Changes and Enhancements to the Quarterly Workforce Indicators

John M. Abowd
Census Bureau
Cornell University

Annual Local Employment Dynamics Workshop
January 20, 2006
Topics

• All suppressions to be eliminated from the Quarterly Workforce Indicators
• The historical NAICS QWI data ready for one-time production run
• Moving to core-based statistical areas and other geography enhancements
• New individual characteristics available
Those Pesky QWI Suppressions

• Did you notice that none of the users of On The Map yesterday mentioned the confidentiality protections?
• Did you notice that aggregations and custom areas produced reports seamlessly?
• Did you think that meant that the protections used in the QWIs were not used in On The Map?
Well, You’re Mistaken

- All QWI protections (noise infusion of the micro-data) are present in the On The Map application
- Additional synthetic data methods were used to protect the household data
- We just figured out how to make those methods more transparent for the user without sacrificing validity or protection
- And now, we’ve done it for the QWIs, too
Theoretical Distribution of the QWI Distortion Factor
Actual Confidentiality Protection Distortion: Employment, Beginning-of-Quarter
Table 8: Distribution of the Error in the First Order Serial Correlation: Two-digit SIC × County, Raw vs. Published Data

\[ \Delta r = r - r^* \]

<table>
<thead>
<tr>
<th>Percentile</th>
<th>Beginning of Quarter</th>
<th>Employment</th>
<th>Accessions</th>
<th>Separations</th>
<th>Full Quarter</th>
<th>Net Job Flows</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>-0.129094</td>
<td>-0.104500</td>
<td>-0.102003</td>
<td></td>
<td>-0.123819</td>
<td>-0.019439</td>
</tr>
<tr>
<td>05</td>
<td>-0.056734</td>
<td>-0.054465</td>
<td>-0.054423</td>
<td></td>
<td>-0.054914</td>
<td>-0.006630</td>
</tr>
<tr>
<td>10</td>
<td>-0.038474</td>
<td>-0.037901</td>
<td>-0.036443</td>
<td></td>
<td>-0.036726</td>
<td>-0.004058</td>
</tr>
<tr>
<td>25</td>
<td>-0.016431</td>
<td>-0.016847</td>
<td>-0.016628</td>
<td></td>
<td>-0.016082</td>
<td>-0.001277</td>
</tr>
<tr>
<td>50</td>
<td>-0.001610</td>
<td>-0.002131</td>
<td>-0.000789</td>
<td></td>
<td>-0.001742</td>
<td>0.000022</td>
</tr>
<tr>
<td>75</td>
<td>0.011486</td>
<td>0.011319</td>
<td>0.013833</td>
<td></td>
<td>0.010231</td>
<td>0.001235</td>
</tr>
<tr>
<td>90</td>
<td>0.029364</td>
<td>0.027751</td>
<td>0.031744</td>
<td></td>
<td>0.026192</td>
<td>0.003639</td>
</tr>
<tr>
<td>95</td>
<td>0.043912</td>
<td>0.039888</td>
<td>0.046670</td>
<td></td>
<td>0.040161</td>
<td>0.005915</td>
</tr>
<tr>
<td>99</td>
<td>0.082596</td>
<td>0.079321</td>
<td>0.098374</td>
<td></td>
<td>0.076498</td>
<td>0.014536</td>
</tr>
</tbody>
</table>
Graph: Distribution of Error in First Order Serial Correlation
## Percentage of Data Items in QWI County-level Release File

**NAICS IL 2001:1-2004:1**

**Employment (Beginning-of-quarter)**

<table>
<thead>
<tr>
<th></th>
<th>Sector</th>
<th>Sub-sector</th>
<th>Industry Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Released</strong></td>
<td>86.45</td>
<td>75.43</td>
<td>70.08</td>
</tr>
<tr>
<td><strong>Not significantly distorted</strong></td>
<td>70.06</td>
<td>58.96</td>
<td>57.22</td>
</tr>
<tr>
<td><strong>Significantly distorted</strong></td>
<td>16.39</td>
<td>16.47</td>
<td>12.86</td>
</tr>
<tr>
<td><strong>Suppressed</strong></td>
<td>13.54</td>
<td>24.57</td>
<td>29.91</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sector</th>
<th>Sub-sector</th>
<th>Industry Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Released</strong></td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
<tr>
<td><strong>Not significantly distorted</strong></td>
<td>70.07</td>
<td>58.96</td>
<td>57.23</td>
</tr>
<tr>
<td><strong>Significantly distorted</strong>*</td>
<td>29.93</td>
<td>41.04</td>
<td>42.77</td>
</tr>
<tr>
<td><strong>Suppressed</strong></td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*approximate
Beginning Period Employment in NAICS Sector 52
Men and Women Ages 19-21

Counts

Time

Current QWI
Improved QWI
Full-Quarter New Hires in NAICS 3259
Women Aged 55-64
NAICS QWI All the Way Back

• Kevin McKinney has completed the LEHD assessment of the NAICS historical series.
• Analysis performed for Illinois but other states can be run through the same methods (one-time computation).
Comparison with BLS QCEW Employment Series

- Bureau of Labor Statistics publicly available historical NAICS series
- BLS 1991-2003 Illinois month 1 private firm employment
- QWI 1991-2003 Illinois beginning of period employment
- QWI weights designed such that beginning period employment is equal to BLS month 1 employment for state
- Quarterly totals are thus equal for both series
Additional cross-state inconsistency in historical NAICS coding of NAICS 55

- QCEW reports have two fields for NAICS coding:
  - Field A: BLS-style coding
  - Field B: Census-style coding
- Cross-state inconsistency, some time inconsistency in reporting
- New read-in algorithm for historical series corrects this problem
Construction of the Alternative Series

• Two QWI production runs
  – Run 1: standard ES-202 data read-in and no use of the BLS LDB
  – Run 2: revised ES-202 data read-in and use of the BLS LDB
Run Difference Definitions

- Run 1 uses the LEHD backcoding: Maps reported SIC to NAICS code based on frequency-based imputation model. If SIC changes over time, NAICS changes.
- Run 2 uses the LDB: Provides the LEHD imputation algorithm with LDB NAICS codes, which take precedence if available. By design, less variability: LBD carries back NAICS code at time of conversion, if available. In general, an establishment’s NAICS does not change over time.
Results

- Quarterly state-wide totals are equal by construction
- Most NAICS sector employment totals are within 5% of each other
- General Convergence with BLS over time for both run 1 and run 2
- Use of the LDB has the expected effect, reducing discrepancies between LEHD and BLS employment series
- Occasional large differences remain even after integration of LDB
NAICS Sector Results

- 4 sectors in run 1 (without LDB) show an average difference over the period of greater than 5%
  - 42 Wholesale trade
  - 52 Finance and insurance
  - 55 Management of companies and enterprises
  - 56 Administrative and waste services
NAICS Sector Results (cont.)

• 2 sectors in run 2 (with LDB) show an average difference over the period of greater than 5%
  – 11 Agriculture
  – 55 Management of companies and enterprises

• NAICS 55 is present in both runs:
  – Time inconsistency explains part of Run 1.
  – Systematic difference in Run 2 unexplained
NAICS 42 Wholesale trade

Year and Quarter

Percent Difference

LEHD-BLS no LDB
LEHD-BLS with LDB
NAICS 55 Management of companies and enterprises

Year and Quarter

Percent Difference

LEHD-BLS no LDB
LEHD-BLS with LDB
Geography: Down to the Core

• Moving to Core Based Statistical Areas (CBSA)
• Each CBSA must contain at least one urban area of 10,000 or more population.
• Each metropolitan statistical area must have at least one urbanized area of 50,000 or more inhabitants.
• Each micropolitan statistical area must have at least one urban cluster of at least 10,000 but less than 50,000 population.
Switching to CBSA

- Geography vintage updating began in July 2005.
- The vintage update switches all QWI geography definitions to CBSA basis.
- Testing of the vintage update should be completed in the current cycle.
- First cycle following the vintage update will reflect CBSA.
Race, Ethnicity and Education Finally Arrive

- Enhancements to the individual characteristics available in the LEHD Infrastructure File system
- New race, ethnicity and education variables
- Enhanced residential address information
Race (OMB compliant definitions)

- White Alone
- Black Alone
- American Indian and/or Alaska Native Alone
- Asian Alone
- Native Hawaiian and/or Other Pacific Islander Alone
- Some Other Race Alone
- Two or More Race Groups
Ethnicity

- Hispanic (all designations)
- Non-Hispanic (not Spanish/Hispanic/Latino)
New Education Information

- Highest completed grade as of April 1, 2000
  - Not in universe (less than 3 years old)
  - No schooling completed
  - Nursery school to 4th grade
  - 5th grade or 6th grade
  - 7th grade or 8th grade
  - 9th grade
  - 10th grade
  - 11th grade
  - 12th grade, no diploma
  - High school graduate
  - Some college, but less than 1 year
  - 1 or more years of college, no degree
  - Associate degree
  - Bachelor’s degree
  - Master’s degree
  - Professional degree
  - Doctorate degree
What’s Next?

- Port the confidentiality protection enhancements to production
- Offer LED partners historical NAICS to either with or without LDB backcoding (partner choice)
- Develop release files for CBSA data
- Develop QWI enhancements using new race, ethnicity and education variables