“The Future of Data Collection, Access, and Dissemination: Uses of Administrative Data and Data Matching”

Dennis Fixler, Chief Statistician

Association of Public Data Users 2010 Annual Conference
Data Collection Issues

- Resource constraints
- Costs of expanding surveys
- Respondent burden/response rates

To overcome
- Electronic reporting
- Use of administrative data
- Sharing data
Looking to greater use of administrative data, including more intensive and more effective use of:

- tax data
- micro health insurance data
- business financial and accounting records
- financial and other regulatory data
- state, local and federal administrative data
- international banking and other financial data
Benefits of Administrative Data

- Lower Collection and processing cost
- Near universe level coverage with few sampling or coverage problems
- No additional burden on businesses
- Administrative, or tax scrutiny usually results in high rates of compliance and accuracy (for data used for administrative purposes)
- Identification information, especially EIN, offers lots of possibilities for linking with other administrative data
Challenges of Administrative Data

- Collected for non-statistical purposes and virtually no control over concepts, definitions, and methods.
- Inconsistencies across “similar” administrative data.
- Concerns about timeliness of updating, consistency over time.
- Limited or no editing for items unrelated to administrative purpose.
- Tax avoidance and other “biases” in reporting.
- Legal limits on access impede ability of BEA, Census, and BLS to share data derived from administrative data.
Different Measures of Wages and Salaries

Growth Rate of Wages and Salaries

- Census, County Business Patterns (CBP)
- BLS Qtrly Census of Employment & Wages (QCEW)
- IRS, SOI Wages & Salaries
Benefits of Exchanging and Matching Existing Data

- Identify gaps and duplication
- Improve accuracy:
  - By using stronger source of data: Bank reported deposits vs. depositors reports; W-2s instead of household reported salary income
  - By identifying and correcting misreporting (e.g. units, reporting level, concept, etc)
- Improve consistency in allocations by industry and region
- Additional data with no response burden and low cost
Challenges of Exchanging and Matching Existing Data

- Legal barriers in decentralized U.S. system restrict ability to harmonize data
- Significant differences across source data that need to be resolved
  - Resolution is at firm-level and can be labor intensive
  - Identifiers for matching are imperfect
  - Differences may need to be resolved by statistical agency adjustments
- IT requirements may be significant
Example 1: Foreign-Direct Investment
Detail and Geography

- Foreign-Direct Investment Match: International Investment, Finance, and Trade in Services Act
- Match dramatically expanded data, with no additional respondent burden and relatively low processing cost (BEA-Census-BLS)
  - Three agency in the 1990s; BEA and Census about every 5 years since 1987
  - BEA Enterprise to Census Establishment: from all U.S. total for 200 industries to data for all 50 states for over 1000 industries
Example 2: Medical Care Output and Prices

- Studies of heart disease, depression, and cataracts
  - Making the broader case: Insurance data, provider-based survey data, or household survey data
- Initial results suggest more accurate measure of measured medical care prices, real output (GDP), and productivity.
- Data by disease, by type of medical care provider, by geography
Example 3: Globalization of Financial Markets

- Rise of direct transactions, not through U.S. banks, large gaps and growing gaps.
- Data exchange with foreign banks (macro data)
  - Their claims are our liabilities and their liabilities are their claims
- Very large improvement in accuracy
  - Depending on type of financial asset, between 33 and 70% of data are from financial partners
- Opportunities for access to additional financial data should increase as the Financial Reform bill gives BEA access to data collected by Office of Financial Research