Using Administrative Data to Calculate Export Price Indexes

Don Fast and Susan E. Fleck, PhD
International Price Program
Office of Prices and Living Conditions
U.S. Bureau of Labor Statistics
Improving Economic Price Statistics through the Use of Alternative Data
Session sponsored by
AEA Committee on Economics Statistics
January 5, 2020
Official Import & Export Price Indexes – The Basics
Official Import & Export Price Indexes

- Challenges
  - Staff reduction → Sample size constraints → Half of detailed price indexes not publication quality
  - Concentration of large companies → High impact on nonresponse rate

- Opportunities
  - Administrative Trade Data → Exponentially more items and prices → Expand number and deepen coverage of indexes
Published Share of All 5-digit MXPI

5-digit BEA End Use MXPI

- Published XPI, 50
- Published MPI, 76
- Unpublished XPI, 77
- Unpublished MPI, 61
Using Administrative Trade Data for Price Indexes

Recent prototype of 2 UV indexes addresses:

- Not unique items, just similar items
- Not able to track across months
- Use an ‘average price’ concept
- Shipment records don’t provide detailed product information

Not yet addressed:

- Not timely enough for news release
Research Questions

- Can average prices and unit value indexes be used in MXPI?
- If so, which ones?
Unit Value Bias

Matched Model – Actual Price of Unique Item

Unit Value – Average Price of Similar Groups of Items
MXPI Survey vs. Admin Data

**MXPI Survey**
- Matched model
- Current actual Price
- Sample size limits representative coverage
- 20k items/month
- Nonresponse rate and participation

**Administrative Trade Data**
- Unit values – Total $ value & Quantity
- Current avge price and Quantity
- No constraint on representative coverage
- Millions of items/month
- Unit value bias and outliers
Research Questions

- How do we select unit value indexes without unit value bias?
- What is the impact of administrative trade data source and new BLS methods on GDP measurement thru Net Trade?
Import & Export Price Indexes

10-digit Harmonized System
Product Category

BEA End Use 5-digit

HTSA (MPI)
Schedule B (XPI) 4-digit

NAICS 6-digit

Items priced monthly

Upper Level Strata

Lower Level Strata

Entry Level Item
Unit Value Calculation - ELI

\[ p_{(j,t),H}^{(j,t),H} = \frac{\sum_{z \in i} p_{K_i,z}^{(j,t),H}}{|z|} \]

\[ p_{(j,t)}^{H} = \exp \left( \frac{\sum_{i \in j} \left[ w_{K_i}^{(j,t),H} \cdot ln \left( p_{K_i}^{(j,t),H} \right) \right]}{\sum_{i \in K} w_{K_i}^{(j,t),H}} \right) \]
Lower Level Strata Calculation

- Tornquist

\[ I_{t,0}^T = \left\{ \prod_i \left( \frac{P_{i,t}}{P_{i,0}} \right)^{\frac{(W_{i,0} + W_{i,t})}{2}} \right\} \times 100 \]

- where \( W_{i,t} = \left( \frac{P_{i,t}Q_{i,t}}{\sum_i P_{i,t}Q_{i,t}} \right) \) \( W_{i,0} = \left( \frac{P_{i,0}Q_{i,0}}{\sum_i P_{i,0}Q_{i,0}} \right) \)
Upper Level Strata Calculation

Laspeyres

\[ L_t,0 = \left( \frac{\sum_i P_{i,t} Q_{i,0}}{\sum_i P_{i,0} Q_{i,0}} \right) \times 100 \]
Admin Data Address Criticisms of Current Methodology

- Lower Level Substitution Bias ✔️
- Upper Level Substitution Bias ✔️
- Product Bias ✔️
- Country Substitution Bias ✔️
- Quality Bias
- Outsourcing Bias N/A
Research Approach

- Jan 2012-Dec 2017
- 200 million trade records
- Create monthly STRs w/ new methodology
- Create LTRs w/ current methodology
- Unique ELIs share
  - 10-digit Harmonized System code
  - Employer ID
  - Domestic/Foreign Content
  - State of Origin
  - Country of Destination
  - Unit of Measure
  - Intercompany Trade
Research Approach

- Calculate 127 5-digit BEA End Use unit value indexes
- Test for homogeneity
- Test other ‘best fit’ characteristics vs. official comparable price index
- Group unit value indexes by quality
- Determine impact on real value of exports
Homogeneity “Floor” – Vegetable Price STRs Coefficient of Variation

Note: STRs of ELIS based on concatenations of price related characteristics: Domestic/Foreign (F), EIN (E), State of Origin (S), Country of Destination (C), Unit of Measure (Q), Related Transaction (R) and HS.
Homogeneous and Heterogeneous based upon the CV Test \( n=127 \)

- 73 Heterogeneous 5-Digit BEA Indexes
- 54 Homogeneous 5-Digit BEA Indexes
Coefficient of Variation for “Good" Indexes based upon the CV test

N=24
Coefficient of Variation for “Undecided” Indexes based upon the CV test

N=28
## 5-digit Export UV Indexes by Quality

<table>
<thead>
<tr>
<th>Homogeneous/Heterogeneous</th>
<th>Index Quality</th>
<th>Number of 5-digit BEA End Use U.V. Indexes</th>
<th>Trade Dollar Value, 2015 In millions</th>
<th>Percent Trade Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homogeneous</td>
<td>Good</td>
<td>24</td>
<td>$328,869</td>
<td>22.5</td>
</tr>
<tr>
<td>Homogeneous</td>
<td>Undecided</td>
<td>19</td>
<td>$150,099</td>
<td>10.3</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>Undecided</td>
<td>11</td>
<td>$136,100</td>
<td>9.3</td>
</tr>
<tr>
<td>Homogeneous</td>
<td>Poor</td>
<td>9</td>
<td>$68,781</td>
<td>4.7</td>
</tr>
<tr>
<td>Heterogeneous</td>
<td>Poor</td>
<td>64</td>
<td>$777,116</td>
<td>53.2</td>
</tr>
<tr>
<td><strong>ALL INDEXES</strong></td>
<td></td>
<td><strong>127</strong></td>
<td><strong>$1,460,964</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>
All Export Indexes by Category

n=127

- 24 Homogeneous Good: 17.39%
- 30 Homogeneous Undecided: 21.74%
- 9 Homogeneous Poor: 6.52%
- 64 Heterogeneous Poor: 46.38%
- 11 Heterogeneous Undecided: 7.97%
Quality Groups of the 5-Digit BEA Research Indexes

- Good - Homogeneous products that pass all mean and SD tests, and at least one of the three statistical tests.
- Undecided - HM and HT products that demonstrate potential bias, but with changes in methods may produce a quality index.
- Poor – There is not enough detail in the item characteristics to validate it is a homogenous item.
### Average Index Value Variability by Quality Groups

**(January 2012-December 2017)**

Jan. 2012 = 100

<table>
<thead>
<tr>
<th>Quality Group</th>
<th>Mean Official XPI Values</th>
<th>Mean Unit Value XPI</th>
<th>Std. Dev. Official XPI</th>
<th>Std. Dev. Unit Value XPI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good</td>
<td>89.6</td>
<td>90.9</td>
<td>13.4</td>
<td>11.2</td>
</tr>
<tr>
<td>Undecided</td>
<td>98.2</td>
<td>106.0</td>
<td>8.6</td>
<td>13.9</td>
</tr>
<tr>
<td>Poor</td>
<td>100.7</td>
<td>220.6</td>
<td>4.4</td>
<td>158.9</td>
</tr>
</tbody>
</table>
Distribution of Correlation Coefficients by Quality Groups

-0.6  -0.4  -0.2  0.0  0.2  0.4  0.6  0.8  1.0

Correlation Coefficient

Total n=125
Distribution of Root Mean Square Errors by Quality Groups

Root Mean Square Error

Total n=125
Partial 5-digit BEA XPI, Dec 2017
(Jan 2012=100)

Annual Average Difference in Price Levels

- **Poor (N=73)**: 99.47 (Official Partial Price Index) and 307.47 (Unit Value Partial Index)
- **Undecided (N=30)**: 99.75 (Official Partial Price Index) and 98.47 (Unit Value Partial Index)
- **Good (N=24)**: 80.07 (Official Partial Price Index) and 83.06 (Unit Value Partial Index)

**Percentage Differences:**
- 41.82%
- -0.26%
- 0.75%
Top Level XPI, Dec 2017
(Jan 2012=100)

Average Annual Difference in Price Levels

<table>
<thead>
<tr>
<th></th>
<th>Official Aggregate Price Index</th>
<th>Unit Value Partial Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poor (N=73)</td>
<td>99.69</td>
<td>220.57</td>
</tr>
<tr>
<td>Undecided (N=30)</td>
<td>99.95  99.69</td>
<td>-.05%</td>
</tr>
<tr>
<td>Good (N=24)</td>
<td>95.38  96.08</td>
<td>.15%</td>
</tr>
</tbody>
</table>
Official and Unit Value Price Indexes
Homogeneous and Heterogeneous

HM-Meat&Poultry

HT-Computer Parts
Official and Unit Value Index Comparison – Synthetic Rubber-Primary

- IPP Index
- Unit Value Index

Index Value

Year
Tornquist Index Formula Bias?

- By using Tornquist
  - Flattens index trends in both + and - directions
  - Changing weight values increase STR variability
Potential Impact on 5-digit BEA End Use Price Indexes

Export 5-digit N = 127
Expand + Replace = 24-54

Import 5-digit N = 137
Expand + Replace = 52

- Foods, Feeds, & Beverages
- Industrial Supplies
- Capital Goods
- Automotive Vehicles
- Consumer Goods
Conclusion

- Homogeneous unit value indexes can be used in price indexes
  - Create items that approximate matched model
  - Intra-Item Substitutability bounds homogeneity
  - Homogeneity minimizes unit value bias
  - Need detailed and consistent item keys for homogeneity
  - The Coefficient of Variation Test performs well at identifying homogenous areas
  - Homogeneity is defined judgmentally
Conclusion

- Improvements
  - SOME Similar items are unique enough
  - Addressed all calculation problems except for timeliness of data availability
  - The Tornqvist index formula corrects for new goods/substitution/volatility of trade.
Conclusion

- Challenges
  - Variable monthly Q creates systemic flattening bias.
  - Greater variation of “Good” UV indexes than for “Undecided” UV indexes
  - Refining ‘Undecided’ UV index methods and definitions
Next Steps

Before finalizing an approach, we hope to:

- Use hedonic linear regressions to determine the ideal item key for 5-digit BEA indexes
- Explore the use of time-dummy hedonic models at the 10-digit based Harmonized level of classification
- Measure chain drift and investigate alternative aggregation methods
Next Steps

Critical path before deciding whether to operationalize:

- Research import unit value indexes
- Partial month data for preliminary measures
Contact Information

Don Fast
202-691-7147
Fast.Don@bls.gov

Susan E. Fleck
202-691-6043
Fleck.Susan@bls.gov

Use of the export trade data are subject to Agreement No. 2067-2018-001, Memorandum of Understanding (MOU) between the U.S. Census Bureau and the Bureau of Labor Statistics (BLS). The BLS has received prior approval from the U.S. Census Bureau, which affirms that the research results do not present disclosure risks and approve the publication of the results.