, all records sourced exclusively from Title 26 data are removed from the main GAL dataset, and stored separately (`gal_ZZ_t26.sas7bdat`).

8.4   ADDITIONAL DETAILS

The following document was prepared several years ago by Marc Roemer, U.S. Census Bureau, describing the GAL (version 3). It provides a general overview of how the original GAL files are created. Note that the Census Bureau continually improves the processing, and the current GAL processing differs in some details.

The Geocoded Address List (GAL) is a data set containing unique commercial and residential addresses in a state geocoded to the Census Block and latitude/longitude coordinates. The file encompasses addresses from

the state ES202 data, the Business Register, the Census Bureau's Master Address File (MAF), the American Community Survey Place of Work file (ACS-POW), and others. Addresses from these source files go through Code1, Vality standardizer, Vality matching for unduplication, and several other steps in SAS. This document refers to one year's data from a source file as a file-year (for example, the 1995 ES202).

The job stream follows the steps below using the indicated software.

- Step 1: Create input (SAS).

- Step 2: Standardize and geocode addresses (Code1).

- Step 3: Parse and standardize address elements (Vality Standardize).

- Step 4: Match addresses, flag masters and duplicates (Vality Unduplicate).

- Step 5: Create preliminary crosswalk and unique address list with address identifier (SAS).

- Step 6: Set file-year flags, create GAL Crosswalks containing the input identifier and address identifier (SAS).

- Step 7: Retrieve and derive block codes and coordinates from the MAF (SAS).

- Step 8: Impute block within known tract (SAS).

- Step 9: Create GAL by adding higher-level geocodes by block (SAS).

- Step 10: Delete intermediate data files and create links.

The final output consists of the address list and a crosswalk for each processed file-year. The GAL contains each address, its geocodes, a flag for each file-year in which it appears, data quality indicators, and data processing information. The GAL Crosswalk contains the ID of each input entity and the ID of its address. The following section describes the GAL's content.

8.4.1   Important Variables

Unique identifier   The variable `galid` is the unique address identifier on the GAL, a 26-character string consisting of the letter 'A' in the first column followed by the 2-character state FIPS code and a zero-padded sequential number. The `galid` is created each time a GAL is created. There's no consistency in the `galid` between versions or vintages of the GAL.

Geographic vintage   The release date (year) of the GRF identifies the geographic vintage. In the GAL the vintage becomes the variable `a_vintage`.

Geographic codes   The variable `a_geocode` is $FIPS-state(2)\|FIPS-county(3)\|Census-tract(6)$, and it uniquely identifies the Census tract in the U.S. The tract is the lowest level of geography recommended for analysis. The Census block within the tract is `a_block`. The uncertainties in block-coding make block-level analysis questionable. However, geocoding to the block allows us to add all the higher-level geocodes to the addresses. The variable `a_block_src` generated in Steps 7 and 8 describes the source of the block-code.

Figure 8.1: GAL Processing

Flowchart for Adding WIBs to GAL/LEG Processing - 09/12/05

Source: Longitudinal Employer-Household Dynamics (LEHD) Program, Census Bureau
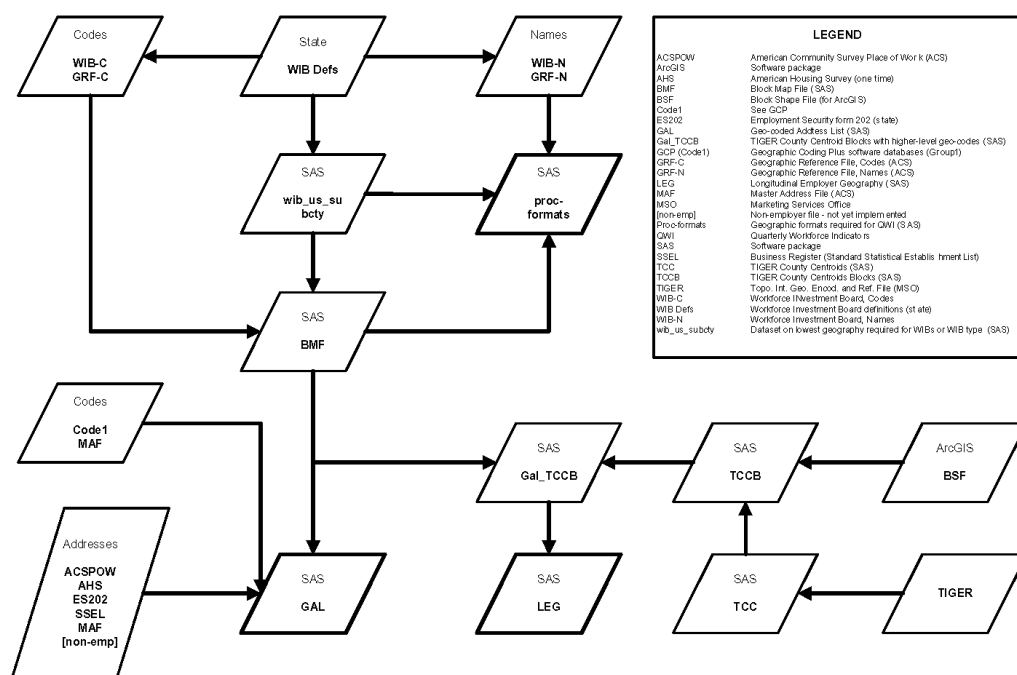
Value	Typical Percent	Meaning
C	12.20	Code1, or the address matches an address for which Code1 supplied the block code
M	81.86	The MAF - the address is a MAF address or matches a MAF address
E	0.00	The MAF, the street address is exactly the same as a MAF address in the same tract
W	0.03	The MAF, the street address is between 2 MAF addresses on the same block face
O	1.23	Imputed by the distribution of commercial addresses in the tract
S	1.17	Imputed by the distribution of residential addresses in the tract
I	0.01	Imputed by the distribution of mixed-use addresses in the tract
D	0.00	Imputed by the distribution of all addresses in the tract
missing	3.50	Block code is missing

In all states observed so far except California, no address required the 'D' method. That is, almost every tract where an address lacks a block code contains commercial, residential, and mixed-use addresses.

The Census Bureau splits blocks to accommodate changes in political boundaries. Most commonly, these are place boundaries (a place is a city, village, or similar municipality). The resulting block parts are identified by 2 suffixes, each taking a value from A to Z. The GAL assigns the block part directly from the MAF, or by adopting the one whose internal point is closest to the address by the straight-line distance. The variables `a_block_suf1` and `a_block_suf2` identify the block part, and `a_block_suf_src` generated in Step 9 describes the method used to assign it.

Value	Typical Percent	Meaning
A	1.50	Assigned by distance
M	4.18	The MAF - the address is a MAF address or matches a MAF address
missing	94.32	Not a split block

The GAL also provides the following basic geographic variables:

a_ssccc	FIPS-state(2)‖FIPS-county (3)
a_st	FIPS state (2)
a_cty	FIPS county within the state (3)
a_tract	Census tract within the county (6)

Higher-level geographic codes originate from the Block Map File (BMF) and attach to the GAL in Step 9. The BMF is an extract of the GRF-C (Geographic Reference File - Codes). All these geocodes are character variables. FIPS (Federal Information Processing Standard) codes are unique within the U.S.; Census codes are not.

a_fipsmcd	5-digit FIPS Minor Civil Division (a division of a county)
a_mcd	3-digit Census Minor Civil Division (a division of a county)
a_fipspl	5-digit FIPS Place
a_place	4-digit Census Place
a_msapmsa	Metropolitan-Statistical-Area(4)‖Primary-Metropolitan-Statistical-Area(4)
a_wib	6-digit Workforce Investment Board area

Geographic coordinates The coordinates of each address are in the variables `a_latitude` and `a_longitude`. These variables are numeric with 6 implied decimals (divide by 1,000,000 to convert them). The coordinates are not as accurate as 6 decimal places implies. An indication of their quality is in the variable `a_geoqual`, a numeric variable taking values from 1 to 9 and generated in Steps 7, 8, and 9:

Value	Typical Percent	Meaning
1	80.15	Rooftop or MAF (most accurate)
2	1.59	ZIP4 or block face, block face is certain
3	10.12	Block group is certain
4	4.65	Tract is certain
9	3.50	Coordinates are missing

The format 'agqual' provided by 'format_geo.sas' in '/programs/projects/auxiliary/Formats' contains the meanings of the `a_geoqual` values listed above.

Two other variables give information about the coordinates. The flag `a_latlong_src` indicates their source:

Value	Typical Percent	Meaning
B	14.77	Block (or block part) internal point
C	70.04	Code1
D	0.03	Derived
M	11.66	the MAF
missing	3.50	Coordinates are missing

Few addresses have `a_latlong_src` equal to 'D'. Deriving coordinates occurs only if they're still missing after Code1 processing and direct extraction from the MAF, but the tract is known. In this case, the flag `a_latlong_drv` generated in Step 7 describes the derivation method:

Value	Typical Percent	Meaning
F	0.00	Adopted from the only address on the block face
P	0.04	Extrapolated between 2 addresses on the block face
missing	99.96	Derivation not performed

In GAL Version 1, deriving coordinates and block codes by these methods was an important means of block-coding. It rarely operates now, since Code1 began providing block codes. Nevertheless, GAL Version 3 still exhausts all methods of assigning block-codes and coordinates before resorting to imputation.

File-year flags A set of flags generated in Step 6 indicates what file-years an address appears in. The names of the flags conform to the naming convention `[f][yyyy]` for the source file `[f]` and year `[yyyy]`, where `[f]` takes the following values:

Business Register	f = b
ES202	f = e
Master Address File	f = m
American Community Survey - Place of Work	f = p
American Housing Survey	f = h

For example, the flag variable `b1997` equals 1 if the address is on the 1997 Business Register; otherwise it equals 0. Note that if a [LEHD] state partner supplies 1991 ES202 data with no address information, `e1991` will be 0 for all addresses. Typically, the `e[yyyy]` flags equal 1 for between 3 and 6 percent of addresses, the `b[yyyy]` flags equal 1 for between 4 and 10 percent, and the `m[yyyy]` flag is 1 for between 80 and 90 percent. The `p[yyyy]` and `h[yyyy]` flags equal 1 for less than 1 percent of addresses because the ACS-POW and AHS data are sample surveys.

8.4.2   Other Variables

occupant_type   The variable `occupant_type`, recoded from the file-year flags in Step 8, indicates whether an address is commercial, residential, or mixed-use.

bigsrcid   The tracking ID `bigsrcid`, created in Step 1, uniquely identifies the entity that supplied the address. It consists of `[f]`, `[yyyy]`, the unique ID from the input file, zero-padding, and for some source files, a flag indicating which set of variables supplied the address. For addresses originating in the Business Register, another flag indicates the single-unit data set or the multi-unit data set. This tracking ID variable is useful for debugging.
   This variable is only available GAL_ZZ_2003_T26FLAGS.

srcmast   A diagnostic variable `srcmast` contains `[f][yyyy]`, indicating the file-year that supplied this address. Bear in mind that it's often arbitrary which observation becomes the master address for a set of duplicates in Step 1 and Step 4, so `bigsrcid` and `srcmast` don't indicate anything special about an address or an entity. They simply identify the origin of an address that became a master address in unduplication.
   This variable is only available GAL_ZZ_2003_T26FLAGS.

Code 1 variables   The names of Code1 variables contain the prefix `c1_`. They impart mostly diagnostic information from Code1 processing. They could be useful for development work or address research.
   For records sourced exclusively from the BR, these variables are available on GAL_ZZ_2003_T26. For records sourced exclusively from the ES-202, some variables were blanked on GAL_ZZ_2003 and are available on GAL_ZZ_2003_ES202ONLY to Census personnel only. Code1 diagnostic codes remain available to all researchers.

Vality variables   The parsed address elements from Step 3 sit in the variables named with the prefix `v_`. They could be useful for development work, particularly in improving the parsing routine.
   For records sourced exclusively from the BR, these variables are available on GAL_ZZ_2003_T26. For records sourced exclusively from the ES-202, some variables were blanked on GAL_ZZ_2003 and are available on GAL_ZZ_2003_ES202ONLY to Census personnel only. Vality diagnostic codes remain available to all researchers.

8.4.3   Accessing the GAL: the GAL Crosswalks

The GAL Crosswalks allow you to extract geographic and address information about any entity whose address went into the GAL. Each crosswalk contains the identifiers of the entity, its galid, and sometimes flags. To attach geocodes, coordinates, or address information to an entity, merge the GAL Crosswalk to the GAL by galid, outputting only observations existing on the GAL Crosswalk. Then merge the resulting file to the entities of interest using the entity identifiers. An entity whose address wasn't processed (because it's out of state or lacks address information) will have blank GAL data.

- For the AHS, the entity ID variables are control and year.

- For the ES202, the entity ID variables are sein, seinunit, year, and quarter. The flag variable `e_flag` indicates whether the address came from the `address_street1`, `address_state`, and `address_zip9` variables (`e_flag=P` for physical address) or from the `ui_address_street1`, `ui_address_state`, and `ui_address_zip9` variables (`e_flag=M` for mailing address).

- For the ACS-POW data, the entity ID variables are `acsfileseq, cmid, seq,` and `pnum`.

- For the Business Register, the entity ID variables are `cfn, year`, and `singmult`. The flag variable `singmult` indicates whether the entity resides in the single-unit (su) or the multi-unit (mu) data set. Another flag variable `b_flag` indicates whether the address originated from the variables `pstreet, pplce, pst`, and `pzip` (`b_flag=P` for physical address) or `street, plce, st`, and `zip` (`b_flag=M` for mailing address). .

- For the MAF, `mafid` and `year` identify entities.

8.4.4 Resources for geographic information

The best place for information about Census geography is

<div align="center">

http://www.census.gov/geo/www/reference.html.

</div>

Especially informative is the Geographic Areas Reference Manual (GARM), at

<div align="center">

http://www.census.gov/geo/www/garm.html

</div>

8.5 DATA SET DESCRIPTIONS

8.5.1 Naming scheme

All GAL files are labelled with the geovintage used in the creation, i.e., *2003* in S2004, *2006* in S2008, and *2010* in S2011, and except for the main dataset, a suffix, composed of a dataset abbreviation and a calendar year. SAS datasets with zero observations are attached to this document:[1]

- gal/gal_zz_2010.sas7bdat

- gal/gal_zz_2010_tccb.sas7bdat

- gal/gal_zz_2010_xwalk_yyyy.sas7bdat

- galt26/gal_zz_2010_t26flags.sas7bdat

- galt26/gal_zz_2010_t26.sas7bdat

`ZZ` stands for the state postal abbreviation, and `YYYY` for a calendar year. Not all files are available for all states. In particular, LEHD-related crosswalks are only available for states actively participating with LEHD at the time of creation of the GAL.

Suffix	Crosswalk to:	Availability
acspow	American Community Survey Place-of-Work Coding	2001-2005
ahs	American Housing Survey	as of 2002
br	Business Register (ex-SSEL)	1990-2001
maf	(Census) Master Address File	as of 2004
xwalk	LEHD ES-202	varies by state; consult LEHD-ES-202 documentation

Files with suffixes `t26`, `t26flags`, and `tccb` are not cross-walks. Consult Sections 8.5.4, 8.5.5, and 8.5.6, respectively.

8.5.2 Data location

The files are stored in two main directories, with state-specific subdirectories:

```
gal/ZZ/       for most files
galt26/ZZ     for files with Title 26 protected content
```

On the RDC network, both directories can be found under

```
/mixed/lehd/s2011
```

For the exact location, consult with the RDC administrator.

---

1. Also visible on the attachment tab - Adobe Reader may be required.

8.5.3   Main dataset: GAL_ZZ_2010

This file does not contain data protected exclusively under Title 26. Consult Section 8.5.4 and 8.5.5. This file also does not report any address data sourced exclusively from ES-202. If a field contains address data sourced exclusively from ES-202, the values have been blanked on this file.

Record identifier: GALID

Sort order: GALID

File indexes: none

Entity unique address

Unique Entity Key GALID

Field name	Data dictionary reference name	Starting position	Field size	Data type
4-digit Census Place	A_PLACE	00016	8	N
5-digit Core-Based Statistical Area	A_CBSA	00589	5	A/N
5-digit FIPS MCD	A_FIPSMCD	00599	5	A/N
5-digit FIPS Place	A_FIPSPL	00604	5	A/N
6-character Traffic Analysis Zone (leading blanks)	A_TAZ	00581	8	A/N
CBSA Type 1=Metro, 2=Micro, Else=9	A_CBSA_MEMI	00609	1	A/N
Census Block suffix 1	A_BLOCK_SUF1	00570	1	A/N
Census Block suffix 2	A_BLOCK_SUF2	00571	1	A/N
Census block within tract	A_BLOCK	00566	4	A/N
Census tract within county	A_TRACT	00575	6	A/N
Code1 Census block id 3 digit	C1_BLOCK	00183	3	A/N
Code1 Census block id 4 digit	C1_BLOCK4	00047	4	A/N
Code1 Census geocode (tract)	C1_GEOCODE	00171	12	A/N
Code1 USPS record type	C1_USPSRECTYPE	00161	1	A/N
Code1 ZIP	C1_ZIP	00151	5	A/N
Code1 ZIP code status	C1_ZIP_STATUS	00165	1	A/N
Code1 ZIP return code	C1_ZIP_RC	00041	1	A/N
Code1 ZIP source	C1_ZIP_SRC	00160	1	A/N
Code1 ZIP+4 code	C1_ZIP4	00156	4	A/N
Code1 ZIP4 return code	C1_ZIP4_RC	00042	1	A/N
Code1 address correctness score	C1_ADDRESS_CS	00163	1	A/N
Code1 address return code	C1_ADDRESS_RC	00037	1	A/N
Code1 address w/apt	C1_ADDRESS	00051	70	A/N
Code1 alias/base return code	C1_ALIAS_RC	00038	1	A/N
Code1 apartment return code	C1_APT_RC	00046	1	A/N
Code1 carrier route return code	C1_CARRTE_RC	00043	1	A/N
Code1 city name	C1_CITY	00121	28	A/N
Code1 city/state return code	C1_CITYSTATE_RC	00040	1	A/N
Code1 directional return code	C1_DIRECTIONAL_RC	00044	1	A/N
Code1 dropped information code	C1_DROPPEDINFO_RC	00039	1	A/N
Code1 general return code	C1_GENERAL_RC	00036	1	A/N
Code1 geocode return code	C1_GEO_RC	00170	1	A/N
Code1 lat/long coordinate	C1_LATLONG	00186	20	A/N
Code1 lat/long level	C1_LATLONG_RC	00206	1	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Code1 master file vintage	C1_VINTDATE	00166	4	A/N
Code1 overall correctness	C1_OVERALL_RC	00162	1	A/N
Code1 state abbrev	C1_STATE	00149	2	A/N
Code1 street name correctness score	C1_STREETNAME_CS	00164	1	A/N
Code1 suffix return code	C1_SUFFIX_RC	00045	1	A/N
Commercial, Mixed, or Residential	OCCUPANT_TYPE	00553	1	A/N
Contains only ES202-sourced information	FLAG_ES202	00033	3	N
Contains only T26-sourced information	FLAG_T26	00030	3	N
Describes source of block coding	A_BLOCK_SRC	00572	1	A/N
FIPS county within state	A_CTY	00596	3	A/N
FIPS state	A_ST	00594	2	A/N
FIPS state——FIPS county	A_SSCCC	00616	5	A/N
FIPS state——FIPS county——Census tract	A_GEOCODE	00554	11	A/N
Flag for output contribution status from ACS Place of Work in 2001	P2001	00546	1	A/N
Flag for output contribution status from ACS Place of Work in 2002	P2002	00547	1	A/N
Flag for output contribution status from ACS Place of Work in 2003	P2003	00548	1	A/N
Flag for output contribution status from ACS Place of Work in 2004	P2004	00549	1	A/N
Flag for output contribution status from ACS Place of Work in 2005	P2005	00550	1	A/N
Flag for output contribution status from ACS Place of Work in 2006	P2006	00551	1	A/N
Flag for output contribution status from ACS Place of Work in 2007	P2007	00552	1	A/N
Flag for output contribution status from ES202 in 2000	E2000	00531	1	A/N
Flag for output contribution status from ES202 in 2001	E2001	00532	1	A/N
Flag for output contribution status from ES202 in 2002	E2002	00533	1	A/N
Flag for output contribution status from ES202 in 2003	E2003	00534	1	A/N
Flag for output contribution status from ES202 in 2004	E2004	00535	1	A/N
Flag for output contribution status from ES202 in 2005	E2005	00536	1	A/N
Flag for output contribution status from ES202 in 2006	E2006	00537	1	A/N
Flag for output contribution status from ES202 in 2007	E2007	00538	1	A/N
Flag for output contribution status from ES202 in 2008	E2008	00539	1	A/N
Flag for output contribution status from ES202 in 2009	E2009	00540	1	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flag for output contribution status from ES202 in 2010	E2010	00541	1	A/N
Flag for output contribution status from ES202 in 2011	E2011	00542	1	A/N
Flag for output contribution status from ES202 in 2012	E2012	00543	1	A/N
Flag for output contribution status from MAF	M2010	00545	1	A/N
Flag for output contribution status from the AHS	H2002	00544	1	A/N
Latitude, 6 implied decimal places	A_LATITUDE	00000	8	N
Longitude, 6 implied decimal places	A_LONGITUDE	00008	8	N
Maf, Code1, Derived, Block (or part) internal point	A_LATLONG_SRC	00565	1	A/N
Maf; Assigned by distance	A_BLOCK_SUF_SRC	00574	1	A/N
Quality of lat/long	A_GEOQUAL	00027	3	N
Street Number	DQ_STNO	00207	15	A/N
Unique GAL address ID A——YYYYMMDD_HHMM_ST——nnnnnnnnnnn	GALID	00502	29	A/N
Vintage of Census geography (GRF)	A_VINTAGE	00024	3	N
Workforce Investment Board area	A_WIB	00610	6	A/N
only addr on blockFace; extraPolation	A_LATLONG_DRV	00573	1	A/N
street extension	DQ_STEXT	00472	15	A/N
street extension number	DQ_STEXTN	00487	15	A/N
street name	DQ_STNAME	00237	200	A/N
street post direction	DQ_STPOST	00457	15	A/N
street prefix	DQ_STPRE	00222	15	A/N
street type	DQ_STTYPE	00437	20	A/N

8.5.4 Auxiliary dataset: GAL_ZZ_2010_T26

This file has the same column structure as the main file, but contains all records sourced exclusively from Title 26-protected information. The columns are described in Section 8.5.3.

Record identifier: GALID

Sort order: GALID

File indexes: none

Entity unique address

Unique Entity Key GALID

8.5.5 Auxiliary dataset: GAL_ZZ_2010_T26flags

This file contains all Business Register-related flags, for all GAL records. The variables are not sourced themselves from T26-protected files, but are suppressed because they reveal fact-of-filing.

Record identifier: GALID

Sort order: GALID

File indexes: none

Entity unique address

Unique Entity Key GALID

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flag for output contribution status from the Standard Statistical Establishment List in 1990	B1990	00060	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1991	B1991	00061	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1992	B1992	00062	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1993	B1993	00063	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1994	B1994	00064	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1995	B1995	00065	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1996	B1996	00066	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1997	B1997	00067	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1998	B1998	00068	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 1999	B1999	00069	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 2000	B2000	00070	1	A/N
Flag for output contribution status from the Standard Statistical Establishment List in 2001	B2001	00071	1	A/N
Source file of this address	SRCMAST	00000	5	A/N
Tracking ID	BIGSRCID	00005	26	A/N
Unique GAL address ID A——YYYYMMDD_HHMM_ST——nnnnnnnnnnn	GALID	00031	29	A/N

8.5.6 Auxiliary dataset: GAL_ZZ_2010_TCCB

The TCCB file provides county centroids in a structure similar to the main GAL file.

Record identifier: GALID

Sort order: GALID

File indexes: none

Entity unique address

Unique Entity Key GALID (merge or concat?)

Field name	Data dictionary reference name	Starting position	Field size	Data type
4-digit Census Place	A_PLACE	00095	4	A/N
5-digit Core-Based Statistical Area	A_CBSA	00070	5	A/N
5-digit FIPS MCD	A_FIPSMCD	00082	5	A/N
5-digit FIPS Place	A_FIPSPL	00087	5	A/N
A=Arcview	A_BLOCK_SRC	00068	1	A/N
CBSA Type 1=Metro, 2=Micro, Else=9	A_CBSA_MEMI	00075	1	A/N
Census block suffix 1	A_BLOCK_SUF1	00099	1	A/N
Census block suffix 2	A_BLOCK_SUF2	00100	1	A/N
Census block within tract	A_BLOCK	00064	4	A/N
FIPS state——FIPS county——Census tract	A_GEOCODE	00053	11	A/N
Latitude, 6 implied decimal places	A_LATITUDE	00000	8	N
Longitude, 6 implied decimal places	A_LONGITUDE	00008	8	N
MSA-PMSA	A_MSAPMSA	00069	1	A/N
Metropolitan Civil Division	A_MCD	00092	3	A/N
Quality of lat/long	A_GEOQUAL	00016	8	N
Unique GAL address ID A——YYYYMMDD_HHMM_SS——nnnnnnnnnnnn	GALID	00024	29	A/N
Workforce Investment Board area	A_WIB	00076	6	A/N

8.5.7   ES202 Crosswalk: GAL_ZZ_2010_XWALK_YYYY

There is one ES202 Crosswalk per year `YYYY` of input data. The files are called `gal_ZZ_2010_xwalk_YYYY`.

Record identifier: sein, seinunit, year, quarter

Sort order: sein seinunit

File indexes: none

Entity Reporting unit (SESA)

Unique Entity Key sein, seinunit, year, quarter

Field name	Data dictionary reference name	Starting position	Field size	Data type
Address ID	GALID	00006	29	A/N
P=physical,M=mailing	E_FLAG	00052	1	A/N
Quarter (numeric)	QUARTER	00003	3	N
State Employer ID Number	SEIN	00035	12	A/N
State UI Reporting Unit Number	SEINUNIT	00047	5	A/N
Year YYYY	YEAR	00000	3	N

8.5.8 Summary information on datasets

Table 8.5: Number of observations for GAL

Group	Number of datafiles	Records (1000s)	Filesize (GB)
GAL	989	747,614	74

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Due to the very large number of GAL files, the usual list of files is not printed. The interested reader is referred to the attached CSV file for more details.

8.6   NOTES

Chapter 9.
Individual Characteristics File (ICF)

9.1 OVERVIEW

The National Individual Characteristics File (ICF) contains one record for every person who is ever employed in any LEHD state over the time period spanned by states's unemployment insurance records, conditional on participation in the LEHD program (the equivalent file for persons employed at some point by the federal government in positions covered by OPM is constructed independently, and documented separately). It consolidates information from multiple input sources on gender, age, place-of-birth, race, ethnicity, and education. Additional information on yearly place-of-residence since 1999 is also available. Information on age, gender, place-of-birth, race, ethnicity, and education is imputed ten times when missing.

9.1.1 Details of the Construction of the ICF Variables

9.1.1.1 Overview

Each variable on the new ICF is derived from at least one of the following three files: the PCF, the HCEF, or the SCEF. The PCF is the Person Characteristics file, which is our conduit to Census' version of the SSA Numident file. The Numident contains information recorded from every transaction with SSA, including the initial application. The HCEF is the Hundred Percent Census Edited File and the SCEF is the Sample Census Edited File, commonly known as the short and long form respectively. Response variables on the HCEF and the SCEF generally begin with a Q and their corresponding flag begins with an F.

Many variables such as date of birth have multiple sources, but some such as education only have one source. If more than one source exists, one is designated as the primary and the other is the secondary. If the variable is missing then the standard SAS missing value for that type is used, a single blank space for character variables and a dot for numeric.

9.1.1.2 Imputation process for demographic variables

The three source data files follow a monotone data pattern. This type of data pattern allows us to implement a hierarchical approach. We start with the variables with the least amount of missing data, the variables sourced from the PCF. In stage A we complete (replace a missing value with an implicate (the actual value returned from an imputation)) DOB, gender, and POB. With these variables completed we move on to the HCEF (stage B) and impute race and ethnicity conditional on the DOB, gender, and POB values from stage A. In the final stage we complete education conditional on the variables imputed in both stage A and B. By looking at the percent column in the table above you can get an idea of the amount of completed data for each variable. For example POB has about 5% completed data, while education has about 88% completed or 12% as reported (see Table 9.1.1.3 on page 9-2).

Table 9.1: Distribution of data sources for the ICF

PCF	HCEF	SCEF	Percent
in	in	in	12%
in	in	not	61%
in	not	not	22%
not	not	not	5%

9.1.1.3 Updating records in the LEHD Production process

The full-information update is occassionally done by LEHD Research staff (at least once upon a state joining the LEHD Program). During regular production, workers not present on the (previous quarter's) National ICF show up every quarter (mostly new entrants to the US labor force in partner states). Thus, during the production process, a "new worker impute" is performed, using less information (and less computational resources) than the full-information impute. End-of-quarter processing unduplicates the state-level updates, and adds the new workers to the National ICF for the next production cycle.

Thus, the National ICF available to researchers in this snapshot will contain both records from one or more full-information process runs, and separate reduced-information runs at the state level. The SOURCE and VINTAGE variables identify the source of each record.

Quantifying cumulative updates. Across all available states, and cumulatively over a 4-quarter period, only about 0.03% of all PIKs are added by the state-level process. This fraction is somewhat higher among observations with imputed values (lag in availability of updated PCF, which itself has a lag with respect to the first-time appearance of workers on the labor market).

- Overall in the base data, about 10% of PIKs have a age or sex or POB impute. Among the updated (unique) records, this proportion is 35%.

- Overall in the base data, slightly more than 30% of PIKs have a race or ethnicity impute. Among the updated (unique) records, this proportion is slightly more than 55%.

- Overall in the base data, about 92% of PIKs have a education impute. Among the updated (unique) records, this proportion is over 99%.

Overlap between state-level updates. Because each state-level process identifies a PIK with missing data separately, workers who (over the course of the 4 quarters analyzed) appear in multiple states have independent imputes in each state. Naturally, these imputes are not identical. Over the four quarters, about 2% of PIKs appear in at least one other state (about 3.5% of those appear in more than 2 states). All imputes condition on the same type of information (all new workers, by definition, do not have a work history), and only one (randomly selected) record (and its imputes) is retained when updating the National ICF at the end of the production cycle.

9.1.1.4 Place of residence imputation

Place of residence information on the ICF is derived from the StARS (Statistical Administrative Records System), as provided historically to LEHD as the Composite Person Record (CPR) (Chapter 4). The vast majority of the individuals found in the UI wage records have information in these data on the place of residence down to the exact geographical coordinates. However, in more than 1 percent of all cases the geography information is incomplete or missing. The QWI estimation relies on completed place of residence information. Because this

information is a critical conditioning variable in the Unit-to-Worker Impute (U2W) (Chapter 14) imputation model, all missing residential addresses are imputed.

County of residence is imputed based on a categorical model of the data that is a fully-saturated contingency table. Separately for each year, unique combinations of categories of sex, age, race, income and state/county of work are used to form $i = 1, \ldots, I$ populations. For each sample i, the probability of residing in a particular county, π_{ij}, is estimated by the sample proportion, $p_{ij} = n_{ij}/n_i$, where $j = 1, \ldots, J$ indexes all the counties in the country.

County of residence is then imputed based on

$$county = j \;\; \text{if} \;\; P_{ij-1} \leq u_k < P_{ij}$$

where P_i is the CDF corresponding to p_i for the ith population and $\mu_{kl} \sim U[0,1]$ is one of $k = 1, \ldots, 10$ independent draws for the lth individual belonging to the ith population.

No geography below the county level is imputed and in those cases where exact geographical coordinates are incomplete the centroid of the finest geographical area is used. Thus, in cases where no geography information is available this amounts to the centroid of the imputed county.

9.1.1.5 Understanding the differences between the S2011 ICF geography files and pre-S2011 data

In order to better understand the changes made to the ICF4 (in S2011) home address algorithm, it is necessary to understand the ICF3 implementation (as included in previous LEHD snapshots), which has been unchanged throughout the previous history of the QWI.

- The ICF3 attempts to find an address on the 1999 CPR for all workers that had worked in a particular state at any in-scope time (as determined through the EHF).

- If an address was not found on the 1999 CPR, an imputation algorithm was used. Race, sex, age (3 categories), position in the earnings distribution (3 categories), and place of work contribute to the imputation of place of residence. (see S2008 documentation for more details)

- To classifiy the worker's position in the earnings distribution, the 1999 earnings distribution for the entire state is computed. If a state was not part of QWI in 1999, the closest year is chosen instead.

- A worker's employment record in the year closest to the CPR year (=1999) in that particular state is chosen, and earnings classified according to the state's earnings distribution. If the worker did not work in the CPR year, earnings are not used for that worker in the impute.

- Race is simplified to 'white' and 'non-white'. Since race was not completed on the ICF3, a simple completion algorithm is used to ensure presence of the 'white' / 'non-white' indicator. This is only used in the imputation algorithm, and not anywhere else.

- If the observed residence county is out-of-state, then county is recoded to 'OSt'. This becomes an assignment in the imputation.

- The conditional distribution of county-of-residence is computed by race/sex/agegroup/earningsgroup/county-of-work.

- 10 implicates are drawn from this distribution (only one is available on the S2011 file). If there are less than 100 workers in a particular county of work, then for any workers with missing place-of-residence, the county-of-work is assigned as the county-of-residence.

- Subsequently, place of residence is coded to the latitude/longitude level. For imputed records, the county centroid is used if in-state, and latitude/longitude is missing for those imputed into the 'out-of-state' category.

- Imputation rates vary across states, and range from 10% to 18%, overall, and from below 2% to 20%, by year.

The entire algorithm was executed every time a ICF is run, for all workers with jobs ever held in that state. No cross-state information was used, and workers with missing residence information in multiple states have different draws from the distribution in each state.

The ICF3 place of residence was only used in the U2W, and only one implicate was used. OnTheMap used its own lookup of residence location in the full time series of the available CPR files.

9.1.1.6 Summary of modifications for ICF4

While the generic imputation algorithm (based on distributions of county-of-residence by race/ sex/ agegroup/ earningsgroup/ county-of-work) does not change in the ICF4, the following changes have been made to improve the place-of-residence algorithm.

- The database has been upgraded to account for time-varying address information. Thus, for all years 1999 and forward, if the worker was observed working in some LEHD state, a separate database entry is created.

- Databasing of imputes. Imputes are drawn only the first time a worker shows up in any state. Under development is currently the algorithm under which subsequent updates to geographic information may replace imputes.

- The database is national in scope. Thus, each worker has only one address on file (per year in which a job was observed to be held in some LEHD state), and when imputed, only one set of draws from the distribution is used and stored. While this does not affect the probability of finding a record in the CPR for place of residence, the ICF3 treated all out-of-state as a generic "OSt" class. The nationally dominant job is used to impute residence, rather than a per-state total earnings measure. This is likely to change the imputations for secondary jobs in cross-border metropolitan jobs (daytime in New York City, evening job in Newark).

- No 'out-of-state' coarsening is used.

- Enhanced CPR lookup. For new workers, all available CPR files are used to attempt to find a known place of residence. The algorithm used in ICF3 only considered the 1999 CPR. The ICF4 considers contemporaneous CPR first, then widens the search, using an iterative algorithm that starts with a forward search of one year, then a backward search of one year, then a forward search of two years, and finally a backward search of two years. For now, information available farther away than two years is not used. Failing the widened lookup, an impute is run. In general, the new algorithm substantially reduces the imputation ratio. Historical impute rates for years after 1999 drop from up to 10% to around 2%.

- Earnings distributions for all years are used, and a worker's earnings are classified by the earnings of the year s/he first appeared. To the extent that the earnings distribution changes over the years, many young workers will now be correctly classified for the purpose of imputation.

9.1.2 Variable Details

Imputation flags Every variable has a corresponding imputation flag variable that identifies the status (observed or imputed) of the main variable. The impute flag name is always of the format <varname>_imputed. For example for DOB the flag is named DOB_imputed.

- All impute flags can take on the following values:

  - "1" = "observed"

- "2" = "imputed"
- "3" = "imputed but not replaced, implicates 1-10 on implicate file"

The third value for the imputation flag is assigned when the observed value of <varname> fails consistency checks, and is deemed implausible, for instance, when observed age at the beginning of the first quarter of labor market activity is less than 12 or greater than 85.

Source flags Every variable also has a corresponding flag variable that tells the user the source and status (reported or missing) of the main variable. The flag name is always of the format <varname>_flag. For example for DOB the flag is named DOB_flag. Two flag values are reserved: 0 indicates an "as reported" value and 9 indicates a missing value. These variables are primarily of use for internal processing, and only available in the RDC on the `icf_us_nonworker` (Section 9.3.6) file.

Date of Birth

- Primary Source File: PCF

    - Variables: DOBYYO, DOBMMO, and DOBDDO

- Secondary Source File:HCEF

    - Variables: QDB and FDB

- Output Variable Name: DOB

- Variable Construction: The year, month, and day variables from the PCF and HCEF are first cleaned in preparation for conversion to a valid SAS date. We begin by marking as ineligible for further processing any values that contain less than 4 numeric characters (0-9). Next, each variable still eligible is converted to numeric. The year value is checked first and if a valid year (year > current year - 126, which is a max age of 125) is present then processing continues. If month is between 1 and 12 and day is between 1 and 31 inclusive then a date is potentially complete to the day (our finest resolution). Each type is processed separately: if year only is present then the month and day are imputed, if year and month are present then the day is imputed and finally if all 3 are available then the day is checked to insure it is valid and if not it is replaced with the closest valid value (28, 29, or 30). The end result is a SAS date for each HCEF and PCF value with a valid year. In the final step, the information is combined with the PCF information taking precedence unless the PCF is not available or the HCEF is clearly superior (valid year, month, and day reported).

- Notes:

    - The SAS functions year(), month(), and day() can always be used to create a Gregorian calendar year of birth from the SAS date.
    - When calculating age, please use the following formula: age=(reference SAS date - DOB)/365.2425).
    - SAS does not consistently handle non-integer DOB values, which is a known issue, see http://support.sas.com/kb/24/808.html.

- DOB values:

    - Date of birth will be stored as a SAS date on the file. The SAS System stores date values as an offset in days from January 1, 1960. (SAS numeric 4)

- DOB_flag Values:

    - 0=PCF DOB valid and complete non-corrected

  – 1=PCF DOB valid and complete once day is corrected

  – 2=PCF DOB valid year and month, day is imputed

  – 3=PCF DOB valid year, month and day are imputed

  – 4=HCEF DOB of type 0 replaces a missing or type 2,3 PCF DOB

  – 5=HCEF DOB of type 1 replaces a missing or type 2,3 PCF DOB

  – 9=DOB missing

Gender

- Primary Source File: PCF

  – Variable: gender

- Secondary Source File: HCEF

  – Variables: qsex and fsex

- Output Variable Name: gender (internal) sex (production)

- Variable Construction: PCF gender takes precedence over HCEF gender. HCEF qsex is used when PCF gender is missing and qsex is either reported or imputed based on the first name of the respondent.

- sex values:

  – M=Male

  – F=Female

- sex_flag values:

  – 0=PCF sex is a M or F

  – 1=HCEF sex (PCF sex is not a M or F and fsex is a 0 (reported) or a 1 (allocated based on first name))

  – 9=gender missing

Place of Birth

- Primary Source File: PCF

  – Variables: POBST and POBFIN

- Secondary Source File: SCEF

  – Variables: qpobst and fpob

- Output Variable Name: POB

- Variable Construction: The PCF and HCEF variables are passed through formats, assigning each country code to either a new standardized country code or region. The individual country codes represent the top 23 immigrant source countries (including Puerto Rico) among all PIK records with a valid POBST and POBFIN and at least one quarter of positive earnings. Together, over 70% of the foreign born emigrated from one of the 23 source countries. In addition, the list contains every source country with at least 1% of the U.S. foreign born population.

- POB values:

  - A = US or territory (not Puerto Rico)
  - B = Mexico
  - C = Philippines
  - D = Vietnam
  - E = India
  - F = Germany
  - G = Puerto Rico
  - H = El Salvador
  - I = Cuba
  - J = United Kingdom
  - K = Canada
  - L = China
  - M = South Korea
  - N = Taiwan
  - O = Guatemala
  - P = Japan
  - Q = Haiti
  - R = USSR Core
  - S = Jamaica
  - T = Columbia
  - U = Poland
  - V = Iran
  - W = Dominican Republic
  - X = Italy
  - Y = Former Socialist Europe
  - Z = Western Europe
  - 1 = Central Asia
  - 2 = South East Asia
  - 3 = Middle East and North Africa
  - 4 = Caribbean
  - 5 = Central America
  - 6 = South America
  - 7 = Africa
  - 8 = Oceania

- POB_flag values:

  - 0=PCF POB is valid and complete
  - 1=HCEF POB is valid and complete as reported
  - 9=POB missing (Born abroad of an unknown country are included here, POBFIN=* and POBST=" " are removed)

Race

- Primary Source File: HCEF

    - Variables: imprace, frace, and fimprace

- Output Variable Name: race

- Variable Construction: A collapsed version of imprace is the primary source. This variable was chosen after an exhaustive analysis of both the recorded responses from the HCEF (inrace1-inrace21) and the variables qrace1-qrace8 that capture the 8 "best" responses. First, the inrace variables were shown to be mapped sensibly into the qrace variables. A variable was created using qrace, but the values were consistent with imprace, once processing as described in the flags was applied. This rendered the variable constructed from qrace obsolete and imprace was used directly. The interaction of the flag variables frace and fimprace is represented in race_flag. Generally when the Census has information not available to our imputer, the HCEF edit or allocation was retained. For example, if a person's race response is missing but at least one other member of the household reports a valid race, then the allocation is retained. However, this rule was not blindly applied, the quality of the allocation was confirmed using bestrace from the Numident. The correspondence with bestrace must be relatively high for the allocation to be retained. No hot (cold) deck allocations were retained.

- race values:

    - 1=White Alone
    - 2=Black or African American Alone
    - 3=American Indian or Alaska Native Alone
    - 4=Asian Alone
    - 5=Native Hawaiian or Other Pacific Islander Alone
    - 7=Two or More Race Groups

- race_flag values:

    - 0=(frace=0 and fimprace=0) race is as reported
    - 1=(frace=0 and fimprace=1) For multiple race respondents, the write-in "some other" race value is dropped and the checkbox (valid) race value is retained
    - 2=(frace=0 and fimprace=4) For respondents with only a "some other" race value, a new valid race was allocated from within the household
    - 3=(frace=1 and fimprace=0) Code changed through a consistency check
    - 4=(frace=1 and fimprace=1) For multiple race respondents, the code changed through consistency check write-in "some other" race value is dropped and the checkbox (valid) race value is retained
    - 5=(frace=3 and fimprace=0) The classified from race response in the Hispanic question value is retained
    - 6=(frace=3 and fimprace=4) The classified from race response in the Hispanic question "some other" race value is allocated a new valid race from within the household.
    - 7=(frace=4 and fimprace=0) The allocated from within the household value is retained
    - 8=(frace=4 and fimprace=1) The allocated from within the household multiple race value is adjusted. The write-in "some other" race value is dropped and the checkbox (valid) race value is retained.
    - 9=race is missing
    - 10=(frace=4 and fimprace=4) The allocated "some other" race value is replaced with a non "some other" race value from another member of the household.

Ethnicity

- Primary Source File: HCEF

    - Variable: qspan and fspan

- Output Variable Name: ethnicity

- Variable Construction: The variable qspan is passed through a format to assign the 3 digit codes to a simple Hispanic or non-Hispanic. See the variable flag for details of the values retained. The values retained were determined using a similar logic as the race variable. The main exception is that some hot deck values are retained. In this case, the surname hot deck is retained due to its relatively close correspondence with bestrace.

- ethnicity values:

- N=Not Hispanic or Latino

- H=Hispanic or Latino

- ethnicity_flag values:

    - 0=as reported
    - 1=Multiple responses were a given a unique Hispanic or non-Hispanic code
    - 2=Assigned Hispanic from the race code
    - 3=Allocated from within the household
    - 4=Allocated using a hot deck conditioning on surname
    - 9=ethnicity is missing

Education (educ_c)

- Primary Source File: SCEF

    - Source Variable: qhigh and fhigh

- Output Variable Name: educ_c

- Variable Construction: The SCEF qhigh values were collapsed using a format. If present, the variable DOB was used to calculate the respondents age on April 1, 2000. If the age was greater than or equal to 25 then the as reported education value was retained. EDUC_C is derived from EDUC_F, which is not available on Production or RDC files.

- educ_c values: (values from the full education coding are in parenthesis):

    - 1 = Less than high school (1-8)
    - 2 = High school or equivalent, no college (9)
    - 3 = Some college or Associate degree (10-12)
    - 4 = Bachelor's degree or advanced degree (13-16)

- educ_c_flag values:

    - 0=as reported education, DOB available, and calculated age greater than or equal to 25.
    - 9=education missing

9.1.3 Changes in this Snapshot

Completely new structure Since the last snapshot (S2008), the ICF has been completely restructured. There now is a single national ICF, rather than state-level ICFs, and missing data is imputed (multiply) only once for any individual, then stored until observed data becomes available (in a later production cycle).

Users wishing to subset by person can condition on selected two-digit (numeric) PIK substrings (`substr(PIK,1,2)`). A separate file contains the longitudinal address information.

Access rules and conditions The National ICF is constructed based on data from the Census Numident (derived from SSA data), Decennial Census 2000 (HCEF for race/ethnicity, and SCEF for education), as well as imputation models which leverage all of the above, plus information on coworkers and neighbors, where the links are inferred from the LEHD Infrastructure and the CPR respectively. The longitudinal address information is derived from CPR information, and is subject to Title 26 restrictions. Address information is completed from 1999 to the most current CPR date, using longitudinal edits and imputation models that condition on contemporaneous coworker information.

Use of the National ICF is thus

- subject to approval by SSA

- subject to approval by IRS when using longitudinal address information

- incompatible with simultaneous access to swapped Decennial (100 Percent Detail File (HDF) and Sample Edited Detail File (SEDF))

- subject to additional conditions for the (planned) release of results, above and beyond general RDC and LEHD conditions.

The most recent version of these restrictions and rules are available from the RDC administrators or in the CES Researcher Handbook. We discuss the release restrictions in the next paragraph.

Disclosure avoidance rules for ICF Special rules apply for Census 2000 and ACS tabulations in general, and transfer to the ICF. Note that the National ICF (S2011) itself does not contain or use ACS information. The following is an extract from a memo to LEHD staff by LEHD Senior Management, which was first issued in 2003, and is continuously updated. The text below is from a draft 2013 version, and provided here for reference only. The latest memo always applies, and can be obtained through the RDC Administrator or the LEHD Research Branch Chief.

a. A research project is deemed to use Census 2000 data if any variable used in the production of the tables or research results comes from the HCEF/SCEF Decennial Census file system in use at LEHD.

[...]

c. A research project uses a "special tabulation" from Census 2000 or the ACS if it produces a table of results using input files that contain a variable from Census 2000 (definition 3.a) or ACS (definition 3.b). All special tabulations from Census 2000 or ACS must be directly reviewed by the Disclosure Review Board, except as noted below. See the attached memos for guidelines in preparing such tables. Note, in particular, the population definition rules, the rounding rules, and the required methodology for computing percentiles.

d. The finest level of detail that may be shown for Group Quarters data is Institutional/ Noninstitutional. There are no exceptions to this rule.

e. Special tabulations with geographic detail that is national or state-level may be released without prior DRB approval. LEHD disclosure review is still required.

f. Model-based statistical results (coefficients, standard errors) that were prepared from national or state-level geography may be released without prior DRB approval. If the model includes geographic controls at the sub-state level, the coefficients on these controls may not be released without DRB approval. It is OK to note on the table of coefficients: includes controls for [insert geography]."

The gist is that if researchers do state or national tabulations, they are OK, anything else will require DRB review. Researchers do not need approval by individual states, but the use of the ICF is subject to approval by SSA.

Dropping link variables to SIPP and CPS Furthermore, the ICF's function as a crosswalk to SIPP and CPS was no longer being actively maintained, and has been dropped - no crosswalked identifiers are stored on the ICF anymore, and must be obtained separately by researchers.

9.2 DATA CITATION

> U.S. Census Bureau. 2014. *Geo-coded Address List (GAL) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

9.3 DATA SET DESCRIPTIONS

9.3.1 Unique record identifier

The unique record identifier within each ICF file is the **P!** (**P!**)IK.

9.3.2 Naming scheme

There are 10 files in the ICF/ICFT26 group: SAS datasets with zero observations are attached to this document:[1]

- icf/icf_us_earn_pctile.sas7bdat

- icf/icf_us_geoprobs.sas7bdat

- icf/icf_us_implicates_age_sex_pob.sas7bdat

- icf/icf_us_implicates_education.sas7bdat

- icf/icf_us_implicates_race_ethnicity.sas7bdat

- icf/icf_us_nonworkers.sas7bdat

- icf/icf_us.sas7bdat

- icf/icf_us_wide.sas7bvew

- icf/nicf_us_wide.sas7bvew

- icft26/icf_us_addresses.sas7bdat

9.3.3 Data location

The files are stored in two main directories:

```
icf/us/        for most files
icft26/us      for files with Title 26 protected content
```

On the RDC network, both directories can be found under

```
/mixed/lehd/s2011
```

9.3.4 Main dataset: ICF_us

This is the core dataset, containing all observed non-FTI and the first implicate for imputed variables.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

---

1. Also visible on the attachment tab - Adobe Reader may be required.

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date and Time of File Creation	VINTAGE	00024	13	A/N
Date of birth	DOB	00000	4	N
Ethnicity	ETHNICITY	00007	1	A/N
Gender	SEX	00004	1	A/N
Highest educational attainment (age 25+)	EDUC_C	00008	1	A/N
Imputation status for DOB	DOB_IMPUTED	00019	1	A/N
Imputation status for POB	POB_IMPUTED	00020	1	A/N
Imputation status for race	RACE_IMPUTED	00018	1	A/N
Imputation status for sex	SEX_IMPUTED	00023	1	A/N
Imputation status of educ_c	EDUC_C_IMPUTED	00021	1	A/N
Imputation status of ethnicity	ETHNICITY_IMPUTED	00022	1	A/N
Place of birth	POB	00005	1	A/N
Protected Identification Key	PIK	00009	9	A/N
Race	RACE	00006	1	A/N
us=created by NICF process, if state abbreviation=new worker process in state	SOURCE_PROCESS	00037	2	A/N

9.3.5 Utility dataset (view): ICF_us_wide and NICF_us_wide

This is a SAS view (views do not work in Stata). For the utility of users wishing a wide file, this view combines all variables on all implicates (<varname>[n]) and the variables from the core ICF file (<varname>) into single dataset. Note that a view performs the merge "on the fly". The only difference between the two views is the naming of the sex/gender variables ("sex" on `icf_us_wide`, "gender" on `nicf_us_wide`). Only the variables from `icf_us_wide` are listed below. In general, researchers should use `icf_us_wide.sas7bvew`.

Usage: A SAS view is read the same way a regular SAS dataset is read:

```
libname icf '/mixed/lehd/s2011/icf/us' access=readonly;
data mysample;
set icf.icf_us_wide (where=(substr(pik,1,2)='01'));
/* further processing steps */
run;
```

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date and Time of File Creation	VINTAGE	00064	13	A/N
Date of birth	DOB	00000	4	N
Date of birth Implicate 1	DOB1	00004	4	N
Date of birth Implicate 10	DOB10	00040	4	N
Date of birth Implicate 2	DOB2	00008	4	N
Date of birth Implicate 3	DOB3	00012	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date of birth Implicate 4	DOB4	00016	4	N
Date of birth Implicate 5	DOB5	00020	4	N
Date of birth Implicate 6	DOB6	00024	4	N
Date of birth Implicate 7	DOB7	00028	4	N
Date of birth Implicate 8	DOB8	00032	4	N
Date of birth Implicate 9	DOB9	00036	4	N
Ethnicity	ETHNICITY	00047	1	A/N
Ethnicity Implicate 1	ETHNICITY1	00099	1	A/N
Ethnicity Implicate 10	ETHNICITY10	00117	1	A/N
Ethnicity Implicate 2	ETHNICITY2	00101	1	A/N
Ethnicity Implicate 3	ETHNICITY3	00103	1	A/N
Ethnicity Implicate 4	ETHNICITY4	00105	1	A/N
Ethnicity Implicate 5	ETHNICITY5	00107	1	A/N
Ethnicity Implicate 6	ETHNICITY6	00109	1	A/N
Ethnicity Implicate 7	ETHNICITY7	00111	1	A/N
Ethnicity Implicate 8	ETHNICITY8	00113	1	A/N
Ethnicity Implicate 9	ETHNICITY9	00115	1	A/N
Gender	SEX	00044	1	A/N
Gender Implicate 1	SEX1	00079	1	A/N
Gender Implicate 10	SEX10	00097	1	A/N
Gender Implicate 2	SEX2	00081	1	A/N
Gender Implicate 3	SEX3	00083	1	A/N
Gender Implicate 4	SEX4	00085	1	A/N
Gender Implicate 5	SEX5	00087	1	A/N
Gender Implicate 6	SEX6	00089	1	A/N
Gender Implicate 7	SEX7	00091	1	A/N
Gender Implicate 8	SEX8	00093	1	A/N
Gender Implicate 9	SEX9	00095	1	A/N
Highest educational attainment (age 25+)	EDUC_C	00048	1	A/N
Highest educational attainment (age 25+) Implicate 1	EDUC_C1	00119	1	A/N
Highest educational attainment (age 25+) Implicate 10	EDUC_C10	00128	1	A/N
Highest educational attainment (age 25+) Implicate 2	EDUC_C2	00120	1	A/N
Highest educational attainment (age 25+) Implicate 3	EDUC_C3	00121	1	A/N
Highest educational attainment (age 25+) Implicate 4	EDUC_C4	00122	1	A/N
Highest educational attainment (age 25+) Implicate 5	EDUC_C5	00123	1	A/N
Highest educational attainment (age 25+) Implicate 6	EDUC_C6	00124	1	A/N
Highest educational attainment (age 25+) Implicate 7	EDUC_C7	00125	1	A/N
Highest educational attainment (age 25+) Implicate 8	EDUC_C8	00126	1	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Highest educational attainment (age 25+) Implicate 9	EDUC_C9	00127	1	A/N
Imputation status for DOB	DOB_IMPUTED	00059	1	A/N
Imputation status for POB	POB_IMPUTED	00060	1	A/N
Imputation status for race	RACE_IMPUTED	00058	1	A/N
Imputation status for sex	SEX_IMPUTED	00063	1	A/N
Imputation status of educ_c	EDUC_C_IMPUTED	00061	1	A/N
Imputation status of ethnicity	ETHNICITY_IMPUTED	00062	1	A/N
Place of birth	POB	00045	1	A/N
Place of birth Implicate 1	POB1	00080	1	A/N
Place of birth Implicate 10	POB10	00098	1	A/N
Place of birth Implicate 2	POB2	00082	1	A/N
Place of birth Implicate 3	POB3	00084	1	A/N
Place of birth Implicate 4	POB4	00086	1	A/N
Place of birth Implicate 5	POB5	00088	1	A/N
Place of birth Implicate 6	POB6	00090	1	A/N
Place of birth Implicate 7	POB7	00092	1	A/N
Place of birth Implicate 8	POB8	00094	1	A/N
Place of birth Implicate 9	POB9	00096	1	A/N
Protected Identification Key	PIK	00049	9	A/N
Race	RACE	00046	1	A/N
Race Implicate 1	RACE1	00100	1	A/N
Race Implicate 10	RACE10	00118	1	A/N
Race Implicate 2	RACE2	00102	1	A/N
Race Implicate 3	RACE3	00104	1	A/N
Race Implicate 4	RACE4	00106	1	A/N
Race Implicate 5	RACE5	00108	1	A/N
Race Implicate 6	RACE6	00110	1	A/N
Race Implicate 7	RACE7	00112	1	A/N
Race Implicate 8	RACE8	00114	1	A/N
Race Implicate 9	RACE9	00116	1	A/N
base —— age —— race —— educ	_MERGE	00129	4	A/N
us=created by NICF process, if state abbreviation=new worker process in state	SOURCE_PROCESS	00077	2	A/N

9.3.6   Auxiliary dataset: ICF_us_nonworkers

This dataset contains observed values from all basedata, for records that were not completed in the first vintage of the National ICF research file. Due to updates from both the full-information and new-worker processes, there is overlap between this file and the universe of icf_us. This file contains no imputes, and only those variables from the basedata necessary for processing the full-information process.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date and Time of File Creation	VINTAGE	00024	13	A/N
Date of birth	DOB	00000	4	N
Ethnicity	ETHNICITY	00007	1	A/N
Gender	SEX	00004	1	A/N
Highest educational attainment (age 25+)	EDUC_C	00008	1	A/N
Imputation status for DOB	DOB_IMPUTED	00019	1	A/N
Imputation status for POB	POB_IMPUTED	00020	1	A/N
Imputation status for race	RACE_IMPUTED	00018	1	A/N
Imputation status for sex	SEX_IMPUTED	00023	1	A/N
Imputation status of educ_c	EDUC_C_IMPUTED	00021	1	A/N
Imputation status of ethnicity	ETHNICITY_IMPUTED	00022	1	A/N
Place of birth	POB	00005	1	A/N
Protected Identification Key	PIK	00009	9	A/N
Race	RACE	00006	1	A/N
us=created by NICF process, if state abbreviation=new worker process in state	SOURCE_PROCESS	00037	2	A/N

9.3.7 Age, sex, and place-of-birth implicates: ICF_us_implicates_age_sex

The first implicates for date of birth, sex, and place-of-birth are stored on the main ICF file as DOB, SEX, and POB. Imputed values are flagged by the appropriate flag. Other implicates are found in this file, and can be merged on when required.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Date of birth Implicate 1	DOB1	00000	4	N
Date of birth Implicate 10	DOB10	00036	4	N
Date of birth Implicate 2	DOB2	00004	4	N
Date of birth Implicate 3	DOB3	00008	4	N
Date of birth Implicate 4	DOB4	00012	4	N
Date of birth Implicate 5	DOB5	00016	4	N
Date of birth Implicate 6	DOB6	00020	4	N
Date of birth Implicate 7	DOB7	00024	4	N
Date of birth Implicate 8	DOB8	00028	4	N
Date of birth Implicate 9	DOB9	00032	4	N
Gender Implicate 1	SEX1	00049	1	A/N
Gender Implicate 10	SEX10	00067	1	A/N
Gender Implicate 2	SEX2	00051	1	A/N
Gender Implicate 3	SEX3	00053	1	A/N
Gender Implicate 4	SEX4	00055	1	A/N
Gender Implicate 5	SEX5	00057	1	A/N
Gender Implicate 6	SEX6	00059	1	A/N
Gender Implicate 7	SEX7	00061	1	A/N
Gender Implicate 8	SEX8	00063	1	A/N
Gender Implicate 9	SEX9	00065	1	A/N
Place of birth Implicate 1	POB1	00050	1	A/N
Place of birth Implicate 10	POB10	00068	1	A/N
Place of birth Implicate 2	POB2	00052	1	A/N
Place of birth Implicate 3	POB3	00054	1	A/N
Place of birth Implicate 4	POB4	00056	1	A/N
Place of birth Implicate 5	POB5	00058	1	A/N
Place of birth Implicate 6	POB6	00060	1	A/N
Place of birth Implicate 7	POB7	00062	1	A/N
Place of birth Implicate 8	POB8	00064	1	A/N
Place of birth Implicate 9	POB9	00066	1	A/N
Protected Identification Key	PIK	00040	9	A/N
Source of Impute	SOURCE_PROCESS	00082	2	A/N
Vintage of Impute	VINTAGE	00069	13	A/N

9.3.8   Education implicates: ICF_us_implicates_education

The first implicate is stored on the main ICF file as EDUC_C. Imputed values are flagged by the appropriate flag. Other implicates are found in this file, and can be merged on when required.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Highest educational attainment (age 25+) Implicate 1	EDUC_C1	00009	1	A/N
Highest educational attainment (age 25+) Implicate 10	EDUC_C10	00018	1	A/N
Highest educational attainment (age 25+) Implicate 2	EDUC_C2	00010	1	A/N
Highest educational attainment (age 25+) Implicate 3	EDUC_C3	00011	1	A/N
Highest educational attainment (age 25+) Implicate 4	EDUC_C4	00012	1	A/N
Highest educational attainment (age 25+) Implicate 5	EDUC_C5	00013	1	A/N
Highest educational attainment (age 25+) Implicate 6	EDUC_C6	00014	1	A/N
Highest educational attainment (age 25+) Implicate 7	EDUC_C7	00015	1	A/N
Highest educational attainment (age 25+) Implicate 8	EDUC_C8	00016	1	A/N
Highest educational attainment (age 25+) Implicate 9	EDUC_C9	00017	1	A/N
Protected Identification Key	PIK	00000	9	A/N
Source of Impute	SOURCE_PROCESS	00032	2	A/N
Vintage of Impute	VINTAGE	00019	13	A/N

9.3.9   Race and ethnicity implicates: ICF_us_implicates_race_ethnicity

The first implicates are stored on the main ICF file as RACE and ETHNICITY. Imputed values are flagged by the appropriate flag. Other implicates are found in this file, and can be merged on when required.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Ethnicity Implicate 1	ETHNICITY1	00009	1	A/N
Ethnicity Implicate 10	ETHNICITY10	00027	1	A/N
Ethnicity Implicate 2	ETHNICITY2	00011	1	A/N
Ethnicity Implicate 3	ETHNICITY3	00013	1	A/N
Ethnicity Implicate 4	ETHNICITY4	00015	1	A/N
Ethnicity Implicate 5	ETHNICITY5	00017	1	A/N
Ethnicity Implicate 6	ETHNICITY6	00019	1	A/N
Ethnicity Implicate 7	ETHNICITY7	00021	1	A/N
Ethnicity Implicate 8	ETHNICITY8	00023	1	A/N
Ethnicity Implicate 9	ETHNICITY9	00025	1	A/N
Protected Identification Key	PIK	00000	9	A/N
Race Implicate 1	RACE1	00010	1	A/N
Race Implicate 10	RACE10	00028	1	A/N
Race Implicate 2	RACE2	00012	1	A/N
Race Implicate 3	RACE3	00014	1	A/N
Race Implicate 4	RACE4	00016	1	A/N
Race Implicate 5	RACE5	00018	1	A/N
Race Implicate 6	RACE6	00020	1	A/N
Race Implicate 7	RACE7	00022	1	A/N
Race Implicate 8	RACE8	00024	1	A/N
Race Implicate 9	RACE9	00026	1	A/N
Source of Impute	SOURCE_PROCESS	00042	2	A/N
Vintage of Impute	VINTAGE	00029	13	A/N

9.3.10   Title 26 information: ICF_us_addresses

FTI has been removed from the core ICF, and stored separately. Note that in the RDC network, this file is stored under a separate set of permissions, and if users require access to this information, need to request access to an additional group. T26 variables are starred below. In contrast to previous snapshots, only a single impute is provided.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
Admin record huid	HUID	00045	35	A/N
County of Residence imputation flag	COUNTYLIVEIMPUTED	00085	1	A/N
FIPS State(2) —— FIPS County (3) as of address_year	COUNTY_LIVE	00080	5	A/N
Flag quality of latitude/longitude of residence	FLAG_LATLONG	00000	8	N
Latitude of residence, 6 implied decimal places	LATITUDE_LIVE	00008	8	N
Longitude of residence, 6 implied decimal places	LONGITUDE_LIVE	00016	8	N
Protected Identification Key	PIK	00036	9	A/N
Source process (state name=ICF for that state)	SOURCE	00099	3	A/N
Vintage in which record was created	VINTAGE	00086	13	A/N
Year of address record - worker worked in this year	ADDRESS_YEAR	00024	8	N
Years away from observed CPR value (edit flag)	FLAG_DISTANCE	00032	4	N

9.3.11   Summary information on datasets

Table 9.9: Number of observations for ICF

Group	Number of datafiles	Records (1000s)	Filesize (GB)
ICF	9	1,364,805	102

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 9.10: Number of observations for ICFT26

Group	Number of datafiles	Records (1000s)	Filesize (GB)
ICFT26	1	2,018,825	197

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 9.11: List of data files for ICF, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
National (us)					
icf_us	1985Q2	2012Q1	262,106	10	1:445356+MADC
icf_us_earn_pctile	1999Q.	2010Q.	< 1	< 1	1:445360
icf_us_geoprobs	1990Q1	2008Q4	1,051	48	1:239272
icf_us_implicates_age_sex_pob	1985Q2	2012Q1	29,524	2	1:445357+MADC
icf_us_implicates_education	1985Q2	2012Q1	242,793	8	1:445358+MADC
icf_us_implicates_race_ethnicity	1985Q2	2012Q1	85,545	4	1:445359+MADC
icf_us_nonworkers	1990Q1	2009Q2	219,573	11	1:239973
icf_us_wide	1985Q2	2012Q1	262,106	10	1:44535x+MADC+v1
nicf_us_wide	1985Q2	2012Q1	262,106	10	1:44535x+MADC+v2

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:" to obtain the full SnapshotID.

Table 9.12: List of data files for ICFT26, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
National (us)					
icf_us_addresses	1990Q1	2008Q4	2,018,825	197	1:239271

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:" to obtain the full SnapshotID.

9.4 HELPFUL PROGRAMS

The following programs might be found to be useful when using the data.

9.4.1 Recombining T26 data with the core ICF

The following program allows users to combine the Title 26 variables with the core ICF. This program was used in slightly modified form for quality assurance during the preparation of the data for the RDC environment.

9.4.2 Selecting a random subsample of persons

The following program allows users to select a random sample of approximately one percent of individuals on the ICF. It relies on the fact that the first two characters of the PIK are approximately uniformly distributed on $[00, 99]$. Note that 'AA' is a valid value for the first two characters and denotes individuals for whom no valid SSN was on file. Occurrence of such "pseudo-PIKs" varies by state.

```
libname INLIB "/mixed/lehd/s2011/icf/us/";

data my_icf;
   set INLIB.icf_us(where=(substr(PIK,1,2)='01'));
run;
```

9.5 NOTES

Chapter 10.
Office of Personnel Management files (OPM)

10.1 OVERVIEW

In 2006, the Office of Personnel Management (OPM) signed a MOU to incorporate information on the federal workforce into the Census Bureau's data infrastructure, in order to "improve economic and demographic censuses, surveys, and intercensal population estimates." Research using these files is intended to further support the "Master Address File Program, current demographic and economic survey and census operations."

We note that OPM also releases quarterly data on the federal government's workforce at `http://www.fedscope.opm.gov`, and allows access to individual-level (de-identified) data underlying the FedScope data at `http://www.opm.gov/data/Index.aspx?tag=FedScope`. The data provided to the Census Bureau are extracted from the same Enterprise Human Resources Integration-Statistical Data Mart (EHRI-SDM) that feeds FedScope and the raw data at the above location.

The present chapter describes how federal workers are added to the QWI infrastructure, in order to create the new data product "Quarterly Workforce Indicators for OPM" (short: QWI-OPM). The core data is provided by OPM. We highlight the differences between the structure and content of the data provided by OPM and the data provided by state UI systems, and the efforts undertaken to make the data comparable.

The OPM data create some challenges. In contrast to the data from the state UI systems, which record cumulative employment over a quarter, OPM data are provided as a database extract, with a true point-in-time stock of employees at the end of a calendar year quarter, and a separate file providing for information on status changes. Whereas the UI systems record cumulative earnings, the OPM system only records the nominal annual salary, plus an indicator of whether or not a particular employee is full-time, part-time, or seasonal; neither system records actual hours worked. Finally, work location is not collected in the same manner as in the QCEW, and industry is not collected at all.

We have implemented solutions for all of these shortcomings. Federal workplaces do report their workplace employment in the QCEW. We leverage this information both to address the absence of precise workplace location in the OPM-provided data, and to assess coverage.

10.1.1 Data Sources and Definitions

10.1.1.1 Office of Personnel Management input data files

The OPM data provided to LEHD is composed of four types of files:

- Dynamics file: A personnel action file describing personnel actions for federal workers that took place during the quarter (and sometimes, took place in previous quarters but didnt show up in the file until later). In addition to basic characteristics of the workers largely included also on the status file (described below), the dynamics file records personnel actions for each federal worker. Personnel actions include accessions, separations, promotions, movements between different work schedules, adjustments in locality or basic pay, etc. The date of each action is recorded at daily precision.

- Status file: A status 'snapshot' of the federal workforce on a particular date (the last day of the last pay period in the calendar year quarter). Most of the variables on the status and dynamics file overlap, but not all. A worker will appear in the status file but not in the dynamics file if no personnel actions took place for this worker in that quarter. A worker will appear in the dynamics file, but not the status file, if that worker's attachment to the federal workforce was terminated during the quarter. Other, more complex situations, may also occur. Lags sometimes occur between actions described in the Dynamics file and their reflection in subsequent Status files. This will be addressed dynamically during processing.

- Standard Code Table (SCT) file: The Standard Code Table file is a lookup table for values in the fields in the Dynamics and Status files.

- Point of information (POI) file: A personnel office address file that gives a street address for the personnel officer contact.

In addition, a Duty Station File is available from the OPM website, mapping duty station codes to CBSA. This is used for QA purposes.

Detailed information on data elements can be found at http://www.opm.gov/feddata/guidance.asp. Note that not all elements are available on the files provided to LEHD.

Overall, OPM provides data on 543 agencies (402 Cabinet Level Agencies from 18 departments, 141 independent agencies).[1]

10.1.1.2 Available data and definitions

For the LEHD infrastructure, and to be compatible with the UI wage records, the key variable is quarterly earnings. We use the variable `totpay` (Total pay). All employment statistics are constructed for periods where positive earnings were received. `totpay` is computed by OPM, and includes

- basic pay

- locality adjustment

- supervisory differential

- retention allowance

- cost of living allowance (COLA)

We further use demographic data contained within the file, for tabulation of OPM-specific data, as well as to enhance other LEHD tabulations. Treatment of demographic data are described in Section 10.1.2.8.

10.1.1.3 Missing elements

For the LEHD production system, the critical elements missing are the realized quarterly earnings of employees, the roof-top address of the worker's workplace, and the industry coding of the agency. We address these issues in Section 10.1.2.

Furthermore, while it is possible to identify part-time or intermittent workers from data fields, no hours are reported, and the `totpay` variable contains the full-year-equivalent earnings. We describe how this is addressed in Section 10.1.2.

---

1. Data derived from public-use OPM files, current as of Sept 2011.

10.1.1.4 Problematic issues

Personnel actions with effective dates in quarter n may not appear in the Dynamics file until quarter $m > n$. Since agencies can delay Dynamics file submissions, and resubmit incorrect submissions up to two years later, a retrospective processing may be needed. We address this through a search of the entire Dynamics file history, and allow for revisions up to four quarters before the most current quarter being processed.

Personnel actions that lead to termination of employment with a particular agency in quarter n lead to no records (and thus no earnings) appearing in the Status file in quarter n . The Status file only reflects point-in-time employment at the end of the quarter. Any termination of employment within the quarter means the worker has no record at that job in that quarter. We address this by identifying such situations from the Dynamics file, and adding employment records in quarter n. *(Not implemented in 2012Q2 release)*

Award amounts not included in total pay We use only the `totpay` variable to define (regular) earnings. To completely compute earnings, any award amount (`awardamt`) from the Dynamics file (e.g., cash bonus, separation incentive, student loan repayment, etc.) needs to be added. The newer OPM system (Enterprise Human Resources Integration (EHRI)) has a variable that computes such complete earnings automatically, but only since 2009. Furthermore, the LEHD MOU pre-dating the introduction of EHRI, that variable is not transmitted to LEHD. In 2006, approximately 30% of workers received such awards.

Departmental reorganizations In 2002, the Department of Homeland Security (DHS) was created as a Cabinet-level department by Congress, with an effective date of March 1, 2003.[2] Multiple transfers of agencies, including complex re-assignment of personnel and responsabilities, were implemented (Table 10.9), which complicate flows. QWI tabulations uses a flow-based approach to capturing such reorganizations as they relate to separations and hires, but by design will not capture partial splits and mergers (Benedetto et al. 2007). Table XX shows the post-merger distribution of DHS employment across states, as a fraction of overall OPM employment.

10.1.1.5 Exclusions

Certain agencies provide no data to OPM (U.S. Office of Personal Management 2012)and are thus excluded from the data universe for LEHD processing (see Table 10.4). We also chose to exclude several agencies at this time due to particular processing restrictions:

- The concordance of the geographic location (county) of many Department of Defense bases was weak between OPM records and QCEW records. Until this matter can be resolved, all Department of Defense (civilian) employees are excluded from the QWI-OPM universe. (Department codes `DD, AR, AF, NV`)

- OPM, in its official publications, does not disclose the location (at the state level) of employees of several security-related agencies, other than that they may be working in the general Washington D.C. area. Because LEHD tabulations rely on sub-state geography, we have excluded these agencies from the QWI-OPM universe (see Table 10.5).

- The Federal Bureau of Investigation (FBI) provided Status file data to OPM until FY 2007, but did not provide Dynamics file data. Because we rely on the Dynamics file for critical timing information, all FBI-provided information is excluded from the universe for all periods. (Agency code: `DJ02`)

---

2. Homeland Security Act, November 2002, see http://www.dhs.gov/xabout/history/editorial_0133.shtm

- The State Department no longer provides Status and Dynamics file data on workers in the Foreign Service after 2006Q2, and those workers are thus out of scope from that point onwards (drop in employment). Other State Department employees, however, are included.

10.1.2 Integration Methodology

10.1.2.1 Creating wage-record like files from OPM

In this section, we describe how the wage record files are created that match the structure of the UI files used in the rest of the LEHD system of files. In particular, we address creation of quarterly status snapshots (Section 10.1.2.2), adjustments of earnings (Section 10.1.2.4), and attaching detailed geographic and industry, consistent with QCEW coding, to duty stations (Section 10.1.2.5).

10.1.2.2 Selecting records

We first exclude data that are out-of-universe, suffer from known data quality issues, or other issues, as noted earlier (see Tables 10.4 and 10.5 as well as release notes). Once exclusions have been processed, a first pass through the data creates a "pseudo-UI" wage record: a record for any job (employment relation with an agency) during a quarter.

10.1.2.3 Adjusting records

We parse the dynamics files for each quarter, and for each employee, an additional determination is made as to whether the employee was employed by some agency during the quarter. Because the status file records only active jobs in the last pay period of the quarter, jobs that end within a quarter, but were active at the start of the quarter, are not reflected in the status file - in contrast to UI wage records, where separations are reflected as wage records in the quarter they occur. Some additional adjustments also occur.

Impute missing dynamics records to improve consistency of dynamics and status file. If the status file reports an individual is on unpaid status in-between consecutive accessions or recall actions reported on the dynamics file, a separation is imputed to have occurred either in the middle of the quarter, or halfway in-between the last accession and the end of the quarter, whichever is later. Conversely, if the status file reports an individual is on paid status in-between consecutive separation actions reported on the dynamics file, an appointment is generated in the middle of the quarter, or halfway in-between the last separation and the end of the quarter, whichever is later.

Adjusting earnings. Earnings are computed from `totpay`, which comprise basic pay, locality adjustments, supervisory differential, retention allowance, as well as COLA. `totpay` reflects a full-time, full-year equivalent salary, not actual earnings. When accessions or separations occur within a quarter, the exact number of days in the quarter that a worker was in pay status is computed, and stored as a fraction. Earnings are adjusted at a later stage (after processing of seasonal, intermittent, and part-time workers, see next section).

10.1.2.4 Processing seasonal, intermittent, and part-time workers

OPM data identify whether a worker is part-time (PT), seasonally, or intermittently (SI) employed. However, the data do not identify the actual number of hours during any time period, and do not directly identify the number of days worked during a quarter. About 6.2% of workers are classified as either part-time, seasonally, or intermittently employed.

Because the standard UI data records exclusively record the actual earnings (or wage-like payments) received by a worker, an adjustment is required.[3] In order to make OPM data compatible with UI earnings concept, we adjust the earnings for non-fulltime workers. For seasonally or intermittently employed workers, we compute the exact dates when a worker is in pay status throughout a calendar-year quarter from the dynamics file, and adjust the full-time, full-year equivalent earnings accordingly. For instance, if a worker was employed on a seasonal basis, and worked until April 15, then entered a non-pay status, returning to pay-status again on June 15, then the adjustment ratio for that worker is 30/91, and 33% of the full-time, full-year equivalent pay for that quarter is recorded as earnings on the pseudo-UI record.[4]

For workers identified as part-time workers, we impute hours worked during a quarter. Using a model based on CPS and Decennial Census (DC) data, we condition on demographic characteristics, job characteristics, (employment history information: number of employers) and draw 10 imputes of hours worked. Actual quarterly earnings are then computed by multiplying the full-time, full-year pay rate by the ratio of imputed hours to potential quarterly hours.

10.1.2.5   Attaching geographic and industry classification to OPM records

OPM records do not provide industrial classification (NAICS) of the agency's activity. Furthermore, geographic precision of the agency's location is limited to the city the workers dutystation is located in. While in principle, the QWI tabulations only require county-level precision, OnTheMap requires roof-top precision. In order to compensate for these two data shortcomings, the OPM agencies are matched to their corresponding QCEW reports. However, the OPM and QCEW share no common identification variables. We probabilistically match the two universes by name, higher-level geography, and other attributes to obtain a correspondence between the two sets of identifiers (OPM agency ID, and SEIN on the QCEW records. The following section describes how that matching is done. In an ideal world, with consistent naming and high-level geography (county) on both OPM and QCEW records, this would be a straightforward exercise with little if any uncertainty. Unfortunately, real world data is not ideal. Across the 50 states and the District of Columbia, the 543 agencies known to OPM expand to 647 name variations on QCEW. Some agencies are not reported in the QCEW with positive employment in the same county that OPM personnel records show active employees' duty stations. Others, in particular military bases, are recorded in different counties altogether. We describe our attempts at addressing them below.

10.1.2.6   Selecting agency/establishment records

As mentioned, some agencies are excluded from the LEHD-OPM universe (see Section 10.1.2.2). These agencies must also be excluded as possible match candidates from the QCEW (see Table 10.5). We establish a master exclusion list, based on standardized OPM names. Agencies' records in the QCEW do not necessarily have time consistency or spatial consistency. Put differently, there are substantial variations in names of agencies and departments within any given state's historical QCEW records, and significant variations across states for the same agency. Thus, name matches are not exact, and probabilistic matching is used. [5] Records where the reliability of the probabilistic match is too low are clerically edited. All matches are databased, and are re-used for subsequent quarterly processing.

We note that in regular processing, the matches obtained (and validated) in previous periods are re-used. A database is maintained of all historical matches (SEIN and SEINUNIT assigned to specific agencies in each county), and a lookup against this database is performed prior to all name-matching. If the lookup is successful, and the establishment still exists in the corresponding QCEW data for that quarter, then no further matching is attempted.[6]

---

3. Note, however, that UI wage records typically contain no information on hours, weeks, or days worked, and no information on part-time or intermittent status.

4. Note that we assume that seasonal or intermittent workers are full-time when they do work.

5. The Census Bureau uses SAS Data Quality Server for this purpose.

6. Under certain conditions, the establishment need not exist in current QCEW records for the lookup result to be deemed valid.

An additional complication is that the level of detail (multi-unit breakouts) may differ between QCEW and OPM: QCEW may show more or fewer establishments in a particular county than OPM, for different states and different time periods.

The end result of this step are two lists: OPM agencies that are in-scope and define the universe that will be used as a baseline in all subsequent processing and reporting, and QCEW establishment records that are approximately equivalent in terms of the covered entities. The next step then consists in finding the QCEW reported establishment that corresponds to each reported agency, thus acquiring the QCEW record's industry classification and roof-top geocodes.

10.1.2.7 Name-based matching to obtain SEIN and SEINUNIT

Because there are no common identifiers in the QCEW and OPM firm-level datasets, we resort to a probabilistic matching strategy that relies on exact and fuzzy matches by name, size, and location. To this end, we aggregate the compatible pseudo-UI records created in a previous step up to the department-agency-county level. These records, which are equivalent to the establishment-level records in the QCEW, are then matched to in-scope QCEW records, using a variety of matching criteria. The exact matching strategy is outlined in Table 10.7. Detailed results are available in a separate document.

The Duty Station code contains the city and county. We retain the county as a matching and blocking criterion. In general, we attempt to find the QCEW agency within the same county, but in some instances, there are persistent mismatches between OPM and QCEW counts within county, and the county criterion is relaxed, allowing agencies to match in other counties.

If the algorithm fails to find a likely establishment, a firm-level identifier (SEIN) is attached based on an impute that takes into account the employment-weighted distribution of establishments within the state (match passes 81-96, 4.8 percent). Note that for these records, the standard Unit-to-Worker Impute (U2W) algorithm in the LEHD infrastructure (Abowd et al. 2009) will be used to perform a probabilistic allocation of workplaces, and thus industry and geography, to jobs.

Finally, about 4.2 percent of matches are made during clerical review, either by visual inspection of candidate records, or through a series of custom algorithmic edits.

10.1.2.8 Demographic information

OPM records contain demographic information for OPM workers. The information is provided each quarter, and the underlying personnel records are updated based on certain triggers. For the purposes of LEHD processing, the following data elements are required:

- Gender

- Age

- Race

- Ethnicity

- Education

However, some records are either incomplete (have missing information) or are not coded to the Office of Management and Budget (OMB) standard that LEHD uses currently (in particular, race and ethnicity from federal employees who were hired in their most recent position before 2006). We complete the information using two types of imputation models:

1. for variables with very few missing records, we derive the likelihood for the imputation model from the observed values for all ever-observed OPM workers, and construct a non-Bayesian draw from the likelihood;

2. for variables with high-levels of missingness, essentially race and ethnicity, we use a likelihood derived from the private-sector population data underlying the ICF, derived from 2000 Decennial information. The imputation conditions on reported non-OMB-compliant race and ethnicity, among other items, and jointly imputes OMB-compliant race and ethnicity combinations, and thus is better described as a probabilistic recode than a pure impute. For workers with a recent personnel action that involved re-declaring race and ethnicity, no impute is necessary.

Impute rates are generally low, with the exception of OMB-compliant race and ethnicity:

Variable	Impute rate
Education (collapsed)	$< 5\%$
Gender	$< 0.01\%$
Date of birth	$< 0.01\%$
Race and ethnicity	$< 61\%$

10.1.3 Changes in this Snapshot

LEHD has been working on integrating OPM data on Federal workers. The current efforts have been contributed to the Snapshot. The value-added to these data are labelled "beta". Data available will complement the EHF, ECF, ICF, U2W, the new JHF, and the QWI SEINUNIT-level file, in direct analogies of the existing file structures. RDC users should be able to access these files by requesting a "OPM" dataset. Access to the OPM data do not require state permissions.

10.2 DATA CITATION

> U.S. Census Bureau. 2014. *Individual Characteristics Files (ICF) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

10.3 DATA SET DESCRIPTIONS

10.3.1 Naming scheme

The OPM naming scheme is somewhat inconsistent with the remaining infrastructure files, reflecting the early-access nature of the data files. All files start with `opm`, followed by `_us` as the geographic indicator, and then the generic name of the file that is being supplied. This differs slightly from the naming conventions of UI-derived Infrastructure files. For instance, the EHF for OPM records is called:

`opm_us_ehf.sas7bdat`

and not `ehf_opm.sas7bdat`, and the ECF SEINUNIT file is called

`opm_us_ecf_seinunit.sas7bdat`

and not `ecf_opm_seinunit.sas7bdat`. Future snapshots should have a consistent naming convention. However, all files are compatible in structure to the regular UI-based EHF. In what follows, EHF, ECF, etc. are referred to as "processes". SAS datasets with zero observations are attached to this document:[7]

- **ecf**/opm/opm_us_ecf_sein_aux.sas7bdat

- **ecf**/opm/opm_us_ecf_sein.sas7bdat

- **ecf**/opm/opm_us_ecf_seinunit_aux.sas7bdat

- **ecf**/opm/opm_us_ecf_seinunit.sas7bdat

- **ecft26**/opm/opm_us_ecf_t26.sas7bdat

- **ehf**/opm/opm_us_ehf.sas7bdat

- **ehf**/opm/opm_us_ehf_phf.sas7bdat

- **ehf**/opm/opm_us_jhf.sas7bdat

- **ehf**/opm/opm_us_ehf_controltotals.sas7bdat

- **ehf**/opm/opm_us_ehf_sein_employment.sas7bdat

- **ehf**/opm/opm_us_ehf_shf.sas7bdat

- **ehf**/opm/opm_us_ehf_uhf.sas7bdat

- **ehf**/opm/opm_us_ehf_uniqpik.sas7bdat

- **icf**/opm/opm_us_icf_aux.sas7bdat

- **icf**/opm/opm_us_icf.sas7bdat

- **qwi**/opm/opm_us_qwi_seinunit_rh.sas7bdat

- **qwi**/opm/opm_us_qwi_seinunit_se.sas7bdat

- **qwi**/opm/opm_us_qwi_seinunit_wia.sas7bdat

- **u2w**/opm/opm_us_u2w.sas7bdat

---

7. Also visible on the attachment tab - Adobe Reader may be required.

10.3.2   Data location

The core files are stored in a `opm` sub-directories of the process-specific directories, similar to how state-specific files are stored

```
{PROCESS}/opm/
```

Some California-sourced QCEW data from the OPM-ECF files are covered under Title 26, and can be found in

```
ecft26/opm/
```

On the RDC network, the directories can be found under

```
/mixed/lehd/s2011
```

10.3.3   Available processes

Generally, OPM records are identified by the same variables (PIK, SEIN, SEINUNIT) as UI-wage-record derived records. SEIN and SEINUNIT have been attached to the records by a name-and-address match to QCEW files. A mapping of SEIN to agency identifiers is available on demand.

10.3.3.1   ECF

The OPM-ECF is processed the same way as the regular ECF, based on "pseudo-UI" records and matched QCEW records. Only establishments for federal agencies are in-scope to have a record in the OPM-ECF. The current version of OPM does not, however, include the firm-age and firm-size (or agency-age and agency-size) variables that were computed for the ECF. OPM entities that link to California records on the QCEW are available under Title 26 permissions, the same way as for the regular ECF. For more details, see Chapter 5.

10.3.3.2   EHF

The OPM-EHF is designed to store the complete federal government work history, for each individual that appears in the OPM source records, subject to the coverage limitations of OPM (see Table 10.5). The data has been restructured to resemble UI wage records. This involves some coarsening of the data (OPM source records contain the exact termination date of a job, whereas UI wage records do not). Otherwise, the structure of all EHF-like files is identical to those provided by the EHF process, and are documented in Chapter 6 (this includes the JHF). The establishment identifier corresponds to the SEIN-SEINUNIT found when matching to the QCEW.

10.3.3.3   ICF

While the structure of the OPM-ICF is identical to the ICF documented in Chapter 9, none of the contents are derived from the data sources noted there. Rather, all demographic information is derived entirely from the OPM-provided input files. Section 10.1.2.8 describes how any missing data is filled in, and how race in particular is standardized across the two different coding schemata present in the OPM data. Due to the low incidence of missingness in the data, only a single implicate was generated for any imputes; thus, no implicate files are provided. In addition to the standard ICF, an auxiliary file with variables not usually present on the ICF is provided, see Section 10.3.4.1

10.3.3.4   QWI-SEINUNIT files

Establishment-level files from the QWI series are provided, and are identical in structure to the files described in Chapter 11. Note that the SEIN-SEINUNIT identifier is used, not the OPM agency identifier.

10.3.3.5   U2W

Because there is some mismatch between establishment-level reports in the QCEW and the equivalent agency-within-county or agency-within-city reporting in the duty station file provided by OPM, the Unit-to-Worker Impute process is used to allocated to QCEW-reported establishments when a unique allocation is not feasible. The structure of the file is otherwise identical to the file(s) described in Chapter 14.

10.3.4 Dataset documentation on files unique to the OPM process

10.3.4.1 Auxiliary dataset for ICF: OPM_us_ICF_aux

This contains variables not otherwise present on the ICF, derived from OPM information. No effort has been made to fill in any missing information, the file is provided as-is to researchers.

Record identifier PIK

Sort order PIK

Entity PIK

Unique Entity Key PIK

Field name	Data dictionary reference name	Starting position	Field size	Data type
BEST_DOB has a value	FLAG_BEST_DOB	00165	3	N
BEST_EDLVL has a value	FLAG_BEST_EDLVL	00162	3	N
BEST_ERIRACE has a value	FLAG_BEST_ERIRACE	00159	3	N
BEST_RACE has a value	FLAG_BEST_RACE	00156	3	N
CITIZEN is missing a value	MISSING_CITIZEN	00080	8	N
CREDITMILSRV is missing a value	MISSING_CREDITMILSRV	00120	8	N
Creditable Military Service	OPM_CREDITMILSRV	00196	4	A/N
DEGREEYR is missing a value	MISSING_DEGREEYR	00088	8	N
DOB has a value	FLAG_DOB	00032	8	N
DOB is missing a value	MISSING_DOB	00040	8	N
Date of Arrival (SAS date value)	DOA	00152	4	N
Date of Birth	OPM_DOB	00183	6	A/N
EDLVL has a value	FLAG_EDLVL	00024	8	N
EDLVL is missing a value	MISSING_EDLVL	00048	8	N
EDUC_C has a value	FLAG_EDUC_C	00174	3	N
ERIRACE has a value	FLAG_ERIRACE	00008	8	N
ERIRACE is missing a value	MISSING_ERIRACE	00064	8	N
Education Level	OPM_EDLVL	00200	2	A/N
Either ERIRACE or RACE have a value	FLAG_ANYRACE	00168	3	N
Either best_erirace or best_RACE have a value	MISSING_ANYRACE	00144	8	N
Ethnicity/race ind(eri)	OPM_ERIRACE	00214	6	A/N
FROZENSERV is missing a value	MISSING_FROZENSERV	00112	8	N
Foreign-born status	FB	00245	1	A/N
Frozen Service	OPM_FROZENSERV	00202	4	A/N
HANDICAP is missing a value	MISSING_HANDICAP	00136	8	N
Handicap	OPM_HANDICAP	00206	2	A/N
Highest grade achieved	EDUC_F	00171	3	N
OCCUPATION is missing a value	MISSING_OCCUPATION	00096	8	N
OPM (OMB-compliant) erirace: bested	BEST_ERIRACE	00230	6	A/N
OPM (old-style) race: bested	BEST_RACE	00229	1	A/N
OPM Date of Birth (YYYYMM): bested	BEST_DOB	00238	6	A/N
OPM edlvl: bested	BEST_EDLVL	00236	2	A/N
Occupation	OPM_OCCUPATION	00189	4	A/N
Occupational Category(PATCO)	OPM_PATCO	00193	1	A/N
PATCO is missing a value	MISSING_PATCO	00104	8	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
PCF-compliant (not sourced) race information	PCF_RACE	00244	1	A/N
PIK	PIK	00220	9	A/N
RACE has a value	FLAG_RACE	00016	8	N
RACE is missing a value	MISSING_RACE	00072	8	N
RACE_ETHNICITY has a value	FLAG_RACE_ETHNICITY	00180	3	N
Race or National Orgin	OPM_RACE	00194	1	A/N
SEX has a value	FLAG_SEX	00177	3	N
SEX is missing a value	MISSING_SEX	00056	8	N
Source of OPM record (qtime) run	OPM_SRC	00000	8	N
US Citizenship	OPM_CITIZEN	00208	1	A/N
VSTATUS is missing a value	MISSING_VSTATUS	00128	8	N
Veterans Status	OPM_VSTATUS	00209	1	A/N
Year of Degree	OPM_DEGREEYR	00210	4	A/N
sex	OPM_SEX	00195	1	A/N

10.3.5 Summary information on datasets

Table 10.2: Number of observations for OPM

Group	Number of datafiles	Records (1000s)	Filesize (GB)
OPM	21	43,091	16

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

The exception to the provenance description in Section 1.8 are the OPM files, which stem from an experimental pre-production process, and had not been assigned unique identifiers at the time of S2011 data preparation.

Table 10.3: List of data files for OPM, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
National (us)					
opm_us_ecf_cc	2000Q1	2011Q4	1,752	< 1	1:opmbeta2
opm_us_ecf_leg	2000Q1	2011Q4	1,752	< 1	1:opmbeta2
opm_us_ecf_sein	2000Q1	2011Q4	3	< 1	1:opmbeta2
opm_us_ecf_sein_aux	2000Q1	2011Q4	3	< 1	1:opmbeta2
opm_us_ecf_seinunit	2000Q1	2011Q4	1,752	< 1	1:opmbeta2
opm_us_ecf_seinunit_aux	2000Q1	2011Q4	1,752	1	1:opmbeta2
opm_us_ecf_u2w_break	2000Q1	2011Q4	< 1	< 1	1:opmbeta2
opm_us_ehf	2000Q1	2011Q4	15,263	1	1:opmbeta2
opm_us_ehf_controltotals	2000Q1	2011Q4	9	< 1	1:opmbeta2
opm_us_ehf_phf	2000Q1	2011Q4	4,107	3	1:opmbeta2
opm_us_ehf_sein_employment	2000Q1	2011Q4	92	< 1	1:opmbeta2
opm_us_ehf_shf	2000Q1	2011Q4	10	< 1	1:opmbeta2
opm_us_ehf_uhf	2000Q1	2011Q4	143	< 1	1:opmbeta2
opm_us_ehf_uniqpik	2000Q1	2011Q4	2,672	< 1	1:opmbeta2
opm_us_icf	2000Q1	2011Q4	4,063	< 1	1:opmbeta2
opm_us_icf_aux	2000Q1	2011Q4	4,063	1	1:opmbeta2
opm_us_jhf	2000Q1	2011Q4	4,112	1	1:opmbeta2
opm_us_qwi_seinunit_rh	2000Q1	2011Q4	458	3	1:opmbeta2
opm_us_qwi_seinunit_se	2000Q1	2011Q4	457	2	1:opmbeta2
opm_us_qwi_seinunit_wia	2000Q1	2011Q4	458	3	1:opmbeta2
opm_us_u2w	2000Q1	2011Q4	169	< 1	1:opmbeta2

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

10.4   NOTES

10.5   TABLES

Table 10.4: Non-reporting agencies

Notable non-reporting agencies to OPM

Some Executive Branch agencies:

- Office of the President and of the Vice President

- White House

- Office of Policy Development

- multiple Department of Defense agencies, including military personnel

- most Intelligence agencies

- Board of Governors of the Federal Reserve

- Tennessee Valley Authority

- State Department Regional Personnel Centers

- U.S. Postal Service

- Other smaller federal agencies or commissions.

Most of the Legislative Branch, such as:

- Congress

- Congressional Budget Office

- Government Accountability Office

- Library of Congress

- Office of Compliance

Judicial branch completely, including

- Supreme Court

- U.S. Courts

Table 10.5: Exclusions from federal worker universe

Excluded from OPM because of missing geography

- DJ02-FEDERAL BUREAU OF INVESTIGATION

- DJ06-DRUG ENFORCEMENT ADMINISTRATION

- DJ15-BUREAU OF ALCOHOL, TOBACCO, FIREARMS, AND EXPLO-SIVES (ATF)

- HSAD-U.S. SECRET SERVICE

- TR40-ALCOHOL AND TOBACCO TAX AND TRADE BUREAU

- TRAC-U.S. SECRET SERVICE

- TRAD-U.S. MINT

Table 10.6: Employment in agencies that do not report geography

Agency	DC area	Other US
DJ02	10468	10468
DJ02	23018	23018
DJ06	2269	2269
DJ06	7354	7354
DJ15	1118	1118
DJ15	4001	4001
HSAD	3948	3948
HSAD	2854	2854
TR40	159	159
TR40	357	357
TRAD	364	364
TRAD	1429	1429

Note: Data from public-use OPM, 2009Q4.

Table 10.7: Matching stragegy

Match pass	Match type	Match description	Blocking vars	Matching vars	
Automated matches					
1	DQ95F5	Automated Matches on agency names within dept/county, fuzzy employment	state county dept_full	agency aggemp_post(95)	(95)
2	DQ85F5	Automated Matches on agency names within dept/county, fuzzy employment	state county dept_full	agency aggemp_post(95)	(85)
3	DQ95F15	Automated Matches on agency names within dept/county, fuzzy employment	state county dept_full	agency aggemp_post(85)	(95)
4	DQ85F15	Automated Matches on agency names within dept/county, fuzzy employment	state county dept_full	agency aggemp_post(85)	(85)
5	DQ95F25	Automated Matches on agency names within dept/county, fuzzy employment	state county dept_full	agency aggemp_post(75)	(95)
6	DQ95	Automated Matches on agency names within dept/county, no employment	state county dept_full	agency (95)	
7	UNIQ1	Uniques Finds agencies that are the only agency for DEPT within COUNTY	DEPT COUNTY count(agency)=1	n.a.	
Matches subject to clerical review					
61	SEIN F5	PreClerical Find SEINUNITs with assumed SEIN from x-county editing, ignore agency name	state county dept_full sein	aggemp_post(95)	
62	SEIN F15	PreClerical Find SEINUNITs with assumed SEIN from x-county editing, ignore agency name	state county dept_full sein	aggemp_post(85)	
63	DQ95	PreClerical Relax department, match on agency name alone	state county	agency(95)	
64	DQ95F5	PreClerical Ignore agency name, use fuzzy employment and dept	state county dept_full	aggemp_post(95)	
65	DQ95F15	PreClerical Ignore agency name, use fuzzy employment and dept	state county dept_full	aggemp_post(85)	
66	DQ95F25	PreClerical Ignore agency name, use fuzzy employment and dept	state county dept_full	aggemp_post(85)	
67	DQ95F5	PreClerical Ignore county, use fuzzy employment and agency	state dept_full	agency(95) aggemp_post(95)	
68	DQ95F15	PreClerical Ignore county, use fuzzy employment and agency	state dept_full	agency(95) aggemp_post(85)	
71-78		PreClerical Repeat above sequence, but allow for re-matches	see above	see above	
SEIN Imputes					
81	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state county dept_full sizeclass	random (employment distribution)	
82	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state county dept_full	random (employment distribution)	
83	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state dept_full sizeclass	random (employment distribution)	
84	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state dept_full	random (employment distribution)	
85	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state county	random (employment distribution)	
86	IMP	PreImpute Impute SEIN (not SEINUNIT) based on conditional distributions These units will need to go through U2W!!	state	random (employment distribution)	
91-96		PreImpute Repeat above sequence, but allow for re-matches			

Table 10.8: Fedscope availability, by year and quarter

Year	Q1	Q2	Q3	Q4
1998			x	
1999			x	
2000			x	
2001			x	
2002			x	
2003			x	
2004			x	
2005			x	
2006			x	
2007			x	x
2008	x	x	x	x
2009	x	x	x	x
2010	x	x	x	x
2011	x	x	x	

Table 10.9: DHS Reorganization 2003

Original Agency	Original Department	Current Agency or Office (in DHS)
U.S. Customs Service	Treasury	U.S. Customs and Border Protection
		U.S. Immigration and Customs Enforcement
Immigration and Naturalization Service	Justice	U.S. Customs and Border Protection
		U.S. Immigration and Customs Enforcement
		U.S. Citizenship and Immigration Services
Federal Protective Service	General Services Administration (GSA)	National Protection and Programs Directorate
Transportation Security Administration	Transportation	Transportation Security Administration
Federal Law Enforcement Training Center	Treasury	Federal Law Enforcement Training Center
Animal and Plant Health Inspection Service (part)	Agriculture	U.S. Customs and Border Protection
Federal Emergency Management Agency (FEMA)	none	Federal Emergency Management Agency
Office for Domestic Preparedness	Justice	Responsibilities distributed within FEMA
Strategic National Stockpile, National Disaster Medical System	Health and Human Services (HHS)	Returned to HHS, July, 2004
Nuclear Incident Response Team	Energy	
Domestic Emergency Support Teams	Justice	Responsibilities distributed within FEMA
National Domestic Preparedness Office	FBI	
CBRN Countermeasures Programs	Energy	
Environmental Measurements Laboratory	Energy	Science & Technology Directorate
National Biological Warfare, Defense Analysis Center	Defense	
Plum Island Animal Disease Center	Agriculture	
Federal Computer Incident Response Center	GSA	US-CERT, Office of Cybersecurity and Communications
		National Protection and Programs Directorate
National Communications System	Defense	Office of Cybersecurity and Communications
		National Protection and Programs Directorate
National Infrastructure Protection Center	FBI	Office of Operations Coordination
		Office of Infrastructure Protection
Energy Security and Assurance Program	Energy	Office of Infrastructure Protection
U.S. Coast Guard	Transportation	U.S. Coast Guard
U.S. Secret Service	Treasury	U.S. Secret Service

Source: https://en.wikipedia.org/wiki/United_States_Department_of_Homeland_Security, accessed 2012-04-17, and http://www.dhs.gov/xabout/history/editorial_0133.shtm

Chapter 11.
Quarterly Workforce Indicators - SEINUNIT file (QWI)

11.1 OVERVIEW

The Quarterly Workforce Indicators (QWI) establishment file contains quarterly measures of workforce composition and worker turnover at the establishment level. The LEHD establishment-level measures are created from longitudinally integrated person and establishment-level data. Establishment-level measures include: (i) Worker and Job Flows: accessions, separations, job creation, job destruction by age and gender of workforce; (ii) Worker composition by gender and age, (iii) Worker compensation for stocks and flows by gender and age; (iv) Dynamic worker compensation summary statistics for stocks and flows by gender and age. The QWI may be used in combination with the Chapter 3LBD Bridge or the Chapter 5ECF to match to other Census micro business databases, and can be linked by firm-establishment identifiers to other LEHD Infrastructure files.

11.1.1 Changes in this Snapshot

The QWI_SEINUNIT files (internally known as UFF_B) have been expanded. Each file contains the statistics known from the public-use QWI, for each interaction of demographic characteristics. Prior to S2011, only the "WIA" tabulations were available, and the files were simply called "QWI_SEINUNIT". With the release of race, ethnicity, and education tabulations, two additional files have been created, and one file modified:

- QWI_SEINUNIT_WIA is the new name of the previously available file for *age x sex* statistics

- QWI_SEINUNIT_RH contains the same statistics for *race x ethnicity* groups

- QWI_SEINUNIT_SE contains the same statistics for *sex x education* groups

In addition, for the convenience of researchers, a smaller file containing only the marginal categories (i.e., no breakouts by specific groups) was created, as QWI_SEINUNIT_estabtots.

Note that the use of the QWI_SEINUNIT files is incompatible with the use of the QWI public-use files also now part of the S2011 snapshot. Researchers must choose one or the other.

Further note that since release R2013Q2 of the public-use QWI, the shorthand for demographic characteristics "sex-age" has changed from `WIA` to `SA`. This is *not* reflected in the S2011 snapshot, which is based on earlier data.

11.2 DATA CITATION

> U.S. Census Bureau. 2014. *Office of Personal Management (OPM) files in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

11.3 DATA SET DESCRIPTIONS

11.3.1 Coverage of QWI

QWI data are available for all states that are LED-state partners, however, not every state is currently a LED-state partner. The QWI are built upon wage records in the UI system and information from state ES-202 data. The universe of QWI data is UI-covered earnings. UI coverage is broad, covering over 90% of total wage and salary civilian jobs.

When QWI private industry employment numbers are compared with other employment data, exclusions to UI coverage should be taken into account. Federal government employment is not generally included. Exempted employment varies slightly from state to state due to variations in state unemployment laws, but generally also excludes many farmers and agricultural employees, domestic workers, self-employed non-agricultural workers, members of the Armed Services, some state and local government employees as well as certain types of nonprofit employers and religious organizations (which are given a choice of coverage or noncoverage in a number of states). See Stevens (2007) for a more detailed discussion.

11.3.2 Naming scheme

SAS datasets with zero observations are attached to this document:[1]

- qwi/qwi_zz_seinunit_wia.sas7bdat

- qwi/qwi_zz_seinunit_rh.sas7bdat

- qwi/qwi_zz_seinunit_se.sas7bdat

- qwi/qwi_zz_seinunit_estabtot.sas7bdat

`ZZ` stands for the state postal abbreviation, and `YYYY` for a calendar year. `_wia` identifies the sex-age tabulations, `_se` the sex-education tabulations, and `_rh` the race-ethnicity tabulations.

11.3.3 Data location

The files are stored in a main directory, with state-specific subdirectories:

`qwi/ZZ/`

On the RDC network, the directory can be found under

`/mixed/lehd/s2011`

---

1. Also visible on the attachment tab - Adobe Reader may be required.

11.3.4 Main dataset: QWI_ZZ_SEINUNIT

The `QWI_ZZ_SEINUNIT_AGG` file (LEHD internal name: UFFb) is a file at the SEINUNIT level, providing detailed statistics for an establishment (SEIN + SEINUNIT) at every combination (AGG) of (WIA) SEX x AGEGROUP or (SE) SEX x EDUCGROUP or (RH) RACE x ETHNICITY. Age groups are defined using the WIA categorization. The different margins are represented as variable arrays in the UFFb. Due to the very large number of variables, we only list exhaustively the variables for the WIA file. Zero-obs datasets are attached to this PDF for all files. The generic variable name is constructed as <`STATISTIC`>_<`MARGIN1`><`MARGIN2`> where <STATISTIC> is one of the statistics described in Abowd et al. (2009), and the values for the margins are taken from two "legal" combinations of codes from Table 11.1 on page 11-4 (taken from Vilhuber and Schmutte 2012). Thus, `A_1A02` are accessions A for $sex = 1$ (men) of $agegroup = A02$ (ages 19-21) (on file `qwi_zz_seinunit_wia`), whereas `S_R3H2` are separations S for Asians ($race = R3$) of non-Hispanic ethnicity ($ethnicity = H2$) (on file `qwi_zz_seinunit_rh`).

Table 11.1: QWI coding

QWI coding: Sex

sex	sexfm
0	Male and Female
1	Male
2	Female

QWI coding: Age groups

agegrp	agegrpfm
A00	14-99
A01	14-18
A02	19-21
A03	22-24
A04	25-34
A05	35-44
A06	45-54
A07	55-64
A08	65-99

QWI coding: Education groups

education	educationfm
E0	All Education Groups Together (1-4)
E1	Less than high school
E2	High school or equivalent, no college
E3	Some college or Associate degree
E4	Bachelor's degree or advanced degree

QWI coding: Ethnicity

ethnicity	ethnicityfm
H0	All (Any Ethnicity)
H1	Hispanic or Latino
H2	Not Hispanic or Latino

QWI coding: Race

race	racefm
R0	All Races Together (1-6)
R1	American Indian or Alaska Native Alone
R2	Asian Alone
R3	Black or African American Alone
R4	Native Hawaiian or Other Pacific Islander Alone
R5	White Alone
R6	Two or More Race Groups

Record identifier YEAR QUARTER SEIN SEINUNIT

Sort order YEAR QUARTER SEIN SEINUNIT

Entity Establishment

Unique Entity Key SEIN SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
=0 from ECF_SEIN, =1 if from ECF_SEINUNIT, =z not found	UNIT_DETAIL_FLAG	05787	1	A/N
Accessions for Female and age 14-18	A_2A01	00548	4	N
Accessions for Female and age 14-99	A_2A00	00544	4	N
Accessions for Female and age 19-21	A_2A02	00552	4	N
Accessions for Female and age 22-24	A_2A03	00556	4	N
Accessions for Female and age 25-34	A_2A04	00560	4	N
Accessions for Female and age 35-44	A_2A05	00564	4	N
Accessions for Female and age 45-54	A_2A06	00568	4	N
Accessions for Female and age 55-64	A_2A07	00572	4	N
Accessions for Female and age 65-99	A_2A08	00576	4	N
Accessions for Male and Female and age 14-18	A_0A01	00476	4	N
Accessions for Male and Female and age 14-99	A_0A00	00472	4	N
Accessions for Male and Female and age 19-21	A_0A02	00480	4	N
Accessions for Male and Female and age 22-24	A_0A03	00484	4	N
Accessions for Male and Female and age 25-34	A_0A04	00488	4	N
Accessions for Male and Female and age 35-44	A_0A05	00492	4	N
Accessions for Male and Female and age 45-54	A_0A06	00496	4	N
Accessions for Male and Female and age 55-64	A_0A07	00500	4	N
Accessions for Male and Female and age 65-99	A_0A08	00504	4	N
Accessions for Male and age 14-18	A_1A01	00512	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
	A_1A00			
Accessions for Male and age 14-99		00508	4	N
	A_1A02			
Accessions for Male and age 19-21		00516	4	N
	A_1A03			
Accessions for Male and age 22-24		00520	4	N
	A_1A04			
Accessions for Male and age 25-34		00524	4	N
	A_1A05			
Accessions for Male and age 35-44		00528	4	N
	A_1A06			
Accessions for Male and age 45-54		00532	4	N
	A_1A07			
Accessions for Male and age 55-64		00536	4	N
	A_1A08			
Accessions for Male and age 65-99		00540	4	N
	BDOT_2A01			
Alternate definition of B that does not reflect flow suppression for Female and age 14-18		05516	4	N
	BDOT_2A00			
Alternate definition of B that does not reflect flow suppression for Female and age 14-99		05512	4	N
	BDOT_2A02			
Alternate definition of B that does not reflect flow suppression for Female and age 19-21		05520	4	N
	BDOT_2A03			
Alternate definition of B that does not reflect flow suppression for Female and age 22-24		05524	4	N
	BDOT_2A04			
Alternate definition of B that does not reflect flow suppression for Female and age 25-34		05528	4	N
	BDOT_2A05			
Alternate definition of B that does not reflect flow suppression for Female and age 35-44		05532	4	N
	BDOT_2A06			
Alternate definition of B that does not reflect flow suppression for Female and age 45-54		05536	4	N
	BDOT_2A07			
Alternate definition of B that does not reflect flow suppression for Female and age 55-64		05540	4	N
	BDOT_2A08			
Alternate definition of B that does not reflect flow suppression for Female and age 65-99		05544	4	N
	BDOT_0A01			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 14-18		05444	4	N
	BDOT_0A00			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Alternate definition of B that does not reflect flow suppression for Male and Female and age 14-99		05440	4	N
	BDOT_0A02			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 19-21		05448	4	N
	BDOT_0A03			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 22-24		05452	4	N
	BDOT_0A04			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 25-34		05456	4	N
	BDOT_0A05			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 35-44		05460	4	N
	BDOT_0A06			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 45-54		05464	4	N
	BDOT_0A07			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 55-64		05468	4	N
	BDOT_0A08			
Alternate definition of B that does not reflect flow suppression for Male and Female and age 65-99		05472	4	N
	BDOT_1A01			
Alternate definition of B that does not reflect flow suppression for Male and age 14-18		05480	4	N
	BDOT_1A00			
Alternate definition of B that does not reflect flow suppression for Male and age 14-99		05476	4	N
	BDOT_1A02			
Alternate definition of B that does not reflect flow suppression for Male and age 19-21		05484	4	N
	BDOT_1A03			
Alternate definition of B that does not reflect flow suppression for Male and age 22-24		05488	4	N
	BDOT_1A04			
Alternate definition of B that does not reflect flow suppression for Male and age 25-34		05492	4	N
	BDOT_1A05			
Alternate definition of B that does not reflect flow suppression for Male and age 35-44		05496	4	N
	BDOT_1A06			
Alternate definition of B that does not reflect flow suppression for Male and age 45-54		05500	4	N
	BDOT_1A07			
Alternate definition of B that does not reflect flow suppression for Male and age 55-64		05504	4	N
	BDOT_1A08			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Alternate definition of B that does not reflect flow suppression for Male and age 65-99		05508	4	N
	EDOT_2A01			
Alternate definition of E that does not reflect flow suppression for Female and age 14-18		05624	4	N
	EDOT_2A00			
Alternate definition of E that does not reflect flow suppression for Female and age 14-99		05620	4	N
	EDOT_2A02			
Alternate definition of E that does not reflect flow suppression for Female and age 19-21		05628	4	N
	EDOT_2A03			
Alternate definition of E that does not reflect flow suppression for Female and age 22-24		05632	4	N
	EDOT_2A04			
Alternate definition of E that does not reflect flow suppression for Female and age 25-34		05636	4	N
	EDOT_2A05			
Alternate definition of E that does not reflect flow suppression for Female and age 35-44		05640	4	N
	EDOT_2A06			
Alternate definition of E that does not reflect flow suppression for Female and age 45-54		05644	4	N
	EDOT_2A07			
Alternate definition of E that does not reflect flow suppression for Female and age 55-64		05648	4	N
	EDOT_2A08			
Alternate definition of E that does not reflect flow suppression for Female and age 65-99		05652	4	N
	EDOT_0A01			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 14-18		05552	4	N
	EDOT_0A00			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 14-99		05548	4	N
	EDOT_0A02			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 19-21		05556	4	N
	EDOT_0A03			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 22-24		05560	4	N
	EDOT_0A04			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 25-34		05564	4	N
	EDOT_0A05			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 35-44		05568	4	N
	EDOT_0A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Alternate definition of E that does not reflect flow suppression for Male and Female and age 45-54		05572	4	N
	EDOT_0A07			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 55-64		05576	4	N
	EDOT_0A08			
Alternate definition of E that does not reflect flow suppression for Male and Female and age 65-99		05580	4	N
	EDOT_1A01			
Alternate definition of E that does not reflect flow suppression for Male and age 14-18		05588	4	N
	EDOT_1A00			
Alternate definition of E that does not reflect flow suppression for Male and age 14-99		05584	4	N
	EDOT_1A02			
Alternate definition of E that does not reflect flow suppression for Male and age 19-21		05592	4	N
	EDOT_1A03			
Alternate definition of E that does not reflect flow suppression for Male and age 22-24		05596	4	N
	EDOT_1A04			
Alternate definition of E that does not reflect flow suppression for Male and age 25-34		05600	4	N
	EDOT_1A05			
Alternate definition of E that does not reflect flow suppression for Male and age 35-44		05604	4	N
	EDOT_1A06			
Alternate definition of E that does not reflect flow suppression for Male and age 45-54		05608	4	N
	EDOT_1A07			
Alternate definition of E that does not reflect flow suppression for Male and age 55-64		05612	4	N
	EDOT_1A08			
Alternate definition of E that does not reflect flow suppression for Male and age 65-99		05616	4	N
	FDOT_2A01			
Alternate definition of F that does not reflect flow suppression for Female and age 14-18		05732	4	N
	FDOT_2A00			
Alternate definition of F that does not reflect flow suppression for Female and age 14-99		05728	4	N
	FDOT_2A02			
Alternate definition of F that does not reflect flow suppression for Female and age 19-21		05736	4	N
	FDOT_2A03			
Alternate definition of F that does not reflect flow suppression for Female and age 22-24		05740	4	N
	FDOT_2A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Alternate definition of F that does not reflect flow suppression for Female and age 25-34		05744	4	N
	FDOT_2A05			
Alternate definition of F that does not reflect flow suppression for Female and age 35-44		05748	4	N
	FDOT_2A06			
Alternate definition of F that does not reflect flow suppression for Female and age 45-54		05752	4	N
	FDOT_2A07			
Alternate definition of F that does not reflect flow suppression for Female and age 55-64		05756	4	N
	FDOT_2A08			
Alternate definition of F that does not reflect flow suppression for Female and age 65-99		05760	4	N
	FDOT_0A01			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 14-18		05660	4	N
	FDOT_0A00			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 14-99		05656	4	N
	FDOT_0A02			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 19-21		05664	4	N
	FDOT_0A03			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 22-24		05668	4	N
	FDOT_0A04			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 25-34		05672	4	N
	FDOT_0A05			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 35-44		05676	4	N
	FDOT_0A06			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 45-54		05680	4	N
	FDOT_0A07			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 55-64		05684	4	N
	FDOT_0A08			
Alternate definition of F that does not reflect flow suppression for Male and Female and age 65-99		05688	4	N
	FDOT_1A01			
Alternate definition of F that does not reflect flow suppression for Male and age 14-18		05696	4	N
	FDOT_1A00			
Alternate definition of F that does not reflect flow suppression for Male and age 14-99		05692	4	N
	FDOT_1A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Alternate definition of F that does not reflect flow suppression for Male and age 19-21		05700	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 22-24	FDOT_1A03	05704	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 25-34	FDOT_1A04	05708	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 35-44	FDOT_1A05	05712	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 45-54	FDOT_1A06	05716	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 55-64	FDOT_1A07	05720	4	N
Alternate definition of F that does not reflect flow suppression for Male and age 65-99	FDOT_1A08	05724	4	N
Average accession rate for Female and age 14-18	AR_2A01	05084	4	N
Average accession rate for Female and age 14-99	AR_2A00	05080	4	N
Average accession rate for Female and age 19-21	AR_2A02	05088	4	N
Average accession rate for Female and age 22-24	AR_2A03	05092	4	N
Average accession rate for Female and age 25-34	AR_2A04	05096	4	N
Average accession rate for Female and age 35-44	AR_2A05	05100	4	N
Average accession rate for Female and age 45-54	AR_2A06	05104	4	N
Average accession rate for Female and age 55-64	AR_2A07	05108	4	N
Average accession rate for Female and age 65-99	AR_2A08	05112	4	N
Average accession rate for Male and Female and age 14-18	AR_0A01	05012	4	N
Average accession rate for Male and Female and age 14-99	AR_0A00	05008	4	N
Average accession rate for Male and Female and age 19-21	AR_0A02	05016	4	N
	AR_0A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average accession rate for Male and Female and age 22-24	AR_0A04	05020	4	N
Average accession rate for Male and Female and age 25-34	AR_0A05	05024	4	N
Average accession rate for Male and Female and age 35-44	AR_0A06	05028	4	N
Average accession rate for Male and Female and age 45-54	AR_0A07	05032	4	N
Average accession rate for Male and Female and age 55-64	AR_0A08	05036	4	N
Average accession rate for Male and Female and age 65-99	AR_1A01	05040	4	N
Average accession rate for Male and age 14-18	AR_1A00	05048	4	N
Average accession rate for Male and age 14-99	AR_1A02	05044	4	N
Average accession rate for Male and age 19-21	AR_1A03	05052	4	N
Average accession rate for Male and age 22-24	AR_1A04	05056	4	N
Average accession rate for Male and age 25-34	AR_1A05	05060	4	N
Average accession rate for Male and age 35-44	AR_1A06	05064	4	N
Average accession rate for Male and age 45-54	AR_1A07	05068	4	N
Average accession rate for Male and age 55-64	AR_1A08	05072	4	N
Average accession rate for Male and age 65-99	Ebar_2A01	05076	4	N
Average employment for Female and age 14-18	Ebar_2A00	03680	4	N
Average employment for Female and age 14-99	Ebar_2A02	03676	4	N
Average employment for Female and age 19-21	Ebar_2A03	03684	4	N
Average employment for Female and age 22-24	Ebar_2A04	03688	4	N
Average employment for Female and age 25-34	Ebar_2A05	03692	4	N
Average employment for Female and age 35-44	Ebar_2A06	03696	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average employment for Female and age 45-54 EBAR_2A07		03700	4	N
Average employment for Female and age 55-64 EBAR_2A08		03704	4	N
Average employment for Female and age 65-99 EBAR_0A01		03708	4	N
Average employment for Male and Female and age 14-18 EBAR_0A00		03608	4	N
Average employment for Male and Female and age 14-99 EBAR_0A02		03604	4	N
Average employment for Male and Female and age 19-21 EBAR_0A03		03612	4	N
Average employment for Male and Female and age 22-24 EBAR_0A04		03616	4	N
Average employment for Male and Female and age 25-34 EBAR_0A05		03620	4	N
Average employment for Male and Female and age 35-44 EBAR_0A06		03624	4	N
Average employment for Male and Female and age 45-54 EBAR_0A07		03628	4	N
Average employment for Male and Female and age 55-64 EBAR_0A08		03632	4	N
Average employment for Male and Female and age 65-99 EBAR_1A01		03636	4	N
Average employment for Male and age 14-18 EBAR_1A00		03644	4	N
Average employment for Male and age 14-99 EBAR_1A02		03640	4	N
Average employment for Male and age 19-21 EBAR_1A03		03648	4	N
Average employment for Male and age 22-24 EBAR_1A04		03652	4	N
Average employment for Male and age 25-34 EBAR_1A05		03656	4	N
Average employment for Male and age 35-44 EBAR_1A06		03660	4	N
Average employment for Male and age 45-54 EBAR_1A07		03664	4	N
Average employment for Male and age 55-64		03668	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average employment for Male and age 65-99	Ebar_1A08	03672	4	N
Average full-quarter employment for Female and age 14-18	Fbar_2A01	04328	4	N
Average full-quarter employment for Female and age 14-99	Fbar_2A00	04324	4	N
Average full-quarter employment for Female and age 19-21	Fbar_2A02	04332	4	N
Average full-quarter employment for Female and age 22-24	Fbar_2A03	04336	4	N
Average full-quarter employment for Female and age 25-34	Fbar_2A04	04340	4	N
Average full-quarter employment for Female and age 35-44	Fbar_2A05	04344	4	N
Average full-quarter employment for Female and age 45-54	Fbar_2A06	04348	4	N
Average full-quarter employment for Female and age 55-64	Fbar_2A07	04352	4	N
Average full-quarter employment for Female and age 65-99	Fbar_2A08	04356	4	N
Average full-quarter employment for Male and Female and age 14-18	Fbar_0A01	04256	4	N
Average full-quarter employment for Male and Female and age 14-99	Fbar_0A00	04252	4	N
Average full-quarter employment for Male and Female and age 19-21	Fbar_0A02	04260	4	N
Average full-quarter employment for Male and Female and age 22-24	Fbar_0A03	04264	4	N
Average full-quarter employment for Male and Female and age 25-34	Fbar_0A04	04268	4	N
Average full-quarter employment for Male and Female and age 35-44	Fbar_0A05	04272	4	N
	Fbar_0A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter employment for Male and Female and age 45-54		04276	4	N
	Fʙᴀʀ_0A07			
Average full-quarter employment for Male and Female and age 55-64		04280	4	N
	Fʙᴀʀ_0A08			
Average full-quarter employment for Male and Female and age 65-99		04284	4	N
	Fʙᴀʀ_1A01			
Average full-quarter employment for Male and age 14-18		04292	4	N
	Fʙᴀʀ_1A00			
Average full-quarter employment for Male and age 14-99		04288	4	N
	Fʙᴀʀ_1A02			
Average full-quarter employment for Male and age 19-21		04296	4	N
	Fʙᴀʀ_1A03			
Average full-quarter employment for Male and age 22-24		04300	4	N
	Fʙᴀʀ_1A04			
Average full-quarter employment for Male and age 25-34		04304	4	N
	Fʙᴀʀ_1A05			
Average full-quarter employment for Male and age 35-44		04308	4	N
	Fʙᴀʀ_1A06			
Average full-quarter employment for Male and age 45-54		04312	4	N
	Fʙᴀʀ_1A07			
Average full-quarter employment for Male and age 55-64		04316	4	N
	Fʙᴀʀ_1A08			
Average full-quarter employment for Male and age 65-99		04320	4	N
	FG_2A01			
Average full-quarter employment growth rate for Female and age 14-18		04544	4	N
	FG_2A00			
Average full-quarter employment growth rate for Female and age 14-99		04540	4	N
	FG_2A02			
Average full-quarter employment growth rate for Female and age 19-21		04548	4	N
	FG_2A03			
Average full-quarter employment growth rate for Female and age 22-24		04552	4	N
	FG_2A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter employment growth rate for Female and age 25-34		04556	4	N
	FG_2A05			
Average full-quarter employment growth rate for Female and age 35-44		04560	4	N
	FG_2A06			
Average full-quarter employment growth rate for Female and age 45-54		04564	4	N
	FG_2A07			
Average full-quarter employment growth rate for Female and age 55-64		04568	4	N
	FG_2A08			
Average full-quarter employment growth rate for Female and age 65-99		04572	4	N
	FG_0A01			
Average full-quarter employment growth rate for Male and Female and age 14-18		04472	4	N
	FG_0A00			
Average full-quarter employment growth rate for Male and Female and age 14-99		04468	4	N
	FG_0A02			
Average full-quarter employment growth rate for Male and Female and age 19-21		04476	4	N
	FG_0A03			
Average full-quarter employment growth rate for Male and Female and age 22-24		04480	4	N
	FG_0A04			
Average full-quarter employment growth rate for Male and Female and age 25-34		04484	4	N
	FG_0A05			
Average full-quarter employment growth rate for Male and Female and age 35-44		04488	4	N
	FG_0A06			
Average full-quarter employment growth rate for Male and Female and age 45-54		04492	4	N
	FG_0A07			
Average full-quarter employment growth rate for Male and Female and age 55-64		04496	4	N
	FG_0A08			
Average full-quarter employment growth rate for Male and Female and age 65-99		04500	4	N
	FG_1A01			
Average full-quarter employment growth rate for Male and age 14-18		04508	4	N
	FG_1A00			
Average full-quarter employment growth rate for Male and age 14-99		04504	4	N
	FG_1A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter employment growth rate for Male and age 19-21		04512	4	N
	FG_1A03			
Average full-quarter employment growth rate for Male and age 22-24		04516	4	N
	FG_1A04			
Average full-quarter employment growth rate for Male and age 25-34		04520	4	N
	FG_1A05			
Average full-quarter employment growth rate for Male and age 35-44		04524	4	N
	FG_1A06			
Average full-quarter employment growth rate for Male and age 45-54		04528	4	N
	FG_1A07			
Average full-quarter employment growth rate for Male and age 55-64		04532	4	N
	FG_1A08			
Average full-quarter employment growth rate for Male and age 65-99		04536	4	N
	FJCR_2A01			
Average full-quarter job creation rate for Female and age 14-18		04760	4	N
	FJCR_2A00			
Average full-quarter job creation rate for Female and age 14-99		04756	4	N
	FJCR_2A02			
Average full-quarter job creation rate for Female and age 19-21		04764	4	N
	FJCR_2A03			
Average full-quarter job creation rate for Female and age 22-24		04768	4	N
	FJCR_2A04			
Average full-quarter job creation rate for Female and age 25-34		04772	4	N
	FJCR_2A05			
Average full-quarter job creation rate for Female and age 35-44		04776	4	N
	FJCR_2A06			
Average full-quarter job creation rate for Female and age 45-54		04780	4	N
	FJCR_2A07			
Average full-quarter job creation rate for Female and age 55-64		04784	4	N
	FJCR_2A08			
Average full-quarter job creation rate for Female and age 65-99		04788	4	N
	FJCR_0A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter job creation rate for Male and Female and age 14-18		04688	4	N
	FJCR_0A00			
Average full-quarter job creation rate for Male and Female and age 14-99		04684	4	N
	FJCR_0A02			
Average full-quarter job creation rate for Male and Female and age 19-21		04692	4	N
	FJCR_0A03			
Average full-quarter job creation rate for Male and Female and age 22-24		04696	4	N
	FJCR_0A04			
Average full-quarter job creation rate for Male and Female and age 25-34		04700	4	N
	FJCR_0A05			
Average full-quarter job creation rate for Male and Female and age 35-44		04704	4	N
	FJCR_0A06			
Average full-quarter job creation rate for Male and Female and age 45-54		04708	4	N
	FJCR_0A07			
Average full-quarter job creation rate for Male and Female and age 55-64		04712	4	N
	FJCR_0A08			
Average full-quarter job creation rate for Male and Female and age 65-99		04716	4	N
	FJCR_1A01			
Average full-quarter job creation rate for Male and age 14-18		04724	4	N
	FJCR_1A00			
Average full-quarter job creation rate for Male and age 14-99		04720	4	N
	FJCR_1A02			
Average full-quarter job creation rate for Male and age 19-21		04728	4	N
	FJCR_1A03			
Average full-quarter job creation rate for Male and age 22-24		04732	4	N
	FJCR_1A04			
Average full-quarter job creation rate for Male and age 25-34		04736	4	N
	FJCR_1A05			
Average full-quarter job creation rate for Male and age 35-44		04740	4	N
	FJCR_1A06			
Average full-quarter job creation rate for Male and age 45-54		04744	4	N
	FJCR_1A07			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter job creation rate for Male and age 55-64		04748	4	N
	FJCR_1A08			
Average full-quarter job creation rate for Male and age 65-99		04752	4	N
	FJDR_2A01			
Average full-quarter job destruction rate for Female and age 14-18		04976	4	N
	FJDR_2A00			
Average full-quarter job destruction rate for Female and age 14-99		04972	4	N
	FJDR_2A02			
Average full-quarter job destruction rate for Female and age 19-21		04980	4	N
	FJDR_2A03			
Average full-quarter job destruction rate for Female and age 22-24		04984	4	N
	FJDR_2A04			
Average full-quarter job destruction rate for Female and age 25-34		04988	4	N
	FJDR_2A05			
Average full-quarter job destruction rate for Female and age 35-44		04992	4	N
	FJDR_2A06			
Average full-quarter job destruction rate for Female and age 45-54		04996	4	N
	FJDR_2A07			
Average full-quarter job destruction rate for Female and age 55-64		05000	4	N
	FJDR_2A08			
Average full-quarter job destruction rate for Female and age 65-99		05004	4	N
	FJDR_0A01			
Average full-quarter job destruction rate for Male and Female and age 14-18		04904	4	N
	FJDR_0A00			
Average full-quarter job destruction rate for Male and Female and age 14-99		04900	4	N
	FJDR_0A02			
Average full-quarter job destruction rate for Male and Female and age 19-21		04908	4	N
	FJDR_0A03			
Average full-quarter job destruction rate for Male and Female and age 22-24		04912	4	N
	FJDR_0A04			
Average full-quarter job destruction rate for Male and Female and age 25-34		04916	4	N
	FJDR_0A05			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average full-quarter job destruction rate for Male and Female and age 35-44		04920	4	N
	FJDR_0A06			
Average full-quarter job destruction rate for Male and Female and age 45-54		04924	4	N
	FJDR_0A07			
Average full-quarter job destruction rate for Male and Female and age 55-64		04928	4	N
	FJDR_0A08			
Average full-quarter job destruction rate for Male and Female and age 65-99		04932	4	N
	FJDR_1A01			
Average full-quarter job destruction rate for Male and age 14-18		04940	4	N
	FJDR_1A00			
Average full-quarter job destruction rate for Male and age 14-99		04936	4	N
	FJDR_1A02			
Average full-quarter job destruction rate for Male and age 19-21		04944	4	N
	FJDR_1A03			
Average full-quarter job destruction rate for Male and age 22-24		04948	4	N
	FJDR_1A04			
Average full-quarter job destruction rate for Male and age 25-34		04952	4	N
	FJDR_1A05			
Average full-quarter job destruction rate for Male and age 35-44		04956	4	N
	FJDR_1A06			
Average full-quarter job destruction rate for Male and age 45-54		04960	4	N
	FJDR_1A07			
Average full-quarter job destruction rate for Male and age 55-64		04964	4	N
	FJDR_1A08			
Average full-quarter job destruction rate for Male and age 65-99		04968	4	N
	JCR_2A01			
Average job creation rate for Female and age 14-18		04004	4	N
	JCR_2A00			
Average job creation rate for Female and age 14-99		04000	4	N
	JCR_2A02			
Average job creation rate for Female and age 19-21		04008	4	N
	JCR_2A03			
Average job creation rate for Female and age 22-24		04012	4	N
	JCR_2A04			
Average job creation rate for Female and age 25-34		04016	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average job creation rate for Female and age 35-44	JCR_2A05	04020	4	N
Average job creation rate for Female and age 45-54	JCR_2A06	04024	4	N
Average job creation rate for Female and age 55-64	JCR_2A07	04028	4	N
Average job creation rate for Female and age 65-99	JCR_2A08	04032	4	N
Average job creation rate for Male and Female and age 14-18	JCR_0A01	03932	4	N
Average job creation rate for Male and Female and age 14-99	JCR_0A00	03928	4	N
Average job creation rate for Male and Female and age 19-21	JCR_0A02	03936	4	N
Average job creation rate for Male and Female and age 22-24	JCR_0A03	03940	4	N
Average job creation rate for Male and Female and age 25-34	JCR_0A04	03944	4	N
Average job creation rate for Male and Female and age 35-44	JCR_0A05	03948	4	N
Average job creation rate for Male and Female and age 45-54	JCR_0A06	03952	4	N
Average job creation rate for Male and Female and age 55-64	JCR_0A07	03956	4	N
Average job creation rate for Male and Female and age 65-99	JCR_0A08	03960	4	N
Average job creation rate for Male and age 14-18	JCR_1A01	03968	4	N
Average job creation rate for Male and age 14-99	JCR_1A00	03964	4	N
Average job creation rate for Male and age 19-21	JCR_1A02	03972	4	N
Average job creation rate for Male and age 22-24	JCR_1A03	03976	4	N
Average job creation rate for Male and age 25-34	JCR_1A04	03980	4	N
Average job creation rate for Male and age 35-44	JCR_1A05	03984	4	N
	JCR_1A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average job creation rate for Male and age 45-54	JCR_1A07	03988	4	N
Average job creation rate for Male and age 55-64	JCR_1A08	03992	4	N
Average job creation rate for Male and age 65-99	JDR_2A01	03996	4	N
Average job destruction rate for Female and age 14-18	JDR_2A00	04220	4	N
Average job destruction rate for Female and age 14-99	JDR_2A02	04216	4	N
Average job destruction rate for Female and age 19-21	JDR_2A03	04224	4	N
Average job destruction rate for Female and age 22-24	JDR_2A04	04228	4	N
Average job destruction rate for Female and age 25-34	JDR_2A05	04232	4	N
Average job destruction rate for Female and age 35-44	JDR_2A06	04236	4	N
Average job destruction rate for Female and age 45-54	JDR_2A07	04240	4	N
Average job destruction rate for Female and age 55-64	JDR_2A08	04244	4	N
Average job destruction rate for Female and age 65-99	JDR_0A01	04248	4	N
Average job destruction rate for Male and Female and age 14-18	JDR_0A00	04148	4	N
Average job destruction rate for Male and Female and age 14-99	JDR_0A02	04144	4	N
Average job destruction rate for Male and Female and age 19-21	JDR_0A03	04152	4	N
Average job destruction rate for Male and Female and age 22-24	JDR_0A04	04156	4	N
Average job destruction rate for Male and Female and age 25-34	JDR_0A05	04160	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average job destruction rate for Male and Female and age 35-44		04164	4	N
	JDR_0A06			
Average job destruction rate for Male and Female and age 45-54		04168	4	N
	JDR_0A07			
Average job destruction rate for Male and Female and age 55-64		04172	4	N
	JDR_0A08			
Average job destruction rate for Male and Female and age 65-99		04176	4	N
	JDR_1A01			
Average job destruction rate for Male and age 14-18		04184	4	N
	JDR_1A00			
Average job destruction rate for Male and age 14-99		04180	4	N
	JDR_1A02			
Average job destruction rate for Male and age 19-21		04188	4	N
	JDR_1A03			
Average job destruction rate for Male and age 22-24		04192	4	N
	JDR_1A04			
Average job destruction rate for Male and age 25-34		04196	4	N
	JDR_1A05			
Average job destruction rate for Male and age 35-44		04200	4	N
	JDR_1A06			
Average job destruction rate for Male and age 45-54		04204	4	N
	JDR_1A07			
Average job destruction rate for Male and age 55-64		04208	4	N
	JDR_1A08			
Average job destruction rate for Male and age 65-99		04212	4	N
	FAR_2A01			
Average rate of flow into full-quarter employment for Female and age 14-18		05300	4	N
	FAR_2A00			
Average rate of flow into full-quarter employment for Female and age 14-99		05296	4	N
	FAR_2A02			
Average rate of flow into full-quarter employment for Female and age 19-21		05304	4	N
	FAR_2A03			
Average rate of flow into full-quarter employment for Female and age 22-24		05308	4	N
	FAR_2A04			
Average rate of flow into full-quarter employment for Female and age 25-34		05312	4	N
	FAR_2A05			
Average rate of flow into full-quarter employment for Female and age 35-44		05316	4	N
	FAR_2A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average rate of flow into full-quarter employment for Female and age 45-54	FAR_2A07	05320	4	N
Average rate of flow into full-quarter employment for Female and age 55-64	FAR_2A08	05324	4	N
Average rate of flow into full-quarter employment for Female and age 65-99	FAR_0A01	05328	4	N
Average rate of flow into full-quarter employment for Male and Female and age 14-18	FAR_0A00	05228	4	N
Average rate of flow into full-quarter employment for Male and Female and age 14-99	FAR_0A02	05224	4	N
Average rate of flow into full-quarter employment for Male and Female and age 19-21	FAR_0A03	05232	4	N
Average rate of flow into full-quarter employment for Male and Female and age 22-24	FAR_0A04	05236	4	N
Average rate of flow into full-quarter employment for Male and Female and age 25-34	FAR_0A05	05240	4	N
Average rate of flow into full-quarter employment for Male and Female and age 35-44	FAR_0A06	05244	4	N
Average rate of flow into full-quarter employment for Male and Female and age 45-54	FAR_0A07	05248	4	N
Average rate of flow into full-quarter employment for Male and Female and age 55-64	FAR_0A08	05252	4	N
Average rate of flow into full-quarter employment for Male and Female and age 65-99	FAR_1A01	05256	4	N
Average rate of flow into full-quarter employment for Male and age 14-18	FAR_1A00	05264	4	N
Average rate of flow into full-quarter employment for Male and age 14-99	FAR_1A02	05260	4	N
Average rate of flow into full-quarter employment for Male and age 19-21	FAR_1A03	05268	4	N
Average rate of flow into full-quarter employment for Male and age 22-24	FAR_1A04	05272	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average rate of flow into full-quarter employment for Male and age 25-34		05276	4	N
	FAR_1A05			
Average rate of flow into full-quarter employment for Male and age 35-44		05280	4	N
	FAR_1A06			
Average rate of flow into full-quarter employment for Male and age 45-54		05284	4	N
	FAR_1A07			
Average rate of flow into full-quarter employment for Male and age 55-64		05288	4	N
	FAR_1A08			
Average rate of flow into full-quarter employment for Male and age 65-99		05292	4	N
	FSR_2A01			
Average rate of flow out of full-quarter employment for Female and age 14-18		05408	4	N
	FSR_2A00			
Average rate of flow out of full-quarter employment for Female and age 14-99		05404	4	N
	FSR_2A02			
Average rate of flow out of full-quarter employment for Female and age 19-21		05412	4	N
	FSR_2A03			
Average rate of flow out of full-quarter employment for Female and age 22-24		05416	4	N
	FSR_2A04			
Average rate of flow out of full-quarter employment for Female and age 25-34		05420	4	N
	FSR_2A05			
Average rate of flow out of full-quarter employment for Female and age 35-44		05424	4	N
	FSR_2A06			
Average rate of flow out of full-quarter employment for Female and age 45-54		05428	4	N
	FSR_2A07			
Average rate of flow out of full-quarter employment for Female and age 55-64		05432	4	N
	FSR_2A08			
Average rate of flow out of full-quarter employment for Female and age 65-99		05436	4	N
	FSR_0A01			
Average rate of flow out of full-quarter employment for Male and Female and age 14-18		05336	4	N
	FSR_0A00			
Average rate of flow out of full-quarter employment for Male and Female and age 14-99		05332	4	N
	FSR_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average rate of flow out of full-quarter employment for Male and Female and age 19-21	FSR_0A03	05340	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 22-24	FSR_0A04	05344	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 25-34	FSR_0A05	05348	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 35-44	FSR_0A06	05352	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 45-54	FSR_0A07	05356	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 55-64	FSR_0A08	05360	4	N
Average rate of flow out of full-quarter employment for Male and Female and age 65-99	FSR_1A01	05364	4	N
Average rate of flow out of full-quarter employment for Male and age 14-18	FSR_1A00	05372	4	N
Average rate of flow out of full-quarter employment for Male and age 14-99	FSR_1A02	05368	4	N
Average rate of flow out of full-quarter employment for Male and age 19-21	FSR_1A03	05376	4	N
Average rate of flow out of full-quarter employment for Male and age 22-24	FSR_1A04	05380	4	N
Average rate of flow out of full-quarter employment for Male and age 25-34	FSR_1A05	05384	4	N
Average rate of flow out of full-quarter employment for Male and age 35-44	FSR_1A06	05388	4	N
Average rate of flow out of full-quarter employment for Male and age 45-54	FSR_1A07	05392	4	N
Average rate of flow out of full-quarter employment for Male and age 55-64	FSR_1A08	05396	4	N
Average rate of flow out of full-quarter employment for Male and age 65-99	SR_2A01	05400	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average separation rate for Female and age 14-18		05192	4	N
	SR_2A00			
Average separation rate for Female and age 14-99		05188	4	N
	SR_2A02			
Average separation rate for Female and age 19-21		05196	4	N
	SR_2A03			
Average separation rate for Female and age 22-24		05200	4	N
	SR_2A04			
Average separation rate for Female and age 25-34		05204	4	N
	SR_2A05			
Average separation rate for Female and age 35-44		05208	4	N
	SR_2A06			
Average separation rate for Female and age 45-54		05212	4	N
	SR_2A07			
Average separation rate for Female and age 55-64		05216	4	N
	SR_2A08			
Average separation rate for Female and age 65-99		05220	4	N
	SR_0A01			
Average separation rate for Male and Female and age 14-18		05120	4	N
	SR_0A00			
Average separation rate for Male and Female and age 14-99		05116	4	N
	SR_0A02			
Average separation rate for Male and Female and age 19-21		05124	4	N
	SR_0A03			
Average separation rate for Male and Female and age 22-24		05128	4	N
	SR_0A04			
Average separation rate for Male and Female and age 25-34		05132	4	N
	SR_0A05			
Average separation rate for Male and Female and age 35-44		05136	4	N
	SR_0A06			
Average separation rate for Male and Female and age 45-54		05140	4	N
	SR_0A07			
Average separation rate for Male and Female and age 55-64		05144	4	N
	SR_0A08			
Average separation rate for Male and Female and age 65-99		05148	4	N
	SR_1A01			
Average separation rate for Male and age 14-18		05156	4	N
	SR_1A00			
Average separation rate for Male and age 14-99		05152	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Average separation rate for Male and age 19-21	SR_1A02	05160	4	N
Average separation rate for Male and age 22-24	SR_1A03	05164	4	N
Average separation rate for Male and age 25-34	SR_1A04	05168	4	N
Average separation rate for Male and age 35-44	SR_1A05	05172	4	N
Average separation rate for Male and age 45-54	SR_1A06	05176	4	N
Average separation rate for Male and age 55-64	SR_1A07	05180	4	N
Average separation rate for Male and age 65-99	SR_1A08	05184	4	N
Beginning-of-period employment for Female and age 14-18	B_2A01	00116	4	N
Beginning-of-period employment for Female and age 14-99	B_2A00	00112	4	N
Beginning-of-period employment for Female and age 19-21	B_2A02	00120	4	N
Beginning-of-period employment for Female and age 22-24	B_2A03	00124	4	N
Beginning-of-period employment for Female and age 25-34	B_2A04	00128	4	N
Beginning-of-period employment for Female and age 35-44	B_2A05	00132	4	N
Beginning-of-period employment for Female and age 45-54	B_2A06	00136	4	N
Beginning-of-period employment for Female and age 55-64	B_2A07	00140	4	N
Beginning-of-period employment for Female and age 65-99	B_2A08	00144	4	N
Beginning-of-period employment for Male and Female and age 14-18	B_0A01	00044	4	N
Beginning-of-period employment for Male and Female and age 14-99	B_0A00	00040	4	N
	B_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Beginning-of-period employment for Male and Female and age 19-21		00048	4	N
	B_0A03			
Beginning-of-period employment for Male and Female and age 22-24		00052	4	N
	B_0A04			
Beginning-of-period employment for Male and Female and age 25-34		00056	4	N
	B_0A05			
Beginning-of-period employment for Male and Female and age 35-44		00060	4	N
	B_0A06			
Beginning-of-period employment for Male and Female and age 45-54		00064	4	N
	B_0A07			
Beginning-of-period employment for Male and Female and age 55-64		00068	4	N
	B_0A08			
Beginning-of-period employment for Male and Female and age 65-99		00072	4	N
	B_1A01			
Beginning-of-period employment for Male and age 14-18		00080	4	N
	B_1A00			
Beginning-of-period employment for Male and age 14-99		00076	4	N
	B_1A02			
Beginning-of-period employment for Male and age 19-21		00084	4	N
	B_1A03			
Beginning-of-period employment for Male and age 22-24		00088	4	N
	B_1A04			
Beginning-of-period employment for Male and age 25-34		00092	4	N
	B_1A05			
Beginning-of-period employment for Male and age 35-44		00096	4	N
	B_1A06			
Beginning-of-period employment for Male and age 45-54		00100	4	N
	B_1A07			
Beginning-of-period employment for Male and age 55-64		00104	4	N
	B_1A08			
Beginning-of-period employment for Male and age 65-99		00108	4	N
	FIRMAGE			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Best firm age		00008	8	N
	DWA_2A01			
Change in total earnings for accessions for Female and age 14-18		01412	4	N
	DWA_2A00			
Change in total earnings for accessions for Female and age 14-99		01408	4	N
	DWA_2A02			
Change in total earnings for accessions for Female and age 19-21		01416	4	N
	DWA_2A03			
Change in total earnings for accessions for Female and age 22-24		01420	4	N
	DWA_2A04			
Change in total earnings for accessions for Female and age 25-34		01424	4	N
	DWA_2A05			
Change in total earnings for accessions for Female and age 35-44		01428	4	N
	DWA_2A06			
Change in total earnings for accessions for Female and age 45-54		01432	4	N
	DWA_2A07			
Change in total earnings for accessions for Female and age 55-64		01436	4	N
	DWA_2A08			
Change in total earnings for accessions for Female and age 65-99		01440	4	N
	DWA_0A01			
Change in total earnings for accessions for Male and Female and age 14-18		01340	4	N
	DWA_0A00			
Change in total earnings for accessions for Male and Female and age 14-99		01336	4	N
	DWA_0A02			
Change in total earnings for accessions for Male and Female and age 19-21		01344	4	N
	DWA_0A03			
Change in total earnings for accessions for Male and Female and age 22-24		01348	4	N
	DWA_0A04			
Change in total earnings for accessions for Male and Female and age 25-34		01352	4	N
	DWA_0A05			
Change in total earnings for accessions for Male and Female and age 35-44		01356	4	N
	DWA_0A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for accessions for Male and Female and age 45-54		01360	4	N
	DWA_0A07			
Change in total earnings for accessions for Male and Female and age 55-64		01364	4	N
	DWA_0A08			
Change in total earnings for accessions for Male and Female and age 65-99		01368	4	N
	DWA_1A01			
Change in total earnings for accessions for Male and age 14-18		01376	4	N
	DWA_1A00			
Change in total earnings for accessions for Male and age 14-99		01372	4	N
	DWA_1A02			
Change in total earnings for accessions for Male and age 19-21		01380	4	N
	DWA_1A03			
Change in total earnings for accessions for Male and age 22-24		01384	4	N
	DWA_1A04			
Change in total earnings for accessions for Male and age 25-34		01388	4	N
	DWA_1A05			
Change in total earnings for accessions for Male and age 35-44		01392	4	N
	DWA_1A06			
Change in total earnings for accessions for Male and age 45-54		01396	4	N
	DWA_1A07			
Change in total earnings for accessions for Male and age 55-64		01400	4	N
	DWA_1A08			
Change in total earnings for accessions for Male and age 65-99		01404	4	N
	DWFS_2A01			
Change in total earnings for full-quarter separations for Female and age 14-18		02168	4	N
	DWFS_2A00			
Change in total earnings for full-quarter separations for Female and age 14-99		02164	4	N
	DWFS_2A02			
Change in total earnings for full-quarter separations for Female and age 19-21		02172	4	N
	DWFS_2A03			
Change in total earnings for full-quarter separations for Female and age 22-24		02176	4	N
	DWFS_2A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for full-quarter separations for Female and age 25-34	DWFS_2A05	02180	4	N
Change in total earnings for full-quarter separations for Female and age 35-44	DWFS_2A06	02184	4	N
Change in total earnings for full-quarter separations for Female and age 45-54	DWFS_2A07	02188	4	N
Change in total earnings for full-quarter separations for Female and age 55-64	DWFS_2A08	02192	4	N
Change in total earnings for full-quarter separations for Female and age 65-99	DWFS_0A01	02196	4	N
Change in total earnings for full-quarter separations for Male and Female and age 14-18	DWFS_0A00	02096	4	N
Change in total earnings for full-quarter separations for Male and Female and age 14-99	DWFS_0A02	02092	4	N
Change in total earnings for full-quarter separations for Male and Female and age 19-21	DWFS_0A03	02100	4	N
Change in total earnings for full-quarter separations for Male and Female and age 22-24	DWFS_0A04	02104	4	N
Change in total earnings for full-quarter separations for Male and Female and age 25-34	DWFS_0A05	02108	4	N
Change in total earnings for full-quarter separations for Male and Female and age 35-44	DWFS_0A06	02112	4	N
Change in total earnings for full-quarter separations for Male and Female and age 45-54	DWFS_0A07	02116	4	N
Change in total earnings for full-quarter separations for Male and Female and age 55-64	DWFS_0A08	02120	4	N
Change in total earnings for full-quarter separations for Male and Female and age 65-99	DWFS_1A01	02124	4	N
Change in total earnings for full-quarter separations for Male and age 14-18	DWFS_1A00	02132	4	N
Change in total earnings for full-quarter separations for Male and age 14-99	DWFS_1A02	02128	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for full-quarter separations for Male and age 19-21		02136	4	N
	DWFS_1A03			
Change in total earnings for full-quarter separations for Male and age 22-24		02140	4	N
	DWFS_1A04			
Change in total earnings for full-quarter separations for Male and age 25-34		02144	4	N
	DWFS_1A05			
Change in total earnings for full-quarter separations for Male and age 35-44		02148	4	N
	DWFS_1A06			
Change in total earnings for full-quarter separations for Male and age 45-54		02152	4	N
	DWFS_1A07			
Change in total earnings for full-quarter separations for Male and age 55-64		02156	4	N
	DWFS_1A08			
Change in total earnings for full-quarter separations for Male and age 65-99		02160	4	N
	DWS_2A01			
Change in total earnings for separations for Female and age 14-18		01952	4	N
	DWS_2A00			
Change in total earnings for separations for Female and age 14-99		01948	4	N
	DWS_2A02			
Change in total earnings for separations for Female and age 19-21		01956	4	N
	DWS_2A03			
Change in total earnings for separations for Female and age 22-24		01960	4	N
	DWS_2A04			
Change in total earnings for separations for Female and age 25-34		01964	4	N
	DWS_2A05			
Change in total earnings for separations for Female and age 35-44		01968	4	N
	DWS_2A06			
Change in total earnings for separations for Female and age 45-54		01972	4	N
	DWS_2A07			
Change in total earnings for separations for Female and age 55-64		01976	4	N
	DWS_2A08			
Change in total earnings for separations for Female and age 65-99		01980	4	N
	DWS_0A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for separations for Male and Female and age 14-18	DWS_0A00	01880	4	N
Change in total earnings for separations for Male and Female and age 14-99	DWS_0A02	01876	4	N
Change in total earnings for separations for Male and Female and age 19-21	DWS_0A03	01884	4	N
Change in total earnings for separations for Male and Female and age 22-24	DWS_0A04	01888	4	N
Change in total earnings for separations for Male and Female and age 25-34	DWS_0A05	01892	4	N
Change in total earnings for separations for Male and Female and age 35-44	DWS_0A06	01896	4	N
Change in total earnings for separations for Male and Female and age 45-54	DWS_0A07	01900	4	N
Change in total earnings for separations for Male and Female and age 55-64	DWS_0A08	01904	4	N
Change in total earnings for separations for Male and Female and age 65-99	DWS_1A01	01908	4	N
Change in total earnings for separations for Male and age 14-18	DWS_1A00	01916	4	N
Change in total earnings for separations for Male and age 14-99	DWS_1A02	01912	4	N
Change in total earnings for separations for Male and age 19-21	DWS_1A03	01920	4	N
Change in total earnings for separations for Male and age 22-24	DWS_1A04	01924	4	N
Change in total earnings for separations for Male and age 25-34	DWS_1A05	01928	4	N
Change in total earnings for separations for Male and age 35-44	DWS_1A06	01932	4	N
Change in total earnings for separations for Male and age 45-54	DWS_1A07	01936	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for separations for Male and age 55-64	DWS_1A08	01940	4	N
Change in total earnings for separations for Male and age 65-99	DWFA_2A01	01944	4	N
Change in total earnings for transits to full-quarter status for Female and age 14-18	DWFA_2A00	01628	4	N
Change in total earnings for transits to full-quarter status for Female and age 14-99	DWFA_2A02	01624	4	N
Change in total earnings for transits to full-quarter status for Female and age 19-21	DWFA_2A03	01632	4	N
Change in total earnings for transits to full-quarter status for Female and age 22-24	DWFA_2A04	01636	4	N
Change in total earnings for transits to full-quarter status for Female and age 25-34	DWFA_2A05	01640	4	N
Change in total earnings for transits to full-quarter status for Female and age 35-44	DWFA_2A06	01644	4	N
Change in total earnings for transits to full-quarter status for Female and age 45-54	DWFA_2A07	01648	4	N
Change in total earnings for transits to full-quarter status for Female and age 55-64	DWFA_2A08	01652	4	N
Change in total earnings for transits to full-quarter status for Female and age 65-99	DWFA_0A01	01656	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 14-18	DWFA_0A00	01556	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 14-99	DWFA_0A02	01552	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 19-21	DWFA_0A03	01560	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 22-24	DWFA_0A04	01564	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 25-34	DWFA_0A05	01568	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Change in total earnings for transits to full-quarter status for Male and Female and age 35-44	DWFA_0A06	01572	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 45-54	DWFA_0A07	01576	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 55-64	DWFA_0A08	01580	4	N
Change in total earnings for transits to full-quarter status for Male and Female and age 65-99	DWFA_1A01	01584	4	N
Change in total earnings for transits to full-quarter status for Male and age 14-18	DWFA_1A00	01592	4	N
Change in total earnings for transits to full-quarter status for Male and age 14-99	DWFA_1A02	01588	4	N
Change in total earnings for transits to full-quarter status for Male and age 19-21	DWFA_1A03	01596	4	N
Change in total earnings for transits to full-quarter status for Male and age 22-24	DWFA_1A04	01600	4	N
Change in total earnings for transits to full-quarter status for Male and age 25-34	DWFA_1A05	01604	4	N
Change in total earnings for transits to full-quarter status for Male and age 35-44	DWFA_1A06	01608	4	N
Change in total earnings for transits to full-quarter status for Male and age 45-54	DWFA_1A07	01612	4	N
Change in total earnings for transits to full-quarter status for Male and age 55-64	DWFA_1A08	01616	4	N
Change in total earnings for transits to full-quarter status for Male and age 65-99	LEG_COUNTY	01620	4	N
Cleaned GEO FIPS County CCC	ES_OWNER_CODE	05804	3	A/N
Cleaned OWNER_CODE O	ES_SIC	05811	1	A/N
Cleaned SIC Code IIII	ES_STATE	05807	4	A/N
ES202 FIPS State SS	M_2A01	05818	2	A/N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Employment any time during the period for Female and age 14-18		00332	4	N
	M_2A00			
Employment any time during the period for Female and age 14-99		00328	4	N
	M_2A02			
Employment any time during the period for Female and age 19-21		00336	4	N
	M_2A03			
Employment any time during the period for Female and age 22-24		00340	4	N
	M_2A04			
Employment any time during the period for Female and age 25-34		00344	4	N
	M_2A05			
Employment any time during the period for Female and age 35-44		00348	4	N
	M_2A06			
Employment any time during the period for Female and age 45-54		00352	4	N
	M_2A07			
Employment any time during the period for Female and age 55-64		00356	4	N
	M_2A08			
Employment any time during the period for Female and age 65-99		00360	4	N
	M_0A01			
Employment any time during the period for Male and Female and age 14-18		00260	4	N
	M_0A00			
Employment any time during the period for Male and Female and age 14-99		00256	4	N
	M_0A02			
Employment any time during the period for Male and Female and age 19-21		00264	4	N
	M_0A03			
Employment any time during the period for Male and Female and age 22-24		00268	4	N
	M_0A04			
Employment any time during the period for Male and Female and age 25-34		00272	4	N
	M_0A05			
Employment any time during the period for Male and Female and age 35-44		00276	4	N
	M_0A06			
Employment any time during the period for Male and Female and age 45-54		00280	4	N
	M_0A07			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Employment any time during the period for Male and Female and age 55-64	M_0A08	00284	4	N
Employment any time during the period for Male and Female and age 65-99	M_1A01	00288	4	N
Employment any time during the period for Male and age 14-18	M_1A00	00296	4	N
Employment any time during the period for Male and age 14-99	M_1A02	00292	4	N
Employment any time during the period for Male and age 19-21	M_1A03	00300	4	N
Employment any time during the period for Male and age 22-24	M_1A04	00304	4	N
Employment any time during the period for Male and age 25-34	M_1A05	00308	4	N
Employment any time during the period for Male and age 35-44	M_1A06	00312	4	N
Employment any time during the period for Male and age 45-54	M_1A07	00316	4	N
Employment any time during the period for Male and age 55-64	M_1A08	00320	4	N
Employment any time during the period for Male and age 65-99	E_2A01	00324	4	N
End-of-period employment for Female and age 14-18	E_2A00	00224	4	N
End-of-period employment for Female and age 14-99	E_2A02	00220	4	N
End-of-period employment for Female and age 19-21	E_2A03	00228	4	N
End-of-period employment for Female and age 22-24	E_2A04	00232	4	N
End-of-period employment for Female and age 25-34	E_2A05	00236	4	N
End-of-period employment for Female and age 35-44	E_2A06	00240	4	N
End-of-period employment for Female and age 45-54	E_2A07	00244	4	N
End-of-period employment for Female and age 55-64		00248	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
	E_2A08			
End-of-period employment for Female and age 65-99		00252	4	N
	E_0A01			
End-of-period employment for Male and Female and age 14-18		00152	4	N
	E_0A00			
End-of-period employment for Male and Female and age 14-99		00148	4	N
	E_0A02			
End-of-period employment for Male and Female and age 19-21		00156	4	N
	E_0A03			
End-of-period employment for Male and Female and age 22-24		00160	4	N
	E_0A04			
End-of-period employment for Male and Female and age 25-34		00164	4	N
	E_0A05			
End-of-period employment for Male and Female and age 35-44		00168	4	N
	E_0A06			
End-of-period employment for Male and Female and age 45-54		00172	4	N
	E_0A07			
End-of-period employment for Male and Female and age 55-64		00176	4	N
	E_0A08			
End-of-period employment for Male and Female and age 65-99		00180	4	N
	E_1A01			
End-of-period employment for Male and age 14-18		00188	4	N
	E_1A00			
End-of-period employment for Male and age 14-99		00184	4	N
	E_1A02			
End-of-period employment for Male and age 19-21		00192	4	N
	E_1A03			
End-of-period employment for Male and age 22-24		00196	4	N
	E_1A04			
End-of-period employment for Male and age 25-34		00200	4	N
	E_1A05			
End-of-period employment for Male and age 35-44		00204	4	N
	E_1A06			
End-of-period employment for Male and age 45-54		00208	4	N
	E_1A07			
End-of-period employment for Male and age 55-64		00212	4	N
	E_1A08			
End-of-period employment for Male and age 65-99		00216	4	N
	ES_NAICS_FNL2007			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Final 2007 NAICS Code NNNNNN		05812	6	A/N
	CA_2A01			
Flow into consecutive quarter employment for Female and age 14-18		03032	4	N
	CA_2A00			
Flow into consecutive quarter employment for Female and age 14-99		03028	4	N
	CA_2A02			
Flow into consecutive quarter employment for Female and age 19-21		03036	4	N
	CA_2A03			
Flow into consecutive quarter employment for Female and age 22-24		03040	4	N
	CA_2A04			
Flow into consecutive quarter employment for Female and age 25-34		03044	4	N
	CA_2A05			
Flow into consecutive quarter employment for Female and age 35-44		03048	4	N
	CA_2A06			
Flow into consecutive quarter employment for Female and age 45-54		03052	4	N
	CA_2A07			
Flow into consecutive quarter employment for Female and age 55-64		03056	4	N
	CA_2A08			
Flow into consecutive quarter employment for Female and age 65-99		03060	4	N
	CA_0A01			
Flow into consecutive quarter employment for Male and Female and age 14-18		02960	4	N
	CA_0A00			
Flow into consecutive quarter employment for Male and Female and age 14-99		02956	4	N
	CA_0A02			
Flow into consecutive quarter employment for Male and Female and age 19-21		02964	4	N
	CA_0A03			
Flow into consecutive quarter employment for Male and Female and age 22-24		02968	4	N
	CA_0A04			
Flow into consecutive quarter employment for Male and Female and age 25-34		02972	4	N
	CA_0A05			
Flow into consecutive quarter employment for Male and Female and age 35-44		02976	4	N
	CA_0A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow into consecutive quarter employment for Male and Female and age 45-54		02980	4	N
	CA_0A07			
Flow into consecutive quarter employment for Male and Female and age 55-64		02984	4	N
	CA_0A08			
Flow into consecutive quarter employment for Male and Female and age 65-99		02988	4	N
	CA_1A01			
Flow into consecutive quarter employment for Male and age 14-18		02996	4	N
	CA_1A00			
Flow into consecutive quarter employment for Male and age 14-99		02992	4	N
	CA_1A02			
Flow into consecutive quarter employment for Male and age 19-21		03000	4	N
	CA_1A03			
Flow into consecutive quarter employment for Male and age 22-24		03004	4	N
	CA_1A04			
Flow into consecutive quarter employment for Male and age 25-34		03008	4	N
	CA_1A05			
Flow into consecutive quarter employment for Male and age 35-44		03012	4	N
	CA_1A06			
Flow into consecutive quarter employment for Male and age 45-54		03016	4	N
	CA_1A07			
Flow into consecutive quarter employment for Male and age 55-64		03020	4	N
	CA_1A08			
Flow into consecutive quarter employment for Male and age 65-99		03024	4	N
	FA_2A01			
Flow into full-quarter employment for Female and age 14-18		00764	4	N
	FA_2A00			
Flow into full-quarter employment for Female and age 14-99		00760	4	N
	FA_2A02			
Flow into full-quarter employment for Female and age 19-21		00768	4	N
	FA_2A03			
Flow into full-quarter employment for Female and age 22-24		00772	4	N
	FA_2A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow into full-quarter employment for Female and age 25-34		00776	4	N
	FA_2A05			
Flow into full-quarter employment for Female and age 35-44		00780	4	N
	FA_2A06			
Flow into full-quarter employment for Female and age 45-54		00784	4	N
	FA_2A07			
Flow into full-quarter employment for Female and age 55-64		00788	4	N
	FA_2A08			
Flow into full-quarter employment for Female and age 65-99		00792	4	N
	FA_0A01			
Flow into full-quarter employment for Male and Female and age 14-18		00692	4	N
	FA_0A00			
Flow into full-quarter employment for Male and Female and age 14-99		00688	4	N
	FA_0A02			
Flow into full-quarter employment for Male and Female and age 19-21		00696	4	N
	FA_0A03			
Flow into full-quarter employment for Male and Female and age 22-24		00700	4	N
	FA_0A04			
Flow into full-quarter employment for Male and Female and age 25-34		00704	4	N
	FA_0A05			
Flow into full-quarter employment for Male and Female and age 35-44		00708	4	N
	FA_0A06			
Flow into full-quarter employment for Male and Female and age 45-54		00712	4	N
	FA_0A07			
Flow into full-quarter employment for Male and Female and age 55-64		00716	4	N
	FA_0A08			
Flow into full-quarter employment for Male and Female and age 65-99		00720	4	N
	FA_1A01			
Flow into full-quarter employment for Male and age 14-18		00728	4	N
	FA_1A00			
Flow into full-quarter employment for Male and age 14-99		00724	4	N
	FA_1A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow into full-quarter employment for Male and age 19-21		00732	4	N
	FA_1A03			
Flow into full-quarter employment for Male and age 22-24		00736	4	N
	FA_1A04			
Flow into full-quarter employment for Male and age 25-34		00740	4	N
	FA_1A05			
Flow into full-quarter employment for Male and age 35-44		00744	4	N
	FA_1A06			
Flow into full-quarter employment for Male and age 45-54		00748	4	N
	FA_1A07			
Flow into full-quarter employment for Male and age 55-64		00752	4	N
	FA_1A08			
Flow into full-quarter employment for Male and age 65-99		00756	4	N
	CS_2A01			
Flow out of consecutive quarter employment for Female and age 14-18		03140	4	N
	CS_2A00			
Flow out of consecutive quarter employment for Female and age 14-99		03136	4	N
	CS_2A02			
Flow out of consecutive quarter employment for Female and age 19-21		03144	4	N
	CS_2A03			
Flow out of consecutive quarter employment for Female and age 22-24		03148	4	N
	CS_2A04			
Flow out of consecutive quarter employment for Female and age 25-34		03152	4	N
	CS_2A05			
Flow out of consecutive quarter employment for Female and age 35-44		03156	4	N
	CS_2A06			
Flow out of consecutive quarter employment for Female and age 45-54		03160	4	N
	CS_2A07			
Flow out of consecutive quarter employment for Female and age 55-64		03164	4	N
	CS_2A08			
Flow out of consecutive quarter employment for Female and age 65-99		03168	4	N
	CS_0A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow out of consecutive quarter employment for Male and Female and age 14-18		03068	4	N
	CS_0A00			
Flow out of consecutive quarter employment for Male and Female and age 14-99		03064	4	N
	CS_0A02			
Flow out of consecutive quarter employment for Male and Female and age 19-21		03072	4	N
	CS_0A03			
Flow out of consecutive quarter employment for Male and Female and age 22-24		03076	4	N
	CS_0A04			
Flow out of consecutive quarter employment for Male and Female and age 25-34		03080	4	N
	CS_0A05			
Flow out of consecutive quarter employment for Male and Female and age 35-44		03084	4	N
	CS_0A06			
Flow out of consecutive quarter employment for Male and Female and age 45-54		03088	4	N
	CS_0A07			
Flow out of consecutive quarter employment for Male and Female and age 55-64		03092	4	N
	CS_0A08			
Flow out of consecutive quarter employment for Male and Female and age 65-99		03096	4	N
	CS_1A01			
Flow out of consecutive quarter employment for Male and age 14-18		03104	4	N
	CS_1A00			
Flow out of consecutive quarter employment for Male and age 14-99		03100	4	N
	CS_1A02			
Flow out of consecutive quarter employment for Male and age 19-21		03108	4	N
	CS_1A03			
Flow out of consecutive quarter employment for Male and age 22-24		03112	4	N
	CS_1A04			
Flow out of consecutive quarter employment for Male and age 25-34		03116	4	N
	CS_1A05			
Flow out of consecutive quarter employment for Male and age 35-44		03120	4	N
	CS_1A06			
Flow out of consecutive quarter employment for Male and age 45-54		03124	4	N
	CS_1A07			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow out of consecutive quarter employment for Male and age 55-64		03128	4	N
	CS_1A08			
Flow out of consecutive quarter employment for Male and age 65-99		03132	4	N
	FSNx_2A01			
Flow out of full-quarter employment - next quarter for Female and age 14-18		00980	4	N
	FSNx_2A00			
Flow out of full-quarter employment - next quarter for Female and age 14-99		00976	4	N
	FSNx_2A02			
Flow out of full-quarter employment - next quarter for Female and age 19-21		00984	4	N
	FSNx_2A03			
Flow out of full-quarter employment - next quarter for Female and age 22-24		00988	4	N
	FSNx_2A04			
Flow out of full-quarter employment - next quarter for Female and age 25-34		00992	4	N
	FSNx_2A05			
Flow out of full-quarter employment - next quarter for Female and age 35-44		00996	4	N
	FSNx_2A06			
Flow out of full-quarter employment - next quarter for Female and age 45-54		01000	4	N
	FSNx_2A07			
Flow out of full-quarter employment - next quarter for Female and age 55-64		01004	4	N
	FSNx_2A08			
Flow out of full-quarter employment - next quarter for Female and age 65-99		01008	4	N
	FSNx_0A01			
Flow out of full-quarter employment - next quarter for Male and Female and age 14-18		00908	4	N
	FSNx_0A00			
Flow out of full-quarter employment - next quarter for Male and Female and age 14-99		00904	4	N
	FSNx_0A02			
Flow out of full-quarter employment - next quarter for Male and Female and age 19-21		00912	4	N
	FSNx_0A03			
Flow out of full-quarter employment - next quarter for Male and Female and age 22-24		00916	4	N
	FSNx_0A04			
Flow out of full-quarter employment - next quarter for Male and Female and age 25-34		00920	4	N
	FSNx_0A05			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow out of full-quarter employment - next quarter for Male and Female and age 35-44		00924	4	N
	FSnx_0A06			
Flow out of full-quarter employment - next quarter for Male and Female and age 45-54		00928	4	N
	FSnx_0A07			
Flow out of full-quarter employment - next quarter for Male and Female and age 55-64		00932	4	N
	FSnx_0A08			
Flow out of full-quarter employment - next quarter for Male and Female and age 65-99		00936	4	N
	FSnx_1A01			
Flow out of full-quarter employment - next quarter for Male and age 14-18		00944	4	N
	FSnx_1A00			
Flow out of full-quarter employment - next quarter for Male and age 14-99		00940	4	N
	FSnx_1A02			
Flow out of full-quarter employment - next quarter for Male and age 19-21		00948	4	N
	FSnx_1A03			
Flow out of full-quarter employment - next quarter for Male and age 22-24		00952	4	N
	FSnx_1A04			
Flow out of full-quarter employment - next quarter for Male and age 25-34		00956	4	N
	FSnx_1A05			
Flow out of full-quarter employment - next quarter for Male and age 35-44		00960	4	N
	FSnx_1A06			
Flow out of full-quarter employment - next quarter for Male and age 45-54		00964	4	N
	FSnx_1A07			
Flow out of full-quarter employment - next quarter for Male and age 55-64		00968	4	N
	FSnx_1A08			
Flow out of full-quarter employment - next quarter for Male and age 65-99		00972	4	N
	FS_2A01			
Flow out of full-quarter employment for Female and age 14-18		00872	4	N
	FS_2A00			
Flow out of full-quarter employment for Female and age 14-99		00868	4	N
	FS_2A02			
Flow out of full-quarter employment for Female and age 19-21		00876	4	N
	FS_2A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow out of full-quarter employment for Female and age 22-24		00880	4	N
	FS_2A04			
Flow out of full-quarter employment for Female and age 25-34		00884	4	N
	FS_2A05			
Flow out of full-quarter employment for Female and age 35-44		00888	4	N
	FS_2A06			
Flow out of full-quarter employment for Female and age 45-54		00892	4	N
	FS_2A07			
Flow out of full-quarter employment for Female and age 55-64		00896	4	N
	FS_2A08			
Flow out of full-quarter employment for Female and age 65-99		00900	4	N
	FS_0A01			
Flow out of full-quarter employment for Male and Female and age 14-18		00800	4	N
	FS_0A00			
Flow out of full-quarter employment for Male and Female and age 14-99		00796	4	N
	FS_0A02			
Flow out of full-quarter employment for Male and Female and age 19-21		00804	4	N
	FS_0A03			
Flow out of full-quarter employment for Male and Female and age 22-24		00808	4	N
	FS_0A04			
Flow out of full-quarter employment for Male and Female and age 25-34		00812	4	N
	FS_0A05			
Flow out of full-quarter employment for Male and Female and age 35-44		00816	4	N
	FS_0A06			
Flow out of full-quarter employment for Male and Female and age 45-54		00820	4	N
	FS_0A07			
Flow out of full-quarter employment for Male and Female and age 55-64		00824	4	N
	FS_0A08			
Flow out of full-quarter employment for Male and Female and age 65-99		00828	4	N
	FS_1A01			
Flow out of full-quarter employment for Male and age 14-18		00836	4	N
	FS_1A00			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Flow out of full-quarter employment for Male and age 14-99		00832	4	N
	FS_1A02			
Flow out of full-quarter employment for Male and age 19-21		00840	4	N
	FS_1A03			
Flow out of full-quarter employment for Male and age 22-24		00844	4	N
	FS_1A04			
Flow out of full-quarter employment for Male and age 25-34		00848	4	N
	FS_1A05			
Flow out of full-quarter employment for Male and age 35-44		00852	4	N
	FS_1A06			
Flow out of full-quarter employment for Male and age 45-54		00856	4	N
	FS_1A07			
Flow out of full-quarter employment for Male and age 55-64		00860	4	N
	FS_1A08			
Flow out of full-quarter employment for Male and age 65-99		00864	4	N
	F_2A01			
Full-quarter employment for Female and age 14-18		00440	4	N
	F_2A00			
Full-quarter employment for Female and age 14-99		00436	4	N
	F_2A02			
Full-quarter employment for Female and age 19-21		00444	4	N
	F_2A03			
Full-quarter employment for Female and age 22-24		00448	4	N
	F_2A04			
Full-quarter employment for Female and age 25-34		00452	4	N
	F_2A05			
Full-quarter employment for Female and age 35-44		00456	4	N
	F_2A06			
Full-quarter employment for Female and age 45-54		00460	4	N
	F_2A07			
Full-quarter employment for Female and age 55-64		00464	4	N
	F_2A08			
Full-quarter employment for Female and age 65-99		00468	4	N
	F_0A01			
Full-quarter employment for Male and Female and age 14-18		00368	4	N
	F_0A00			
Full-quarter employment for Male and Female and age 14-99		00364	4	N
	F_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter employment for Male and Female and age 19-21		00372	4	N
	F_0A03			
Full-quarter employment for Male and Female and age 22-24		00376	4	N
	F_0A04			
Full-quarter employment for Male and Female and age 25-34		00380	4	N
	F_0A05			
Full-quarter employment for Male and Female and age 35-44		00384	4	N
	F_0A06			
Full-quarter employment for Male and Female and age 45-54		00388	4	N
	F_0A07			
Full-quarter employment for Male and Female and age 55-64		00392	4	N
	F_0A08			
Full-quarter employment for Male and Female and age 65-99		00396	4	N
	F_1A01			
Full-quarter employment for Male and age 14-18		00404	4	N
	F_1A00			
Full-quarter employment for Male and age 14-99		00400	4	N
	F_1A02			
Full-quarter employment for Male and age 19-21		00408	4	N
	F_1A03			
Full-quarter employment for Male and age 22-24		00412	4	N
	F_1A04			
Full-quarter employment for Male and age 25-34		00416	4	N
	F_1A05			
Full-quarter employment for Male and age 35-44		00420	4	N
	F_1A06			
Full-quarter employment for Male and age 45-54		00424	4	N
	F_1A07			
Full-quarter employment for Male and age 55-64		00428	4	N
	F_1A08			
Full-quarter employment for Male and age 65-99		00432	4	N
	FJC_2A01			
Full-quarter job creation for Female and age 14-18		04652	4	N
	FJC_2A00			
Full-quarter job creation for Female and age 14-99		04648	4	N
	FJC_2A02			
Full-quarter job creation for Female and age 19-21		04656	4	N
	FJC_2A03			
Full-quarter job creation for Female and age 22-24		04660	4	N
	FJC_2A04			
Full-quarter job creation for Female and age 25-34		04664	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
	FJC_2A05			
Full-quarter job creation for Female and age 35-44		04668	4	N
	FJC_2A06			
Full-quarter job creation for Female and age 45-54		04672	4	N
	FJC_2A07			
Full-quarter job creation for Female and age 55-64		04676	4	N
	FJC_2A08			
Full-quarter job creation for Female and age 65-99		04680	4	N
	FJC_0A01			
Full-quarter job creation for Male and Female and age 14-18		04580	4	N
	FJC_0A00			
Full-quarter job creation for Male and Female and age 14-99		04576	4	N
	FJC_0A02			
Full-quarter job creation for Male and Female and age 19-21		04584	4	N
	FJC_0A03			
Full-quarter job creation for Male and Female and age 22-24		04588	4	N
	FJC_0A04			
Full-quarter job creation for Male and Female and age 25-34		04592	4	N
	FJC_0A05			
Full-quarter job creation for Male and Female and age 35-44		04596	4	N
	FJC_0A06			
Full-quarter job creation for Male and Female and age 45-54		04600	4	N
	FJC_0A07			
Full-quarter job creation for Male and Female and age 55-64		04604	4	N
	FJC_0A08			
Full-quarter job creation for Male and Female and age 65-99		04608	4	N
	FJC_1A01			
Full-quarter job creation for Male and age 14-18		04616	4	N
	FJC_1A00			
Full-quarter job creation for Male and age 14-99		04612	4	N
	FJC_1A02			
Full-quarter job creation for Male and age 19-21		04620	4	N
	FJC_1A03			
Full-quarter job creation for Male and age 22-24		04624	4	N
	FJC_1A04			
Full-quarter job creation for Male and age 25-34		04628	4	N
	FJC_1A05			
Full-quarter job creation for Male and age 35-44		04632	4	N
	FJC_1A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter job creation for Male and age 45-54	FJC_1A07	04636	4	N
Full-quarter job creation for Male and age 55-64	FJC_1A08	04640	4	N
Full-quarter job creation for Male and age 65-99	FJD_2A01	04644	4	N
Full-quarter job destruction for Female and age 14-18	FJD_2A00	04868	4	N
Full-quarter job destruction for Female and age 14-99	FJD_2A02	04864	4	N
Full-quarter job destruction for Female and age 19-21	FJD_2A03	04872	4	N
Full-quarter job destruction for Female and age 22-24	FJD_2A04	04876	4	N
Full-quarter job destruction for Female and age 25-34	FJD_2A05	04880	4	N
Full-quarter job destruction for Female and age 35-44	FJD_2A06	04884	4	N
Full-quarter job destruction for Female and age 45-54	FJD_2A07	04888	4	N
Full-quarter job destruction for Female and age 55-64	FJD_2A08	04892	4	N
Full-quarter job destruction for Female and age 65-99	FJD_0A01	04896	4	N
Full-quarter job destruction for Male and Female and age 14-18	FJD_0A00	04796	4	N
Full-quarter job destruction for Male and Female and age 14-99	FJD_0A02	04792	4	N
Full-quarter job destruction for Male and Female and age 19-21	FJD_0A03	04800	4	N
Full-quarter job destruction for Male and Female and age 22-24	FJD_0A04	04804	4	N
Full-quarter job destruction for Male and Female and age 25-34	FJD_0A05	04808	4	N
Full-quarter job destruction for Male and Female and age 35-44	FJD_0A06	04812	4	N
Full-quarter job destruction for Male and Female and age 45-54	FJD_0A07	04816	4	N
Full-quarter job destruction for Male and Female and age 55-64	FJD_0A08	04820	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter job destruction for Male and Female and age 65-99	FJD_1A01	04824	4	N
Full-quarter job destruction for Male and age 14-18	FJD_1A00	04832	4	N
Full-quarter job destruction for Male and age 14-99	FJD_1A02	04828	4	N
Full-quarter job destruction for Male and age 19-21	FJD_1A03	04836	4	N
Full-quarter job destruction for Male and age 22-24	FJD_1A04	04840	4	N
Full-quarter job destruction for Male and age 25-34	FJD_1A05	04844	4	N
Full-quarter job destruction for Male and age 35-44	FJD_1A06	04848	4	N
Full-quarter job destruction for Male and age 45-54	FJD_1A07	04852	4	N
Full-quarter job destruction for Male and age 55-64	FJD_1A08	04856	4	N
Full-quarter job destruction for Male and age 65-99	H3_2A01	04860	4	N
Full-quarter new hires for Female and age 14-18	H3_2A00	03464	4	N
Full-quarter new hires for Female and age 14-99	H3_2A02	03460	4	N
Full-quarter new hires for Female and age 19-21	H3_2A03	03468	4	N
Full-quarter new hires for Female and age 22-24	H3_2A04	03472	4	N
Full-quarter new hires for Female and age 25-34	H3_2A05	03476	4	N
Full-quarter new hires for Female and age 35-44	H3_2A06	03480	4	N
Full-quarter new hires for Female and age 45-54	H3_2A07	03484	4	N
Full-quarter new hires for Female and age 55-64	H3_2A08	03488	4	N
Full-quarter new hires for Female and age 65-99	H3_0A01	03492	4	N
Full-quarter new hires for Male and Female and age 14-18	H3_0A00	03392	4	N
Full-quarter new hires for Male and Female and age 14-99	H3_0A02	03388	4	N
Full-quarter new hires for Male and Female and age 19-21	H3_0A03	03396	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Full-quarter new hires for Male and Female and age 22-24	H3_0A04	03400	4	N
Full-quarter new hires for Male and Female and age 25-34	H3_0A05	03404	4	N
Full-quarter new hires for Male and Female and age 35-44	H3_0A06	03408	4	N
Full-quarter new hires for Male and Female and age 45-54	H3_0A07	03412	4	N
Full-quarter new hires for Male and Female and age 55-64	H3_0A08	03416	4	N
Full-quarter new hires for Male and Female and age 65-99	H3_1A01	03420	4	N
Full-quarter new hires for Male and age 14-18	H3_1A00	03428	4	N
Full-quarter new hires for Male and age 14-99	H3_1A02	03424	4	N
Full-quarter new hires for Male and age 19-21	H3_1A03	03432	4	N
Full-quarter new hires for Male and age 22-24	H3_1A04	03436	4	N
Full-quarter new hires for Male and age 25-34	H3_1A05	03440	4	N
Full-quarter new hires for Male and age 35-44	H3_1A06	03444	4	N
Full-quarter new hires for Male and age 45-54	H3_1A07	03448	4	N
Full-quarter new hires for Male and age 55-64	H3_1A08	03452	4	N
Full-quarter new hires for Male and age 65-99	JC_2A01	03456	4	N
Job creation for Female and age 14-18	JC_2A00	03896	4	N
Job creation for Female and age 14-99	JC_2A02	03892	4	N
Job creation for Female and age 19-21	JC_2A03	03900	4	N
Job creation for Female and age 22-24	JC_2A04	03904	4	N
Job creation for Female and age 25-34	JC_2A05	03908	4	N
Job creation for Female and age 35-44	JC_2A06	03912	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Job creation for Female and age 45-54	JC_2A07	03916	4	N
Job creation for Female and age 55-64	JC_2A08	03920	4	N
Job creation for Female and age 65-99	JC_0A01	03924	4	N
Job creation for Male and Female and age 14-18	JC_0A00	03824	4	N
Job creation for Male and Female and age 14-99	JC_0A02	03820	4	N
Job creation for Male and Female and age 19-21	JC_0A03	03828	4	N
Job creation for Male and Female and age 22-24	JC_0A04	03832	4	N
Job creation for Male and Female and age 25-34	JC_0A05	03836	4	N
Job creation for Male and Female and age 35-44	JC_0A06	03840	4	N
Job creation for Male and Female and age 45-54	JC_0A07	03844	4	N
Job creation for Male and Female and age 55-64	JC_0A08	03848	4	N
Job creation for Male and Female and age 65-99	JC_1A01	03852	4	N
Job creation for Male and age 14-18	JC_1A00	03860	4	N
Job creation for Male and age 14-99	JC_1A02	03856	4	N
Job creation for Male and age 19-21	JC_1A03	03864	4	N
Job creation for Male and age 22-24	JC_1A04	03868	4	N
Job creation for Male and age 25-34	JC_1A05	03872	4	N
Job creation for Male and age 35-44	JC_1A06	03876	4	N
Job creation for Male and age 45-54	JC_1A07	03880	4	N
Job creation for Male and age 55-64	JC_1A08	03884	4	N
Job creation for Male and age 65-99	JD_2A01	03888	4	N
Job destruction for Female and age 14-18	JD_2A00	04112	4	N
Job destruction for Female and age 14-99	JD_2A02	04108	4	N
Job destruction for Female and age 19-21	JD_2A03	04116	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Job destruction for Female and age 22-24		04120	4	N
	JD_2A04			
Job destruction for Female and age 25-34		04124	4	N
	JD_2A05			
Job destruction for Female and age 35-44		04128	4	N
	JD_2A06			
Job destruction for Female and age 45-54		04132	4	N
	JD_2A07			
Job destruction for Female and age 55-64		04136	4	N
	JD_2A08			
Job destruction for Female and age 65-99		04140	4	N
	JD_0A01			
Job destruction for Male and Female and age 14-18		04040	4	N
	JD_0A00			
Job destruction for Male and Female and age 14-99		04036	4	N
	JD_0A02			
Job destruction for Male and Female and age 19-21		04044	4	N
	JD_0A03			
Job destruction for Male and Female and age 22-24		04048	4	N
	JD_0A04			
Job destruction for Male and Female and age 25-34		04052	4	N
	JD_0A05			
Job destruction for Male and Female and age 35-44		04056	4	N
	JD_0A06			
Job destruction for Male and Female and age 45-54		04060	4	N
	JD_0A07			
Job destruction for Male and Female and age 55-64		04064	4	N
	JD_0A08			
Job destruction for Male and Female and age 65-99		04068	4	N
	JD_1A01			
Job destruction for Male and age 14-18		04076	4	N
	JD_1A00			
Job destruction for Male and age 14-99		04072	4	N
	JD_1A02			
Job destruction for Male and age 19-21		04080	4	N
	JD_1A03			
Job destruction for Male and age 22-24		04084	4	N
	JD_1A04			
Job destruction for Male and age 25-34		04088	4	N
	JD_1A05			
Job destruction for Male and age 35-44		04092	4	N
	JD_1A06			
Job destruction for Male and age 45-54		04096	4	N
	JD_1A07			
Job destruction for Male and age 55-64		04100	4	N
	JD_1A08			
Job destruction for Male and age 65-99		04104	4	N
	FJF_2A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Net change in full-quarter employment for Female and age 14-18		04436	4	N
	FJF_2A00			
Net change in full-quarter employment for Female and age 14-99		04432	4	N
	FJF_2A02			
Net change in full-quarter employment for Female and age 19-21		04440	4	N
	FJF_2A03			
Net change in full-quarter employment for Female and age 22-24		04444	4	N
	FJF_2A04			
Net change in full-quarter employment for Female and age 25-34		04448	4	N
	FJF_2A05			
Net change in full-quarter employment for Female and age 35-44		04452	4	N
	FJF_2A06			
Net change in full-quarter employment for Female and age 45-54		04456	4	N
	FJF_2A07			
Net change in full-quarter employment for Female and age 55-64		04460	4	N
	FJF_2A08			
Net change in full-quarter employment for Female and age 65-99		04464	4	N
	FJF_0A01			
Net change in full-quarter employment for Male and Female and age 14-18		04364	4	N
	FJF_0A00			
Net change in full-quarter employment for Male and Female and age 14-99		04360	4	N
	FJF_0A02			
Net change in full-quarter employment for Male and Female and age 19-21		04368	4	N
	FJF_0A03			
Net change in full-quarter employment for Male and Female and age 22-24		04372	4	N
	FJF_0A04			
Net change in full-quarter employment for Male and Female and age 25-34		04376	4	N
	FJF_0A05			
Net change in full-quarter employment for Male and Female and age 35-44		04380	4	N
	FJF_0A06			
Net change in full-quarter employment for Male and Female and age 45-54		04384	4	N
	FJF_0A07			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Net change in full-quarter employment for Male and Female and age 55-64		04388	4	N
	FJF_0A08			
Net change in full-quarter employment for Male and Female and age 65-99		04392	4	N
	FJF_1A01			
Net change in full-quarter employment for Male and age 14-18		04400	4	N
	FJF_1A00			
Net change in full-quarter employment for Male and age 14-99		04396	4	N
	FJF_1A02			
Net change in full-quarter employment for Male and age 19-21		04404	4	N
	FJF_1A03			
Net change in full-quarter employment for Male and age 22-24		04408	4	N
	FJF_1A04			
Net change in full-quarter employment for Male and age 25-34		04412	4	N
	FJF_1A05			
Net change in full-quarter employment for Male and age 35-44		04416	4	N
	FJF_1A06			
Net change in full-quarter employment for Male and age 45-54		04420	4	N
	FJF_1A07			
Net change in full-quarter employment for Male and age 55-64		04424	4	N
	FJF_1A08			
Net change in full-quarter employment for Male and age 65-99		04428	4	N
	JF_2A01			
Net job flows for Female and age 14-18		03788	4	N
	JF_2A00			
Net job flows for Female and age 14-99		03784	4	N
	JF_2A02			
Net job flows for Female and age 19-21		03792	4	N
	JF_2A03			
Net job flows for Female and age 22-24		03796	4	N
	JF_2A04			
Net job flows for Female and age 25-34		03800	4	N
	JF_2A05			
Net job flows for Female and age 35-44		03804	4	N
	JF_2A06			
Net job flows for Female and age 45-54		03808	4	N
	JF_2A07			
Net job flows for Female and age 55-64		03812	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Net job flows for Female and age 65-99	JF_2A08	03816	4	N
Net job flows for Male and Female and age 14-18	JF_0A01	03716	4	N
Net job flows for Male and Female and age 14-99	JF_0A00	03712	4	N
Net job flows for Male and Female and age 19-21	JF_0A02	03720	4	N
Net job flows for Male and Female and age 22-24	JF_0A03	03724	4	N
Net job flows for Male and Female and age 25-34	JF_0A04	03728	4	N
Net job flows for Male and Female and age 35-44	JF_0A05	03732	4	N
Net job flows for Male and Female and age 45-54	JF_0A06	03736	4	N
Net job flows for Male and Female and age 55-64	JF_0A07	03740	4	N
Net job flows for Male and Female and age 65-99	JF_0A08	03744	4	N
Net job flows for Male and age 14-18	JF_1A01	03752	4	N
Net job flows for Male and age 14-99	JF_1A00	03748	4	N
Net job flows for Male and age 19-21	JF_1A02	03756	4	N
Net job flows for Male and age 22-24	JF_1A03	03760	4	N
Net job flows for Male and age 25-34	JF_1A04	03764	4	N
Net job flows for Male and age 35-44	JF_1A05	03768	4	N
Net job flows for Male and age 45-54	JF_1A06	03772	4	N
Net job flows for Male and age 55-64	JF_1A07	03776	4	N
Net job flows for Male and age 65-99	JF_1A08	03780	4	N
New Hires into Continuous Quarter Employment for Female and age 14-18	CH_2A01	03356	4	N
New Hires into Continuous Quarter Employment for Female and age 14-99	CH_2A00	03352	4	N
New Hires into Continuous Quarter Employment for Female and age 19-21	CH_2A02	03360	4	N
	CH_2A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
New Hires into Continuous Quarter Employment for Female and age 22-24		03364	4	N
CH_2A04				
New Hires into Continuous Quarter Employment for Female and age 25-34		03368	4	N
CH_2A05				
New Hires into Continuous Quarter Employment for Female and age 35-44		03372	4	N
CH_2A06				
New Hires into Continuous Quarter Employment for Female and age 45-54		03376	4	N
CH_2A07				
New Hires into Continuous Quarter Employment for Female and age 55-64		03380	4	N
CH_2A08				
New Hires into Continuous Quarter Employment for Female and age 65-99		03384	4	N
CH_0A01				
New Hires into Continuous Quarter Employment for Male and Female and age 14-18		03284	4	N
CH_0A00				
New Hires into Continuous Quarter Employment for Male and Female and age 14-99		03280	4	N
CH_0A02				
New Hires into Continuous Quarter Employment for Male and Female and age 19-21		03288	4	N
CH_0A03				
New Hires into Continuous Quarter Employment for Male and Female and age 22-24		03292	4	N
CH_0A04				
New Hires into Continuous Quarter Employment for Male and Female and age 25-34		03296	4	N
CH_0A05				
New Hires into Continuous Quarter Employment for Male and Female and age 35-44		03300	4	N
CH_0A06				
New Hires into Continuous Quarter Employment for Male and Female and age 45-54		03304	4	N
CH_0A07				
New Hires into Continuous Quarter Employment for Male and Female and age 55-64		03308	4	N
CH_0A08				
New Hires into Continuous Quarter Employment for Male and Female and age 65-99		03312	4	N
CH_1A01				
New Hires into Continuous Quarter Employment for Male and age 14-18		03320	4	N
CH_1A00				

Field name	Data dictionary reference name	Starting position	Field size	Data type
New Hires into Continuous Quarter Employment for Male and age 14-99		03316	4	N
	CH_1A02			
New Hires into Continuous Quarter Employment for Male and age 19-21		03324	4	N
	CH_1A03			
New Hires into Continuous Quarter Employment for Male and age 22-24		03328	4	N
	CH_1A04			
New Hires into Continuous Quarter Employment for Male and age 25-34		03332	4	N
	CH_1A05			
New Hires into Continuous Quarter Employment for Male and age 35-44		03336	4	N
	CH_1A06			
New Hires into Continuous Quarter Employment for Male and age 45-54		03340	4	N
	CH_1A07			
New Hires into Continuous Quarter Employment for Male and age 55-64		03344	4	N
	CH_1A08			
New Hires into Continuous Quarter Employment for Male and age 65-99		03348	4	N
	H_2A01			
New hires for Female and age 14-18		02384	4	N
	H_2A00			
New hires for Female and age 14-99		02380	4	N
	H_2A02			
New hires for Female and age 19-21		02388	4	N
	H_2A03			
New hires for Female and age 22-24		02392	4	N
	H_2A04			
New hires for Female and age 25-34		02396	4	N
	H_2A05			
New hires for Female and age 35-44		02400	4	N
	H_2A06			
New hires for Female and age 45-54		02404	4	N
	H_2A07			
New hires for Female and age 55-64		02408	4	N
	H_2A08			
New hires for Female and age 65-99		02412	4	N
	H_0A01			
New hires for Male and Female and age 14-18		02312	4	N
	H_0A00			
New hires for Male and Female and age 14-99		02308	4	N
	H_0A02			
New hires for Male and Female and age 19-21		02316	4	N
	H_0A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
New hires for Male and Female and age 22-24		02320	4	N
	H_0A04			
New hires for Male and Female and age 25-34		02324	4	N
	H_0A05			
New hires for Male and Female and age 35-44		02328	4	N
	H_0A06			
New hires for Male and Female and age 45-54		02332	4	N
	H_0A07			
New hires for Male and Female and age 55-64		02336	4	N
	H_0A08			
New hires for Male and Female and age 65-99		02340	4	N
	H_1A01			
New hires for Male and age 14-18		02348	4	N
	H_1A00			
New hires for Male and age 14-99		02344	4	N
	H_1A02			
New hires for Male and age 19-21		02352	4	N
	H_1A03			
New hires for Male and age 22-24		02356	4	N
	H_1A04			
New hires for Male and age 25-34		02360	4	N
	H_1A05			
New hires for Male and age 35-44		02364	4	N
	H_1A06			
New hires for Male and age 45-54		02368	4	N
	H_1A07			
New hires for Male and age 55-64		02372	4	N
	H_1A08			
New hires for Male and age 65-99		02376	4	N
	FIRMSIZE			
Noise infused value of firmsize		00016	8	N
	QWI_WCF			
QWI weight correction factor		00032	8	N
	QUARTER			
Quarter QQ		05767	3	N
	R_2A01			
Recalls for Female and age 14-18		02492	4	N
	R_2A00			
Recalls for Female and age 14-99		02488	4	N
	R_2A02			
Recalls for Female and age 19-21		02496	4	N
	R_2A03			
Recalls for Female and age 22-24		02500	4	N
	R_2A04			
Recalls for Female and age 25-34		02504	4	N
	R_2A05			
Recalls for Female and age 35-44		02508	4	N
	R_2A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Recalls for Female and age 45-54	R_2A07	02512	4	N
Recalls for Female and age 55-64	R_2A08	02516	4	N
Recalls for Female and age 65-99	R_0A01	02520	4	N
Recalls for Male and Female and age 14-18	R_0A00	02420	4	N
Recalls for Male and Female and age 14-99	R_0A02	02416	4	N
Recalls for Male and Female and age 19-21	R_0A03	02424	4	N
Recalls for Male and Female and age 22-24	R_0A04	02428	4	N
Recalls for Male and Female and age 25-34	R_0A05	02432	4	N
Recalls for Male and Female and age 35-44	R_0A06	02436	4	N
Recalls for Male and Female and age 45-54	R_0A07	02440	4	N
Recalls for Male and Female and age 55-64	R_0A08	02444	4	N
Recalls for Male and Female and age 65-99	R_1A01	02448	4	N
Recalls for Male and age 14-18	R_1A00	02456	4	N
Recalls for Male and age 14-99	R_1A02	02452	4	N
Recalls for Male and age 19-21	R_1A03	02460	4	N
Recalls for Male and age 22-24	R_1A04	02464	4	N
Recalls for Male and age 25-34	R_1A05	02468	4	N
Recalls for Male and age 35-44	R_1A06	02472	4	N
Recalls for Male and age 45-54	R_1A07	02476	4	N
Recalls for Male and age 55-64	R_1A08	02480	4	N
Recalls for Male and age 65-99	CR_2A01	02484	4	N
Recalls into Continuous Quarter Employment for Female and age 14-18	CR_2A00	02600	4	N
Recalls into Continuous Quarter Employment for Female and age 14-99	CR_2A02	02596	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Recalls into Continuous Quarter Employment for Female and age 19-21		02604	4	N
	CR_2A03			
Recalls into Continuous Quarter Employment for Female and age 22-24		02608	4	N
	CR_2A04			
Recalls into Continuous Quarter Employment for Female and age 25-34		02612	4	N
	CR_2A05			
Recalls into Continuous Quarter Employment for Female and age 35-44		02616	4	N
	CR_2A06			
Recalls into Continuous Quarter Employment for Female and age 45-54		02620	4	N
	CR_2A07			
Recalls into Continuous Quarter Employment for Female and age 55-64		02624	4	N
	CR_2A08			
Recalls into Continuous Quarter Employment for Female and age 65-99		02628	4	N
	CR_0A01			
Recalls into Continuous Quarter Employment for Male and Female and age 14-18		02528	4	N
	CR_0A00			
Recalls into Continuous Quarter Employment for Male and Female and age 14-99		02524	4	N
	CR_0A02			
Recalls into Continuous Quarter Employment for Male and Female and age 19-21		02532	4	N
	CR_0A03			
Recalls into Continuous Quarter Employment for Male and Female and age 22-24		02536	4	N
	CR_0A04			
Recalls into Continuous Quarter Employment for Male and Female and age 25-34		02540	4	N
	CR_0A05			
Recalls into Continuous Quarter Employment for Male and Female and age 35-44		02544	4	N
	CR_0A06			
Recalls into Continuous Quarter Employment for Male and Female and age 45-54		02548	4	N
	CR_0A07			
Recalls into Continuous Quarter Employment for Male and Female and age 55-64		02552	4	N
	CR_0A08			
Recalls into Continuous Quarter Employment for Male and Female and age 65-99		02556	4	N
	CR_1A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Recalls into Continuous Quarter Employment for Male and age 14-18		02564	4	N
	CR_1A00			
Recalls into Continuous Quarter Employment for Male and age 14-99		02560	4	N
	CR_1A02			
Recalls into Continuous Quarter Employment for Male and age 19-21		02568	4	N
	CR_1A03			
Recalls into Continuous Quarter Employment for Male and age 22-24		02572	4	N
	CR_1A04			
Recalls into Continuous Quarter Employment for Male and age 25-34		02576	4	N
	CR_1A05			
Recalls into Continuous Quarter Employment for Male and age 35-44		02580	4	N
	CR_1A06			
Recalls into Continuous Quarter Employment for Male and age 45-54		02584	4	N
	CR_1A07			
Recalls into Continuous Quarter Employment for Male and age 55-64		02588	4	N
	CR_1A08			
Recalls into Continuous Quarter Employment for Male and age 65-99		02592	4	N
	S_2A01			
Separations for Female and age 14-18		00656	4	N
	S_2A00			
Separations for Female and age 14-99		00652	4	N
	S_2A02			
Separations for Female and age 19-21		00660	4	N
	S_2A03			
Separations for Female and age 22-24		00664	4	N
	S_2A04			
Separations for Female and age 25-34		00668	4	N
	S_2A05			
Separations for Female and age 35-44		00672	4	N
	S_2A06			
Separations for Female and age 45-54		00676	4	N
	S_2A07			
Separations for Female and age 55-64		00680	4	N
	S_2A08			
Separations for Female and age 65-99		00684	4	N
	S_0A01			
Separations for Male and Female and age 14-18		00584	4	N
	S_0A00			
Separations for Male and Female and age 14-99		00580	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Separations for Male and Female and age 19-21	S_0A02	00588	4	N
Separations for Male and Female and age 22-24	S_0A03	00592	4	N
Separations for Male and Female and age 25-34	S_0A04	00596	4	N
Separations for Male and Female and age 35-44	S_0A05	00600	4	N
Separations for Male and Female and age 45-54	S_0A06	00604	4	N
Separations for Male and Female and age 55-64	S_0A07	00608	4	N
Separations for Male and Female and age 65-99	S_0A08	00612	4	N
Separations for Male and age 14-18	S_1A01	00620	4	N
Separations for Male and age 14-99	S_1A00	00616	4	N
Separations for Male and age 19-21	S_1A02	00624	4	N
Separations for Male and age 22-24	S_1A03	00628	4	N
Separations for Male and age 25-34	S_1A04	00632	4	N
Separations for Male and age 35-44	S_1A05	00636	4	N
Separations for Male and age 45-54	S_1A06	00640	4	N
Separations for Male and age 55-64	S_1A07	00644	4	N
Separations for Male and age 65-99	S_1A08	00648	4	N
State Employer ID Number	SEIN	05770	12	A/N
State UI Reporting Unit Number	SEINUNIT	05782	5	A/N
Sub-county geocode	LEG_SUBCTYGEO	05794	10	A/N
Total earnings of separations for Female and age 14-18	WS_2A01	01844	4	N
Total earnings of separations for Female and age 14-99	WS_2A00	01840	4	N
Total earnings of separations for Female and age 19-21	WS_2A02	01848	4	N
	WS_2A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total earnings of separations for Female and age 22-24		01852	4	N
Total earnings of separations for Female and age 25-34	WS_2A04	01856	4	N
Total earnings of separations for Female and age 35-44	WS_2A05	01860	4	N
Total earnings of separations for Female and age 45-54	WS_2A06	01864	4	N
Total earnings of separations for Female and age 55-64	WS_2A07	01868	4	N
Total earnings of separations for Female and age 65-99	WS_2A08	01872	4	N
Total earnings of separations for Male and Female and age 14-18	WS_0A01	01772	4	N
Total earnings of separations for Male and Female and age 14-99	WS_0A00	01768	4	N
Total earnings of separations for Male and Female and age 19-21	WS_0A02	01776	4	N
Total earnings of separations for Male and Female and age 22-24	WS_0A03	01780	4	N
Total earnings of separations for Male and Female and age 25-34	WS_0A04	01784	4	N
Total earnings of separations for Male and Female and age 35-44	WS_0A05	01788	4	N
Total earnings of separations for Male and Female and age 45-54	WS_0A06	01792	4	N
Total earnings of separations for Male and Female and age 55-64	WS_0A07	01796	4	N
Total earnings of separations for Male and Female and age 65-99	WS_0A08	01800	4	N
Total earnings of separations for Male and age 14-18	WS_1A01	01808	4	N
Total earnings of separations for Male and age 14-99	WS_1A00	01804	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
	WS_1A02			
Total earnings of separations for Male and age 19-21		01812	4	N
	WS_1A03			
Total earnings of separations for Male and age 22-24		01816	4	N
	WS_1A04			
Total earnings of separations for Male and age 25-34		01820	4	N
	WS_1A05			
Total earnings of separations for Male and age 35-44		01824	4	N
	WS_1A06			
Total earnings of separations for Male and age 45-54		01828	4	N
	WS_1A07			
Total earnings of separations for Male and age 55-64		01832	4	N
	WS_1A08			
Total earnings of separations for Male and age 65-99		01836	4	N
	WFS_2A01			
Total earnings of separations from full-quarter status for Female and age 14-18		02060	4	N
	WFS_2A00			
Total earnings of separations from full-quarter status for Female and age 14-99		02056	4	N
	WFS_2A02			
Total earnings of separations from full-quarter status for Female and age 19-21		02064	4	N
	WFS_2A03			
Total earnings of separations from full-quarter status for Female and age 22-24		02068	4	N
	WFS_2A04			
Total earnings of separations from full-quarter status for Female and age 25-34		02072	4	N
	WFS_2A05			
Total earnings of separations from full-quarter status for Female and age 35-44		02076	4	N
	WFS_2A06			
Total earnings of separations from full-quarter status for Female and age 45-54		02080	4	N
	WFS_2A07			
Total earnings of separations from full-quarter status for Female and age 55-64		02084	4	N
	WFS_2A08			
Total earnings of separations from full-quarter status for Female and age 65-99		02088	4	N
	WFS_0A01			
Total earnings of separations from full-quarter status for Male and Female and age 14-18		01988	4	N
	WFS_0A00			
Total earnings of separations from full-quarter status for Male and Female and age 14-99		01984	4	N
	WFS_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total earnings of separations from full-quarter status for Male and Female and age 19-21		01992	4	N
	WFS_0A03			
Total earnings of separations from full-quarter status for Male and Female and age 22-24		01996	4	N
	WFS_0A04			
Total earnings of separations from full-quarter status for Male and Female and age 25-34		02000	4	N
	WFS_0A05			
Total earnings of separations from full-quarter status for Male and Female and age 35-44		02004	4	N
	WFS_0A06			
Total earnings of separations from full-quarter status for Male and Female and age 45-54		02008	4	N
	WFS_0A07			
Total earnings of separations from full-quarter status for Male and Female and age 55-64		02012	4	N
	WFS_0A08			
Total earnings of separations from full-quarter status for Male and Female and age 65-99		02016	4	N
	WFS_1A01			
Total earnings of separations from full-quarter status for Male and age 14-18		02024	4	N
	WFS_1A00			
Total earnings of separations from full-quarter status for Male and age 14-99		02020	4	N
	WFS_1A02			
Total earnings of separations from full-quarter status for Male and age 19-21		02028	4	N
	WFS_1A03			
Total earnings of separations from full-quarter status for Male and age 22-24		02032	4	N
	WFS_1A04			
Total earnings of separations from full-quarter status for Male and age 25-34		02036	4	N
	WFS_1A05			
Total earnings of separations from full-quarter status for Male and age 35-44		02040	4	N
	WFS_1A06			
Total earnings of separations from full-quarter status for Male and age 45-54		02044	4	N
	WFS_1A07			
Total earnings of separations from full-quarter status for Male and age 55-64		02048	4	N
	WFS_1A08			
Total earnings of separations from full-quarter status for Male and age 65-99		02052	4	N
	WA_2A01			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of accessions for Female and age 14-18 WA_2A00		01304	4	N
Total payroll of accessions for Female and age 14-99 WA_2A02		01300	4	N
Total payroll of accessions for Female and age 19-21 WA_2A03		01308	4	N
Total payroll of accessions for Female and age 22-24 WA_2A04		01312	4	N
Total payroll of accessions for Female and age 25-34 WA_2A05		01316	4	N
Total payroll of accessions for Female and age 35-44 WA_2A06		01320	4	N
Total payroll of accessions for Female and age 45-54 WA_2A07		01324	4	N
Total payroll of accessions for Female and age 55-64 WA_2A08		01328	4	N
Total payroll of accessions for Female and age 65-99 WA_0A01		01332	4	N
Total payroll of accessions for Male and Female and age 14-18 WA_0A00		01232	4	N
Total payroll of accessions for Male and Female and age 14-99 WA_0A02		01228	4	N
Total payroll of accessions for Male and Female and age 19-21 WA_0A03		01236	4	N
Total payroll of accessions for Male and Female and age 22-24 WA_0A04		01240	4	N
Total payroll of accessions for Male and Female and age 25-34 WA_0A05		01244	4	N
Total payroll of accessions for Male and Female and age 35-44 WA_0A06		01248	4	N
Total payroll of accessions for Male and Female and age 45-54 WA_0A07		01252	4	N
Total payroll of accessions for Male and Female and age 55-64 WA_0A08		01256	4	N
Total payroll of accessions for Male and Female and age 65-99 WA_1A01		01260	4	N
Total payroll of accessions for Male and age 14-18 WA_1A00		01268	4	N
Total payroll of accessions for Male and age 14-99		01264	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
	WA_1A02			
Total payroll of accessions for Male and age 19-21		01272	4	N
	WA_1A03			
Total payroll of accessions for Male and age 22-24		01276	4	N
	WA_1A04			
Total payroll of accessions for Male and age 25-34		01280	4	N
	WA_1A05			
Total payroll of accessions for Male and age 35-44		01284	4	N
	WA_1A06			
Total payroll of accessions for Male and age 45-54		01288	4	N
	WA_1A07			
Total payroll of accessions for Male and age 55-64		01292	4	N
	WA_1A08			
Total payroll of accessions for Male and age 65-99		01296	4	N
	W1_2A01			
Total payroll of all employees for Female and age 14-18		02924	4	N
	W1_2A00			
Total payroll of all employees for Female and age 14-99		02920	4	N
	W1_2A02			
Total payroll of all employees for Female and age 19-21		02928	4	N
	W1_2A03			
Total payroll of all employees for Female and age 22-24		02932	4	N
	W1_2A04			
Total payroll of all employees for Female and age 25-34		02936	4	N
	W1_2A05			
Total payroll of all employees for Female and age 35-44		02940	4	N
	W1_2A06			
Total payroll of all employees for Female and age 45-54		02944	4	N
	W1_2A07			
Total payroll of all employees for Female and age 55-64		02948	4	N
	W1_2A08			
Total payroll of all employees for Female and age 65-99		02952	4	N
	W1_0A01			
Total payroll of all employees for Male and Female and age 14-18		02852	4	N
	W1_0A00			
Total payroll of all employees for Male and Female and age 14-99		02848	4	N
	W1_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of all employees for Male and Female and age 19-21		02856	4	N
	W1_0A03			
Total payroll of all employees for Male and Female and age 22-24		02860	4	N
	W1_0A04			
Total payroll of all employees for Male and Female and age 25-34		02864	4	N
	W1_0A05			
Total payroll of all employees for Male and Female and age 35-44		02868	4	N
	W1_0A06			
Total payroll of all employees for Male and Female and age 45-54		02872	4	N
	W1_0A07			
Total payroll of all employees for Male and Female and age 55-64		02876	4	N
	W1_0A08			
Total payroll of all employees for Male and Female and age 65-99		02880	4	N
	W1_1A01			
Total payroll of all employees for Male and age 14-18		02888	4	N
	W1_1A00			
Total payroll of all employees for Male and age 14-99		02884	4	N
	W1_1A02			
Total payroll of all employees for Male and age 19-21		02892	4	N
	W1_1A03			
Total payroll of all employees for Male and age 22-24		02896	4	N
	W1_1A04			
Total payroll of all employees for Male and age 25-34		02900	4	N
	W1_1A05			
Total payroll of all employees for Male and age 35-44		02904	4	N
	W1_1A06			
Total payroll of all employees for Male and age 45-54		02908	4	N
	W1_1A07			
Total payroll of all employees for Male and age 55-64		02912	4	N
	W1_1A08			
Total payroll of all employees for Male and age 65-99		02916	4	N
	W2_2A01			
Total payroll of end-of-period employees for Female and age 14-18		01088	4	N
	W2_2A00			
Total payroll of end-of-period employees for Female and age 14-99		01084	4	N
	W2_2A02			
Total payroll of end-of-period employees for Female and age 19-21		01092	4	N
	W2_2A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of end-of-period employees for Female and age 22-24		01096	4	N
	W2_2A04			
Total payroll of end-of-period employees for Female and age 25-34		01100	4	N
	W2_2A05			
Total payroll of end-of-period employees for Female and age 35-44		01104	4	N
	W2_2A06			
Total payroll of end-of-period employees for Female and age 45-54		01108	4	N
	W2_2A07			
Total payroll of end-of-period employees for Female and age 55-64		01112	4	N
	W2_2A08			
Total payroll of end-of-period employees for Female and age 65-99		01116	4	N
	W2_0A01			
Total payroll of end-of-period employees for Male and Female and age 14-18		01016	4	N
	W2_0A00			
Total payroll of end-of-period employees for Male and Female and age 14-99		01012	4	N
	W2_0A02			
Total payroll of end-of-period employees for Male and Female and age 19-21		01020	4	N
	W2_0A03			
Total payroll of end-of-period employees for Male and Female and age 22-24		01024	4	N
	W2_0A04			
Total payroll of end-of-period employees for Male and Female and age 25-34		01028	4	N
	W2_0A05			
Total payroll of end-of-period employees for Male and Female and age 35-44		01032	4	N
	W2_0A06			
Total payroll of end-of-period employees for Male and Female and age 45-54		01036	4	N
	W2_0A07			
Total payroll of end-of-period employees for Male and Female and age 55-64		01040	4	N
	W2_0A08			
Total payroll of end-of-period employees for Male and Female and age 65-99		01044	4	N
	W2_1A01			
Total payroll of end-of-period employees for Male and age 14-18		01052	4	N
	W2_1A00			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of end-of-period employees for Male and age 14-99		01048	4	N
	W2_1A02			
Total payroll of end-of-period employees for Male and age 19-21		01056	4	N
	W2_1A03			
Total payroll of end-of-period employees for Male and age 22-24		01060	4	N
	W2_1A04			
Total payroll of end-of-period employees for Male and age 25-34		01064	4	N
	W2_1A05			
Total payroll of end-of-period employees for Male and age 35-44		01068	4	N
	W2_1A06			
Total payroll of end-of-period employees for Male and age 45-54		01072	4	N
	W2_1A07			
Total payroll of end-of-period employees for Male and age 55-64		01076	4	N
	W2_1A08			
Total payroll of end-of-period employees for Male and age 65-99		01080	4	N
	W3_2A01			
Total payroll of full-quarter employees for Female and age 14-18		01196	4	N
	W3_2A00			
Total payroll of full-quarter employees for Female and age 14-99		01192	4	N
	W3_2A02			
Total payroll of full-quarter employees for Female and age 19-21		01200	4	N
	W3_2A03			
Total payroll of full-quarter employees for Female and age 22-24		01204	4	N
	W3_2A04			
Total payroll of full-quarter employees for Female and age 25-34		01208	4	N
	W3_2A05			
Total payroll of full-quarter employees for Female and age 35-44		01212	4	N
	W3_2A06			
Total payroll of full-quarter employees for Female and age 45-54		01216	4	N
	W3_2A07			
Total payroll of full-quarter employees for Female and age 55-64		01220	4	N
	W3_2A08			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of full-quarter employees for Female and age 65-99		01224	4	N
	W3_0A01			
Total payroll of full-quarter employees for Male and Female and age 14-18		01124	4	N
	W3_0A00			
Total payroll of full-quarter employees for Male and Female and age 14-99		01120	4	N
	W3_0A02			
Total payroll of full-quarter employees for Male and Female and age 19-21		01128	4	N
	W3_0A03			
Total payroll of full-quarter employees for Male and Female and age 22-24		01132	4	N
	W3_0A04			
Total payroll of full-quarter employees for Male and Female and age 25-34		01136	4	N
	W3_0A05			
Total payroll of full-quarter employees for Male and Female and age 35-44		01140	4	N
	W3_0A06			
Total payroll of full-quarter employees for Male and Female and age 45-54		01144	4	N
	W3_0A07			
Total payroll of full-quarter employees for Male and Female and age 55-64		01148	4	N
	W3_0A08			
Total payroll of full-quarter employees for Male and Female and age 65-99		01152	4	N
	W3_1A01			
Total payroll of full-quarter employees for Male and age 14-18		01160	4	N
	W3_1A00			
Total payroll of full-quarter employees for Male and age 14-99		01156	4	N
	W3_1A02			
Total payroll of full-quarter employees for Male and age 19-21		01164	4	N
	W3_1A03			
Total payroll of full-quarter employees for Male and age 22-24		01168	4	N
	W3_1A04			
Total payroll of full-quarter employees for Male and age 25-34		01172	4	N
	W3_1A05			
Total payroll of full-quarter employees for Male and age 35-44		01176	4	N
	W3_1A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of full-quarter employees for Male and age 45-54		01180	4	N
	W3_1A07			
Total payroll of full-quarter employees for Male and age 55-64		01184	4	N
	W3_1A08			
Total payroll of full-quarter employees for Male and age 65-99		01188	4	N
	WH3_2A01			
Total payroll of new hires to full-quarter status for Female and age 14-18		03572	4	N
	WH3_2A00			
Total payroll of new hires to full-quarter status for Female and age 14-99		03568	4	N
	WH3_2A02			
Total payroll of new hires to full-quarter status for Female and age 19-21		03576	4	N
	WH3_2A03			
Total payroll of new hires to full-quarter status for Female and age 22-24		03580	4	N
	WH3_2A04			
Total payroll of new hires to full-quarter status for Female and age 25-34		03584	4	N
	WH3_2A05			
Total payroll of new hires to full-quarter status for Female and age 35-44		03588	4	N
	WH3_2A06			
Total payroll of new hires to full-quarter status for Female and age 45-54		03592	4	N
	WH3_2A07			
Total payroll of new hires to full-quarter status for Female and age 55-64		03596	4	N
	WH3_2A08			
Total payroll of new hires to full-quarter status for Female and age 65-99		03600	4	N
	WH3_0A01			
Total payroll of new hires to full-quarter status for Male and Female and age 14-18		03500	4	N
	WH3_0A00			
Total payroll of new hires to full-quarter status for Male and Female and age 14-99		03496	4	N
	WH3_0A02			
Total payroll of new hires to full-quarter status for Male and Female and age 19-21		03504	4	N
	WH3_0A03			
Total payroll of new hires to full-quarter status for Male and Female and age 22-24		03508	4	N
	WH3_0A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of new hires to full-quarter status for Male and Female and age 25-34		03512	4	N
	WH3_0A05			
Total payroll of new hires to full-quarter status for Male and Female and age 35-44		03516	4	N
	WH3_0A06			
Total payroll of new hires to full-quarter status for Male and Female and age 45-54		03520	4	N
	WH3_0A07			
Total payroll of new hires to full-quarter status for Male and Female and age 55-64		03524	4	N
	WH3_0A08			
Total payroll of new hires to full-quarter status for Male and Female and age 65-99		03528	4	N
	WH3_1A01			
Total payroll of new hires to full-quarter status for Male and age 14-18		03536	4	N
	WH3_1A00			
Total payroll of new hires to full-quarter status for Male and age 14-99		03532	4	N
	WH3_1A02			
Total payroll of new hires to full-quarter status for Male and age 19-21		03540	4	N
	WH3_1A03			
Total payroll of new hires to full-quarter status for Male and age 22-24		03544	4	N
	WH3_1A04			
Total payroll of new hires to full-quarter status for Male and age 25-34		03548	4	N
	WH3_1A05			
Total payroll of new hires to full-quarter status for Male and age 35-44		03552	4	N
	WH3_1A06			
Total payroll of new hires to full-quarter status for Male and age 45-54		03556	4	N
	WH3_1A07			
Total payroll of new hires to full-quarter status for Male and age 55-64		03560	4	N
	WH3_1A08			
Total payroll of new hires to full-quarter status for Male and age 65-99		03564	4	N
	WCA_2A01			
Total payroll of transits to consecutive-quarter status for Female and age 14-18		03248	4	N
	WCA_2A00			
Total payroll of transits to consecutive-quarter status for Female and age 14-99		03244	4	N
	WCA_2A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of transits to consecutive-quarter status for Female and age 19-21	WCA_2A03	03252	4	N
Total payroll of transits to consecutive-quarter status for Female and age 22-24	WCA_2A04	03256	4	N
Total payroll of transits to consecutive-quarter status for Female and age 25-34	WCA_2A05	03260	4	N
Total payroll of transits to consecutive-quarter status for Female and age 35-44	WCA_2A06	03264	4	N
Total payroll of transits to consecutive-quarter status for Female and age 45-54	WCA_2A07	03268	4	N
Total payroll of transits to consecutive-quarter status for Female and age 55-64	WCA_2A08	03272	4	N
Total payroll of transits to consecutive-quarter status for Female and age 65-99	WCA_0A01	03276	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 14-18	WCA_0A00	03176	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 14-99	WCA_0A02	03172	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 19-21	WCA_0A03	03180	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 22-24	WCA_0A04	03184	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 25-34	WCA_0A05	03188	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 35-44	WCA_0A06	03192	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 45-54	WCA_0A07	03196	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 55-64	WCA_0A08	03200	4	N
Total payroll of transits to consecutive-quarter status for Male and Female and age 65-99	WCA_1A01	03204	4	N

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of transits to consecutive-quarter status for Male and age 14-18		03212	4	N
	WCA_1A00			
Total payroll of transits to consecutive-quarter status for Male and age 14-99		03208	4	N
	WCA_1A02			
Total payroll of transits to consecutive-quarter status for Male and age 19-21		03216	4	N
	WCA_1A03			
Total payroll of transits to consecutive-quarter status for Male and age 22-24		03220	4	N
	WCA_1A04			
Total payroll of transits to consecutive-quarter status for Male and age 25-34		03224	4	N
	WCA_1A05			
Total payroll of transits to consecutive-quarter status for Male and age 35-44		03228	4	N
	WCA_1A06			
Total payroll of transits to consecutive-quarter status for Male and age 45-54		03232	4	N
	WCA_1A07			
Total payroll of transits to consecutive-quarter status for Male and age 55-64		03236	4	N
	WCA_1A08			
Total payroll of transits to consecutive-quarter status for Male and age 65-99		03240	4	N
	WFA_2A01			
Total payroll of transits to full-quarter status for Female and age 14-18		01520	4	N
	WFA_2A00			
Total payroll of transits to full-quarter status for Female and age 14-99		01516	4	N
	WFA_2A02			
Total payroll of transits to full-quarter status for Female and age 19-21		01524	4	N
	WFA_2A03			
Total payroll of transits to full-quarter status for Female and age 22-24		01528	4	N
	WFA_2A04			
Total payroll of transits to full-quarter status for Female and age 25-34		01532	4	N
	WFA_2A05			
Total payroll of transits to full-quarter status for Female and age 35-44		01536	4	N
	WFA_2A06			
Total payroll of transits to full-quarter status for Female and age 45-54		01540	4	N
	WFA_2A07			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of transits to full-quarter status for Female and age 55-64		01544	4	N
	WFA_2A08			
Total payroll of transits to full-quarter status for Female and age 65-99		01548	4	N
	WFA_0A01			
Total payroll of transits to full-quarter status for Male and Female and age 14-18		01448	4	N
	WFA_0A00			
Total payroll of transits to full-quarter status for Male and Female and age 14-99		01444	4	N
	WFA_0A02			
Total payroll of transits to full-quarter status for Male and Female and age 19-21		01452	4	N
	WFA_0A03			
Total payroll of transits to full-quarter status for Male and Female and age 22-24		01456	4	N
	WFA_0A04			
Total payroll of transits to full-quarter status for Male and Female and age 25-34		01460	4	N
	WFA_0A05			
Total payroll of transits to full-quarter status for Male and Female and age 35-44		01464	4	N
	WFA_0A06			
Total payroll of transits to full-quarter status for Male and Female and age 45-54		01468	4	N
	WFA_0A07			
Total payroll of transits to full-quarter status for Male and Female and age 55-64		01472	4	N
	WFA_0A08			
Total payroll of transits to full-quarter status for Male and Female and age 65-99		01476	4	N
	WFA_1A01			
Total payroll of transits to full-quarter status for Male and age 14-18		01484	4	N
	WFA_1A00			
Total payroll of transits to full-quarter status for Male and age 14-99		01480	4	N
	WFA_1A02			
Total payroll of transits to full-quarter status for Male and age 19-21		01488	4	N
	WFA_1A03			
Total payroll of transits to full-quarter status for Male and age 22-24		01492	4	N
	WFA_1A04			
Total payroll of transits to full-quarter status for Male and age 25-34		01496	4	N
	WFA_1A05			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total payroll of transits to full-quarter status for Male and age 35-44		01500	4	N
	WFA_1A06			
Total payroll of transits to full-quarter status for Male and age 45-54		01504	4	N
	WFA_1A07			
Total payroll of transits to full-quarter status for Male and age 55-64		01508	4	N
	WFA_1A08			
Total payroll of transits to full-quarter status for Male and age 65-99		01512	4	N
	NA_2A01			
Total periods of non-employment for accessions for Female and age 14-18		01736	4	N
	NA_2A00			
Total periods of non-employment for accessions for Female and age 14-99		01732	4	N
	NA_2A02			
Total periods of non-employment for accessions for Female and age 19-21		01740	4	N
	NA_2A03			
Total periods of non-employment for accessions for Female and age 22-24		01744	4	N
	NA_2A04			
Total periods of non-employment for accessions for Female and age 25-34		01748	4	N
	NA_2A05			
Total periods of non-employment for accessions for Female and age 35-44		01752	4	N
	NA_2A06			
Total periods of non-employment for accessions for Female and age 45-54		01756	4	N
	NA_2A07			
Total periods of non-employment for accessions for Female and age 55-64		01760	4	N
	NA_2A08			
Total periods of non-employment for accessions for Female and age 65-99		01764	4	N
	NA_0A01			
Total periods of non-employment for accessions for Male and Female and age 14-18		01664	4	N
	NA_0A00			
Total periods of non-employment for accessions for Male and Female and age 14-99		01660	4	N
	NA_0A02			
Total periods of non-employment for accessions for Male and Female and age 19-21		01668	4	N
	NA_0A03			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for accessions for Male and Female and age 22-24		01672	4	N
	NA_0A04			
Total periods of non-employment for accessions for Male and Female and age 25-34		01676	4	N
	NA_0A05			
Total periods of non-employment for accessions for Male and Female and age 35-44		01680	4	N
	NA_0A06			
Total periods of non-employment for accessions for Male and Female and age 45-54		01684	4	N
	NA_0A07			
Total periods of non-employment for accessions for Male and Female and age 55-64		01688	4	N
	NA_0A08			
Total periods of non-employment for accessions for Male and Female and age 65-99		01692	4	N
	NA_1A01			
Total periods of non-employment for accessions for Male and age 14-18		01700	4	N
	NA_1A00			
Total periods of non-employment for accessions for Male and age 14-99		01696	4	N
	NA_1A02			
Total periods of non-employment for accessions for Male and age 19-21		01704	4	N
	NA_1A03			
Total periods of non-employment for accessions for Male and age 22-24		01708	4	N
	NA_1A04			
Total periods of non-employment for accessions for Male and age 25-34		01712	4	N
	NA_1A05			
Total periods of non-employment for accessions for Male and age 35-44		01716	4	N
	NA_1A06			
Total periods of non-employment for accessions for Male and age 45-54		01720	4	N
	NA_1A07			
Total periods of non-employment for accessions for Male and age 55-64		01724	4	N
	NA_1A08			
Total periods of non-employment for accessions for Male and age 65-99		01728	4	N
	NH_2A01			
Total periods of non-employment for new hires for Female and age 14-18		02708	4	N
	NH_2A00			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for new hires for Female and age 14-99		02704	4	N
	NH_2A02			
Total periods of non-employment for new hires for Female and age 19-21		02712	4	N
	NH_2A03			
Total periods of non-employment for new hires for Female and age 22-24		02716	4	N
	NH_2A04			
Total periods of non-employment for new hires for Female and age 25-34		02720	4	N
	NH_2A05			
Total periods of non-employment for new hires for Female and age 35-44		02724	4	N
	NH_2A06			
Total periods of non-employment for new hires for Female and age 45-54		02728	4	N
	NH_2A07			
Total periods of non-employment for new hires for Female and age 55-64		02732	4	N
	NH_2A08			
Total periods of non-employment for new hires for Female and age 65-99		02736	4	N
	NH_0A01			
Total periods of non-employment for new hires for Male and Female and age 14-18		02636	4	N
	NH_0A00			
Total periods of non-employment for new hires for Male and Female and age 14-99		02632	4	N
	NH_0A02			
Total periods of non-employment for new hires for Male and Female and age 19-21		02640	4	N
	NH_0A03			
Total periods of non-employment for new hires for Male and Female and age 22-24		02644	4	N
	NH_0A04			
Total periods of non-employment for new hires for Male and Female and age 25-34		02648	4	N
	NH_0A05			
Total periods of non-employment for new hires for Male and Female and age 35-44		02652	4	N
	NH_0A06			
Total periods of non-employment for new hires for Male and Female and age 45-54		02656	4	N
	NH_0A07			
Total periods of non-employment for new hires for Male and Female and age 55-64		02660	4	N
	NH_0A08			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for new hires for Male and Female and age 65-99		02664	4	N
	NH_1A01			
Total periods of non-employment for new hires for Male and age 14-18		02672	4	N
	NH_1A00			
Total periods of non-employment for new hires for Male and age 14-99		02668	4	N
	NH_1A02			
Total periods of non-employment for new hires for Male and age 19-21		02676	4	N
	NH_1A03			
Total periods of non-employment for new hires for Male and age 22-24		02680	4	N
	NH_1A04			
Total periods of non-employment for new hires for Male and age 25-34		02684	4	N
	NH_1A05			
Total periods of non-employment for new hires for Male and age 35-44		02688	4	N
	NH_1A06			
Total periods of non-employment for new hires for Male and age 45-54		02692	4	N
	NH_1A07			
Total periods of non-employment for new hires for Male and age 55-64		02696	4	N
	NH_1A08			
Total periods of non-employment for new hires for Male and age 65-99		02700	4	N
	NR_2A01			
Total periods of non-employment for recalls for Female and age 14-18		02816	4	N
	NR_2A00			
Total periods of non-employment for recalls for Female and age 14-99		02812	4	N
	NR_2A02			
Total periods of non-employment for recalls for Female and age 19-21		02820	4	N
	NR_2A03			
Total periods of non-employment for recalls for Female and age 22-24		02824	4	N
	NR_2A04			
Total periods of non-employment for recalls for Female and age 25-34		02828	4	N
	NR_2A05			
Total periods of non-employment for recalls for Female and age 35-44		02832	4	N
	NR_2A06			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for recalls for Female and age 45-54		02836	4	N
	NR_2A07			
Total periods of non-employment for recalls for Female and age 55-64		02840	4	N
	NR_2A08			
Total periods of non-employment for recalls for Female and age 65-99		02844	4	N
	NR_0A01			
Total periods of non-employment for recalls for Male and Female and age 14-18		02744	4	N
	NR_0A00			
Total periods of non-employment for recalls for Male and Female and age 14-99		02740	4	N
	NR_0A02			
Total periods of non-employment for recalls for Male and Female and age 19-21		02748	4	N
	NR_0A03			
Total periods of non-employment for recalls for Male and Female and age 22-24		02752	4	N
	NR_0A04			
Total periods of non-employment for recalls for Male and Female and age 25-34		02756	4	N
	NR_0A05			
Total periods of non-employment for recalls for Male and Female and age 35-44		02760	4	N
	NR_0A06			
Total periods of non-employment for recalls for Male and Female and age 45-54		02764	4	N
	NR_0A07			
Total periods of non-employment for recalls for Male and Female and age 55-64		02768	4	N
	NR_0A08			
Total periods of non-employment for recalls for Male and Female and age 65-99		02772	4	N
	NR_1A01			
Total periods of non-employment for recalls for Male and age 14-18		02780	4	N
	NR_1A00			
Total periods of non-employment for recalls for Male and age 14-99		02776	4	N
	NR_1A02			
Total periods of non-employment for recalls for Male and age 19-21		02784	4	N
	NR_1A03			
Total periods of non-employment for recalls for Male and age 22-24		02788	4	N
	NR_1A04			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for recalls for Male and age 25-34		02792	4	N
	NR_1A05			
Total periods of non-employment for recalls for Male and age 35-44		02796	4	N
	NR_1A06			
Total periods of non-employment for recalls for Male and age 45-54		02800	4	N
	NR_1A07			
Total periods of non-employment for recalls for Male and age 55-64		02804	4	N
	NR_1A08			
Total periods of non-employment for recalls for Male and age 65-99		02808	4	N
	NS_2A01			
Total periods of non-employment for separations for Female and age 14-18		02276	4	N
	NS_2A00			
Total periods of non-employment for separations for Female and age 14-99		02272	4	N
	NS_2A02			
Total periods of non-employment for separations for Female and age 19-21		02280	4	N
	NS_2A03			
Total periods of non-employment for separations for Female and age 22-24		02284	4	N
	NS_2A04			
Total periods of non-employment for separations for Female and age 25-34		02288	4	N
	NS_2A05			
Total periods of non-employment for separations for Female and age 35-44		02292	4	N
	NS_2A06			
Total periods of non-employment for separations for Female and age 45-54		02296	4	N
	NS_2A07			
Total periods of non-employment for separations for Female and age 55-64		02300	4	N
	NS_2A08			
Total periods of non-employment for separations for Female and age 65-99		02304	4	N
	NS_0A01			
Total periods of non-employment for separations for Male and Female and age 14-18		02204	4	N
	NS_0A00			
Total periods of non-employment for separations for Male and Female and age 14-99		02200	4	N
	NS_0A02			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Total periods of non-employment for separations for Male and Female and age 19-21		02208	4	N
	NS_0A03			
Total periods of non-employment for separations for Male and Female and age 22-24		02212	4	N
	NS_0A04			
Total periods of non-employment for separations for Male and Female and age 25-34		02216	4	N
	NS_0A05			
Total periods of non-employment for separations for Male and Female and age 35-44		02220	4	N
	NS_0A06			
Total periods of non-employment for separations for Male and Female and age 45-54		02224	4	N
	NS_0A07			
Total periods of non-employment for separations for Male and Female and age 55-64		02228	4	N
	NS_0A08			
Total periods of non-employment for separations for Male and Female and age 65-99		02232	4	N
	NS_1A01			
Total periods of non-employment for separations for Male and age 14-18		02240	4	N
	NS_1A00			
Total periods of non-employment for separations for Male and age 14-99		02236	4	N
	NS_1A02			
Total periods of non-employment for separations for Male and age 19-21		02244	4	N
	NS_1A03			
Total periods of non-employment for separations for Male and age 22-24		02248	4	N
	NS_1A04			
Total periods of non-employment for separations for Male and age 25-34		02252	4	N
	NS_1A05			
Total periods of non-employment for separations for Male and age 35-44		02256	4	N
	NS_1A06			
Total periods of non-employment for separations for Male and age 45-54		02260	4	N
	NS_1A07			
Total periods of non-employment for separations for Male and age 55-64		02264	4	N
	NS_1A08			
Total periods of non-employment for separations for Male and age 65-99		02268	4	N
	LEG_WIB			

Field name	Data dictionary reference name	Starting position	Field size	Data type
WIB code, wwwwww		05788	6	A/N
	QWI_UNIT_WEIGHT			
Weight such that weighted sum of B_UI = sum(month1_BLS)		00000	8	N
	YEAR			
Year YYYY		05764	3	N
	QWI_FINAL_WEIGHT			
qwi_wcf*qwi_unit_weight		00024	8	N

11.3.5   Summary information on datasets

Table 11.3: Number of observations for QWI

Group	Number of datafiles	Records (1000s)	Filesize (GB)
QWI	200	1,894,666	1738

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 11.4: List of data files for QWI, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Alaska (ak)					
qwi_ak_seinunit_estabtot	2000Q1	2012Q1	884	< 1	1:424617
qwi_ak_seinunit_rh	2000Q1	2012Q1	884	1	1:424618
qwi_ak_seinunit_se	2000Q1	2012Q1	852	1	1:424619
qwi_ak_seinunit_wia	2000Q1	2012Q1	884	1	1:424620
Alabama (al)					
qwi_al_seinunit_estabtot	2001Q1	2012Q1	4,882	1	1:423762
qwi_al_seinunit_rh	2001Q1	2012Q1	4,882	5	1:423763
qwi_al_seinunit_se	2001Q1	2012Q1	4,811	5	1:423764
qwi_al_seinunit_wia	2001Q1	2012Q1	4,882	7	1:423765
Arkansas (ar)					
qwi_ar_seinunit_estabtot	2002Q3	2012Q1	2,784	1	1:423947
qwi_ar_seinunit_rh	2002Q3	2012Q1	2,784	3	1:423948
qwi_ar_seinunit_se	2002Q3	2012Q1	2,739	3	1:423949
qwi_ar_seinunit_wia	2002Q3	2012Q1	2,784	4	1:423950
Arizona (az)					
qwi_az_seinunit_estabtot	2004Q1	2012Q1	4,047	1	1:421868
qwi_az_seinunit_rh	2004Q1	2012Q1	4,047	4	1:421869
qwi_az_seinunit_se	2004Q1	2012Q1	3,982	4	1:421870
qwi_az_seinunit_wia	2004Q1	2012Q1	4,047	6	1:421871
California (ca)					
qwi_ca_seinunit_estabtot	1991Q3	2012Q1	81,984	18	1:433949
qwi_ca_seinunit_rh	1991Q3	2012Q1	81,984	86	1:433950
qwi_ca_seinunit_se	1991Q3	2012Q1	79,295	73	1:433951
qwi_ca_seinunit_wia	1991Q3	2012Q1	81,984	111	1:433952
Colorado (co)					
qwi_co_seinunit_estabtot	1993Q2	2012Q1	10,144	2	1:445200
qwi_co_seinunit_rh	1993Q2	2012Q1	10,144	11	1:445201
qwi_co_seinunit_se	1993Q2	2012Q1	9,964	10	1:445202
qwi_co_seinunit_wia	1993Q2	2012Q1	10,144	14	1:445203
Connecticut (ct)					
qwi_ct_seinunit_estabtot	1996Q1	2012Q1	6,350	1	1:425927

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 11.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
qwi_ct_seinunit_rh	1996Q1	2012Q1	6,350	7	1:425928
qwi_ct_seinunit_se	1996Q1	2012Q1	6,224	6	1:425929
qwi_ct_seinunit_wia	1996Q1	2012Q1	6,350	9	1:425930
District of Columbia (dc)					
qwi_dc_seinunit_estabtot	2005Q2	2012Q1	671	< 1	1:424107
qwi_dc_seinunit_rh	2005Q2	2012Q1	671	1	1:424108
qwi_dc_seinunit_se	2005Q2	2012Q1	662	1	1:424109
qwi_dc_seinunit_wia	2005Q2	2012Q1	671	1	1:424110
Delaware (de)					
qwi_de_seinunit_estabtot	1998Q3	2012Q1	1,290	< 1	1:431814
qwi_de_seinunit_rh	1998Q3	2012Q1	1,290	1	1:431818
qwi_de_seinunit_se	1998Q3	2012Q1	1,273	1	1:431822
qwi_de_seinunit_wia	1998Q3	2012Q1	1,290	2	1:431826
Florida (fl)					
qwi_fl_seinunit_estabtot	1992Q4	2012Q1	33,694	8	1:433875
qwi_fl_seinunit_rh	1992Q4	2012Q1	33,694	36	1:433877
qwi_fl_seinunit_se	1992Q4	2012Q1	33,270	32	1:433879
qwi_fl_seinunit_wia	1992Q4	2012Q1	33,694	48	1:433881
Georgia (ga)					
qwi_ga_seinunit_estabtot	1998Q1	2012Q1	12,253	3	1:424929
qwi_ga_seinunit_rh	1998Q1	2012Q1	12,253	14	1:424930
qwi_ga_seinunit_se	1998Q1	2012Q1	12,089	12	1:424931
qwi_ga_seinunit_wia	1998Q1	2012Q1	12,253	18	1:424932
Hawaii (hi)					
qwi_hi_seinunit_estabtot	1995Q4	2012Q1	2,108	< 1	1:430751
qwi_hi_seinunit_rh	1995Q4	2012Q1	2,108	3	1:430752
qwi_hi_seinunit_se	1995Q4	2012Q1	2,084	2	1:430753
qwi_hi_seinunit_wia	1995Q4	2012Q1	2,108	3	1:430754
Iowa (ia)					
qwi_ia_seinunit_estabtot	1998Q4	2012Q1	4,751	1	1:431606
qwi_ia_seinunit_rh	1998Q4	2012Q1	4,751	5	1:431607
qwi_ia_seinunit_se	1998Q4	2012Q1	4,653	5	1:431608
qwi_ia_seinunit_wia	1998Q4	2012Q1	4,751	7	1:431609
Idaho (id)					
qwi_id_seinunit_estabtot	1991Q1	2012Q1	3,513	1	1:428137
qwi_id_seinunit_rh	1991Q1	2012Q1	3,513	3	1:428138
qwi_id_seinunit_se	1991Q1	2012Q1	3,418	3	1:428140
qwi_id_seinunit_wia	1991Q1	2012Q1	3,513	5	1:428142
Illinois (il)					
qwi_il_seinunit_estabtot	1990Q1	2012Q1	24,380	6	1:442987
qwi_il_seinunit_rh	1990Q1	2012Q1	24,380	25	1:442988
qwi_il_seinunit_se	1990Q1	2012Q1	23,952	23	1:442989
qwi_il_seinunit_wia	1990Q1	2012Q1	24,380	34	1:442990
Indiana (in)					
qwi_in_seinunit_estabtot	1998Q1	2012Q1	8,042	2	1:430898

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 11.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
qwi_in_seinunit_rh	1998Q1	2012Q1	8,042	9	1:430899
qwi_in_seinunit_se	1998Q1	2012Q1	7,922	8	1:430900
qwi_in_seinunit_wia	1998Q1	2012Q1	8,042	12	1:430901
Kansas (ks)					
qwi_ks_seinunit_estabtot	1993Q1	2012Q1	5,842	1	1:432361
qwi_ks_seinunit_rh	1993Q1	2012Q1	5,842	6	1:432362
qwi_ks_seinunit_se	1993Q1	2012Q1	5,707	5	1:432363
qwi_ks_seinunit_wia	1993Q1	2012Q1	5,842	8	1:432364
Kentucky (ky)					
qwi_ky_seinunit_estabtot	2001Q1	2012Q1	4,164	1	1:425570
qwi_ky_seinunit_rh	2001Q1	2012Q1	4,164	4	1:425571
qwi_ky_seinunit_se	2001Q1	2012Q1	4,106	4	1:425572
qwi_ky_seinunit_wia	2001Q1	2012Q1	4,164	6	1:425573
Louisiana (la)					
qwi_la_seinunit_estabtot	1995Q1	2012Q1	7,137	2	1:443830
qwi_la_seinunit_rh	1995Q1	2012Q1	7,137	8	1:443831
qwi_la_seinunit_se	1995Q1	2012Q1	7,024	7	1:443832
qwi_la_seinunit_wia	1995Q1	2012Q1	7,137	10	1:443833
Maryland (md)					
qwi_md_seinunit_estabtot	1990Q1	2012Q1	11,370	3	1:431315
qwi_md_seinunit_rh	1990Q1	2012Q1	11,370	12	1:431316
qwi_md_seinunit_se	1990Q1	2012Q1	11,160	11	1:431317
qwi_md_seinunit_wia	1990Q1	2012Q1	11,370	16	1:431318
Maine (me)					
qwi_me_seinunit_estabtot	1996Q2	2012Q1	2,663	1	1:434602
qwi_me_seinunit_rh	1996Q2	2012Q1	2,663	3	1:434603
qwi_me_seinunit_se	1996Q2	2012Q1	2,611	2	1:434604
qwi_me_seinunit_wia	1996Q2	2012Q1	2,663	4	1:434605
Michigan (mi)					
qwi_mi_seinunit_estabtot	2000Q3	2012Q1	9,491	2	1:437453
qwi_mi_seinunit_rh	2000Q3	2012Q1	9,491	10	1:437455
qwi_mi_seinunit_se	2000Q3	2012Q1	9,323	9	1:437456
qwi_mi_seinunit_wia	2000Q3	2012Q1	9,491	13	1:437457
Minnesota (mn)					
qwi_mn_seinunit_estabtot	1994Q3	2012Q1	8,541	2	1:423143
qwi_mn_seinunit_rh	1994Q3	2012Q1	8,541	8	1:423145
qwi_mn_seinunit_se	1994Q3	2012Q1	8,333	8	1:423147
qwi_mn_seinunit_wia	1994Q3	2012Q1	8,541	11	1:423149
Missouri (mo)					
qwi_mo_seinunit_estabtot	1995Q1	2012Q1	10,674	2	1:429473
qwi_mo_seinunit_rh	1995Q1	2012Q1	10,674	11	1:429475
qwi_mo_seinunit_se	1995Q1	2012Q1	10,436	10	1:429477
qwi_mo_seinunit_wia	1995Q1	2012Q1	10,674	15	1:429479
Mississippi (ms)					
qwi_ms_seinunit_estabtot	2003Q3	2012Q1	2,206	1	1:428073

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 11.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
qwi_ms_seinunit_rh	2003Q3	2012Q1	2,206	2	1:428075
qwi_ms_seinunit_se	2003Q3	2012Q1	2,173	2	1:428077
qwi_ms_seinunit_wia	2003Q3	2012Q1	2,206	3	1:428079
Montana (mt)					
qwi_mt_seinunit_estabtot	1993Q1	2012Q1	2,613	1	1:432169
qwi_mt_seinunit_rh	1993Q1	2012Q1	2,613	3	1:432171
qwi_mt_seinunit_se	1993Q1	2012Q1	2,549	2	1:432174
qwi_mt_seinunit_wia	1993Q1	2012Q1	2,613	3	1:432177
North Carolina (nc)					
qwi_nc_seinunit_estabtot	1992Q4	2011Q4	16,052	4	1:445348
qwi_nc_seinunit_rh	1992Q4	2011Q4	16,052	18	1:445349
qwi_nc_seinunit_se	1992Q4	2011Q4	15,789	16	1:445350
qwi_nc_seinunit_wia	1992Q4	2011Q4	16,052	24	1:445351
North Dakota (nd)					
qwi_nd_seinunit_estabtot	1998Q1	2012Q1	1,266	< 1	1:432193
qwi_nd_seinunit_rh	1998Q1	2012Q1	1,266	1	1:432195
qwi_nd_seinunit_se	1998Q1	2012Q1	1,236	1	1:432197
qwi_nd_seinunit_wia	1998Q1	2012Q1	1,266	2	1:432199
Nebraska (ne)					
qwi_ne_seinunit_estabtot	1999Q1	2012Q1	2,901	1	1:432204
qwi_ne_seinunit_rh	1999Q1	2012Q1	2,901	3	1:432205
qwi_ne_seinunit_se	1999Q1	2012Q1	2,841	3	1:432206
qwi_ne_seinunit_wia	1999Q1	2012Q1	2,901	4	1:432207
New Hampshire (nh)					
qwi_nh_seinunit_estabtot	2003Q1	2012Q1	1,651	< 1	1:436663
qwi_nh_seinunit_rh	2003Q1	2012Q1	1,651	2	1:436665
qwi_nh_seinunit_se	2003Q1	2012Q1	1,617	2	1:436667
qwi_nh_seinunit_wia	2003Q1	2012Q1	1,651	2	1:436669
New Jersey (nj)					
qwi_nj_seinunit_estabtot	1996Q1	2012Q1	13,829	3	1:422572
qwi_nj_seinunit_rh	1996Q1	2012Q1	13,829	15	1:422573
qwi_nj_seinunit_se	1996Q1	2012Q1	13,617	13	1:422574
qwi_nj_seinunit_wia	1996Q1	2012Q1	13,829	19	1:422575
New Mexico (nm)					
qwi_nm_seinunit_estabtot	1995Q3	2012Q1	3,050	1	1:433216
qwi_nm_seinunit_rh	1995Q3	2012Q1	3,050	3	1:433217
qwi_nm_seinunit_se	1995Q3	2012Q1	2,994	3	1:433218
qwi_nm_seinunit_wia	1995Q3	2012Q1	3,050	4	1:433219
Nevada (nv)					
qwi_nv_seinunit_estabtot	1998Q1	2012Q1	3,266	1	1:443683
qwi_nv_seinunit_rh	1998Q1	2012Q1	3,266	4	1:443684
qwi_nv_seinunit_se	1998Q1	2012Q1	3,217	3	1:443685
qwi_nv_seinunit_wia	1998Q1	2012Q1	3,266	5	1:443686
New York (ny)					
qwi_ny_seinunit_estabtot	2000Q1	2012Q1	24,408	5	1:443443

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

Table 11.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
qwi_ny_seinunit_rh	2000Q1	2012Q1	24,408	26	1:443444
qwi_ny_seinunit_se	2000Q1	2012Q1	24,063	22	1:443445
qwi_ny_seinunit_wia	2000Q1	2012Q1	24,408	34	1:443446
Ohio (oh)					
qwi_oh_seinunit_estabtot	2000Q1	2012Q1	12,980	3	1:444124
qwi_oh_seinunit_rh	2000Q1	2012Q1	12,980	14	1:444125
qwi_oh_seinunit_se	2000Q1	2012Q1	12,788	13	1:444126
qwi_oh_seinunit_wia	2000Q1	2012Q1	12,980	20	1:444127
Oklahoma (ok)					
qwi_ok_seinunit_estabtot	2000Q1	2012Q1	4,359	1	1:422382
qwi_ok_seinunit_rh	2000Q1	2012Q1	4,359	5	1:422383
qwi_ok_seinunit_se	2000Q1	2012Q1	4,287	4	1:422384
qwi_ok_seinunit_wia	2000Q1	2012Q1	4,359	6	1:422385
Oregon (or)					
qwi_or_seinunit_estabtot	1991Q1	2012Q1	9,124	2	1:422813
qwi_or_seinunit_rh	1991Q1	2012Q1	9,124	9	1:422814
qwi_or_seinunit_se	1991Q1	2012Q1	8,923	9	1:422815
qwi_or_seinunit_wia	1991Q1	2012Q1	9,124	13	1:422816
Pennsylvania (pa)					
qwi_pa_seinunit_estabtot	1997Q1	2012Q1	18,766	4	1:426206
qwi_pa_seinunit_rh	1997Q1	2012Q1	18,766	19	1:426207
qwi_pa_seinunit_se	1997Q1	2012Q1	18,406	18	1:426208
qwi_pa_seinunit_wia	1997Q1	2012Q1	18,766	27	1:426209
Rhode Island (ri)					
qwi_ri_seinunit_estabtot	1995Q1	2012Q1	2,060	< 1	1:436987
qwi_ri_seinunit_rh	1995Q1	2012Q1	2,060	2	1:436988
qwi_ri_seinunit_se	1995Q1	2012Q1	2,025	2	1:436989
qwi_ri_seinunit_wia	1995Q1	2012Q1	2,060	3	1:436990
South Carolina (sc)					
qwi_sc_seinunit_estabtot	1998Q1	2012Q1	5,714	1	1:439919
qwi_sc_seinunit_rh	1998Q1	2012Q1	5,714	6	1:439920
qwi_sc_seinunit_se	1998Q1	2012Q1	5,620	6	1:439921
qwi_sc_seinunit_wia	1998Q1	2012Q1	5,714	8	1:439922
South Dakota (sd)					
qwi_sd_seinunit_estabtot	1998Q1	2012Q1	1,509	< 1	1:432803
qwi_sd_seinunit_rh	1998Q1	2012Q1	1,509	1	1:432806
qwi_sd_seinunit_se	1998Q1	2012Q1	1,474	1	1:432809
qwi_sd_seinunit_wia	1998Q1	2012Q1	1,509	2	1:432812
Tennessee (tn)					
qwi_tn_seinunit_estabtot	1998Q1	2012Q1	7,214	2	1:428431
qwi_tn_seinunit_rh	1998Q1	2012Q1	7,214	8	1:428432
qwi_tn_seinunit_se	1998Q1	2012Q1	7,083	7	1:428433
qwi_tn_seinunit_wia	1998Q1	2012Q1	7,214	10	1:428434
Texas (tx)					
qwi_tx_seinunit_estabtot	1995Q1	2012Q1	34,362	8	1:427223

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 11.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
qwi_tx_seinunit_rh	1995Q1	2012Q1	34,362	39	1:427224
qwi_tx_seinunit_se	1995Q1	2012Q1	33,760	35	1:427225
qwi_tx_seinunit_wia	1995Q1	2012Q1	34,362	52	1:427226
Utah (ut)					
qwi_ut_seinunit_estabtot	1999Q3	2012Q1	3,596	1	1:434460
qwi_ut_seinunit_rh	1999Q3	2012Q1	3,596	4	1:434462
qwi_ut_seinunit_se	1999Q3	2012Q1	3,494	3	1:434464
qwi_ut_seinunit_wia	1999Q3	2012Q1	3,596	5	1:434466
Virginia (va)					
qwi_va_seinunit_estabtot	1998Q1	2012Q1	10,840	3	1:439762
qwi_va_seinunit_rh	1998Q1	2012Q1	10,840	12	1:439763
qwi_va_seinunit_se	1998Q1	2012Q1	10,660	11	1:439764
qwi_va_seinunit_wia	1998Q1	2012Q1	10,840	16	1:439765
Vermont (vt)					
qwi_vt_seinunit_estabtot	2000Q1	2012Q1	1,105	< 1	1:432808
qwi_vt_seinunit_rh	2000Q1	2012Q1	1,105	1	1:432811
qwi_vt_seinunit_se	2000Q1	2012Q1	1,083	1	1:432814
qwi_vt_seinunit_wia	2000Q1	2012Q1	1,105	2	1:432816
Washington (wa)					
qwi_wa_seinunit_estabtot	1990Q1	2012Q1	15,654	4	1:428406
qwi_wa_seinunit_rh	1990Q1	2012Q1	15,654	16	1:428408
qwi_wa_seinunit_se	1990Q1	2012Q1	14,973	14	1:428410
qwi_wa_seinunit_wia	1990Q1	2012Q1	15,654	21	1:428412
Wisconsin (wi)					
qwi_wi_seinunit_estabtot	1990Q1	2012Q1	12,247	3	1:423227
qwi_wi_seinunit_rh	1990Q1	2012Q1	12,247	12	1:423228
qwi_wi_seinunit_se	1990Q1	2012Q1	11,960	11	1:423229
qwi_wi_seinunit_wia	1990Q1	2012Q1	12,247	17	1:423230
West Virginia (wv)					
qwi_wv_seinunit_estabtot	1997Q1	2012Q1	2,767	1	1:441521
qwi_wv_seinunit_rh	1997Q1	2012Q1	2,767	3	1:441522
qwi_wv_seinunit_se	1997Q1	2012Q1	2,725	3	1:441523
qwi_wv_seinunit_wia	1997Q1	2012Q1	2,767	4	1:441524
Wyoming (wy)					
qwi_wy_seinunit_estabtot	2001Q1	2012Q1	961	< 1	1:432815
qwi_wy_seinunit_rh	2001Q1	2012Q1	961	1	1:432817
qwi_wy_seinunit_se	2001Q1	2012Q1	940	1	1:432818
qwi_wy_seinunit_wia	2001Q1	2012Q1	961	1	1:432819

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

11.4   NOTES

- Alabama (AL), Kansas (KS), and South Carolina (SC) are currently missing from data archive. A request has been put in to include them.

Chapter 12.
Quarterly Workforce Indicators - Public-use files (QWIPU)

12.1 OVERVIEW

The Quarterly Workforce Indicators (QWI) provide local labor market statistics by industry, worker demographics, employer age and size.[1] Unlike statistics tabulated from firm or person-level data, the QWI source data are unique job-level data that link workers to their employers. Because of this link, labor market data in the QWI is available by worker age, sex, educational attainment, and race/ethnicity. This allows for analysis by demographics of a particular local labor market or industry for instance, identifying industries with aging workforces. Links between workers and firms also allow the QWI to identify worker flows hires, separations, and turnover as well as net employment growth. As most hiring activity is the consequence of worker turnover rather than employment growth, a focus on employment growth alone may misrepresent employment opportunity in the local labor market. Wages by industry and demographics as well as by whether the worker was newly hired are also available. QWI wages for new hires can be compared to wages for continuing workers, and wage growth for similar workers across industries can be compared to identify important local labor market trends.

The source data for the QWI is the Longitudinal Employer-Household Dynamics (LEHD) linked employer-employee microdata. The LEHD data is a longitudinal database covering over 95% of U.S. private sector jobs. Much of this data is collected via a unique federal-state data sharing collaboration, the Local Employment Dynamics (LED) partnership. LED is a cooperative venture between the LEHD program at the U.S. Census Bureau and state agencies of all 50 states, the District of Columbia, Puerto Rico, and the U.S. Virgin Islands. Partner states voluntarily submit quarterly data files from existing administrative record systems, which are combined with a range of other data sources to generate public use products, including QWI and LODES (LEHD Origin-Destination Employment Statistics, presented in OnTheMap) and other new products in development. By integrating data used to administer public programs with existing census and surveys, a new national jobs database is generated at very low cost and with no additional respondent burden.

12.2 DATA AVAILABILITY

Time is reported on the QWI by specifying a year and calendar quarter (Jan-Mar, Apr-Jun, Jul-Sep, Oct-Nov). The extent of the time series available will vary by state, based on the availability of historical data when joining the partnership. The earliest state time series begin in 1990 (see Table 12.1 for details, Figure 12.1 summarizes it graphically). As each state is processed independently, the latest available data across states may also vary at any point in time. Not all measures will be present in all quarters, as most require leading or lagging data for calculation. In addition, [...] data available for firm age and size tabulations will be one quarter behind other tabulations because of additional input data requirements.

---

1. Excerpted from *Quarterly Workforce Indicators 101*, available online at http://lehd.ces.census.gov/doc/QWI_101.pdf.

Figure 12.1: Data availability (QWIPU) by state
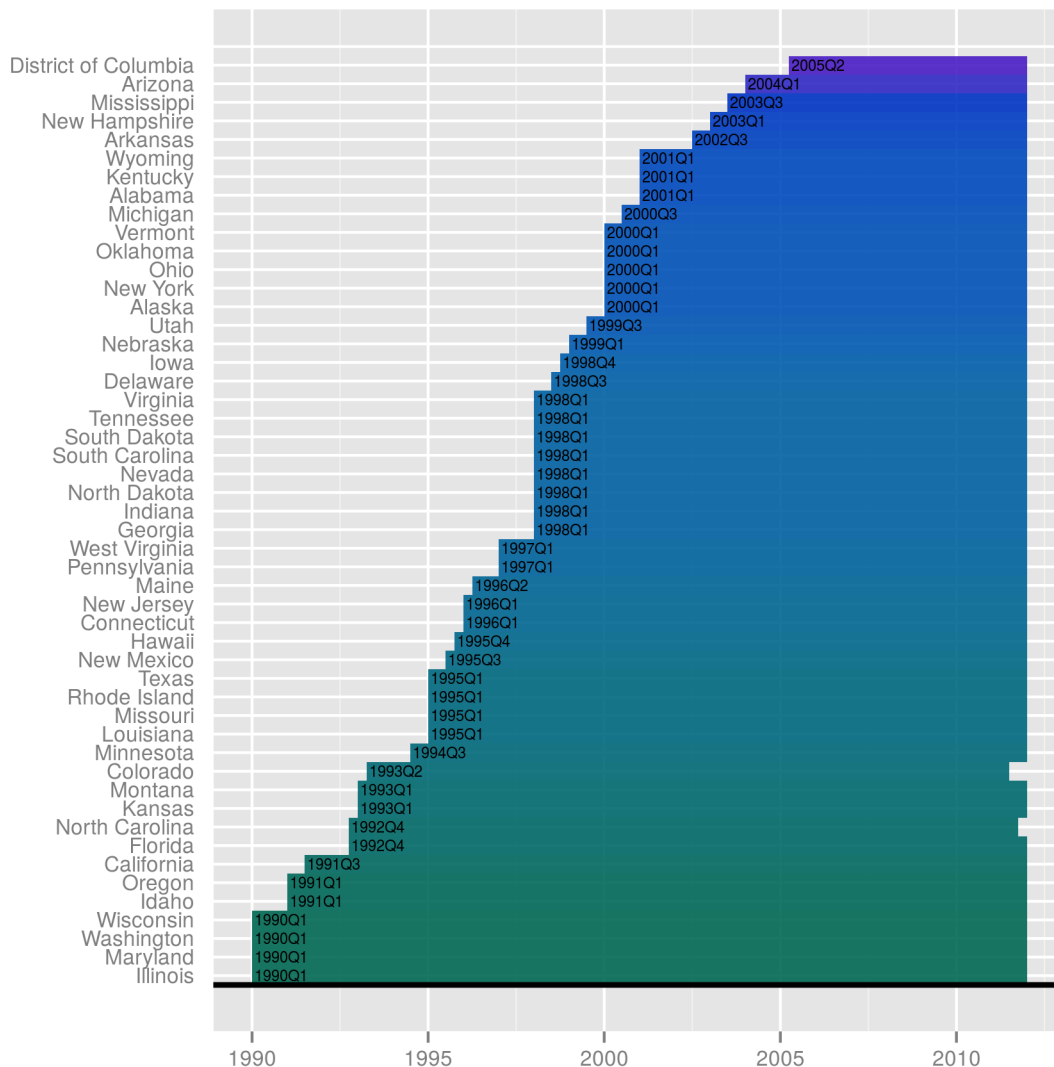
12.3   QWI DATA RELEASES

The QWI are produced on a quarterly schedule. [...] In the event that data submission or data quality issues are encountered, QWI production for a state may be skipped for one or several quarters, until the issue can be resolved. The S2011snapshot version corresponds to R2012Q4 for most states, except where otherwise noted.

12.4   UPDATES AND REVISIONS

The complete QWI time series is recalculated with every release, so numbers may change in any quarter. These changes are due to a number of factors, including:

- Updates to input files (primarily UI and QCEW)

  – States typically make a second submission of the previous quarters data in every quarter, to improve completeness of data reporting. Historical files may be resubmitted to improve data quality.

  – Other input data sources are also periodically updated.

- Modifications to algorithms to develop estimates

  – The data quality of the QWI is continuously reviewed, and the algorithms are periodically modified to improve the results. These modifications may affect measures throughout the time series.

- Stochastic changes to imputations used to complete missing information

  – Random draws are used to generate data that are missing. These draws may change between production runs, though longitudinal consistency is generally maintained within a data release.

For this reason, analyses using the public-use QWIPU should always reference the correct release of the data. The `version.txt` file contains the metadata of each state's release, and identifies the release version of the data, both on the internet as well as on the RDC. Each state will have their own `version.txt` file.

12.4.1   Changes in this Snapshot

The most frequently used files outside of the RDC are the QWIPU tabulations by NAICS sub-sector (NAICS3) and county, by the "classic" age-by-sex ("WIA"), sex-by-education (SE), and race-by-ethnicity (RH) tabulations, as well as the beta-release of firm-age and firm-size tabulations by those same demographic classifications. The files are consistent with the overall snapshot (R2012Q4). The total size is approximately 1TB.

Note that the use of the QWI_SEINUNIT files is incompatible with the use of the QWI public-use files. Researchers must choose one or the other. However, use of the QWI public-use files is not subject to any approvals.

12.5   DATA CITATION

> U.S. Census Bureau. 2014. *Quarterly Workforce Indicators (QWI) for establishments, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

12.6   DATA SET DESCRIPTIONS

The data are provided as compressed CSV files, and need to be read-in by individual researchers. SAS readin programs are made available.

	State	FIPS	Release	Start	End
1	AK	2	R2012Q4	2000Q1	2012Q1
2	AL	1	R2012Q4	2001Q1	2012Q1
3	AR	5	R2012Q4	2002Q3	2012Q1
4	AZ	4	R2012Q4	2004Q1	2012Q1
5	CA	6	R2012Q4	1991Q3	2012Q1
6	CO	8	R2012Q2	1993Q2	2011Q3
7	CT	9	R2012Q4	1996Q1	2012Q1
8	DC	11	R2012Q4	2005Q2	2012Q1
9	DE	10	R2012Q4	1998Q3	2012Q1
10	FL	12	R2012Q4	1992Q4	2012Q1
11	GA	13	R2012Q4	1998Q1	2012Q1
12	HI	15	R2012Q4	1995Q4	2012Q1
13	IA	19	R2012Q4	1998Q4	2012Q1
14	ID	16	R2012Q4	1991Q1	2012Q1
15	IL	17	R2012Q4	1990Q1	2012Q1
16	IN	18	R2012Q4	1998Q1	2012Q1
17	KS	20	R2012Q4	1993Q1	2012Q1
18	KY	21	R2012Q4	2001Q1	2012Q1
19	LA	22	R2012Q4	1995Q1	2012Q1
20	MD	24	R2012Q4	1990Q1	2012Q1
21	ME	23	R2012Q4	1996Q2	2012Q1
22	MI	26	R2012Q4	2000Q3	2012Q1
23	MN	27	R2012Q4	1994Q3	2012Q1
24	MO	29	R2012Q4	1995Q1	2012Q1
25	MS	28	R2012Q4	2003Q3	2012Q1
26	MT	30	R2012Q4	1993Q1	2012Q1
27	NC	37	R2012Q3	1992Q4	2011Q4
28	ND	38	R2012Q4	1998Q1	2012Q1
29	NE	31	R2012Q4	1999Q1	2012Q1
30	NH	33	R2012Q4	2003Q1	2012Q1
31	NJ	34	R2012Q4	1996Q1	2012Q1
32	NM	35	R2012Q4	1995Q3	2012Q1
33	NV	32	R2012Q4	1998Q1	2012Q1
34	NY	36	R2012Q4	2000Q1	2012Q1
35	OH	39	R2012Q4	2000Q1	2012Q1
36	OK	40	R2012Q4	2000Q1	2012Q1
37	OR	41	R2012Q4	1991Q1	2012Q1
38	PA	42	R2012Q4	1997Q1	2012Q1
39	RI	44	R2012Q4	1995Q1	2012Q1
40	SC	45	R2012Q4	1998Q1	2012Q1
41	SD	46	R2012Q4	1998Q1	2012Q1
42	TN	47	R2012Q4	1998Q1	2012Q1
43	TX	48	R2012Q4	1995Q1	2012Q1
44	UT	49	R2012Q4	1999Q3	2012Q1
45	VA	51	R2012Q4	1998Q1	2012Q1
46	VT	50	R2012Q4	2000Q1	2012Q1
47	WA	53	R2012Q4	1990Q1	2012Q1
48	WI	55	R2012Q4	1990Q1	2012Q1
49	WV	54	R2012Q4	1997Q1	2012Q1
50	WY	56	R2012Q4	2001Q1	2012Q1

The data underlying this table is attached to this document as CSV.

Table 12.1: Time series example

Chapter 13.
Successor-Predecessor file (SPF)

13.1 OVERVIEW

The Successor-Predecessor File (SPF) is a suite of files providing intertemporal flow-based links based on wage records and administrative links. The file is not fully documented, researchers are advised to use the file with caution.

13.2 DATA CITATION

> U.S. Census Bureau. 2014. *Quarterly Workforce Indicators (QWI),*
> *public-use tabulations, S2011 Version.* [Computer file]. Wash-
> ington,DC: U.S. Census Bureau, Center for Economic Studies,
> Research Data Centers [distributor].

13.3 DETAILED DESCRIPTION

13.3.1 Definition of Successor-Predecessor

The successor-predecessor sequence creates two files, the SPF (Successor-Predecessor File) which has a record for every link (whether that link is identified by employee flows from the UI wage records or from the successor-predecessor data on the ES-202) between SEINs, and the WSLF (Within-SEIN Links File) which has a record for every successor-predecessor link reported on the ES-202 between SEINUNITs within the same SEIN.

13.3.2 Update frequency

Quarterly.

13.3.3 Acquisition process

The Successor-Predecessor sequence waits for the creation of the ES-202 files and the EHF.

13.3.4 Processing description

First, we read the PIK-SEIN work history information from the EHF into simple character strings of 1's or 0's referring to whether or not the PIK has positive earnings at the SEIN in the quarter corresponding to the position in the character string. We then match up each end of job string experienced by a PIK with the beginning of job strings for that PIK at another SEIN which start in the same or subsequent quarter that the first job ends. We then sum up the number of such flows between each SEIN pair in a given quarter. If

the number of transitioning employees and the SEINs involved satisfy certain criteria, then a link is recorded for that SEIN pair in that quarter. We then read in the successor-predecessor information from the ES-202 and divide the data into a within-SEIN links file and an across-SEIN links file. The across-SEIN links file is aggregated to the SEIN-level for comparability to the links formed with the UI wage records. Finally, the UI wage record links and the SEIN-level, ES-202 links are merged into one file.

13.3.5 Changes in this Snapshot

The SPF, which computes worker-flows between firms, and tracks administratively recorded successor-predecessor relationships, is available in this release.

13.4 DATA SET DESCRIPTIONS

13.4.1 Naming scheme

All files start with `spf`. The main **spf!** (**spf!**) file has no suffix, other files have a suffix.
    SAS datasets with zero observations are attached to this document:[1]

- spf/spf_zz.sas7bdat

- spf/spf_zz_wslf.sas7bdat

    `ZZ` stands for the state postal abbreviation.

13.4.2 Data location

The files are stored in state-specific subdirectories of the main SPF directory:

    `spf/ZZ/ for most files`

No files in the `spf` process contain Title 26 data. On the RDC network, the directory can be found under

    `/mixed/lehd/s2011`

---

1. Also visible on the attachment tab - Adobe Reader may be required.

13.4.3 UI-based Output Files

13.4.3.1 SPF

The main SPF stores links between SEINs within a state (no cross-state links). Key variables are `link_ui` and `link_es`.

Record identifier SEIN-SEIN_SUCC

Sort order SEIN-SEIN_SUCC

Entity Link between firms

Unique Entity Key SEIN

Field name	Data dictionary reference name	Starting position	Field size	Data type
Data source of link between firms	SOURCE	00116	4	A/N
First quarter predecessor is active on UI	ACTIVE_BEG_QTR_A	00080	3	N
First quarter successor is active on UI	ACTIVE_BEG_QTR_B	00086	3	N
Last quarter predecessor is active on UI	ACTIVE_END_QTR_A	00083	3	N
Last quarter successor is active on UI	ACTIVE_END_QTR_B	00089	3	N
Number of jobs transitioning between firms	NUM_LEFT	00016	8	N
Percent of jobs at predecessor transitioning to successor (estimated)	RATIO	00040	8	N
Percent of jobs at successor transitioning from predecessor (estimated)	SUCC_RATIO	00048	8	N
Percent of transitions where separation precedes quarter of accession	MATCH_PERIOD	00008	8	N
Pred ES202 Month 1 Employment, max of last three quarters	EMP_ES	00064	8	N
Pred UI B Employment, max of last 3 quarters	BPEMP_MASTER	00032	8	N
Quarter of separation, 1985Q1=1	QTIME	00000	8	N
SEIN - predecessor	SEIN	00092	12	A/N
SEIN - successor	SEIN_SUCC	00104	12	A/N
Type of ES202 based link	LINK_ES	00072	8	N
	LINK_UI			

Field name	Data dictionary reference name	Starting position	Field size	Data type
Type of link for predecessor firm		00024	8	N
	SUCC_LINK_UI			
Type of link for successor firm		00056	8	N

Values taken by UI link variables

```
/*   +--< LEHD-QWI spf 3.1.25 2005-04-21 schwa305           >--+   */
/*   +--< Location: /programs/production/dev1/current/spf   >--+   */
/*   +--< File: library/formats/links_ui.sas                >--+   */
/* Time-stamp: <04/10/20 17:58:32 vilhuber> */
/*BEGINCCC
    Format created to tabulate the variable LINK\_UI.

CCCEND*/

proc format;
    value linkui
1="Pred exits & 80% Pred Employment moves to Succ        "
2="Pred exits & <80% Pred Employment moves to Succ       "
3=""
4="Pred does not exit & 80% Pred Employment moves to Succ"
5="Pred does not exit & <80% Pred Employment moves to Scc"
6="" ;
run;
```

Values taken by Successor UI link variables

```
/*   +--< LEHD-QWI spf 3.1.25 2005-04-21 schwa305           >--+   */
/*   +--< Location: /programs/production/dev1/current/spf   >--+   */
/*   +--< File: library/formats/succ_link_ui.sas            >--+   */
/* Time-stamp: <04/10/20 18:04:26 vilhuber> */
/*BEGINCCC
    Format created to tabulate the variable SUCC\_LINK\_UI.

CCCEND*/

proc format;
    value slinkui
1="Succ enters & 80% Succ Employment comes from Pred        "
2="Succ enters & <80% Succ Employment comes from Pred       "
3=""
4="Succ does not enter & 80% Succ Employment comes from Pred"
5="Succ does not enter & <80% Succ Employment comes from Pred"
6="" ;
run;
```

13.4.3.2 SPF-WSLF

Record identifier PIK-SEIN-SEINUNIT

Sort order PIK-SEIN-SEINUNIT

Entity Job

Unique Entity Key PIK-SEIN-SEINUNIT

Field name	Data dictionary reference name	Starting position	Field size	Data type
Predecessor SEIN	PRED_SEIN	00009	12	A/N
Predecessor SEINUNIT	PRED_SEINUNIT	00021	5	A/N
Quarter QQ	QUARTER	00000	4	N
State Employer ID Number	SEIN	00043	12	A/N
State UI Reporting Unit Number	SEINUNIT	00055	5	A/N
Successor SEIN	SUCC_SEIN	00026	12	A/N
Successor SEINUNIT	SUCC_SEINUNIT	00038	5	A/N
Year YYYY	YEAR	00004	5	N

13.4.4 Summary information on datasets

Table 13.3: Number of observations for SPF

Group	Number of datafiles	Records (1000s)	Filesize (GB)
SPF	100	15,361	2

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 13.4: List of data files for SPF, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Alaska (ak)					
spf_ak	2000Q3	2011Q3	16	< 1	1:422208
spf_ak_wslf	2000Q1	2012Q1	1	< 1	1:422212
Alabama (al)					
spf_al	2001Q3	2011Q3	117	< 1	1:422174
spf_al_wslf	2001Q1	2012Q1	9	< 1	1:422177
Arkansas (ar)					
spf_ar	2003Q1	2011Q3	59	< 1	1:423233
spf_ar_wslf	2002Q3	2012Q1	4	< 1	1:423235
Arizona (az)					
spf_az	2004Q3	2011Q3	112	< 1	1:421594
spf_az_wslf	2004Q1	2012Q1	3	< 1	1:421598
California (ca)					
spf_ca	1991Q3	2011Q3	3,015	< 1	1:421579
spf_ca_wslf	1991Q1	2012Q1	130	< 1	1:421581
Colorado (co)					
spf_co	1990Q3	2011Q3	282	< 1	1:444365
spf_co_wslf	1990Q1	2012Q1	15	< 1	1:444369
Connecticut (ct)					
spf_ct	1996Q3	2011Q3	111	< 1	1:422457
spf_ct_wslf	1996Q1	2012Q1	9	< 1	1:422463
District of Columbia (dc)					
spf_dc	2001Q2	2011Q3	13	< 1	1:423347
spf_dc_wslf	2000Q4	2012Q1	< 1	< 1	1:423349
Delaware (de)					
spf_de	1997Q3	2011Q3	30	< 1	1:430001
spf_de_wslf	1997Q1	2012Q1	< 1	< 1	1:430003
Florida (fl)					
spf_fl	1989Q3	2011Q3	1,160	< 1	1:424206
spf_fl_wslf	1989Q1	2012Q1	74	< 1	1:424208
Georgia (ga)					

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:" to obtain the full SnapshotID.

(cont)

Table 13.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
spf_ga	1998Q3	2011Q3	331	< 1	1:423298
spf_ga_wslf	1998Q1	2012Q1	30	< 1	1:423300
Hawaii (hi)					
spf_hi	1996Q2	2011Q3	21	< 1	1:427179
spf_hi_wslf	1995Q4	2012Q1	7	< 1	1:427183
Iowa (ia)					
spf_ia	1990Q3	2011Q3	99	< 1	1:424124
spf_ia_wslf	1990Q1	2012Q1	35	< 1	1:424126
Idaho (id)					
spf_id	1991Q3	2011Q3	84	< 1	1:426043
spf_id_wslf	1991Q1	2012Q1	4	< 1	1:426045
Illinois (il)					
spf_il	1990Q3	2011Q3	714	< 1	1:442774
spf_il_wslf	1990Q1	2012Q1	19	< 1	1:442776
Indiana (in)					
spf_in	1998Q3	2011Q3	252	< 1	1:425172
spf_in_wslf	1998Q1	2012Q1	11	< 1	1:425174
Kansas (ks)					
spf_ks	1990Q3	2011Q3	122	< 1	1:425256
spf_ks_wslf	1990Q1	2012Q1	12	< 1	1:425264
Kentucky (ky)					
spf_ky	2001Q3	2011Q3	111	< 1	1:423795
spf_ky_wslf	2001Q1	2012Q1	5	< 1	1:423797
Louisiana (la)					
spf_la	1990Q3	2011Q3	549	< 1	1:443290
spf_la_wslf	1990Q1	2012Q1	2	< 1	1:443292
Maryland (md)					
spf_md	1990Q3	2011Q3	176	< 1	1:427192
spf_md_wslf	1990Q1	2012Q1	11	< 1	1:427196
Maine (me)					
spf_me	1996Q3	2011Q3	45	< 1	1:432212
spf_me_wslf	1996Q1	2012Q1	4	< 1	1:432216
Michigan (mi)					
spf_mi	1998Q3	2011Q3	304	< 1	1:433037
spf_mi_wslf	1998Q1	2012Q1	7	< 1	1:433039
Minnesota (mn)					
spf_mn	1995Q1	2011Q3	251	< 1	1:421913
spf_mn_wslf	1994Q3	2012Q1	3	< 1	1:421915
Missouri (mo)					
spf_mo	1990Q3	2011Q3	238	< 1	1:426040
spf_mo_wslf	1990Q1	2012Q1	3	< 1	1:426042
Mississippi (ms)					
spf_ms	2004Q1	2011Q3	41	< 1	1:426046
spf_ms_wslf	2003Q3	2012Q1	3	< 1	1:426048

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 13.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Montana (mt)					
spf_mt	1993Q3	2011Q3	39	< 1	1:429995
spf_mt_wslf	1993Q1	2012Q1	< 1	< 1	1:429997
North Carolina (nc)					
spf_nc	1990Q3	2011Q2	640	< 1	1:444349
spf_nc_wslf	1990Q1	2011Q4	27	< 1	1:444353
North Dakota (nd)					
spf_nd	1998Q3	2011Q3	20	< 1	1:429998
spf_nd_wslf	1998Q1	2012Q1	1	< 1	1:430000
Nebraska (ne)					
spf_ne	1999Q3	2011Q3	51	< 1	1:427800
spf_ne_wslf	1999Q1	2012Q1	5	< 1	1:427802
New Hampshire (nh)					
spf_nh	2003Q3	2011Q3	19	< 1	1:431341
spf_nh_wslf	2003Q1	2012Q1	2	< 1	1:431345
New Jersey (nj)					
spf_nj	1995Q3	2011Q3	250	< 1	1:421909
spf_nj_wslf	1995Q1	2012Q1	3	< 1	1:421912
New Mexico (nm)					
spf_nm	1990Q3	2011Q3	72	< 1	1:431610
spf_nm_wslf	1990Q1	2012Q1	5	< 1	1:431612
Nevada (nv)					
spf_nv	1998Q3	2011Q3	162	< 1	1:443447
spf_nv_wslf	1998Q1	2012Q1	7	< 1	1:443449
New York (ny)					
spf_ny	1990Q3	2011Q3	790	< 1	1:443044
spf_ny_wslf	1990Q1	2012Q1	15	< 1	1:443046
Ohio (oh)					
spf_oh	2000Q3	2011Q3	317	< 1	1:443450
spf_oh_wslf	2000Q1	2012Q1	23	< 1	1:443452
Oklahoma (ok)					
spf_ok	1999Q3	2011Q3	109	< 1	1:421916
spf_ok_wslf	1999Q1	2012Q1	4	< 1	1:421918
Oregon (or)					
spf_or	1990Q3	2011Q3	218	< 1	1:421919
spf_or_wslf	1990Q1	2012Q1	29	< 1	1:421921
Pennsylvania (pa)					
spf_pa	1991Q3	2011Q3	510	< 1	1:424203
spf_pa_wslf	1991Q1	2012Q1	73	< 1	1:424205
Rhode Island (ri)					
spf_ri	1990Q3	2011Q3	85	< 1	1:434245
spf_ri_wslf	1990Q1	2012Q1	< 1	< 1	1:434247
South Carolina (sc)					
spf_sc	1998Q3	2011Q3	143	< 1	1:437421

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 13.4 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
spf_sc_wslf	1998Q1	2012Q1	1	< 1	1:437427
South Dakota (sd)					
spf_sd	1998Q3	2011Q3	25	< 1	1:430010
spf_sd_wslf	1998Q1	2012Q1	2	< 1	1:430012
Tennessee (tn)					
spf_tn	1998Q3	2011Q3	275	< 1	1:426210
spf_tn_wslf	1998Q1	2012Q1	9	< 1	1:426212
Texas (tx)					
spf_tx	1990Q3	2011Q3	1,291	< 1	1:423798
spf_tx_wslf	1990Q1	2012Q1	55	< 1	1:423800
Utah (ut)					
spf_ut	1990Q3	2011Q3	120	< 1	1:433247
spf_ut_wslf	1990Q1	2012Q1	12	< 1	1:433249
Virginia (va)					
spf_va	1996Q1	2011Q3	226	< 1	1:437445
spf_va_wslf	1995Q3	2012Q1	36	< 1	1:437451
Vermont (vt)					
spf_vt	2000Q3	2011Q3	13	< 1	1:427770
spf_vt_wslf	2000Q1	2012Q1	2	< 1	1:427776
Washington (wa)					
spf_wa	1990Q3	2011Q3	604	< 1	1:426220
spf_wa_wslf	1990Q1	2012Q1	29	< 1	1:426222
Wisconsin (wi)					
spf_wi	1990Q3	2011Q3	279	< 1	1:421900
spf_wi_wslf	1990Q1	2012Q1	23	< 1	1:421902
West Virginia (wv)					
spf_wv	1990Q3	2011Q3	54	< 1	1:437475
spf_wv_wslf	1990Q1	2012Q1	13	< 1	1:437477
Wyoming (wy)					
spf_wy	2001Q3	2011Q3	17	< 1	1:427797
spf_wy_wslf	2001Q1	2012Q1	2	< 1	1:427799

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

13.5   NOTES

Chapter 14.
Unit-to-Worker Impute - Job location impute (U2W)

14.1 OVERVIEW

The UI records underlying the LEHD Infrastructure files provide neither establishment identifiers (except for Minnesota), nor industry or geographic detail of the establishment, only a firm identifier. Between 60 and 70 percent of state-level employment is in single-unit employers (employers with only one establishment), for which a link through the firm identifier is sufficent to provide such detail. For the remaining 30 to 40 percent of employment, such links have to be imputed. The Unit-to-Worker Impute (U2W) file contains ten imputed establishments for each employee of a multi-unit employer. The file can be linked to other Census Bureau datasets through the PIK and the LEHD SEIN-SEINUNIT.

14.1.1 Changes in this Snapshot

None.

14.2 DATA CITATION

> U.S. Census Bureau. 2014. *Successor-Predecessor Files (SPF) in LEHD Infrastructure, S2011 Version.* [Computer file]. Washington,DC: U.S. Census Bureau, Center for Economic Studies, Research Data Centers [distributor].

14.3 DETAILED DESCRIPTION

The information in this section draws heavily on Abowd et al. (2009); Stephens (2007).

A primary objective of the QWI is to provide employment, job and worker flows, and wage measures at a very detailed levels of geography (place-of-work) and industry. The structure of the administrative data received by LEHD from state partners, however, poses a challenge to achieving this goal. QWI measures are primarily based on the processing of UI wage records which report, with the exception of Minnesota, only the legal employer (SEIN) of the workers. The ES-202 micro-data, however, are comprised of establishment-level records which provide the geographic and industry detail needed to produce the QWI. For employers operating only one establishment within a state, the assignment of establishment-level characteristics to UI wage records is straightforward because there is no distinction between the employer and the establishment. However, approximately 30 to 40 percent of state-level employment is concentrated in employers that operate more than one establishment in that state. For these multi-unit employers, the SEIN on workers' wage records

identifies the legal employer in the ES-202 data, but not the employing establishment (place-of-work). Thus, establishment level characteristics–geography and industry, in particular–are missing data for these multi-unit job histories.

In order to impute establishment-level characteristics to job histories of multi-unit employers, a non-ignorable missing data model with multiple imputation was developed. The model imputes establishment-of-employment using two key characteristics available in the LEHD Infrastructure Files: 1) distance between place-of-work and place-of-residence and 2) the distribution of employment across establishments of multi-unit employers. The distance to work model is estimated using data from Minnesota, where both the SEIN and SEINUNIT identifiers appear on a UI wage record. Then, the posterior distribution of the parameters from this estimation, combined with the actual SEIN and SEINUNIT employment histories from the ES-202 data, are used for multiple imputation of the SEINUNIT associated with for workers in a given SEIN in the data from states other than Minnesota.[1] Emerging from this process is an output file, called the Unit-to-Worker (U2W) file, containing ten imputed establishments for each worker of a multi-unit employer. These implicates are then used in the downstream processing of the QWI.

The U2W process relies on information from each of the four Infrastructure Files–ECF, GAL, EHF, and ICF–as well as the auxiliary SPF file. Within the ECF, the universe of multi-unit employers is identified. For these employers, the ECF also provides establishment-level employment, date-of-birth, and geocodes (which are acquired from the GAL). The SPF contains information on predecessor relationships which may lead to the revision of date-of-birth implied by the ECF. Finally, job histories in the EHF in conjunction with place-of-residence information stored in the ICF provide the necessary worker information needed to estimate and apply the imputation model.

14.3.1 A probability model for employment location

14.3.1.1 Definitions

Let $i = 1, ..., I$ index workers, $j = 1, ..., J$ index employers (SEINs), and $t = 1, ..., T$ index time (quarters). Let R_{jt} denote the number of active establishments at employer j in quarter t, let $\mathfrak{R} = \max_{j,t} R_{jt}$, and $r = 1, ..., \mathfrak{R}$ index establishments. Note that the index r is nested within j. Let N_{jrt} denote the quarter t employment of establishment r in employer j. Finally, if worker i was employed at employer j in t, denote by y_{ijt} the establishment at which the worker was employed.

Let $\mathcal{J}_t$ denote the set of employers active in quarter t, let $\mathcal{I}_{jt}$ denote the set of individuals employed at employer j in quarter t, let $\mathcal{R}_{jt}$ denote the set of active ($N_{jrt} > 0$) establishments at employer j in t, and let $\mathcal{R}_{jt}^i \subset \mathcal{R}_{jt}$ denote the set of active establishments that are feasible for worker i. Feasibility is defined as follows. An establishment $r \in \mathcal{R}_{jt}^i$ if $N_{jrs} > 0$ for every quarter s that i was employed at j.

14.3.1.2 The probability model

Let $p_{ijrt} = \Pr(y_{ijt} = r)$. At the core of the model is the probability statement:

$$p_{ijrt} = \frac{e^{\alpha_{jrt} + x'_{ijrt}\beta}}{\sum_{s \in \mathcal{R}_{jt}^i} e^{\alpha_{jst} + x'_{ijst}\beta}} \tag{14.1}$$

where α_{jrt} is a establishment- and quarter-specific effect, x_{ijrt} is a time-varying vector of characteristics of the worker and establishment, and β measures the effect of characteristics on the probability of being employed at a particular establishment. In the current implementation, x_{ijrt} is a linear spline in the (great-circle) distance

---

1. The actual SEINUNIT coded on the UI wage records is used for Minnesota, and would be used for any other state that provided such data. Note that there are occasional, and rare, discrepancies between the unit structure on the Minnesota wage records and the unit structure on the Minnesota ES-202 data for the same quarter. These discrepancies are resolved during the initial processing of the Minnesota data in its state-specific readin procedures.

between worker i's residence and the physical location of establishment r. The spline has knots at 25, 50, and 100 miles.

Using (14.1), the following likelihood is defined

$$p\left(y|\alpha,\beta,x\right) = \prod_{t=1}^{T} \prod_{j\in\mathcal{J}_t} \prod_{i\in\mathcal{I}_{jt}} \prod_{r\in\mathcal{R}_{jt}^i} \left(p_{ijrt}\right)^{d_{ijrt}} \tag{14.2}$$

where

$$d_{ijrt} = \left\{ \begin{array}{ll} 1 & \text{if } y_{ijt}=r \\ 0 & \text{otherwise} \end{array} \right. \tag{14.3}$$

and where y is the appropriately-dimensioned vector of the outcome variables y_{ijt}, α is the appropriately-dimensioned vector of the α_{jrt}, and x is the appropriately-dimensioned matrix of characteristics x_{ijrt}. For α_{jrt}, a hierarchical Bayesian model based on employment counts N_{jrt} is specified.

The object of interest is the joint posterior distribution of α and β. A uniform prior on β, $p\left(\beta\right) \propto 1$ is assumed. The characterization of $p\left(\alpha,\beta|x,y,N\right)$ is based on the factorization

$$
\begin{array}{rcl}
p\left(\alpha,\beta|x,y,N\right) & = & p\left(\alpha|N\right)p\left(\beta|\alpha,x,y\right) \\
& \propto & p\left(\alpha|N\right)p\left(\beta\right)p\left(y|\alpha,\beta,x\right) \\
& \propto & p\left(\alpha|N\right)p\left(y|\alpha,\beta,x\right).
\end{array} \tag{14.4}
$$

Thus, the joint posterior (14.4) is completely characterized by the posterior of α and the likelihood of y in (14.2). Note (14.2) and (14.4) assume that the employment counts N affect employment location y only through the parameters α.

14.3.1.3 Estimation

The joint posterior $p\left(\alpha,\beta|x,y,N\right)$ is approximated at the posterior mode. In particular, we estimate the posterior mode of $p\left(\beta|\alpha,x,y\right)$ evaluated at the posterior mode of α. From these we compute the posterior modal values of the α_{jrt}, then, maximize the log posterior density

$$\log p\left(\beta|\alpha,x,y\right) \propto \sum_{t=1}^{T} \sum_{j\in\mathcal{J}_t} \sum_{i\in\mathcal{I}_{jt}} \sum_{r\in\mathcal{R}_{jt}^i} d_{ijrt} \left(\alpha_{jrt} + x'_{ijrt}\beta - \log\left(\sum_{s\in\mathcal{R}_{jt}^i} e^{\alpha_{jst}+x'_{ijst}\beta} \right) \right) \tag{14.5}$$

which is evaluated at the posterior modal values of the α_{jrt}, using a modified Newton-Raphson method. The mode-finding exercise is based on the gradient and Hessian of (14.5). In practice, (14.5) is estimated for three employer employment size classes: 1-100 employees, 101-500 employees, and greater than 500 employees, using data for Minnesota.

14.3.2 Imputing place of work

After estimating the probability model using Minnesota data, the posterior distribution of the estimated β parameters is combined with the entity specific posterior distribution of the α parameters in the imputation process for other states. A brief outline of the imputation method, as it relates to the probability model previously discussed, is provided in this section. Emphasis is placed on not only the imputation process itself, but also the preparation of input data.

14.3.2.1   Sketch of the imputation method

Ignoring temporal considerations, 10 implicates are generated as follows. First, using the posterior mean and variance of β estimated from the Minnesota data, we take 10 draws of β from the normal approximation (at the mode) to $p\left(\beta|\alpha, x, y\right)$. Next, using ES-202 employment counts for the establishments, we compute 10 values of α_{jt} based on the hierarchical model for these parameters. Note that these are draws from the exact posterior distribution of the α_{jrt}. The drawn values of α and β are used to draw 10 imputed values of place of work from the asymptotic approximation to the posterior predictive distribution

$$p\left(\tilde{y}|x, y\right) = \int \int p\left(\tilde{y}|\alpha, \beta, x, y\right) p\left(\alpha|N\right) p\left(\beta|\alpha, x, y\right) d\alpha d\beta. \tag{14.6}$$

14.3.2.2   Implementation

Establishment data   Using state-level micro-data, the set of employers (SEINs) that ever operate more that one establishment in a given quarter is identified; these SEINs represent the set of ever-multi-unit employers defined above as the set $\mathcal{J}_t$. For each of these employers, its establishment-level records are identified. For each establishment, latitude and longitude coordinates, parent employer (SEIN) employment, and ES-202 month-one employment[2] for the entire history of the establishment are retained. Those establishments with positive month-one employment in a given quarter characterize $\mathcal{R}_{jt}$, the set of all active establishments. An establishment birth date is identified and, in most cases, is the first quarter in the ES-202 time series in which the establishment has positive month-one employment. For some employers, predecessor relationships are identified in the SPF; in those instances, the establishment date-of-birth is adjusted to coincided with that of the predecessor's.

Worker data   The EHF provides the earnings histories for employees of the ever-multi-unit employers. For each in-scope job (a worker-employer pair), one observation is generated for the *end* of each job spell, where a job spell is defined as a continuum of quarters of positive earnings for worker at a particular employer during which there are no more than 3 consecutive periods of non-positive earnings.[3] The start date of the job history is identified as the first quarter of positive earnings; the end date is the last date of positive earnings.[4]   These job spells characterize the set $\mathcal{I}_{jt}$

Candidates   Once the universe of establishments and workers is identified, data are combined and *a priori* restrictions and feasibility assumptions are imposed. For each quarter of the date series, the history of every job spell that *ends in that quarter* is compared to the history of *every* active (in terms of ES-202 first month employment) establishment of the employing employer (SEIN). The start date of the job spell is compared to the birth date of each establishment. Establishments that were born after the start of a job spell are immediately discarded from the set of candidate establishments. The remaining establishments constitute the set $\mathcal{R}_{jt}^i \subset \mathcal{R}_{jt}$ for a job spell (worker) at a given employer.[5]

Given the structure of the pairing of job spells with candidate establishments, it is clear that within job spell changes of establishment are ruled-out. An establishment is imputed once for each job spell,[6] thereby creating no spurious labor market transitions.

---

2. In rare instances where no ES-202 employment is available, an alternative employment measure based on UI wage record counts may be used.

3. A new hire is defined in the QWI as a worker who accedes to a firm in the current period but was not employed by the same firm in any of the 4 previous periods. A new job spell is created if, for example, a worker leaves a firm for more than 4 quarters and is subsequently re-employed by the same firm.

4. By definition, an end-date for a job spell is not assigned in cases where a quarter of positive earnings at a firm is succeeded by 4 or fewer quarters of non-employment and subsequent re-employment by the same firm.

5. The sample of UI wage and QCEW data chosen for processing of the QWI is such that the start and end dates are the same. Birth and death dates of establishments are, more precisely, the dates associated with the beginning and ending of employment activity observed in the data. The same is true for the dates assigned to the job spells.

6. More specifically, an establishment is imputed to a job spell only once within each implicate.

Imputation and output data Once the input data are organized, a set of 10 imputed establishment identifiers are generated for each job spell ending in every quarter for which both ES-202 and UI wage records exist. For each quarter, implicate, and size class, $s = 1, 2, 3$, the parameters on the linear spline in distance between place-of-work and place-of-residence $\hat{\beta}^s$ are sampled from the normal approximation of the posterior predictive distribution of β^s conditional on Minnesota (MN)

$$p(\beta^s | \alpha_{MN}, x_{MN}, y_{MN}) \tag{14.7}$$

The draws from this distribution vary across implicates, but not across time, employers, and individuals. Next, for each employer j at time t, a set of $\hat{\alpha}_{jrt}$ are drawn from

$$p(\alpha_{ST} | N_{ST}) \tag{14.8}$$

which are based on the ES-202 month-one employment totals (N_{jrt}) for all candidate establishments $r_{jt} \subset \mathcal{R}_{jt}$ at employer j within the state (ST) being processed. The initial draws of $\hat{\alpha}_{jrt}$ from this distribution vary across time and employers but not across job spells. Combining (14.7) and (14.8) yields

$$
\begin{aligned}
&p(\alpha_{ST} | N_{ST}) \, p(\beta^s | \alpha_{MN}, x_{MN}, y_{MN}) \\
\approx \quad &p(\alpha_{ST} | N_{ST}) \, p(\beta^s | \alpha_{ST}, x_{ST}, y_{ST}) \\
= \quad &p(\alpha_{ST}, \beta_{ST} | x_{ST}, y_{ST}, N_{ST}),
\end{aligned}
\tag{14.9}
$$

an approximation of the joint posterior distribution of α and β^s (14.4) conditional on data from the state being processed.

The draws $\hat{\beta}^s$ and $\hat{\alpha}_{jrt}$ in conjunction with the establishment, employer, and job spell data are used to construct the p_{ijrt} in (14.1) for all candidate establishments $r \in \mathcal{R}^i_{jt}$. For each job spell and candidate establishment combination, the $\hat{\beta}^s$ are applied to the calculated distance between place-of-residence (of the worker holding the job spell) and the location of the establishment, where the choice of $\hat{\beta}^s$ depends on the size class of the establishment's parent employer. For each combination an $\hat{\alpha}_{jrt}$ is drawn which is based primarily on the size (in terms of employment) of the establishment relative to other active establishments at the parent employer. In conjunction, these determine the conditional probability p_{ijrt} of a candidate establishment's assignment to a given job spell. Finally, from this distribution of probabilities is drawn an establishment of employment.

The imputation process yields a data file containing a set of 10 imputed establishment identifiers for each job spell. In a very small set of cases, the model fails to impute an establishment to a job spell. This is often due to unanticipated idiosyncrasies in the underlying administrative data. Furthermore, across states, the proportion of these failures relative to successful imputation is well under 0.5%. For these job spells, a dummy establishment identifier is assigned and in downstream processing, the employment-weighted modal employer-level characteristics are used.

14.4 DATA SET DESCRIPTIONS

14.4.1 Naming scheme

The U2W contains a single file per state:

`u2w_zz.sas7bdat`

`ZZ` stands for the state postal abbreviation. You will find zero-observation SAS datasets attached to this document - see the attachment tab.

14.4.2 Data location

The files are stored in a main directory, with state-specific subdirectories:

`u2w/ZZ/`

On the RDC network, the directory can be found under On the RDC network, the directory can be found under

`/mixed/lehd/s2011`

14.4.3 Main dataset: u2w_zz

This files contain the 10 imputed establishment identifiers are generated for each job spell.

Record identifier PIK SEIN NEW_HIST_FLAG

Sort order PIK SEIN NEW_HIST_FLAG

Entity Job spell

Unique Entity Key PIK SEIN

Field name	Data dictionary reference name	Starting position	Field size	Data type
Start of spell YYYY.F (e.g. 2000Q2 = 2000.25)	FIRST_DATE	00008	3	N
End of spell YYYY.F (e.g. 2000Q4 = 2000.75)	LAST_DATE	00011	3	N
Spell number for same SEIN	NEW_HIST_FLAG	00014	3	N
Protected Identification Key	PIK	00017	9	A/N
State Employer Identification Number	SEIN	00026	12	A/N
State UI Reporting Unit Number (Impute 1)	IMPUTED_UNIT_1	00038	5	A/N
State UI Reporting Unit Number (Impute 10)	IMPUTED_UNIT_10	00083	5	A/N
State UI Reporting Unit Number (Impute 2)	IMPUTED_UNIT_2	00043	5	A/N
State UI Reporting Unit Number (Impute 3)	IMPUTED_UNIT_3	00048	5	A/N
State UI Reporting Unit Number (Impute 4)	IMPUTED_UNIT_4	00053	5	A/N
State UI Reporting Unit Number (Impute 5)	IMPUTED_UNIT_5	00058	5	A/N
State UI Reporting Unit Number (Impute 6)	IMPUTED_UNIT_6	00063	5	A/N
State UI Reporting Unit Number (Impute 7)	IMPUTED_UNIT_7	00068	5	A/N
State UI Reporting Unit Number (Impute 8)	IMPUTED_UNIT_8	00073	5	A/N
State UI Reporting Unit Number (Impute 9)	IMPUTED_UNIT_9	00078	5	A/N

14.4.4 Summary information on datasets

Table 14.2: Number of observations for U2W

Group	Number of datafiles	Records (1000s)	Filesize (GB)
U2W	49	625,220	47

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.

Table 14.3: List of data files for U2W, by state

The list can also be downloaded in CSV format
from the attachments to this document.

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Alaska (ak)					
u2w_ak	2000Q1	2012Q1	1,066	< 1	1:424367
Alabama (al)					
u2w_al	2001Q1	2012Q1	8,157	1	1:423602
Arkansas (ar)					
u2w_ar	2002Q3	2012Q1	4,026	< 1	1:423766
Arizona (az)					
u2w_az	2004Q1	2012Q1	4,325	< 1	1:421705
California (ca)					
u2w_ca	1991Q3	2012Q1	89,546	7	1:425576
Colorado (co)					
u2w_co	1993Q2	2012Q1	13,646	1	1:445035
Connecticut (ct)					
u2w_ct	1996Q1	2012Q1	6,046	< 1	1:424937
District of Columbia (dc)					
u2w_dc	2005Q2	2012Q1	272	< 1	1:423824
Delaware (de)					
u2w_de	1998Q3	2012Q1	1,093	< 1	1:431406
Florida (fl)					
u2w_fl	1992Q4	2012Q1	54,021	4	1:425634
Georgia (ga)					
u2w_ga	1998Q1	2012Q1	20,064	2	1:424213
Hawaii (hi)					
u2w_hi	1995Q4	2012Q1	2,031	< 1	1:429994
Iowa (ia)					
u2w_ia	1998Q4	2012Q1	6,901	1	1:425774
Idaho (id)					
u2w_id	1991Q1	2012Q1	3,285	< 1	1:427442
Illinois (il)					
u2w_il	1990Q1	2012Q1	26,845	2	1:442810
Indiana (in)					
u2w_in	1998Q1	2012Q1	13,899	1	1:429989

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:" to obtain the full SnapshotID.

(cont)

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
Kansas (ks)					
u2w_ks	1993Q1	2012Q1	6,443	< 1	1:429988
Kentucky (ky)					
u2w_ky	2001Q1	2012Q1	6,627	1	1:424754
Louisiana (la)					
u2w_la	1995Q1	2012Q1	10,096	1	1:443516
Maryland (md)					
u2w_md	1990Q1	2012Q1	13,727	1	1:429993
Maine (me)					
u2w_me	1996Q2	2012Q1	2,313	< 1	1:434255
Michigan (mi)					
u2w_mi	2000Q3	2012Q1	10,522	1	1:434271
Missouri (mo)					
u2w_mo	1995Q1	2012Q1	16,447	1	1:427903
Mississippi (ms)					
u2w_ms	2003Q3	2012Q1	2,918	< 1	1:427441
Montana (mt)					
u2w_mt	1993Q1	2012Q1	1,625	< 1	1:431410
North Carolina (nc)					
u2w_nc	1992Q4	2011Q4	28,555	2	1:445204
North Dakota (nd)					
u2w_nd	1998Q1	2012Q1	1,194	< 1	1:431411
Nebraska (ne)					
u2w_ne	1999Q1	2012Q1	3,392	< 1	1:431412
New Hampshire (nh)					
u2w_nh	2003Q1	2012Q1	1,575	< 1	1:433236
New Jersey (nj)					
u2w_nj	1996Q1	2012Q1	14,597	1	1:422196
New Mexico (nm)					
u2w_nm	1995Q3	2012Q1	3,417	< 1	1:432823
Nevada (nv)					
u2w_nv	1998Q1	2012Q1	5,291	< 1	1:443517
New York (ny)					
u2w_ny	2000Q1	2012Q1	26,179	2	1:443289
Ohio (oh)					
u2w_oh	2000Q1	2012Q1	21,518	2	1:443571
Oklahoma (ok)					
u2w_ok	2000Q1	2012Q1	5,799	< 1	1:422197
Oregon (or)					
u2w_or	1991Q1	2012Q1	9,831	1	1:422198
Pennsylvania (pa)					
u2w_pa	1997Q1	2012Q1	27,792	2	1:425579
Rhode Island (ri)					
u2w_ri	1995Q1	2012Q1	1,370	< 1	1:436554
South Carolina (sc)					

Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

(cont)

Table 14.3 – Continued

File name	StartYQ	EndYQ	Obs. (1000s)	Size (GB)	ShortID
u2w_sc	1998Q1	2012Q1	7,184	1	1:437561
South Dakota (sd)					
u2w_sd	1998Q1	2012Q1	1,282	< 1	1:431413
Tennessee (tn)					
u2w_tn	1998Q1	2012Q1	11,184	1	1:427440
Texas (tx)					
u2w_tx	1995Q1	2012Q1	70,305	5	1:426223
Utah (ut)					
u2w_ut	1999Q3	2012Q1	6,242	< 1	1:434191
Virginia (va)					
u2w_va	1998Q1	2012Q1	15,697	1	1:439226
Vermont (vt)					
u2w_vt	2000Q1	2012Q1	701	< 1	1:431425
Washington (wa)					
u2w_wa	1990Q1	2012Q1	14,815	1	1:427443
Wisconsin (wi)					
u2w_wi	1990Q1	2012Q1	17,019	1	1:422578
West Virginia (wv)					
u2w_wv	1997Q1	2012Q1	3,574	< 1	1:441377
Wyoming (wy)					
u2w_wy	2001Q1	2012Q1	771	< 1	1:431447

Number of files for each data set group and state. Aggregate size of all files in GB in parentheses.
Record counts rounded to nearest 1000. Filesize rounded to nearest 1GB.
ShortID must be prepended with "snapshot:s2011:' to obtain the full SnapshotID.

14.5   NOTES

Space for your notes

14.6 ACRONYMS USED

ACS American Community Survey

ACS-POW American Community Survey Place of Work file

AHS American Housing Survey

BED Business Employment Dynamics

BLS Bureau of Labor Statistics

BR Business Register, formerly known as the SSEL

BRB Business Register Bridge

CBSA Core-Based Statistical Area

CEW Covered Employment and Wages

COLA cost of living allowance

CPR Composite Person Record

CPS Current Population Survey

DC Decennial Census

DHS Department of Homeland Security

DRB Disclosure Review Board

DOI Digital Object Identifier

ECF Employer Characteristics File

ES-202 ES-202. An older name for the QCEW program

EHF Employment History Files

EHRI Enterprise Human Resources Integration

EIN (federal) Employer Identification Number

FBI Federal Bureau of Investigation

FEMA Federal Emergency Management Agency

FIPS Federal Information Processing Standards codes issued by National Institute of Standards and Technology (NIST)

FTI Federal Tax Information, typically covered under Title 26, U.S.C.

GAL Geocoded Address List

GSA General Services Administration

HCEF 100 Percent Census Edited File

HDF 100 Percent Detail File

HHS Department of Health and Human Services

ICF Individual Characteristics File

IRS Internal Revenue Service

IRS Internal Revenue Service

JHF Job History File

LBD Longitudinal Business Database

LBDB LBD Bridge

LDB Longitudinal Data Base

LED Local Employment Dynamics

LEHD Longitudinal Employer-Household Dynamics

LMI Labor Market Information

MN Minnesota

MOU Memorandum of Understanding

MSA Metropolitan Statistical Area

NAICS North American Industry Coding System

NIST National Institute of Standards and Technology

OMB Office of Management and Budget

OPM Office of Personnel Management

OTM OnTheMap

PHF Person History File

PIK Protected Identity Key

POI Point of informationfile, one of the OPM data files

QCEW Quarterly Census of Employment and Wages, managed by the Bureau of Labor Statistics (BLS)

QWI Quarterly Workforce Indicators

QWIPU Public-Use QWI

RDC Research Data Center

SCEF Sample Census Edited File

SCT Standard Code Table, one of the OPM data files

SEDF Sample Edited Detail File

SEIN State employer identification number. It is constructed from the state Federal Information Processing Standards (FIPS) code and the UI account number. The BLS refers to the UI account number in combination with the reporting unit number as SESA-ID

SEINUNIT SEIN reporting unit

SESA State Employment Security Agency

SIC Standard Industry Classification

SIPP Survey of Income and Program Participation

SPF Successor-Predecessor File

SSA Social Security Administration

SSN Social Security Number

U2W Unit-to-Worker Impute

UI unemployment insurance

WIB Workforce Investment Board

Bibliography

Abowd, John M., Bryce E. Stephens, and Lars Vilhuber. 2006. *Confidentiality Protection in the Census Bureau's Quarterly Workforce Indicators.* Technical Paper TP-2006-02. LEHD, U.S. Census Bureau. http://econpapers.repec.org/paper/centpaper/2006-02.htm.

Abowd, John M., Bryce E. Stephens, Lars Vilhuber, Fredrik Andersson, Kevin L. McKinney, Marc Roemer, and Simon D. Woodcock. 2006a. *The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators.* Technical report TP-2006-01. U.S. Census Bureau, LEHD and Cornell University. http://econpapers.repec.org/paper/centpaper/2006-01.htm.

——. 2006b. *The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators.* Technical paper TP-2006-01. U.S. Census Bureau, LEHD and Cornell University. http://econpapers.repec.org/paper/centpaper/2006-01.htm.

——. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In *Timothy Dunne, J. Bradford Jensen, and Mark J. Roberts,* edited by Timothy Dunne, J. Brad Jensen, and Mark J. Roberts. University of Chicago Press. http://www.nber.org/chapters/c0485.pdf.

American Economic Association. 2014. "The American Economic Review: Data Availability Policy." Accessed April 16, 2014. http://www.aeaweb.org/aer/data.php.

Benedetto, Gary, John Haltiwanger, Julia Lane, and Kevin McKinney. 2007. "Using Worker Flows in the Analysis of the Firm." *Journal of Business and Economic Statistics* 25, no. 3 (July): 299–313.

Haltiwanger, John, Henry Hyatt, Erika McEntarfer, Liliana Sousa, and Stephen Tibbets. 2014. *Firm Age And Size In The Longitudinal Employer-Household Dynamics Data.* Working Papers 14-16. Center for Economic Studies, U.S. Census Bureau, March. http://ideas.repec.org/p/cen/wpaper/14-16.html.

Journal of Labor Economics. 2009. "Data Policy." Accessed April 16, 2014. http://www.press.uchicago.edu/journals/jole/datapolicy.html?journal=jole.

McKinney, Kevin, and Lars Vilhuber. 2011a. *LEHD Data Documentation LEHD-OVERVIEW-S2008-rev1.* Working Papers 11-43. Center for Economic Studies, U.S. Census Bureau, December. http://ideas.repec.org/p/cen/wpaper/11-43.html.

——. 2011b. *LEHD Infrastructure Files in the Census RDC: Overview of S2004 Snapshot.* Working Papers 11-13. Center for Economic Studies, U.S. Census Bureau, April. http://ideas.repec.org/p/cen/wpaper/11-13.html.

National Science Foundation. 2011. "Data Management Plan." Chap. II.C.2.j in *Grant Proposal Guide,* II–19. NSF, 11-1. National Science Foundation. Accessed April 16, 2014. http://www.nsf.gov/pubs/policydocs/pappguide/nsf11001/gpg_index.jsp.

Review of Economics and Statistics. 2014. "Data Availability Policy." Accessed April 16, 2014. http://www.mitpressjournals.org/page/sub/rest.

BIBLIOGRAPHY

Stephens, Bryce. 2007. "Essays on firm compensation policy and confidentiality protection and imputation in the Quarterly Workforce Indicators." Ph.D., University of Maryland, College Park.

Stevens, David W. 2007. *Employment that is not covered by state unemployment insurance laws.* Technical paper TP-2007-04. LEHD, U.S. Census Bureau.

U.S. Office of Personal Management. 2012. *The Guide to Data Standards - Part B: Payroll.* Technical report. Accessed at http://www.opm.gov/feddata/guidance.asp on May 8, 2012. OPM.

Vilhuber, Lars, and Ian Schmutte. 2012. *QWI Cheatsheet.* Mimeo. VirtualRDC, Cornell University. `http://download.vrdc.cornell.edu/qwipu/QWI-cheatsheet.txt`.

Chapter 15.
Errata

`$Id: overview_errata.tex 11597 2014-06-02 17:15:13Z vilhuber $`

This section will contain a list of any errata found. The revision number (specific to the errata) above will update with later versions.

```
$Id: overview_master.tex 11747 2014-06-20 14:48:21Z vilhuber $
```