

and the Creation of the Quarterly Workforce The LEHD Infrastructure Files

Indicators

John M. Abowd⁴,♣, Bryce E. Stephens♣ and Lars Vilhuber♣

♠ Cornell University

U.S. Census Bureau, LEHD Program



troduction

⇒ What are QWI?

⇒ What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Introduction



ntroduction

▲ Since 2003: publication of Quarterly Workforce Indicators

What are QWI?

⇒ What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Lars Vilhuber, John Abowd and Bryce Stephens



⇒ In this paper ⇒ What is it?

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

▲ The first 21st century statistical system



Introduction

/hat are QWI?

What is it?In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

The first 21st century statistical system

No additional burden



ntroduction

hat are QWI?

OWhat is it? O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

▲ The first 21st century statistical system

No additional burden

Extensive use of modern statistics to integrate and improve the data



ntroduction

What are QWI?

OWhat is it? O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

▲ The first 21st century statistical system

No additional burden

Extensive use of modern statistics to integrate and improve the data

State-of-the-art confidentiality protection methods



ntroduction

/hat are QWI?

OWhat is it? O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

▲ The first 21st century statistical system

No additional burden

Extensive use of modern statistics to integrate and improve the data

State-of-the-art confidentiality protection methods

Innovative use of wage records to constitute a frame to integrate data



ntroduction

What are QWI?

OWhat is it? O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators

▲ The first 21st century statistical system

No additional burden

Extensive use of modern statistics to integrate and improve the data

State-of-the-art confidentiality protection methods

Innovative use of wage records to constitute a frame to integrate data

The first statistical system to use "jobs" as a frame



▲ Combines

⇒ What are QWI?

⇒ What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Lars Vilhuber, John Abowd and Bryce Stephens



ntroduction

⊃What are QWI?

> What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Combines

• (state) administrative records data on workers (UI Wage records)



ntroduction

⇒ What are QWI?

What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Combines

 (state) administrative records data on workers (UI Wage records)

(state) administrative records data on firms (QCEW aka ES-202)



itroduction

⇒ What are QWI?

>What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Combines

 (state) administrative records data on workers (UI Wage records)

 State) administrative records data on firms (QCEW aka ES-202)

administrative information on demographics



ntroduction

⇒ What are QWI?

What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Combines

 (state) administrative records data on workers (UI Wage records)

(state) administrative records data on firms (QCEW aka ES-202)

administrative information on demographics

surveys on people and firms collected by Census Bureau



troduction

⇒ What are QWI?

What is it?

⇒ In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

onclusion

▲ Combines

- (state) administrative records data on workers (UI Wage records)
- (state) administrative records data on firms (QCEW aka ES-202)
- administrative information on demographics
- surveys on people and firms collected by Census Bureau
- careful longitudinal edit of person identifiers and economic firm units



⇒ What are QWI?

What are gwr

Oln this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

onclusion

▲ Combines

 (state) administrative records data on workers (UI Wage records)

(state) administrative records data on firms (QCEW aka ES-202)

administrative information on demographics

surveys on people and firms collected by Census Bureau

▲ careful longitudinal edit of person identifiers and economic firm units

▲ careful longitudinal edit of person and firm characteristics



ntroduction

▲ Describe the construction of the LEHD infrastructure

What are QWI?What is it?

⇒ In this paper

Input Files

Infrastructure Files

•

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

Lars Vilhuber, John Abowd and Bryce Stephens



⇒ What are QWI? ⇒ What is it?

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Describe the construction of the LEHD infrastructure

... in particular the imputation mechanisms used



ntroduction

OWhat are QWI? OWhat is it?

In this pape

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Describe the construction of the LEHD infrastructure

... in particular the imputation mechanisms used

▲ Describe the computation of the QWI statistics



ntroduction

⇒ What are QWI?

OWhat is it?

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Describe the construction of the LEHD infrastructure

... in particular the imputation mechanisms used

▲ Describe the computation of the QWI statistics

... in particular the imputation mechanisms used



ntroduction

OWhat are QWI? OWhat is it?

In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Describe the construction of the LEHD infrastructure

... in particular the imputation mechanisms used

▲ Describe the computation of the QWI statistics

... in particular the imputation mechanisms used

▲ Describe the disclosure-proofing mechanism



itroduction

⇒ What are QWI?
⇒ What is it?

In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Describe the construction of the LEHD infrastructure

... in particular the imputation mechanisms used

▲ Describe the computation of the QWI statistics

... in particular the imputation mechanisms used

Describe the disclosure-proofing mechanism

Describe researcher access to infrastructure files and confidential QWI files



Introduction

put Files

⇒ Wage records: UI

⇒ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Input Files



Introduction

nput Files

Wage records: UI

⇒ Employer reports: ES202

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ report of an individual's UI-covered earnings by an employing entity



Introduction

⇒ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ report of an individual's UI-covered earnings by an employing entity ▲ appears if at least one dollar was earned by that individual during the quarter



Introduction

⇒ Employer reports: ES202 ⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ report of an individual's UI-covered earnings by an employing entity appears if at least one dollar was earned by that individual during the quarter

▲ identifies EARNINGS, EMPLOYER, TIME PERIOD



Introduction

⇒ Employer reports: ES202 ⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Publicly available files

Disclosure-proofing the QWI

Conclusion

▲ report of an individual's UI-covered earnings by an employing entity appears if at least one dollar was earned by that individual during the quarter

▲ identifies EARNINGS, EMPLOYER, TIME PERIOD

some limited other state-dependent information available



Introduction

nput Files Wane records: 11

D Employer reports: ES202D Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

onclusion

▲ report of an individual's UI-covered earnings by an employing entity appears if at least one dollar was earned by that individual during the quarter

▲ identifies EARNINGS, EMPLOYER, TIME PERIOD

some limited other state-dependent information available

▲ in particular, for Minnesota, the ESTABLISHMENT is reported



Introduction

... or QCEW

⇒ Wage records: UI

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Dit Files

⇒ Wage records: UI

≣mployer reports: ES;

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

 collected as part of the Covered Employment and Wages (CEW) (administered by the BLS)



Introduction

⇒ Wage records: UI

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

 collected as part of the Covered Employment and Wages (CEW) (administered by the BLS)

▲ Also used as the inputs to the Business Employment Dynamics (BED)



Introduction

out Files

Nage records: UI

Employer reports: ES2

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

 collected as part of the Covered Employment and Wages (CEW) (administered by the BLS)

▲ Also used as the inputs to the Business Employment Dynamics (BED) collects from employers covered by state unemployment insurance programs:

employment

payroll

geographic information



Introduction

put Files

D Wage records: UI

Employer reports: ES20

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

 collected as part of the Covered Employment and Wages (CEW) (administered by the BLS)

▲ Also used as the inputs to the Business Employment Dynamics (BED) collects from employers covered by state unemployment insurance programs:

employment

payroll

geographic information

fundamental unit: 'reporting unit' (≈ establishment)



Introduction

nput Files

⇒ Wage records: UI

Employer reports: ES2

⇒ Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

nclusion

 collected as part of the Covered Employment and Wages (CEW) (administered by the BLS)

▲ Also used as the inputs to the Business Employment Dynamics (BED) collects from employers covered by state unemployment insurance programs:

employment

payroll

geographic information

▲ fundamental unit: 'reporting unit' (≈ establishment)

One report per establishment per quarter is filed



Demographics

Introduction

put Files

⇒ Wage records: UI

⊃ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Demographics are taken from a number of Census-internal files derived from administrative data:

Person Characteristics File (PCF)

Census Numident



Demographics



Introduction

⇒ Employer reports: ES202 Wage records: UI

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Demographics are taken from a number of Census-internal files derived from administrative data:

Person Characteristics File (PCF)

Census Numident

■ Where available, more detailed data on individuals is also extracted from surveys and censuses:

CPS

SIPP

ACS

1990 Census

2000 Census



Introduction

Input Files

⇒EHF: Employment History

OICF: Individual

Characteristics File ⇒ ECF: Employer

Characteristics File CAL: Geocoded Address

List Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

Infrastructure Files



EHF: Employment History Files

Introduction

Input Files

EHF: Employment History

OICF: Individual

Characteristics File

DECF: Employer

Characteristics File CAAL: Geocoded Address

⇒ Flow so far

Estimates: QWI

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Job-level EHF

 complete in-state work history for each individual on Ulwage records. one record for each employee-employer combination – a <u>qo</u>

earnings and employment patterns



EHF: Employment History Files

Introduction

Input Files

EHF: Employment History

OICF: Individual

Characteristics File

Characteristics File ⇒ECF: Employer

⇒ GAL: Geocoded Address

⇒ Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Job-level EHF

 complete in-state work history for each individual on Ulwage records. one record for each employee-employer combination - a <u>qo</u>

earnings and employment patterns

Employer and establishment-level employment history

 QCEW-based employment-activity history for every SEIN (employer) and SEINUNIT (establishment)



EHF: Employment History Files

Introduction

Input Files

EHF: Employment History

Characteristics File CF: Individual

Characteristics File ⇒ECF: Employer

⇒ GAL: Geocoded Address

⇒ Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Job-level EHF

 complete in-state work history for each individual on Ulwage records. one record for each employee-employer combination - a

earnings and employment patterns

Employer and establishment-level employment history

 QCEW-based employment-activity history for every SEIN (employer) and SEINUNIT (establishment)

 Comparison of employment and activity of SEINs between UI and QCEW files is done for QA purposes, and in preparation of weighting.



Introduction

Input Files

JEHF: Employment History

CE: Individual

⇒ECF: Employer

Characteristics File CAL: Geocoded Address

⊃ Flow so far

Estimates: QWI

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

Introduction

Input Files

JEHF: Employment History

CE: Individual

Characteristics File ⇒ECF: Employer

Characteristics File CAL: Geocoded Address

⇒ Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

Introduction

Input Files

DEHF: Employment History

CE: Individual

Characteristics File

DECF: Employer

⇒ GAL: Geocoded Address Characteristics File

⇒ Flow so far

Estimates: QWI

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

▲ records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.



Introduction

Input Files

DEHF: Employment History

CF: Individual

Characteristics File ⇒ ECF: Employer

OGAL: Geocoded Address Characteristics File

⇒ Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

▲ ... gender, education, and age information from the CPS



Introduction

Input Files

rastructure Files

DEHF: Employment History

Characteristics File

⇒ ECF: Employer
Characteristics File
⇒ GAI : Geocoded Addition

OGAL: Geocoded Address

List ⊃Flow so far Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

▲ ... gender, education, and age information from the CPS

▲ Data completion



Introduction

Input Files

DEHF: Employment History

Characteristics File CF: Individual

Characteristics File ⇒ ECF: Employer

CAL: Geocoded Address

⇒ Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

... gender, education, and age information from the CPS

▲ Data completion

Age



Introduction

Input Files

DEHF: Employment History CF: Individual

Characteristics File ⇒ ECF: Employer

CAL: Geocoded Address Characteristics File

⇒ Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

... gender, education, and age information from the CPS

▲ Data completion

Age

Gender



Introduction

Input Files

rastructure File

⇒ EHF: Employment History

Files

OICF: Individual
Characteristics File
DECF: Employer

Characteristics File Cocoded Address

List

⊃Flow so far

Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

... gender, education, and age information from the CPS

▲ Data completion

Age

Gender

Education



Introduction

Input Files

DEHF: Employment History

CF: Individual

Characteristics File ⇒ ECF: Employer

CAL: Geocoded Address Characteristics File

⇒ Flow so far

Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

... gender, education, and age information from the CPS

▲ Data completion

Age

Gender

Education

County of residence



Introduction

Input Files

ומפוומכומום ו וופפ

DEHF: Employment History

CF: Individu

Characteristics File

⇒ ECF: Employer Characteristics File
⇒ GAL: Geocoded Address

List ⊃Flow so far Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Demographic information from the PCF is merged with universe of PIKs from wage records

records without a valid match flagged

▲ CPS and SIPP identifiers are merged on.

... gender, education, and age information from the CPS

▲ Data completion

Age

Gender

Education

County of residence

are each imputed ten times



Introduction

▲ Two files: firm and establishment level, quarterly records

Input Files

DEHF: Employment History

CE: Individual

Characteristics File

Characteristics File CAAL: Geocoded Address

Forming Aggregated

⊃Flow so far

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Input Files

DEHF: Employment History

CE: Individual

Characteristics File

Characteristics File CAAL: Geocoded Address

⊃Flow so far

Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:



Introduction

Input Files

DEHF: Employment History

CE: Individual

Characteristics File

Characteristics File CAAL: Geocoded Address

⊃Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202



Introduction

Input Files

Structure Files

DEHF: Employment History

DICF: Individual

Characteristics File

ECF: Employer Characteristics Fill

Characteristics File CAAL: Geocoded Address

LIST S Flow so far Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data



Introduction

Input Files

DEHF: Employment History

DICF: Individual

Characteristics File

Characteristics File CAAL: Geocoded Address

⊃ Flow so far

Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes



Introduction

Input Files

DEHF: Employment History

DICF: Individual

Characteristics File

Characteristics File CAL: Geocoded Address

⊃ Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes

4. LDB (BLS) for backfilling NAICS information



Introduction

Input Files

Ī

⇒ EHF: Employment History

OICF: Individual

Characteristics File

Characteristics File

CAL: Geocoded Address

List ⊃Flow so far Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes

4. LDB (BLS) for backfilling NAICS information

Longitudinal edits for consistency and data completion



Introduction

Input Files

astructure File

DEHF: Employment History

Characteristics File

Characteristics File

CAL: Geocoded Address

List ⊃Flow so far Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes

4. LDB (BLS) for backfilling NAICS information

▲ Longitudinal edits for consistency and data completion

▲ Imputation:

impute SIC if NAICS non-missing and vice-versa



Introduction

Input Files

DEHF: Employment History

DICF: Individual

Characteristics File

⇒ GAL: Geocoded Address

⇒ Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes

4. LDB (BLS) for backfilling NAICS information

▲ Longitudinal edits for consistency and data completion

▲ Imputation:

impute SIC if NAICS non-missing and vice-versa

unconditional impute of missing SIC and NAICS codes



Introduction

Input Files

DEHF: Employment History

Files DICF: Individual

Characteristics File

ECF: Employer Characteristics File

Characteristics File
CGAL: Geocoded Address

List ⊃Flow so far Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Two files: firm and establishment level, quarterly records

▲ Inputs:

1. ES202

2. UI: supplement information on the ES202, extend published BLS county-level employment data

3. GAL: establishment geocodes

4. LDB (BLS) for backfilling NAICS information

▲ Longitudinal edits for consistency and data completion

▲ Imputation:

impute SIC if NAICS non-missing and vice-versa

unconditional impute of missing SIC and NAICS codes

geography conditional on industry



Introduction

Input Files

DEHF: Employment History

⇒ICF: Individual

Characteristics File

Characteristics File ⇒ECF: Employer

⇒ Flow so far

Forming Aggregated

Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses



Introduction

Input Files

astructure Files

⇒ EHF: Employment History

CF: Individual

Characteristics File DECF: Employer

Characteristics File

GAL: Geocoded

⊃Flow so far

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ ... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates



Introduction

Input Files

rastructure File

⇒ EHF: Employment History

OICF: Individual

Characteristics File DECF: Employer

Characteristics File

List

⊃ Flow so far

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ ... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:



Introduction

Input Files

astructure Files

DEHF: Employment History

OICF: Individual

Characteristics File DECF: Employer

Characteristics File

GAL: Geocoded Ad

⇒ Flow so far

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data



Introduction

Input Files

astructure Files

⇒ EHF: Employment History

OICF: Individual

Characteristics File DECF: Employer

Characteristics File

LIST D Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Estimates: QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

Census Bureau's Business Register (BR)



Introduction

Input Files

astructure Files

○ EHF: Employment History Files

Characteristics File DECF: Employer

⇒ICF: Individual

Characteristics File

List D Flow so far Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

3. Census Bureau's Master Address File (MAF)



Introduction

Input Files

⇒ EHF: Employment History

CF: Individual

Characteristics File DECF: Employer

Characteristics File

LIST Flow on for

⊃Flow so far

Forming Aggregated Estimates: OWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

3. Census Bureau's Master Address File (MAF)

4. American Community Survey Place of Work file (ACS-POW)



Introduction

Input Files

DEHF: Employment History

Characteristics File **DICF: Individual**

⇒ECF: Employer

Characteristics File

> Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

▲ geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

Census Bureau's Master Address File (MAF)

4. American Community Survey Place of Work file (ACS-POW)

▲ Addresses are



Introduction

Input Files

asiluciule riles

DEHF: Employment History

OICF: Individual

Characteristics File DECF: Employer

Characteristics File

List D Flow so far Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

Census Bureau's Master Address File (MAF)

4. American Community Survey Place of Work file (ACS-POW)

▲ Addresses are

1. geocoded



Introduction

Input Files

asiluciale riles

⇒ EHF: Employment History

⊃ICF: Individual

Characteristics File

OECF: Employer
Characteristics File

Characteristics File GAL: Geocoded Ad

S Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

Census Bureau's Master Address File (MAF)

4. American Community Survey Place of Work file (ACS-POW)

▲ Addresses are

1. geocoded

2. standardized



Introduction

Input Files

DEHF: Employment History

DICF: Individual

Characteristics File DECF: Employer

Characteristics File

GAL: Geocoded List

D Flow so far

Forming Aggregated

Disclosure-proofing the QWI

Publicly available files

Conclusion

... is a data set containing unique commercial and residential addresses

geocoded to the Census Block and latitude/longitude coordinates

▲ Inputs:

1. ES202 data

2. Census Bureau's Business Register (BR)

Census Bureau's Master Address File (MAF)

4. American Community Survey Place of Work file (ACS-POW)

▲ Addresses are

1. geocoded

2. standardized

3. unduplicated (by firm name)



Flow so far

Introduction

Input Files

⇒EHF: Employment History

OICF: Individual

Characteristics File ⇒ ECF: Employer

Characteristics File CAL: Geocoded Address

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion



Main page



Flow so far

Introduction

Input Files

astructure Files

JEHF: Employment History

iles

OICF: Individual Characteristics File

⇒ ECF: Employer
Characteristics File

Characteristics File CAL: Geocoded Address

List

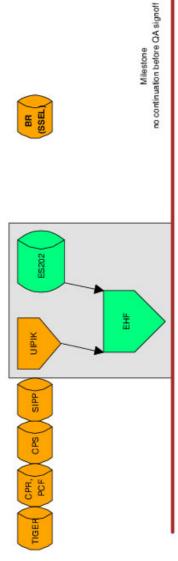
Flow so far

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion



Main page



Flow so far

Introduction

Input Files

astructure Files

JEHF: Employment History

Se

OICF: Individual Characteristics File

DECF: Employer
Characteristics File

Characteristics File CAL: Geocoded Address

List

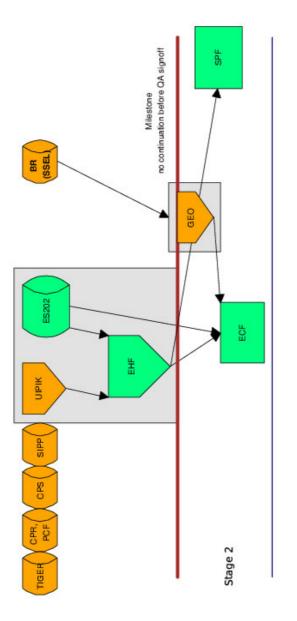
Flow so far

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion



Main page



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs CU2W: Unit to Worker

⇒ Probability Model

⇒ Implementation Umplementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Forming Aggregated Estimates: QWI

Introduction

▲ Firm identifier:

Infrastructure Files

Input Files

Correction of spurious orming Aggregated stimates: QWI

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs CU2W: Unit to Worker

⇒ Probability Model

⇒ Implementation

⇒ Computing the statistics Umplementation

Disclosure-proofing the QWI

Publicly available files

Introduction

▲ Firm identifier: state-specific account number

Infrastructure Files

Input Files

Correction of spurious stimates: QWI

Solution:

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation ⇒ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Lars Vilhuber, John Abowd and Bryce Stephens

Introduction

▲ Firm identifier:

Input Files

▲ Account numbers can and do change:

Infrastructure Files

Correction of spurious stimates: QWI

Successor-Predecessor Solution:

⇒ Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation ⇒ Implementation Computing the statistics

Disclosure-proofing the QWI

Publicly available files



Introduction

Input Files

Infrastructure Files

orming Aggregated stimates: QWI

Correction of spurious Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs ⊃U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation ⊃ Implementation Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Firm identifier:

▲ Account numbers can and do change:

change in legal form



Introduction

Input Files

Infrastructure Files

Correction of spurious orming Aggregated stimates: QWI

Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation ⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Firm identifier:

▲ Account numbers can and do change:

change in legal form

a merger





Introduction

Input Files

Infrastructure Files

Correction of spurious

Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⇒ U2W: Unit to Worker

⇒ Probability Model

> Implementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Firm identifier:

Account numbers can and do change:

change in legal form

a merger

Change in firm identifier



Introduction

Input Files

Infrastructure Files

stimates: QWI

Correction of spurious ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⇒ U2W: Unit to Worker

⇒ Probability Model

Umplementation ⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Firm identifier:

Account numbers can and do change:

change in legal form

a merger

Change in firm identifier is the component determining when a worker changes employers



Introduction

Input Files

Infrastructure Files

stimates: QWI
Correction of spurious

worker flows

Successor-Predecessor

File

Attaching establishment

characteristics to jobs

OU2W: Unit to Worker

⇒ Probability Model

O Implementation
O Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Firm identifier:

▲ Account numbers can and do change:

change in legal form

a merger

▲ Change in <u>firm identifier</u>

▲ → non-economic change in identifier creates spurious flow



Introduction

▲ track large worker movements between SEINs

Input Files

Infrastructure Files

Estimates: QWI

⇒ Correction of spurious

worker flows

Successor-Predecessor

Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation

Computing the statistics ⇒ Implementation

Disclosure-proofing the QWI

Publicly available files



Introduction

Input Files

Infrastructure Files

Estimates: QWI

⇒ Correction of spurious worker flows

Attaching establishment characteristics to jobs

⇒ U2W: Unit to Worker

⇒ Probability Model

Umplementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ track large worker movements between SEINs

→ link entities that have different account numbes, but constitute the same economic entitiy

Lars Vilhuber, John Abowd and Bryce Stephens



Introduction

Input Files

Infrastructure Files

orming Aggregated stimates: QWI

Correction of spurious

worker flows

Successor-Predecessor

- Attaching establishment
 - characteristics to jobs
 - OU2W: Unit to Worker
- ⇒ Probability Model
- D Implementation D Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ track large worker movements between SEINs
- → link entities that have different account numbes, but constitute the same economic entitiy
- relative terms, and the number of workers entering an SEIN, SPF provides a variety of link characteristics, based on the number of workers leaving an SEIN, in both absolute and again in absolute and relative terms.



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

Attaching establishment

DU2W: Unit to Worker

D Probability Model

>Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ track large worker movements between SEINs
- → link entities that have different account numbes, but constitute the same economic entitiy
- relative terms, and the number of workers entering an SEIN, SPF provides a variety of link characteristics, based on the number of workers leaving an SEIN, in both absolute and again in absolute and relative terms.
- observed to move to a single successor, and that successor absorbs 80% of its employees from a single predecessor, QWI: if 80% of an SEIN's workers (the predecessor) are then all flows between those two account numbers are filtered out, and treated as if they had never existed.





Introduction

Input Files

▲ Goal: achieve a high level of accuracy and detail

Estimates: QWI

Infrastructure Files

⇒ Correction of spurious

worker flows

⇒ Solution:

Successor-Predecessor

⇒ U2W: Unit to Worker

⇒ Probability Model

⇒ Implementation

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files



Introduction

Input Files

Problem: no establishment identification on wage record

▲ Goal: achieve a high level of accuracy and detail

Infrastructure Files

- ⇒ Correction of spurious
 - worker flows ⇒ Solution:
- Successor-Predecessor

- **⊃** U2W: Unit to Worker
- ⇒ Probability Model
- Umplementation
- Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files



Introduction

▲ Goal: achieve a high level of accuracy and detail

Input Files

Problem:

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution: Successor-Predecessor

⇒ U2W: Unit to Worker

⇒ Probability Model

⊃ Implementation

Computing the statistics ⇒ Implementation

Disclosure-proofing the QWI

Publicly available files





Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⊃ U2W: Unit to Worker

⇒ Probability Model

Umplementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

DU2W: Unit to Worker

⇒ Probability Model

Umplementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

Solution: probability model for employment location and imputation



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

DU2W: Unit to Worker

⇒ Probability Model Umplementation

Computing the statistics O Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

Solution: probability model for employment location and imputation

▲ Key elements are:





Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation O Implementation Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

Solution: probability model for employment location and imputation

▲ Key elements are:

distance between place-of-work and place-of-residence



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation ⇒ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

Solution: probability model for employment location and imputation

▲ Key elements are:

1. distance between place-of-work and place-of-residence

distribution of employment across establishments of multi-establishment firms.





Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

▲ Solution: probability model for employment location and imputation

▲ Key elements are:

1. distance between place-of-work and place-of-residence

2. distribution of employment across establishments of multi-establishment firms.

Important practical aspects:



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation

Computing the statistics ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

▲ Solution: probability model for employment location and imputation

▲ Key elements are:

1. distance between place-of-work and place-of-residence

2. distribution of employment across establishments of multi-establishment firms.

Important practical aspects:

Non-ignorable missing data imputation





Introduction

Input Files

Infrastructure Files

Correction of spurious

⇒ Solution:

Successor-Predecessor

⊃U2W: Unit to Worker

⇒ Probability Model

⊃ Implementation

Disclosure-proofing the QWI

Umplementation

D Computing the statistics

Publicly available files

Conclusion

Goal: achieve a high level of accuracy and detail

Problem:

30-40% of state-wide employment in multi-establishment firms

▲ Solution: probability model for employment location and imputation

▲ Key elements are:

1. distance between place-of-work and place-of-residence

2. distribution of employment across establishments of multi-establishment firms.

▲ Important practical aspects:

Non-ignorable missing data imputation

Several million imputations every quarter



Introduction

lacktriangle workers i=1,...,I

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

U2W: Unit to Worker

⇒ Probability Model

Umplementation

⇒ Computing the statistics O Implementation

Disclosure-proofing the QWI

Publicly available files

Introduction

 \blacktriangle workers i=1,...,

 $lack firms \ j=1,...,J$

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⇒ Probability Model

Umplementation

Computing the statistics O Implementation

Disclosure-proofing the QWI

Publicly available files



Introduction

Input Files

Infrastructure Files

stimates: QWI

Correction of spurious

worker flows O Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

U2W: Unit to Worker

mpute

⇒ Probability Model

UmplementationUmplementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

lacktriangle workers i=1,...,I

 $\blacktriangle \text{ firms } j=1,...,J$

lacktriangle active establishments at firm j R_{jt}



Introduction

Input Files

Infrastructure Files

stimates: QWI

Correction of spurious

worker flows Solution:

Successor-Predecessor

O Attaching establishment

characteristics to jobs

U2W: Unit to Worker

O Probability Model

O Implementation O Implementation Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $lack workers\ i=1,...,I$

 $\blacktriangle \text{ firms } j=1,\ldots,J$

lacktriangle active establishments at firm j R_{jt}

 \blacksquare quarter t employment of establishment r in firm j N_{jrt}



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

Successor-Predecessor ⇒ Solution:

Attaching establishment characteristics to jobs

S Probability Model

Umplementation

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

lacktriangle workers i=1,...,I

 $\blacktriangle \text{ firms } j=1,...,J$

lacktriangle active establishments at firm j R_{jt}

quarter t employment of establishment r in firm $j \ N_{jrt}$

 y_{ijt} establishment at which i was employed



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious worker flows

Successor-Predecessor ⇒ Solution:

Attaching establishment

characteristics to jobs

S Probability Model

Umplementation

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $lack \operatorname{workers}\, i=1,...,I$

 $\blacktriangle \text{ firms } j=1,...,J$

lacktriangle active establishments at firm j R_{jt}

quarter t employment of establishment r in firm $j N_{jrt}$

 y_{ijt} establishment at which i was employed

▲ \mathcal{J}_t firms active



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

Attaching establishment

characteristics to jobs

S Probability Model

Umplementation ⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $lack \operatorname{workers}\, i=1,...,I$

 $\blacktriangle \text{ firms } j=1,...,J$

lacktriangle active establishments at firm j R_{jt}

quarter t employment of establishment r in firm $j N_{jrt}$

 y_{ijt} establishment at which i was employed

 \mathcal{J}_t firms active

 $lacktriangleright \mathcal{I}_{jt}$ individuals employed at firm j



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

Attaching establishment

characteristics to jobs

S Probability Model

Umplementation

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $lack \operatorname{workers}\ i=1,...,I$

▲ firms j = 1, ..., J

active establishments at firm $j R_{it}$

quarter t employment of establishment r in firm $j N_{jrt}$

 y_{ijt} establishment at which i was employed

 \mathcal{J}_t firms active

 \mathcal{I}_{jt} individuals employed at firm j

 \mathcal{R}_{jt} set of active $(N_{jrt}>0)$ establishments



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

Attaching establishment

characteristics to jobs

D Probability Model

Umplementation

Computing the statistics > Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $lack \operatorname{workers}\ i=1,...,I$

▲ firms j = 1, ..., J

active establishments at firm $j R_{it}$

quarter t employment of establishment r in firm $j N_{jrt}$

 y_{ijt} establishment at which i was employed

 \mathcal{J}_t firms active

 \mathcal{I}_{jt} individuals employed at firm j

 \mathcal{R}_{jt} set of active $(N_{jrt}>0)$ establishments

 $\mathcal{R}^i_{jt} \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for



Introduction

Input Files

Infrastructure Files

timates: QWI

Correction of spurious

worker flows

Successor-Predecessor

O Attaching establishment

characteristics to jobs U2W: Unit to Worker

Impute

Probability Model

Umplementation

Computing the statistics

Disclosure-proofing the QWI

-

Publicly available files

Conclusion

lacktriangle workers i=1,...,I

▲ firms j = 1, ..., J

lacktriangle active establishments at firm j R_{jt}

quarter t employment of establishment r in firm $j N_{jrt}$

 y_{ijt} establishment at which i was employed

▲ J_t firms active

 \blacktriangle \mathcal{I}_{jt} individuals employed at firm j

 \mathcal{R}_{jt} set of active $(N_{jrt}>0)$ establishments

 $lacktriangle \mathcal{R}_{jt}^i \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for

▶ Feasibility: an establishment $r \in \mathcal{R}^i_{jt}$ if $N_{jrs} > 0$ for every quarter s that i was employed at j.



U2W: Unit to Worker Impute

Introduction

Input Files

Infrastructure Files

orming Aggregate stimates: QWI Correction of spurious

worker flows Solution: Successor-Predecessor

Attaching establishment

characteristics to jobs

U2W: Unit to Worker

O Probability Model

Umplementation

⇒ Implementation⇒ Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

lack workers i=1,...,I

 $\blacktriangle \text{ firms } j=1,...,J$

lacktriangle active establishments at firm j R_{jt}

lacktriangle quarter t employment of establishment r in firm j N_{jrt}

 y_{ijt} establishment at which i was employed

 $lack \mathcal{J}_t$ firms active

lacktriangle Individuals employed at firm j

 \mathcal{R}_{jt} set of active $(N_{jrt}>0)$ establishments

 $lacktriangle \mathcal{R}_{jt}^i \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for

▶ Feasibility: an establishment $r \in \mathcal{R}_{jt}^i$ if $N_{jrs} > 0$ for every quarter s that i was employed at j.

Introduction

 $p_{ijrt} = \Pr\left(y_{ijt} = r\right)$

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

Impute

D Implementation O Implementation ⇒ Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs ⇒ U2W: Unit to Worker

Impute

⇒ Implementation

⇒ Computing the statistics ⇒ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

$$e^{lpha_{jrt}+x_{ijrt}'eta}$$

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

Introduction

Input Files

Infrastructure Files

Forming Aggregate Estimates: QWI Correction of spurious

worker flows Solution:

Successor-Predecessor

O Attaching establishment

characteristics to jobs DU2W: Unit to Worker

Uzw: Unit to we Impute

Probability Model

ImplementationImplementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

$$e^{\alpha_{jrt}+x'_{ijrt}\beta}$$

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

 α_{jrt} establishment- and quarter-specific effect

Introduction

Input Files

Infrastructure Files

Estimates: QWI

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

⇒ U2W: Unit to Worker Impute

⊃ Implementation

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

$$e^{\alpha_{jrt}+x'_{ijrt}\beta}$$

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

 x_{ijrt} time-varying vector, worker and establishment α_{jrt} establishment- and quarter-specific effect

Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution: Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

Umplementation

⇒ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

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

 β effect on probability of being employed at a particular x_{ijrt} time-varying vector, worker and establishment α_{jrt} establishment- and quarter-specific effect establishment

Introduction

Input Files

Infrastructure Files

orming Aggregated stimates: QWI

Correction of spurious

worker flows

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

U2W: Unit to Worker

Probability Model

O Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

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

 β effect on probability of being employed at a particular x_{ijrt} time-varying vector, worker and establishment α_{jrt} establishment- and quarter-specific effect establishment

Currently:



Introduction

Input Files

Infrastructure Files

orning Aggregated stimates: QWI Correction of spurious

worker flows Solution:
Successor-Predecessor

File Staching establishment

Attaching establishment characteristics to jobs

DU2W: Unit to Worker

Probability Model

D Probability Model
D Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

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

 β effect on probability of being employed at a particular x_{ijrt} time-varying vector, worker and establishment α_{jrt} establishment- and quarter-specific effect establishment

Currently:

lacktriangle is linear spline in distance between residence and establishment



Introduction

Input Files

Infrastructure Files

orming Aggregated stimates: QWI Correction of spurious

worker flows Solution:

Successor-Predecessor

Attaching establishment characteristics to jobs

characteristics to jobs DU2W: Unit to Worker

> Probability Model

D Probability Model
D Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 $p_{ijrt} = \Pr\left(y_{ijt} = r\right)$

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

 β effect on probability of being employed at a particular x_{ijrt} time-varying vector, worker and establishment α_{jrt} establishment- and quarter-specific effect establishment

Currently:

- \bullet x_{ijrt} is linear spline in distance between residence and establishment
- α_{jrt} is a hierarchical Bayesian model based on N_{jrt} is



Introduction

Using Minnesota data,

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs DU2W: Unit to Worker

⇒ Probability Model Impute

O Implementation

⇒ Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

compute posterior modal value of α_{jrt} Using Minnesota data,

Estimates: QWI

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model Impute

Implementation

⇒ Implementation

⇒ Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

rming Aggregate timates: QWI Correction of spurious worker flows

Solution:

Successor-Predecessor

characteristics to jobs

OU2W: Unit to Worker

O Probability Model

nplementation

© Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Using Minnesota data, compute posterior modal value of α_{jrt}

evaluate the posterior mode of $p(\beta|\alpha,x,y)$



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

Attaching establishment

characteristics to jobs

OU2W: Unit to Worker

⇒ Probability Model

⊃ Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Using Minnesota data,

evaluate the posterior mode of $p(\beta|\alpha,x,y)$ compute posterior modal value of α_{jrt}

maximize

 $\sum_{t=1} \sum_{j \in \mathcal{J}_t} \sum_{i \in \mathcal{I}_{jt}} \sum_{r \in \mathcal{R}_{jt}^i} d_{ijrt}$ 8 $\log p\left(\beta|\alpha,x,y\right)$

 $(\alpha_{jrt} + x'_{ijrt}\beta)$

 $\sum e^{\alpha_{jst} + x'_{ijst}\beta}$ $-\log$



Introduction

▲ use mean and variance of β from Minnesota data

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs CU2W: Unit to Worker

⇒ Probability Model

⇒ Implementation

⇒ Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⇒ U2W: Unit to Worker

⇒ Probability Model

> Implementation

Disclosure-proofing the QWI

⇒ Computing the statistics

Conclusion

Publicly available files

▲ use mean and variance of β from Minnesota data

 \blacktriangle take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$.



Introduction

Input Files

Infrastructure Files

orning Aggregated stimates: QWI Correction of spurious

worker flows Solution:

Successor-Predecessor

Attaching establishment characteristics to jobs

characteristics to jobs

OU2W: Unit to Worker

Probability ModelImplementation

2 Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ use mean and variance of β from Minnesota data

 \blacktriangle take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$. ▲ use QCEW employment counts, compute 10 values of α_{jt}



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

Attaching establishment characteristics to jobs

DU2W: Unit to Worker

Description of the property of the property

Umplementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ use mean and variance of β from Minnesota data

 \blacktriangle take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$. ■ use QCEW employment counts, compute 10 values of \(\alpha_{it}\)

The drawn values of lpha and eta are used to draw 10 imputed values of place of work from to the posterior predictive distribution (2) $p(\tilde{y}|x,y) = \int \int p(\tilde{y}|\alpha,\beta,x,y) p(\alpha|N) p(\beta|\alpha,x,y) d\alpha d\beta$



Introduction

Input Files

Infrastructure Files

Correction of spurious

⇒ Solution:

Successor-Predecessor

Attaching establishment characteristics to jobs

DU2W: Unit to Worker

Description of the property of the property

Umplementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ use mean and variance of β from Minnesota data

 \blacktriangle take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$. ■ use QCEW employment counts, compute 10 values of \(\alpha_{it}\)

The drawn values of α and β are used to draw 10 imputed values of place of work from to the posterior predictive distribution (2) $p(\tilde{y}|x,y) = \int p(\tilde{y}|\alpha,\beta,x,y) p(\alpha|N) p(\beta|\alpha,x,y) d\alpha d\beta$

▲ → 10 establishment identifiers associated with a job spell



Introduction

▲ We now have:

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs CU2W: Unit to Worker

⇒ Probability Model

⇒ Implementation O Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs DU2W: Unit to Worker

⇒ Probability Model

⇒ Implementation ⊃ Implementation

Disclosure-proofing the QWI

Conclusion

Publicly available files

▲ We now have:

Jobs identified

Lars Vilhuber, John Abowd and Bryce Stephens

April 8, 2005 - p. 24/31



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs ⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ We now have:

Jobs identified

Jobholder's demographics (age, gender)



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs CU2W: Unit to Worker

⇒ Probability Model

> Implementation ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ We now have:

Jobs identified

Jobholder's demographics



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics (geography and industry)



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation

⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⊃U2W: Unit to Worker

⇒ Probability Model

Umplementation ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute



Introduction

Input Files

Infrastructure Files

⇒ Correction of spurious

worker flows ⇒ Solution:

Successor-Predecessor

⇒ Attaching establishment

characteristics to jobs

⇒ U2W: Unit to Worker

⇒ Probability Model

Umplementation ⊃ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators)



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows

Successor-Predecessor

⇒ Attaching establishment characteristics to jobs

U2W: Unit to Worker

Impute

Probability Model

O Implementation O Implementation Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level



Introduction

Input Files

Infrastructure Files

Correction of spurious worker flows

Successor-Predecessor ⇒ Solution:

Attaching establishment characteristics to jobs

DU2W: Unit to Worker

⇒ Probability Model

Umplementation ⇒ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level

3. → establishment-level statistics, available in RDC



Introduction

Input Files

Infrastructure Files

stimates: QWI

Correction of spurious

worker flows Solution:
Successor-Predecessor

File Attaching establishmen

⇒ Attaching establishment characteristics to jobs

characteristics to jobs DU2W: Unit to Worker

JUZW: UNIT TO Impute

⇒ Probability Model

O Implementation
O Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level

3. → establishment-level statistics, available in RDC

4. Attach weights to each establishment



Introduction

Input Files

Infrastructure Files

Correction of spurious

worker flows ⇒ Solution: Successor-Predecessor

Attaching establishment characteristics to jobs

DU2W: Unit to Worker

Description of the property of the property

2 Implementation ⇒ Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level

3. → establishment-level statistics, available in RDC

4. Attach weights to each establishment

5. Attach 'fuzz' factors to each establishment



Introduction

Input Files

Infrastructure Files

Correction of spurious

Successor-Predecessor ⇒ Solution:

Attaching establishment

characteristics to jobs

DU2W: Unit to Worker

D Probability Model

D Implementation 2 Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level

3. → establishment-level statistics, available in RDC

4. Attach weights to each establishment

Attach 'fuzz' factors to each establishment

geography-industry-demographic detail Final aggregation to desired



Introduction

Input Files

Infrastructure Files

Correction of spurious

Successor-Predecessor ⇒ Solution:

Attaching establishment

characteristics to jobs

DU2W: Unit to Worker

D Probability Model

D Implementation 2 Implementation

Disclosure-proofing the QWI

Publicly available files

Conclusion

■ We now have:

Jobs identified

Jobholder's demographics

Establishment's characteristics

▲ Now compute

1. For each job, the relevant variables, defined at the person-level (indicators) 2. Aggregate (typically sum) to the establishment level

3. → establishment-level statistics, available in RDC

4. Attach weights to each establishment

5. Attach 'fuzz' factors to each establishment

geography-industry-demographic detail 6. Final aggregation to desired

▲ Disclosure-proof



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI isclosure-proofing the QWI

⇒ Noise-infusion

Utem suppression

Publicly available files

Conclusion

Disclosure-proofing the QWI



Noise-infusion

Introduction

▲ First layer: workplace-level aggregation

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

O Noise-infusion

Stem suppression

Publicly available files

Conclusion



Noise-infusion

Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Noise-infusion

Stem suppression

Publicly available files

Conclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

Noise-infusion

Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI ⊅Noise-infusion

Oltem suppression

Publicly available files

Conclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

 $(b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a]$ $(b-\delta)/(b-a)^2, \delta \in [a,b]$ (3) $p(\delta_j) =$



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI ⇒ Noise-infusion

⇒ Item suppression

Publicly available files

Conclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

 $(b+\delta-2)/(b-a)^2, \ \delta \in [2-b, 2-a]$ $(b-\delta)/(b-a)^2, \ \delta \in \ [a,b]$ (3) $p(\delta_j) =$

Result: random noise factor centered around 1 with distortion of at least a-1 and at most b-1.



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Disclosure-proofing the QWI

⇒ Item suppression

Publicly available files

onclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

 $(b+\delta-2)/(b-a)^2, \ \delta \in [2-b, 2-a]$ $(b-\delta)/(b-a)^2, \delta \in [a,b]$ $(3) p(\delta_j) =$

 Result: random noise factor centered around 1 with distortion of at least a-1 and at most b-1.

▲ Important properties:



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Noise-infusion

Utem suppression

Publicly available files

onclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

 $(b+\delta-2)/(b-a)^2, \ \delta \in [2-b, 2-a]$ $(b-\delta)/(b-a)^2, \delta \in [a,b]$ $(3) p(\delta_j) =$

Result: random noise factor centered around 1 with distortion of at least a-1 and at most b-1.

▲ Important properties:

1. for a given workplace, distortion is always distorted in the same direction (increased or decreased) by the same percentage amount in every period.



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI visclosure-proofing the QWI Noise-infusion

Utem suppression

Publicly available files

nclusion

▲ First layer: workplace-level aggregation

infusion of specially constructed noise:

 $(b+\delta-2)/(b-a)^2, \ \delta \in [2-b, 2-a]$ $(b-\delta)/(b-a)^2, \delta \in [a,b]$ $(3) p(\delta_j) =$

Result: random noise factor centered around 1 with distortion of at least a-1 and at most b-1.

▲ Important properties:

1. for a given workplace, distortion is always distorted in the same direction (increased or decreased) by the same percentage amount in every period.

distortion cancel out for the vast majority of the estimates. 2. when estimates are aggregated, the effects of the



▲ Second layer: after aggregations

Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

⇒ Noise-infusion

Publicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

⇒ Noise-infusion

Publicly available files

Conclusion

 Some estimates are based on fewer than three persons or firms.

▲ Second layer: after aggregations



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

sclosure-proofing the QWI

⇒ Noise-infusion
⇒ Item suppression

Item suppression

Publicly available files

Conclusion

▲ Second layer: after aggregations

 Some estimates are based on fewer than three persons or firms.

ightarrow suppression of these estimates

Lars Vilhuber, John Abowd and Bryce Stephens



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

isclosure-proofing the QWI

⇒ Noise-infusion

Item suppression

Publicly available files

Conclusion

 Some estimates are based on fewer than three persons or ▲ Second layer: after aggregations firms.

Some of the estimates are based on noisy data -> suppression of these estimates

Lars Vilhuber, John Abowd and Bryce Stephens



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

⇒ Noise-infusion

Publicly available files

Conclusion

▲ Second layer: after aggregations

 Some estimates are based on fewer than three persons or firms.

→ suppression of these estimates

Some of the estimates are based on noisy data

→ flagged as "substantially distorted"



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

isclosure-proofing the QWI

⇒ Noise-infusion

5

Publicly available files

Conclusion

▲ Second layer: after aggregations

 Some estimates are based on fewer than three persons or firms.

→ suppression of these estimates

Some of the estimates are based on noisy data

 — flagged as "substantially distorted"



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

isclosure-proofing the QWI

⇒ Noise-infusion

5

Publicly available files

Conclusion

▲ Second layer: after aggregations

 Some estimates are based on fewer than three persons or firms.

→ suppression of these estimates

Some of the estimates are based on noisy data

 — flagged as "substantially distorted"



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

iblicly available files

⇒ Publicly available files

Conclusion

Publicly available files



Introduction

▲ Published QWI

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

⇒ Publicly available files Oublicly available files

Conclusion



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

⇒ Publicly available files ublicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

Lars Vilhuber, John Abowd and Bryce Stephens



Introduction

Input Files

Infrastructure Files

▲ RDC

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

ublicly available files

⇒ Publicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

Lars Vilhuber, John Abowd and Bryce Stephens

April 8, 2005 - p. 29/31



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files D Publicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

▲ RDC

Employer characteristics files ECF → LEHD-ECF

Lars Vilhuber, John Abowd and Bryce Stephens



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Publicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

▲ RDC

■ Employer characteristics files ECF → LEHD-ECF

Establishment level flow files - Firm-level QWI -> LEHD-QWI



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files D Publicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

▲ RDC

▶ Employer characteristics files ECF → LEHD-ECF

Establishment level flow files - Firm-level QWI -> LEHD-QWI

LEHD Business Register Bridge (LEHD-BRB)



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Publicly available files

⊃ublicly availab

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

▲ RDC

Employer characteristics files ECF → LEHD-ECF

Establishment level flow files - Firm-level QWI -> LEHD-QWI

LEHD Business Register Bridge (LEHD-BRB)

Human Capital files LEHD-HCF



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files Publicly available files

Conclusion

▲ Published QWI

→ http://lehd.dsd.census.gov

▲ RDC

Employer characteristics files ECF → LEHD-ECF

Establishment level flow files - Firm-level QWI ightarrowLEHD-QWI

LEHD Business Register Bridge (LEHD-BRB)

Human Capital files LEHD-HCF

→ http://www.ces.census.gov



Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

⇒ Flow so far

Conclusion



Introduction

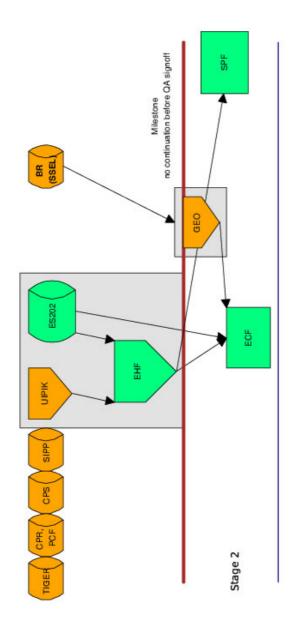
Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

Conclusion D Flow so far



Main page



Introduction

Input Files

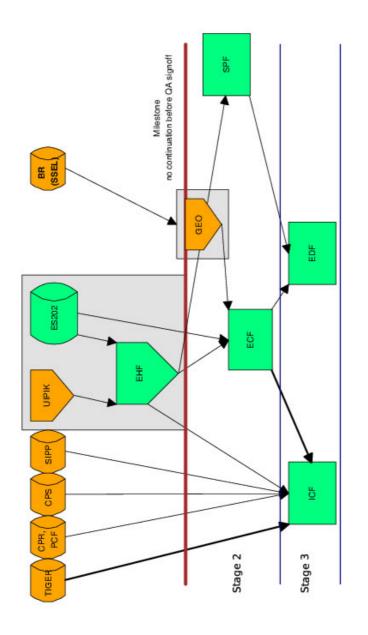
Infrastructure Files

Forming Aggregated Estimates: QWI Disclosure-proofing the QWI

Publicly available files

onclusion

⇒ Flow so far



Main page



Introduction

Input Files

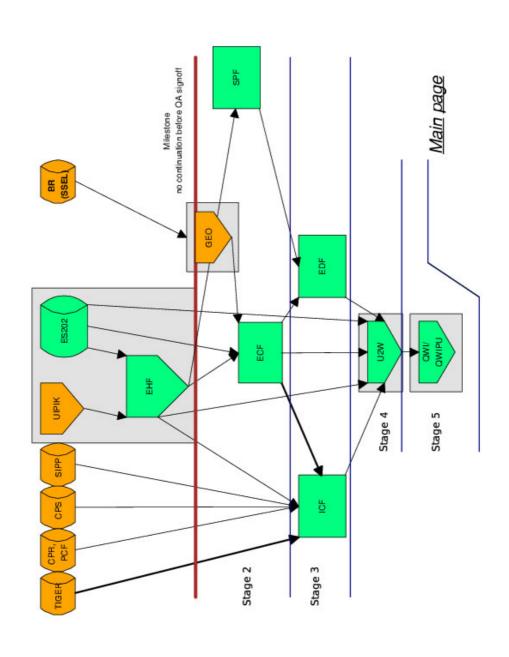
Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

⇒ Flow so far





Introduction

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

⇒ Flow so far

