

Private and Public Data Provision in the U.S.

John M. Abowd
January 2007

Outline

- Information is public good (review)
- Public information suppliers
- Private information suppliers
 - Research firms
 - Information services
- Demand conditions
- Structure of supply
- Example: Standard and Poor's Compustat
- Overview of the US Statistical System
- Organization of the U.S. Census Bureau
- Programs of the U.S. Census Bureau
- Programs of the Bureau of Labor Statistics
- Other Agency Programs

Information is a Public Good

- Two conditions for a public good:
 - No rivalry in consumption
 - No excludability in consumption
- CPI example
 - One person's use of the CPI does not reduce the amount available for another person to use.
 - If the CPI is published (e.g., as a public-use data product) one person cannot stop another from using it.

Pure Private Goods

- Two implicit assumptions we've made about private goods:
 - There is rivalry in consumption.
 - There is feasible excludability.
- Interesting and difficult problems arise if either of these properties is missing.
- A pure public good like information misses both.

Public Information Suppliers

- Statistical agencies
- Researchers who publish in open journals
- Many motivations discussed last week

Private Information Suppliers

- Classical: printed media publishers
- 21st Century: digital media suppliers
- Private suppliers come in two types:
 - Research organizations
 - For-profit information services

How Big is the Market

- Information Services (NAICS 51) \$891 billion in 2002
- Components of many other NAICS sectors

Demand Conditions in the Private Information Market

- What are the customers buying?
- Copyrighted information
 - Unlicensed, via publications
 - Licensed, via information servers (formerly called libraries or data centers)
- Research tools
- Timeliness

Supply Conditions in the Private Information Market

- Research organizations
 - Maintain sampling frames
 - Generate surveys and research
 - Examples: NORC, Westat, ISR/SRC, RTI
- For-profit information services
 - Maintain sampling frames or transactions data bases
 - Generate surveys, indicators, and research
 - Examples: Standard & Poor's Compustat, The Economist Intelligence Unit

Example: COMPUSTAT

- Subscription service of Standard & Poors, Inc.
- General URL:
<http://www.compustat.com/www/>
- CU access: <http://mi.compustat.com/cgi-mi-auth/mihome.cgi?tab=home>
- Micro-data: contact appropriate data librarian (available for use in Census RDCs)

What it Contains

- Universe: Publicly-held companies that are traded on U.S. securities markets
- Frame: corporations, other limited-liability entities required by law to file Securities and Exchange Commission reports or equivalent Canadian reports
- Annual data: 10K and annual reports, other annual filings
- Quarterly data: 10Q and other quarterly filings

How It Is Used

- License the most recent file
- File incorporates entities (mostly U.S. and Canadian corporations) restated to the most recent accounting year
- Restated?
 - The service makes the historical data comparable to the current data by adjusting for “complex business recombinations,” a.k.a. mergers and acquisitions

Value-Added

- The basic data are public, on file at the Securities and Exchange Commission in the EDGAR system
 - URL: <http://www.sec.gov/edgar.shtml>
- Statistical analysis requires:
 - Comparable entities
 - Data editing
 - Geo-coding
 - Activity coding
- COMPUSTAT provides the value-added required for statistical analysis

Advantages

- Entities are corporations, so they conform to the legal entity that issued the security
- Accounting concepts are reported according to U.S. GAAP as defined by the SEC and the Financial Accounting Standards Board (<http://www.fasb.org/>)
- Can be used directly with returns data from traded securities

Disadvantages

- Corporate entities do not have well-defined geography or activity codes
 - Location of the economic activity does not conform to location of the corporate activity
 - Difficult to code primary economic activities
- Complex business recombinations and interlocking ownership structures complicate statistical analyses

Overview of the US Statistical System

- Review from lecture 1b
- Office of Management and Budget
 - Office of the Chief Statistician
- Agencies of the federal government
 - Field operations (Census and BLS have the biggest field operations)
 - Statistical programs (varied and spread all over the government)

U.S. Statistical Standards

- Provided by the Office of Management and Budget
- Overview:
<http://www.whitehouse.gov/omb/inforeg/statpolicy.html>
- Methods and Standards:
<http://www.fedstats.gov/policy/>
- Industry, Geography, Occupation, and Minority/Ethnicity Standards

Organization of the U.S. Census Bureau

- See the organization chart and executive staff
 - <http://www.census.gov/main/www/m-img/orgchart.jpg>
 - <http://www.census.gov/acsd/www/bc245.pdf>
- Major Programmatic Directorates
 - Decennial
 - Economic
 - Demographic

Key Divisions for RDC Users

- Office of the Chief Economist (Center for Economic Studies; research use of most economic data)
- Demographic Surveys (all demographic surveys)
- Data Integration Division (LEHD, Administrative Records)
- Economic Planning and Coordination (Business Register; Economic Censuses)
- American Community Survey Office (ACS)
- Housing and Household Economic Statistics (AHS, SIPP, DEWS, and CPS analysis)
- Policy Office (Memoranda of understanding)
- Population Estimates Division (intercensal population estimates)

Programs of the Census Bureau

- The Decennial Census of Population and Housing
- Other Demographic Surveys
- The Quinquennial Economic Censuses
- Other Economic Surveys

The Decennial Census of Population and Housing

- Brief history
 - Established under the Constitution (Article I)
 - <http://www.law.cornell.edu/constitution/constitution.text.html>
 - First conducted in 1790
 - Required to be an “enumeration” of the population
 - Last conducted in 2000 (at an approximate cost of \$6 billion)

Decennial Census

- Conducted every 10 years
- Historically separated into
 - Short form (enumeration)
 - Long form (1/6 household survey)
- Beginning in 2010
 - Enumeration only
 - Long form replaced by American Community Survey
- <http://www.census.gov/main/www/cen2000.html>

Quinquennial Economic Censuses

- Conducted in the 2 and 7 years by the Bureau of the Census
- Since 1997, excludes agriculture
- Separate Census of Government conducted on the same schedule
- <http://www.census.gov/econ/census02/index.html>

Census Bureau Surveys

- <http://www.census.gov/main/www/surveys.html>
- Main demographic surveys
 - American Community Survey
 - American Housing Survey
 - Current Population Survey (joint with BLS)
 - Survey of Income and Program Participation, now Dynamic Economic Well-being Survey (DEWS, beginning 2008)

Census Bureau Surveys

- <http://www.census.gov/main/www/surveys.html>
- Main economic surveys
 - [Sectors covered in overview](#)
 - Survey of Business Owners
 - Annual Capital Expenditure Survey
 - Business Expenditure Survey
 - Annual Survey of Manufactures
 - Annual Retail and Wholesale Trade Surveys
 - Service Annual Survey

Quinquennial Census of Agriculture

- Conducted by the Department of Agriculture (since 1997) on the 2 and 7 years
- <http://www.nass.usda.gov/census/>

Bureau of Labor Statistics Programs

- Handbook of methods
- <http://www.bls.gov/bls/descriptions.htm>
- <http://www.bls.gov/opub/hom/homtoc.htm>

Quarterly Census of Employment and Wages

- Conducted quarterly by the Bureau of Labor Statistics in collaboration with state departments of employment security
- Provides the basic sampling frame for all BLS business surveys
- The Census Business Register and the BLS QCEW are maintained as two independent establishment frames
- Formerly known as the ES-202
- <http://www.bls.gov/cew/home.htm>

BLS Statistical Programs

- Household statistics from the CPS and specialized longitudinal surveys
- Industry statistics from the Current Employment Statistics program
- [Catalogue](#)

Health Statistics

- Collected by a variety of agencies
- Main responsibility: National Center for Health Statistics
- Vital Statistics
 - Birth, death, marriage and divorce
 - <http://www.cdc.gov/nchs/nvss.htm>
- Survey programs
 - <http://www.cdc.gov/nchs/nhcs.htm>

Education Statistics

- Provided by the National Center for Education Statistics
- Survey programs use educational institutions as the sampling frame
- <http://nces.ed.gov/surveys/>

Crime Statistics

- Provided by the Bureau of Justice Statistics
- Variety of statistical programs based on the Uniform Crime Reports
 - <http://www.ojp.usdoj.gov/bjs/cvict.htm>
- Conviction and sentencing data based on the National Judicial Reporting Program
 - <http://www.ojp.usdoj.gov/bjs/stssent.htm>

Social Security Statistics

- Published by the Office of Policy in SSA
- <http://www.ssa.gov/policy/docs/statcomps/supplement/2006/index.html>
- Based on administrative data collected by SSA

Tax Statistics

- Published by the Statistics of Income Division of the IRS.
- <http://www.irs.gov/taxstats/>