A Balanced System of Industry Accounts for the U.S. and Structural Estimation of Statistical Discrepancy

Baoline Chen
Researcher, NEA Directorate
BEA Advisory Committee Meeting
May 19, 2006

OUTLINE

1. Introduction
   • Motivation
   • Alternative method of reconciliation
   • Application of the proposed method

2. Identification of Major Data Problems

3. GLS Reconciliation Model

4. Reliabilities of Initial Estimates

5. Balanced Industry Accounts and Estimates of Statistical Discrepancy (SD) by Industry

6. Conclusion
1. Introduction

• **Motivation:**

   The desire to find a statistically sound method that can efficiently incorporate information on reliability of initial source data in compiling a consistent set of accounts.

• **Some facts:**

   - Inconsistency exists in the two estimates of GDP.
   - Inconsistency is crucially due to sources of errors in initial source data.
   - Currently there are no estimates of statistical discrepancy by industry or by expenditure category.
   - Lack of such information makes it difficult to identify improvements needed in source data and in estimation methods.

Introduction (cont.)

• **Objective of this study:**

   To propose a Generalized Least Square (GLS) reconciliation method that

   - Can correctly estimate industry distribution of SD according to reliabilities of initial data, and
   - Can accurately reconcile initial I-O and Income-by-industry data with expenditure-based estimate of GDP.

• **Empirical advantages of the proposed GLS method:**

   - Has firm Bayesian foundation.
   - Provides flexibility to balancing process.

• **GLS reconciliation methods have been studied extensively.**
Introduction (cont.)

- Application of GLS reconciliation method

Data to be reconciled:
- 1997 I-O and Income-by-industry data, 2003 benchmark GDP.
- **Aggregation**: 65 industries, 69 commodities, 13 expenditures.

Data:
- Initial output and input data from benchmark I-O accounts database;
- Initial VA from Income-by-industry prior to allocation of aggregate SD;
- Final expenditure estimates from 2003 NIPA benchmark revision;
- Coefficient of variation (CV) from Census and IRS.

---

2. Four Major Data Problems Identified

a. Differences in sources of data used.
   - **I-O accounts**: Economic Census, BES & annual surveys from Census, federal agencies, Federal Reserve, IRS, and trade companies.
   - **Income-by-industry**: IRS, BEA, BLS, Federal Reserve, Census Bureau, regulatory agencies, Federal, state & local governments, and trade companies.

b. **Sampling errors in source data.**

c. **Non-sampling errors:**
   - Double-counting, misallocation, misreporting, misspecification, omission, or simple mistakes.

d. **Official residual errors**: Aggregate SD.
3. GLS Accounts Reconciliation Model

- **The Reconciliation Model:**

  (1) Minimize the sum of reliability weighted squares of adjustments in initial gross output, intermediate inputs, & value added estimates, subject to

  (2)&(3) 65/69 Industry/commodity adding-up constraints,

  (4) VA of 65 industries sum up to GDP.

Initial conditions satisfy:

(5) VA of 65 industries sum up to GDI,

(6) Final expenditures sum up to GDP.

---

4. Reliability of Initial Data

- **Source data often needs adjustment due to non-sampling errors**

- **Decomposition of initial estimates into two components:**
  
  Initial estimate = source data + adjustments

- **Categories of adjustments in descending reliability:**
  
  Adjustment data from
  
  a. Census, IRS, other administrative or regulatory agencies;
  b. Established procedures or from fairly reliable sources;
  c. Incomplete data or analysts’ judgments.

- **Subjective relative standard errors assigned to adjustments:**
  
  \[ CV = (10\%, 20\%, 30\%) \] for categories a, b, c.

- **Measure of reliability of initial data by estimated variances:**
  
  \[ \text{var}(\text{Initial est.}) = \text{var}(\text{source data}) + \text{var}(\text{adjustments}). \]
5. Balanced Estimates and Distribution of SD

- **Two sets of balanced estimates using weights derived from:**
  a. Reliability (var(initial estimate))
  b. Neutral variant (abs(initial estimate))

- **Summary of Results:**
  - GLS model produced balanced estimates & removed aggregate SD.
  - **Using reliability:** Relative reliabilities of initial estimates significantly determine adjustments and industry distribution of SD.
  - **Using neutral variant:** Relative sizes of initial estimates and VA/GDP by industry significantly determine adjustments and industry distribution of SD.
  - Sizes of initial gap between estimates from I-O and Income-by-industry affect sizes of adjustments.

| Table 1: Initial and Balanced Estimates for 65 Industries (Billions of Dollars) |
|---|---|---|---|---|---|---|---|---|
| | Initial Estimates | Balanced Estimates: (Reliability) | Balanced Estimates(Neutral variant) |
| Pub Code | Gross output | Inter. Inputs | VA | (output-inputs)% | Gross Output | Inter. Inputs | VA | Output input | Gross output | Inter. input | VA | Output input |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 322 | 149 | 98 | 51 | 0.32 | 149 | 96 | 53 | 0 | 150 | 98 | 52 | 0 |
| 335 | 109 | 67 | 78 | -33.18 | 109 | 65 | 44 | 0 | 118 | 55 | 63 | 0 |
| 531 | 1260 | 319 | 883 | 4.62 | 1257 | 340 | 917 | 0 | 1248 | 328 | 920 | 0 |
| 532R | 176 | 32 | 73 | 40.31 | 165 | 32 | 133 | 0 | 142 | 43 | 99 | 0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| Sum | 15218 | 6917 | 8258 | 15201 | 6897 | 8304 | 0 | 15184 | 6880 | 8304 | 0 |
| 322: Paper; 335: Elec. Equipment; 531: Real Estate; 532RL: Rental & Leasing |
Table 2: Initial and Balanced Estimates of 69 Commodities
(Billions of Dollars)

<table>
<thead>
<tr>
<th>Com Code</th>
<th>Com. output</th>
<th>Com total use</th>
<th>(output - input) %</th>
<th>Com. output</th>
<th>Com inter. inputs</th>
<th>Final uses</th>
<th>output inputs</th>
<th>Com. output</th>
<th>Com inter. inputs</th>
<th>Final uses</th>
<th>output inputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
<td>... ... ... ... ... ... ... ... ...</td>
</tr>
<tr>
<td>3364OT</td>
<td>148 151 -1.53</td>
<td>149 64 85 0</td>
<td>149 64 85 0</td>
<td>42 736 737 -0.03</td>
<td>745 365 380 0</td>
<td>744 364 380 0</td>
<td>481 124 114 8.65</td>
<td>115 55 60 0</td>
<td>121 61 60 0</td>
<td>487OS 84 89 -6.65</td>
<td>84 67 16 0</td>
</tr>
<tr>
<td>Sum</td>
<td>15218 15221 15201 6897 8304 0</td>
<td>15184 6880 8304 0</td>
<td>42 736 737 -0.03</td>
<td>745 365 380 0</td>
<td>744 364 380 0</td>
<td>481 124 114 8.65</td>
<td>115 55 60 0</td>
<td>121 61 60 0</td>
<td>487OS 84 89 -6.65</td>
<td>84 67 16 0</td>
<td>88 72 16 0</td>
</tr>
</tbody>
</table>

3364OT: Other transportation equipment; 42: Retail; 481: Air transportation; 487OS: Other transportation

Figure 2: Histograms of % Adjustments

Table 3: Statistics of % Adjustments

<table>
<thead>
<tr>
<th>% Adjustments in Initial Gross Output</th>
<th>% Adjustments in Initial Intermediate Inputs</th>
<th>% Adjustments in Initial Value Added</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>-0.23</td>
<td>1.89</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>1.72</td>
</tr>
<tr>
<td>Stdv</td>
<td>1.43</td>
<td>4.08</td>
</tr>
</tbody>
</table>

www.bea.gov
### Table 4: Statistics of % Adjustments

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Stdv</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Adjustments in Initial Gross Output</td>
<td>-0.38</td>
<td>-0.02</td>
<td>3.83</td>
</tr>
<tr>
<td>% Adjustments in Initial Intermediate Inputs</td>
<td>-0.11</td>
<td>0.10</td>
<td>7.81</td>
</tr>
<tr>
<td>% Adjustments in Initial Value Added</td>
<td>0.76</td>
<td>0.71</td>
<td>8.88</td>
</tr>
</tbody>
</table>

### Table 5: Estimates of SD by Industry Using Relative Reliability and Neutral Variant Weights

<table>
<thead>
<tr>
<th>Initial Gap</th>
<th>Estimated SD by Industry (Reliability)</th>
<th>Estimated SD by Industry (Neutral variant)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pub Code</td>
<td>VA(IO)/VA(GDI)</td>
<td>SD'/SD</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------</td>
<td>--------</td>
</tr>
<tr>
<td>324</td>
<td>-43</td>
<td>-42</td>
</tr>
<tr>
<td>334</td>
<td>33</td>
<td>17</td>
</tr>
<tr>
<td>531</td>
<td>58</td>
<td>34</td>
</tr>
<tr>
<td>532RL</td>
<td>72</td>
<td>60</td>
</tr>
<tr>
<td>Sum</td>
<td>42</td>
<td>47</td>
</tr>
</tbody>
</table>

324: Petroleum; 334: Computer & Electronic; 531: Real Estate; 532RL: Rental & Leasing

www.bea.gov
Estimates and Statistics of SD by Industry (Relative Reliability)

Table 6: Analysis of SD Estimates

**OLS Regression of SD by Industry**
Summary statistics: \( R^2 = 0.96 \)

Std Err = 3367

<table>
<thead>
<tr>
<th>Coefficient Estimation</th>
<th>Variable</th>
<th>Coef.</th>
<th>t</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA(I)^0 - VA(GDI)^0</td>
<td>0.749</td>
<td>39</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>var(VA(GDI)/VA(IO))</td>
<td>0.317</td>
<td>3.29</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>VA(GDI)/VA(IO)</td>
<td>0.317</td>
<td>3.29</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>VA/I</td>
<td>-0.053</td>
<td>-0.245</td>
<td>0.771</td>
<td></td>
</tr>
</tbody>
</table>

Estimates and Statistics of SD by Industry (Neutral variant)

Table 7: Analysis of SD Estimates

**OLS Regression of SD Estimates:**
Summary statistics: \( R^2 = 0.97 \)

Std Err = 1805

<table>
<thead>
<tr>
<th>Coefficient Estimation</th>
<th>Variable</th>
<th>Coef.</th>
<th>t</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VA(I)^0 - VA(GDI)^0</td>
<td>0.458</td>
<td>34.8</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>VA(GDI)/VA(IO)</td>
<td>3747</td>
<td>3.90</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>VA/GDP</td>
<td>718</td>
<td>6.27</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>
6. Conclusion

- **Contributions of this study:**
  - Reconciled disaggregated industry accounts with benchmark GDP.
  - Shown that using relative reliability to remove inconsistency produces statistically meaningful results.
  - Identified problems in source data and estimation methods.
  - Demonstrated that reconciling disaggregated accounts is empirically feasible and computationally efficient.

- **Future research:**
  - Continue to improve reliability measures.
  - Allow for unallocated or unobserved initial estimates.
  - Extend to full reconciliation of national and industry accounts.

Questions

1. Given the current budget climate, how do you see the 2007 Economic Census tradeoff between further expansion of final demand information and coverage of detailed industries and inputs?

2. Is the GLS reconciliation method presented here leading to the right direction for BEA to achieve internal integration of national and industry accounts?

3. Quantifying uncertainty in the adjustment data is a real challenge. How can we better improve our assessment of uncertainty in the adjustment data?