INFO 7470/ILRLE 7400
Health Data

John M. Abowd and Lars Vilhuber
March 1, 2011
Defining “health” data

- Data on individuals' health
  - By age (elder, infants)
  - Disability-related
  - Nutrition
  - Etc.

- Data on health providers
  - Insurance access and utilization
  - Access or utilization of doctors, hospitals, etc.
Scope here

- Health data related to disability
- Health data in the RDC network
- Other restricted-access health data
Long list of disability-related data

- **2000 Decennial Census Long Form**: six disability-related questions, and many other questions relevant to participation (e.g., living arrangements, employment, income, transportation utilization, and housing).

- **American Community Survey (ACS)** since 2003, disability questions reworded in 2008 to new OMB-coordinated standard (also CPS, soon SIPP)
Disability-related data (2)

• CPS

Prior to 2008, only work limitation measure, since 2008 same question as ACS.

- Contains extensive information on employment, income, insurance, living arrangements, family status

- Insurance coverage, marital status, and parental status for youth and young adults with work limitations
Health and Retirement Survey (HRS)

- a nationally-representative sample of more than 22,000 Americans *over the age of 50* every two years.

- Subjects:
  - antecedents and consequences of retirement;
  - health, health care, income, and wealth over time;
  - work disability, declining health, and institutionalization
  - chronic illness, functional ability, depression, and self-assessed health status,
HRS (2)

- Restricted-access linkages
  - SSA earnings
  - Detailed geography and industry
  - Medicare claims and summary
- Access and merge to RDC-data is feasible, but needs authorization from HRS
Behavioral Risk Factors Surveillance System (BRFSS)

- annual state-based health survey
  - health risk behaviors,
  - clinical preventative practices,
  - health care use and access focused on chronic disease and injury.
- Collaborative effort between CDC and the state health departments
- cross-sectional telephone survey that collects data on approximately 350,000 non-institutionalized adults (18 and older).
National Health Interview Survey

- annual nationally representative cross-sectional survey of approximately 100,000 non-institutionalized civilians
- conducted by CDC
- Topics/data
  - highly detailed information on health, functional status, and activity limitations.
  - education, employment, marriage, parenting, and family income
Random selection of disability-related research using NHIS:

- Burkhauser et al. (2002)
- Kaye (2002)
- Trupin et al. (1997)
- Horvath-Rose and Stapleton (2004)
NHIS linkages - SSA

- Numident (limited demographics)
- Master Beneficiary Record (MBR),
- Supplemental Security Record (SSR),
- Payment History Update System (PHUS)
- reduced sample size ~ 64,000
- Example analysis: comparisons between self-reports of disability and diagnosis codes contained in the SSA data.
NHIS linkages (CMS)

- CMS Medicare Enrollment and Claims Files
  - Linked to
    - 1994-1998 National Health Interview Survey (NHIS)
    - 1999-2005 National Health Interview Survey (NHIS)
Medical Expenditure Panel Survey (MEPS)

- Set of large-scale surveys of families and individuals, their medical providers (e.g., doctors, hospitals, and pharmacies), and employers
- Household Component (HC) of the MEPS collects data from a sample of families and individuals in selected communities across the United States, drawn from a nationally-representative sub-sample of households that participated in the prior year's NHIS.
- approximately 22,000 individuals
Medical Expenditure Panel Survey (MEPS)

- Topics:
  - specific health services, frequency of use,
  - cost of these services, how they are paid for,
  - cost, scope, and breadth of health insurance held by and available to U.S. workers.
  - demographic characteristics, health conditions, health status, income, and employment.
MEPS-MPC

- MEPS-HC is supplemented by information provided by the Medical Provider Component (MPC)
  - covers hospitals,
  - physicians,
  - home health care providers,
  - pharmacies

identified by household respondents.
NHANES

• National Health and Nutrition Examination Survey, Since 1960s

• Topics/Data:
  − demographic, socioeconomic, dietary, and health-related questions.
  − Examination component consists of medical, dental, and physiological measurements
  − laboratory tests
NHANES linkages

- EPA data (environment on health)
- Using restricted-access data:
  - lower level geography
  - indirect identifiers;
Medicare Current Beneficiary Survey (MCBS)

- non-public-use survey of about 12,000 Medicare beneficiaries at any point in time
- conducted since 1991 by Centers for Medicare and Medicaid Services (CMS)
- overlapping-cohorts structure, 3 contacts/year, any single individual at most four years.
- nationally-representative of all Medicare beneficiaries, including those who are aged, disabled, and institutionalized.
Medicare Current Beneficiary Survey (MCBS)

- Data/topics:
  - health status, health care use and expenditures, and health insurance coverage of Medicare beneficiaries
  - employment and demographic characteristics

- The MCBS is split into two components: Access to Care, and Cost and Use files
Health provider data (example)
State-specific data

- Example: Gruber & Kleiner (2010), NBER WP 15855, “DO STRIKES KILL? EVIDENCE FROM NEW YORK STATE”
  
  http://www.nber.org/papers/w15855

- short-term non-federal hospitals in New York State are required to submit discharge data to the New York State Department of Health through the Statewide Planning and Research Cooperative System (SPARCS).
SPARCS

- patient level, detailed data on
  - patient characteristics (e.g. age, sex, race),
  - diagnoses (several DRG and ICD-9 codes),
  - treatments (several ICD9 codes),
  - services (accommodation), and
  - total charges

- for every hospital discharge in New York State since in 1982.
Gruber & Kleiner (2010)

- Combine with detailed strikes data
- Controlling for hospital specific heterogeneity, patient demographics and disease severity, the results show that nurses’ strikes increase in-hospital mortality by 19.4% and 30-day readmission by 6.5% for patients admitted during a strike, with little change in patient demographics, disease severity or treatment intensity