

The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators

John M. Abowd^{♠,♣}, Bryce E. Stephens[♣] and Lars Vilhuber[♠]

♠ Cornell University

U.S. Census Bureau, LEHD Program



○ Introduction					
• What is it?					
O In this paper					
Input Files					
Infrastructure Files					
Forming Aggregated					

Disclosure-proofing the QWI

Publicly available files

Estimates: QWI

Conclusion

▲ Since 2003: publication of Quarterly Workforce Indicators



Introduction

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Since 2003: publication of Quarterly Workforce Indicators
- ▲ The fi rst 21st century statistical system



C Introduction

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Since 2003: publication of Quarterly Workforce Indicators
- ▲ The first 21st century statistical system
 - No additional burden



_				
	Int	$r \cap c$	luct	-

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



C Introduction

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 confi dentiality protection methods



Introduction

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 confi dentiality protection methods
 - Innovative use of wage records to constitute a frame to integrate data



C Introduction

- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 confi dentiality protection methods
 - Innovative use of wage records to constitute a frame to integrate data
 - The fi rst statistical system to use "jobs" as a frame



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion



Introduction

• What is it?

O 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)



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

- (state) administrative records data on workers (UI Wage records)
- (state) administrative records data on fi rms (QCEW aka ES-202)



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

- (state) administrative records data on workers (UI Wage records)
- (state) administrative records data on fi rms (QCEW aka ES-202)
- administrative information on demographics



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

- (state) administrative records data on workers (UI Wage records)
- (state) administrative records data on fi rms (QCEW aka ES-202)
- administrative information on demographics
- surveys on people and firms collected by Census Bureau



Introduction

• What is it?

O 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 fi rms (QCEW aka ES-202)
- administrative information on demographics
- surveys on people and firms collected by Census Bureau
- careful longitudinal edit of person identifi ers and economic fi rm units



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- Combines
 - (state) administrative records data on workers (UI Wage records)
 - (state) administrative records data on fi rms (QCEW aka ES-202)
 - administrative information on demographics
 - surveys on people and firms collected by Census Bureau
- careful longitudinal edit of person identifi ers and economic fi rm units
- careful longitudinal edit of person and fi rm characteristics



• Introduction

OWhat is it?

O 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



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Describe the construction of the LEHD infrastructure
 - ... in particular the imputation mechanisms used



▶ Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Describe the construction of the LEHD infrastructure
 - ... in particular the imputation mechanisms used
- ▲ Describe the computation of the QWI statistics



IntroductionWhat is it?

On this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



Introduction

OWhat is it?

O In this paper

Input Files

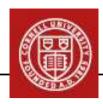
Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



IntroductionWhat is it?

In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 fi les and confi dential QWI fi les



- Introduction
- What is it?
- O In this paper

Input 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
- OWhat is it?
- O In this paper

Input 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
- OWhat is it?
- O In this paper

Input Files

○ Wage records: UI

- ♠ Employer reports: ES202
- Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- OWhat is it?
- O In this paper

Input Files

○ Wage records: UI

- © Employer reports: ES202
- Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- 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
- ▲ identifi es EARNINGS, EMPLOYER, TIME PERIOD



- Introduction
- OWhat is it?
- O In this paper

Input Files

○ Wage records: UI

- ▶ Employer reports: ES202
- Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- ▲ identifi es EARNINGS, EMPLOYER, TIME PERIOD
- ▲ some limited other state-dependent information available



- Introduction
- OWhat is it?
- O In this paper

Input Files

○ Wage records: UI

- ▶ Employer reports: ES202
- Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- ▲ identifi es EARNINGS, EMPLOYER, TIME PERIOD
- ▲ some limited other state-dependent information available
- ▲ in particular, for Minnesota, the ESTABLISHMENT is reported



Introduction

• What is it?

O In this paper

Input Files

O Wage records: UI

○ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

... or QCEW



- Introduction
- OWhat is it?
- O In this paper

Input Files

○ Wage records: UI

○ Employer reports: ES202

Demographics

Infrastructure Files

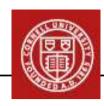
Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

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



- Introduction
- O What is it?
- O In this paper

Input Files

O Wage records: UI

▶ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- What is it?
- O In this paper

Input Files

○ Wage records: UI

▶ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- OWhat is it?
- O In this paper

Input Files

ひ Wage records: UI

▶ Employer reports: ES202

Demographics

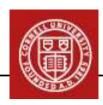
Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- lacktriangle fundamental unit: 'reporting unit' (pprox establishment)



- Introduction
- OWhat is it?
- O In this paper

Input Files

ひ Wage records: UI

▶ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- \blacktriangle fundamental unit: 'reporting unit' (\approx establishment)
- One report per establishment per quarter is fi led



Demographics

- Introduction
- OWhat is it?
- O In this paper

Input Files

- Wage records: UI
- Employer reports: ES202

O Demographics

Infrastructure Files

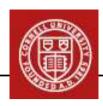
Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

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



Demographics

- Introduction
- OWhat is it?
- O In this paper

Input Files

- Wage records: UI
- C Employer reports: ES202

O Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Demographics are taken from a number of Census-internal files derived from administrative data:
 - Person Characteristics File (PCF)



Demographics

- Introduction
- OWhat is it?
- O In this paper

Input Files

- Wage records: UI
- C Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Demographics are taken from a number of Census-internal files derived from administrative data:
 - Person Characteristics File (PCF)
 - Census Numident



- Introduction
- OWhat is it?
- O In this paper

Input Files

- Wage records: UI
- ▶ Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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:



- Introduction
- OWhat is it?
- O In this paper

Input Files

- O Wage records: UI
- C Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



- Introduction
- OWhat is it?
- O In this paper

Input Files

- O Wage records: UI
- C Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



- Introduction
- OWhat is it?
- O In this paper

Input Files

- O Wage records: UI
- C Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



- Introduction
- OWhat is it?
- O In this paper

Input Files

- O Wage records: UI
- Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



- Introduction
- OWhat is it?
- O In this paper

Input Files

- O Wage records: UI
- Employer reports: ES202

Demographics

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- ♦ What is it?
- O In this paper

Input Files

Infrastructure Files

© EHF: Employment History Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Infrastructure Files



EHF: Employment History Files

- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

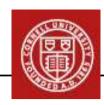
○ EHF: Employment History Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Job-level EHF
 - complete in-state work history for each individual on Ulwage records.
 - one record for each employee-employer combination a job
 - earnings and employment patterns



EHF: Employment History Files

- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

○ EHF: Employment History Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Job-level EHF
 - complete in-state work history for each individual on Ulwage records.
 - one record for each employee-employer combination a job
 - 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

○ EHF: Employment History Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

- ▲ Job-level EHF
 - complete in-state work history for each individual on Ulwage records.
 - one record for each employee-employer combination a job
 - 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
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- **©** Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Forming Aggregated Estimates: QWI



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

• Correction of spurious worker flows

- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- C Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

• Correction of spurious worker flows

- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- C Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

• Correction of spurious worker flows

- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- C Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

• Correction of spurious worker flows

- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- C Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

• Correction of spurious worker flows

- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- C Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ Goal: achieve a high level of accuracy and detail



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ Goal: achieve a high level of accuracy and detail
- Problem: no establishment identifi cation on wage record



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ Goal: achieve a high level of accuracy and detail
- ▲ Problem:



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ Goal: achieve a high level of accuracy and detail
- ▲ Problem:
- ▲ 30-40% of state-wide employment in multi-establishment firms



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

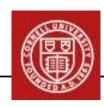
Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

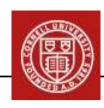
Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 fi rms.



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 fi rms.
- ▲ Important practical aspects:
 - Non-ignorable missing data imputation



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ 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 fi rms.
- ▲ Important practical aspects:
 - Non-ignorable missing data imputation
 - Several million imputations every quarter



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

• Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 \blacktriangle workers i=1,...,I



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j = 1, ..., J



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- lacktriangle active establishments at fi rm j R_{jt}



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j = 1, ..., J
- \blacktriangle active establishments at firm j R_{it}
- lacktriangle quarter t employment of establishment r in fi rm j N_{jrt}



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

• Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- lacktriangle active establishments at fi rm j R_{jt}
- lacktriangle quarter t employment of establishment r in fi rm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

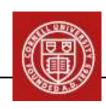
• Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j = 1, ..., J
- lacktriangle active establishments at fi rm j R_{jt}
- lacktriangle quarter t employment of establishment r in firm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

• Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- \blacktriangle active establishments at firm j R_{it}
- lacktriangle quarter t employment of establishment r in fi rm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active
- lacktriangle \mathcal{I}_{jt} individuals employed at fi rm j



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

• Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- lacktriangle active establishments at fi rm j R_{it}
- lacktriangle quarter t employment of establishment r in fi rm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active
- $\blacktriangle \mathcal{I}_{it}$ individuals employed at fi rm j
- \blacktriangle \mathcal{R}_{it} set of active ($N_{irt} > 0$) establishments



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

Attaching establishment characteristics to jobs

- Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j = 1, ..., J
- lacktriangle active establishments at fi rm j R_{jt}
- lacktriangle quarter t employment of establishment r in fi rm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active
- lacktriangle \mathcal{I}_{jt} individuals employed at fi rm j
- \blacktriangle \mathcal{R}_{it} set of active $(N_{irt} > 0)$ establishments
- $lacktriangleq \mathcal{R}_{jt}^i \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for worker i.



Attaching establishment characteristics to jobs

- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

Attaching establishment characteristics to jobs

- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- lacktriangle active establishments at fi rm j R_{jt}
- lacktriangle quarter t employment of establishment r in firm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active
- lacktriangle \mathcal{I}_{jt} individuals employed at fi rm j
- \blacktriangle \mathcal{R}_{jt} set of active $(N_{jrt} > 0)$ establishments
- $\blacktriangle \mathcal{R}^i_{jt} \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for worker i.
- ▲ Feasibility: an establishment $r \in \mathcal{R}^i_{jt}$ if $N_{jrs} > 0$ for every quarter s that i was employed at j.



Attaching establishment characteristics to jobs

- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs

Attaching establishment characteristics to jobs

- C Probability Model
- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle workers i=1,...,I
- \blacktriangle firms j=1,...,J
- lacktriangle active establishments at fi rm j R_{jt}
- lacktriangle quarter t employment of establishment r in firm j N_{jrt}
- \blacktriangle y_{ijt} establishment at which i was employed
- \blacktriangle \mathcal{J}_t firms active
- $lacktriangleq \mathcal{I}_{jt}$ individuals employed at fi rm j
- \blacktriangle \mathcal{R}_{jt} set of active $(N_{jrt} > 0)$ establishments
- $\blacktriangle \mathcal{R}^i_{jt} \subset \mathcal{R}_{jt}$ set of active establishments that are feasible for worker i.
- ▲ Feasibility: an establishment $r \in \mathcal{R}^i_{jt}$ if $N_{jrs} > 0$ for every quarter s that i was employed at j.



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

• Probability Model

- ▶ Implementation
- ▶ Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Correction of spurious worker flows

 Attaching establishment characteristics to jobs

 Attaching establishment characteristics to jobs

• Probability Model

O Implementation

• Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

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

(1)
$$p_{ijrt} = \frac{e^{\alpha_{jrt} + x'_{ijrt}\beta}}{\sum_{s \in \mathcal{R}^i_{jt}} e^{\alpha_{jst} + x'_{ijst}\beta}}$$



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

• Probability Model

- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

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

(1)
$$p_{ijrt} = \frac{e^{\alpha_{jrt} + x'_{ijrt}\beta}}{\sum_{s \in \mathcal{R}_{jt}^i} e^{\alpha_{jst} + x'_{ijst}\beta}}$$

 α_{jrt} establishment- and quarter-specifi c effect



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

C Probability Model

- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

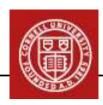
Conclusion

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

(1)
$$p_{ijrt} = \frac{e^{\alpha_{jrt} + x'_{ijrt}\beta}}{\sum_{s \in \mathcal{R}_{jt}^i} e^{\alpha_{jst} + x'_{ijst}\beta}}$$

 α_{jrt} establishment- and quarter-specifi c effect

 x_{ijrt} time-varying vector, worker and establishment



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

• Probability Model

- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

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

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

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



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

C Probability Model

- ▶ Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

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

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

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

Currently:



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

Probability Model

- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

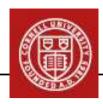
$$p_{ijrt} = \Pr\left(y_{ijt} = r\right)$$

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

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

Currently:

• x_{ijrt} is linear spline in distance between residence and establishment



Introduction

O What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs

Probability Model

- Implementation
- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

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

(1)
$$p_{ijrt} = \frac{e^{\alpha_{jrt} + x'_{ijrt}\beta}}{\sum_{s \in \mathcal{R}_{jt}^i} e^{\alpha_{jst} + x'_{ijst}\beta}}$$

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

Currently:

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



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model

Implementation

- Implementation
- Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

Using Minnesota data,



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model

Implementation

- C Implementation
- Computing the statistics

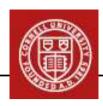
Disclosure-proofing the QWI

Publicly available files

Conclusion

Using Minnesota data,

compute posterior modal value of α_{jrt}



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model

Implementation

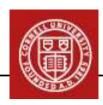
- Implementation
- Computing the statistics

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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model

Implementation

- 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 $p(\beta|\alpha,x,y)$

$$\log p\left(\beta | \alpha, x, y\right) \propto \sum_{t=1}^{T} \sum_{j \in \mathcal{J}_t} \sum_{i \in \mathcal{I}_{jt}} \sum_{r \in \mathcal{R}_{jt}^i} d_{ijrt} \left(\alpha_{jrt} + x'_{ijrt}\beta - \log \left(\sum_{s \in \mathcal{R}_{jt}^i} d_{ijrt}\right)\right)$$

(2)

maximize



Introduction

OWhat is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation

Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

 \blacktriangle use mean and variance of β from Minnesota data



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation

Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle use mean and variance of β from Minnesota data
- ▲ take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$.



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation

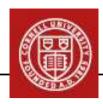
Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

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



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation

Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- \blacktriangle use mean and variance of β from Minnesota data
- ▲ take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$.
- lacktriangle use QCEW employment counts, compute 10 values of $lpha_{jt}$
- ▲ The drawn values of α and β are used to draw 10 imputed values of place of work from to the posterior predictive distribution

(3)
$$p\left(\tilde{y}|x,y\right) = \int \int p\left(\tilde{y}|\alpha,\beta,x,y\right) p\left(\alpha|N\right) p\left(\beta|\alpha,x,y\right) d\alpha d\beta$$



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation

Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

- \blacktriangle use mean and variance of β from Minnesota data
- ▲ take 10 draws of β from the normal approximation (at the mode) to $p(\beta|\alpha, x, y)$.
- lacktriangle use QCEW employment counts, compute 10 values of $lpha_{jt}$
- ▲ The drawn values of α and β are used to draw 10 imputed values of place of work from to the posterior predictive distribution

(3)
$$p(\tilde{y}|x,y) = \int \int p(\tilde{y}|\alpha,\beta,x,y) p(\alpha|N) p(\beta|\alpha,x,y) d\alpha d\beta$$

▲ → 10 establishment identifi ers associated with a job spell



Introduction

• What is it?

O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

Conclusion

▲ We now have:



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- ▶ Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics (age, gender)



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

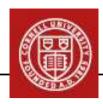
- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics (geography and industry)



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

• Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- Now compute



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- ▲ Now compute
 - 1. For each job, the relevant variables, defi ned at the person-level (indicators)



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- ▲ Now compute
 - 1. For each job, the relevant variables, defi ned at the person-level (indicators)
 - 2. Aggregate (typically sum) to the establishment level



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- ▲ Now compute
 - 1. For each job, the relevant variables, defi ned at the person-level (indicators)
 - 2. Aggregate (typically sum) to the establishment level
 - 3. → establishment-level statistics, available in RDC



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- C Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- ▲ Now compute
 - 1. For each job, the relevant variables, defi ned 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

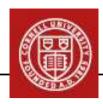
- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- ▲ Now compute
 - 1. For each job, the relevant variables, defi ned 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

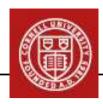
- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- Now compute
 - 1. For each job, the relevant variables, defi ned 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
 - 6. Final aggregation to desired geography-industry-demographic detail



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

- Correction of spurious worker flows
- Attaching establishment characteristics to jobs
- Attaching establishment characteristics to jobs
- Probability Model
- Implementation
- Implementation

Computing the statistics

Disclosure-proofing the QWI

Publicly available files

- ▲ We now have:
 - Jobs identifi ed
 - Jobholder's demographics
 - Establishment's characteristics
- Now compute
 - 1. For each job, the relevant variables, defi ned 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
 - 6. Final aggregation to desired geography-industry-demographic detail
- ▲ Disclosure-proof



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- O Noise-infusion
- Cell suppression

Publicly available files

Conclusion

Disclosure-proofing the QWI



Noise-infusion

- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- Noise-infusion
- Cell suppression

Publicly available files

Conclusion

▲ First layer: workplace-level aggregation



Noise-infusion

- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- ▶ Noise-infusion
- Cell suppression

Publicly available files

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

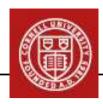
Disclosure-proofing the QWI

- O Noise-infusion
- Cell suppression

Publicly available files

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:

(4)
$$p(\delta_j) = \begin{cases} (b-\delta)/(b-a)^2, \ \delta \in [a,b] \\ (b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a] \end{cases}$$



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Noise-infusion

Cell suppression

Publicly available files

Conclusion

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:

(4)
$$p(\delta_j) = \begin{cases} (b-\delta)/(b-a)^2, \ \delta \in [a,b] \\ (b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a] \end{cases}$$

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



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Noise-infusion

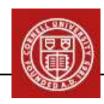
Cell suppression

Publicly available files

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:

(4)
$$p(\delta_j) = \begin{cases} (b-\delta)/(b-a)^2, \ \delta \in [a,b] \\ (b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a] \end{cases}$$

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



- Introduction
- What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Noise-infusion

Cell suppression

Publicly available files

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:

(4)
$$p(\delta_j) = \begin{cases} (b-\delta)/(b-a)^2, \ \delta \in [a,b] \\ (b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a] \end{cases}$$

- 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI Noise-infusion

Cell suppression

Publicly available files

- ▲ First layer: workplace-level aggregation
 - infusion of specially constructed noise:

(4)
$$p(\delta_j) = \begin{cases} (b-\delta)/(b-a)^2, \ \delta \in [a,b] \\ (b+\delta-2)/(b-a)^2, \ \delta \in [2-b,2-a] \end{cases}$$

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



OI	Introdu	uction
----	---------	--------

- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Noise-infusion

• Cell suppression

Publicly available files

Conclusion

▲ Second layer: after aggregations



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Noise-infusion

Cell suppression

Publicly available files

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



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Noise-infusion

Cell suppression

Publicly available files

- ▲ Second layer: after aggregations
 - Some estimates are based on fewer than three persons or firms.
 - → suppression of these estimates



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

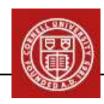
Disclosure-proofing the QWI

Noise-infusion

Cell suppression

Publicly available files

- ▲ 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



- Introduction
- O What is it?
- O In this paper

Input Files

Infrastructure Files

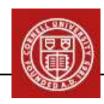
Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- Noise-infusion
- Cell suppression

Publicly available files

- ▲ 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
 - → fagged as "substantially distorted"



- Introduction
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

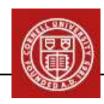
Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- Noise-infusion
- Cell suppression

Publicly available files

- ▲ 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
- OWhat is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

- Noise-infusion
- Cell suppression

Publicly available files

- ▲ 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
- ♦ What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion

Publicly available files



- Introduction
- ♦ What is it?
- O In this paper

Input Files

Infrastructure Files

Forming Aggregated Estimates: QWI

Disclosure-proofing the QWI

Publicly available files

Conclusion