

Data Everywhere

John M. Abowd
January 2005

Outline

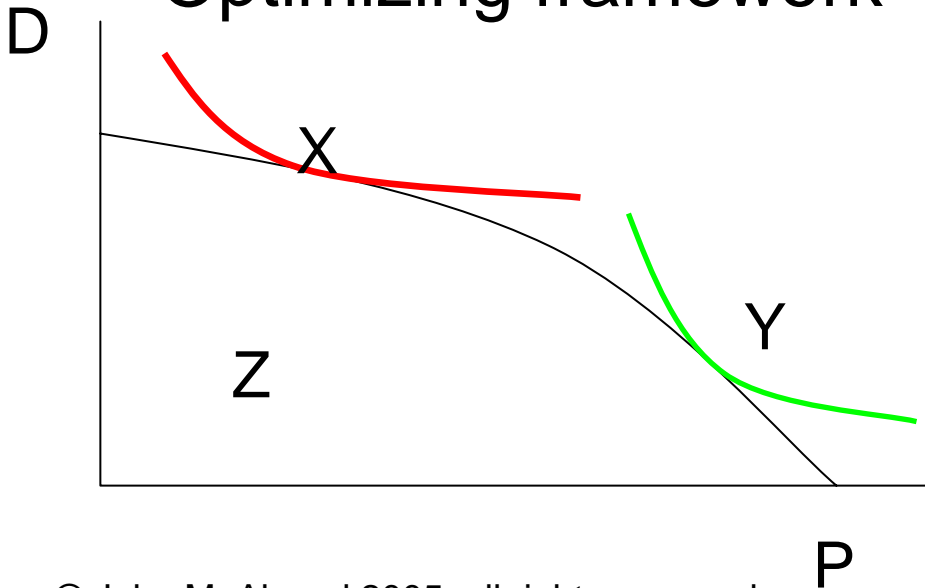
- Information economics basics
- Conceptual framework
- Protection technologies
- Why collect social data?
- The US statistical system

Information Economics Basics

- Information is a public good
- One's person's use does not diminish the availability to another person
- Marginal cost of dissemination is (almost) zero once collected.

Conceptual Framework

- The mission of statistical institutes
 - Disseminate
 - Protect
- Consequences of “mistakes”
- Optimizing framework



Conceptual Framework:

Are Data Under-Disseminated?

- The bargain is mandatory participation in exchange for confidentiality protection and statistical use only
- Data collectors properly use conservative estimates of confidentiality disclosure risk

Under-investment in Protection Technologies

- But data collectors under-invest in protection technologies
 - because they bear the full costs of those investments
 - most of the social benefit accrues from the additional non-agency analyses the new technologies enable
- Identifying “optimal” amount of dissemination depends on identifying benefits and costs associated with a) dissemination and b) dissemination modality

Core Issue

- Statisticians can tell us which techniques dominate in the “risk/utility” tradeoff
- If several approaches pass statistical tests, which is appropriate to use?
 - Public Use Data
 - Licensing
 - Remote Access
 - Research Data Centers

Objectives of Collecting Social Data

- Measurement of the population and household activity
- Measurement of the economy and business activity
- Detailed measurement of selected activities
- Longitudinal measurement of individuals, households, and businesses
- Scientific hypothesis building and testing

US Statistical Infrastructure

- Main players:
 - OMB
 - Census Bureau
 - BEA
 - BLS
 - ETA
 - NIH
 - CDC
 - NCES
 - NASS

Federal Statistical Outlays

- www.fedstats.gov/policy/
- See Tables 1 and 2.

International Infrastructure

- United Nations
- Organization for Economic Cooperation and Development
- Eurostat
- National Statistical Agencies

Components of Statistical Systems

- Registers/Master Frames
 - Households
 - Businesses
 - Addresses
 - Activities
- Interconnections
 - Identifiers
 - Probabilistic linkages

Products of Statistical Systems

- National income and product accounts (SNA)
- Labor force surveys
- Population censuses
- Economic censuses
- Consistent time series
- Detailed tabular data
- Public-use products

A Tour of Some Useful Sites

- www.ces.census.gov (CES)
- www.fedstat.gov (Federal agencies)
- www.census.gov (Census)
- www.bls.gov (BLS)
- www.bea.doc.gov/ (BEA)
- www.cdc.gov/nchs/ (NCHS)