Data Sharing
October 2003 FESAC

- Carol Comisarow (Census)
- Dennis Fixler (BEA)
- Lucia Foster (Census)
- James Spletzer (BLS)

Prepared for presentation to the Federal Economic Statistics Advisory Committee (FESAC) on October 17, 2003. It represents work in progress and does not represent any agency's final positions on issues addressed. The FESAC is a Federal Advisory Committee sponsored jointly by the Bureau of Labor Statistics of the U.S. Department of Labor, and by the Bureau of Economic Analysis and the Bureau of the Census of the U.S. Department of Commerce.

Outline of Presentation

1) Overview of CIPSEA

    > Status of Tax Code Changes

2) Business List Comparison

    > 1998-2000 project, 2003 project

3) Examples of Benefits from Data Sharing

    > BEA, BLS, BoC

4) Questions for the FESAC members

CIPSEA

- Confidentiality Information Protection and Statistical Efficiency Act (CIPSEA) of 2002

- Title V of Public Law 107-347

- Passed by both houses of Congress, and signed by the President on Dec 17, 2002

CIPSEA

- There have been different attempts over the years to pass uniform data confidentiality law and to establish data sharing authority

- In 2002, there was strong and concerted support by the Administration; from OMB, Commerce, Labor, Treasury, and especially the Council of Economic Advisers

CIPSEA

- Major CIPSEA Goals

  ➤ Establish widespread and uniform confidentiality protections

  ➤ Promote statistical efficiency by authorizing (limited) data sharing

CIPSEA

- Data sharing could help resolve several important economic measurement problems
  - Recently, the difference between the income and expenditure measures of GDP approached an historically high $170 billion
  - Nearly 30% of single establishment businesses have inconsistent 4-digit SIC codes in the separate business lists maintained by Census & BLS

CIPSEA

- Data sharing would improve the reliability and accuracy of key business statistics such as GDP, employment, productivity, and industrial production.
  - In the short run, decisions about some data sharing projects will have to be made in terms of impacts on priorities and resources

- Data sharing should also help reduce reporting burdens.

Confidentiality

- Purpose (in CIPSEA text):

  To ensure that information supplied by individuals or organizations to an agency for statistical purposes under a pledge of confidentiality is used exclusively for statistical purposes.

Confidentiality

- Household and business data
- All Federal agencies covered
- Data collected for exclusively statistical purposes & under a pledge of confidentiality
- Provides statutory authority
  - ➤ Up to now, the commitment of federal agencies to protect individually identifiable records rested on a mix of laws, policies, and practices, depending upon the agency.

Confidentiality

- Fines and penalties
  - Knowing and willful disclosure
  - Class E felony
  - Imprisonment for up to 5 years, fined up to $250,000, or both

  - Some, but not all, agencies already had some form of criminal or civil penalties

Data Sharing

Purposes (in CIPSEA text):

1) To authorize the sharing of business data among BoC, BEA, & BLS for exclusively statistical purposes
2) To reduce the paperwork burdens imposed on businesses
3) To improve the comparability and accuracy of Federal economic statistics

Data Sharing

- BEA, BLS, & BoC only
- Business data only
- Exclusively statistical purposes
- Voluntary written agreements
  - ➢What will be shared
  - ➢Why it's being shared
  - ➢Who will have access
  - ➢Physical security arrangements
  - ➢Timelines

Data Sharing

- Issue: Federal Tax Information (FTI)
  - A high proportion of the Census Bureau's business records either contain data taken from tax records or derived from tax records
  - Tax law and IRS regulations put restrictions on the use of tax data
  - If data sharing between the three designated statistical agencies is to reach its full fruition, there must be changes in the tax code and IRS regulations consistent with the aims and procedures of data sharing
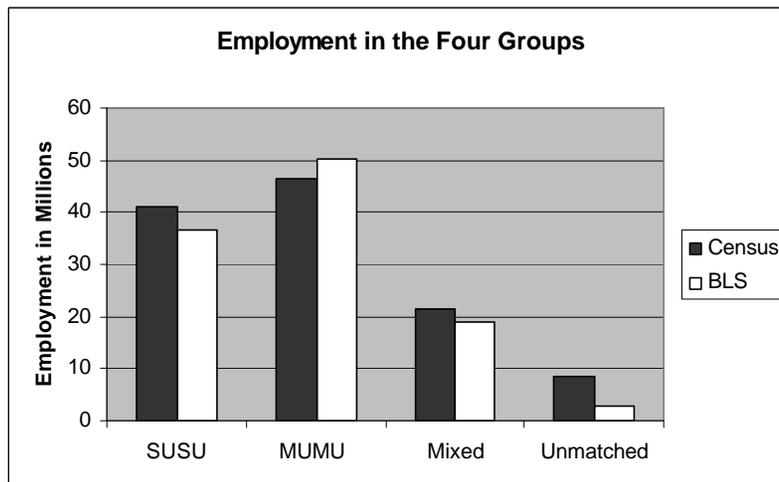
Status of Tax Code Changes

- Although the CIPSEA legislation was passed and signed into law, the accommodating tax code legislation has not been passed.

- Negotiations between the Administration and Congressional tax staff continue.

- Data sharing without tax code changes will be limited to business data uncontaminated by tax data.

Business List Comparison Project Benefits

- Data Producers
  - Suggest areas for improvement.
  - Inform potential future data sharing.
- Data Sources
  - Potential to reduce respondent burden (if data sharing).
- Data Users
  - Potential to improve measures constructed from data based on different lists, for instance detailed industry productivity measures.

Summary Statistics of 1994 Data

- About 10 million EINs & 100 million employees.
  - (EIN: BEL 6m, SSEL 11m. Employees: BEL 109m, SSEL:117m)
- Four Groups of Analysis: SUSU, MUMU, Mixed, and Unmatched.

Employment in the Four Groups

	Census	BLS
SUSU	41	36.5
MUMU	46.5	50.5
Mixed	21.5	19
Unmatched	8.5	2.5

Results of Match

- 5.8 million EINs matched.
  - Accounts for 97% of BEL, 93% of SSEL employment.
- Single Units:
  - 60% of EINs match *both* payroll and employment.
  - 68% of EINs match industry at 4-digit industry level.
  - 92% of EINs match geography at the county level.
- Multi Units:
  - 31% of EINs match: number of establishments in EIN.
  - 65% of EINs match: number of states in an EIN.
  - 53% of EINs match: number of industries in an EIN.
  - 29% of EINs match: employment for an EIN.

Results of Match – Con't

- Mixed:
  - MUSU: Over 14 million employees.
    - Census issue: are these MU who have not been asked to respond to the COS?
  - SUMU: Over 7 million employees.
    - BLS issue: are these MU who have not been asked to respond to the MWR and/or is this an issue with state-level data?
  - Much of these are inactive or out of scope.
- Unmatched:
  - Most of these are SU from the SSEL.
  - Most are inactive or out of scope cases.
  - But some cases where the lack of match is due to differences in timing of births and deaths.

Reviving the Project

- Additional Motivation: CIPSEA
- Building on knowledge gained from original project.
- Focus on 2001.
- Steps to take: MOU, security issues, FTI data issues, staffing, work plan, interagency training, preparing the data.

Examples of Benefits to BEA from Data Sharing

Benefits of sharing non-tax data under CIPSEA

- Projects that link the information from BEA's surveys of foreign-owned companies in the U.S. and U.S owned companies abroad with Census and BLS data can provide meaningful statistics for a variety of purposes; e.g. employment trends for multinational firms.

Examples of Benefits to BEA from Data Sharing

- Current NSF proposed project links BoC survey data on R&D with BEA surveys of international investment

Data sharing benefits with augmenting change in tax law

- Improvement of sampling frames under CIPSEA -- BEA has access to identifying information from Census files that enhances BEA's ability to construct sample frames for its international surveys on services transactions

Examples of Benefits to BEA from Data Sharing

- Identifying anomalies in aggregate statistics provided by Census

- Access to disclosure cells in Census data -- Datasets obtained from Census frequently have undisclosed cells for Census confidentiality reasons.

- Improved estimates arising from reconciliation of Census and BLS business lists.

Examples of Benefits to BLS from Data Sharing

- Business Establishment List

  1) Evaluate benefits of firm information to BLS from obtaining Census alpha codes

  2) Improving consistency of industry codes

     -- Issue: role of States in adjudication process

     -- Issue: cost and workload

  3) Evaluate benefits to BLS from obtaining information on non-employer universe

Examples of Benefits to BLS from Data Sharing

- Producer Price Index (PPI)
  1) Improved sampling efficiency
  2) Improved publishability of planned indexes
  3) Improved ability to publish price indexes that match Census product classes
  4) Reduced reporter burden

Examples of Benefits to BLS from Data Sharing

- Productivity Measures
  1) Improved detailed industry productivity measures
  2) Improved explanations of aggregate and industry productivity changes
  3) Research studies with the goal of developing new or improved measures that could be published on an ongoing basis

Examples of Benefits to BoC from Data Sharing

- Improve BoC's Business Register:

    – obtain BLS employment data for single units

    – obtain industry codes for all records

    – obtain physical location addresses for all records

Examples of Benefits to BoC from Data Sharing

- Combine BoC's Company Organization Survey and BLS' Multiple Worksite Report

Examples of Benefits to BoC from Data Sharing

- With a matched BLS Business Establishment List and BoC Business Register, BoC could develop an effective edit to pinpoint records with large discrepancies.

Questions for the FESAC

- Section 523 of CIPSEA says the BoC, BEA, and BLS shall:

  1) Identify opportunities to eliminate duplication and otherwise reduce reporting burden and cost imposed on the public in providing information for statistical purposes

  2) Enter into joint statistical projects to improve the quality and reduce the cost of statistical programs

Any ideas?

Questions for the FESAC

- ## Any ideas? Specifically:

  -- what improvements in programs, operations, or measures should be high priority

  -- where you believe duplication may exist

  -- research topics that can be addressed with data sharing