

C

Appendix C:

Databases Used

To Develop

Health Reform

Estimates

Many estimates reviewed in this report were developed, in part, using microsimulation models. In microsimulation models, the unit of analysis is an individual person and household. To develop expenditure estimates under health reform, the models usually depict the changes in individual health insurance coverage and health care expenditures between a base case, usually current law, and an alternative scenario with specific policy changes.

The models simulate health expenditures under alternative scenarios by adjusting the data for each individual and family. To estimate aggregate health care expenditures under alternative scenarios, expenditures of individuals and families are simulated and then summed over all units. Data on many different variables are required. Not all of these data are collected in any single database. Consequently, several different databases have served as sources of data for health care reform estimates. Table C-1 provides an overview of the major databases used by various analysts to develop health care reform estimates. The most important types of data and databases are the following:

Demographic and Economic Characteristics of the Population

- Current Population Survey (CPS) March Income Supplement

- National Medical Expenditure Survey (NMES) Household Survey

Health Care Expenditures of Individuals

• NMES Household Survey

- Consumer Expenditure Survey (CES)

Health Care Utilization

- NMES Household Survey

- National Health Interview Survey (NHS)

Aggregate Health Care Expenditures over Time and in the Future

~ National Health Accounts (NHA)

Individual Tax Characteristics and Payments

- Statistics of Income (SOI) Individual Income Tax Returns

Health Insurance Characteristics

- NMES Health Insurance Plans Survey (HIPS)
- Health Insurance Association of America (HIAA) Health Insurance Survey

Firm Characteristics

- County Business Patterns (CBP)
- HIAA Health Insurance Survey

Hospital Utilization and Expenditures

- American Hospital Association (AHA) Annual Survey of Hospitals

TABLE C-1: Databases Used to Develop Health Reform Estimates

Database	Description	Use in models	Models used
Current Population Survey (CPS),	Sponsored by BLS and BC Monthly cross-sectional survey of 60,000 U S households	Serves as host database for HCFA and CBO. In several models, including HCFA, AHCPR, and Lewin-ICF, it is matched with 1987 NMES and provides information on Insurance status, family income, gender, age, poverty classification, family composition, race and ethnicity, sex, and employment status	HCFA's SPAM Model CBO
March Income Supplement	Data collected include labor force status for ages 15 and older March CPS Includes supplementary questions on income and employment status, and health insurance coverage during previous calendar year	Also used to age the NMES data in some models	Lewin-VHI's HBSM model AHCPR AHSIM model
National Medical Expenditure Survey (1987)	Sponsored by AHCPR and HCFA Five rounds of data collection between February 1987 and July 1988 for sample of 14,000 households (Household Survey), plus surveys of physicians and health care facilities providing care to the survey sample households during 1987 (Medical Provider Survey) and of employers and insurance companies responsible for their insurance coverage (Health Insurance Plan Survey) The database also includes a survey of 13,000 residents of institutions such as nursing and personal care homes, psychiatric hospitals, and facilities for mentally retarded persons (institutional Population component). Data collected include utilization, expenditures, and sources of payment for all major forms of medical care, demographic and socioeconomic characteristics of respondents, insurance coverage of respondents, information from medical providers about respondents, and access to medical care.	In the HCFA and AHCPR models, NMES is the source for such key elements as health expenditures and utilization. HIPS, a derivative of NMES, is used in the AHCPR model for such data elements as employer's establishment size, industry, location, and premium sources.	Lewin-VHI's HBSM model AHCPR AHSIM model CBO
Consumer Expenditure Surveys (CES)	Sponsored by BLS. CES are specialized surveys in which the primary purpose is to collect data relating to family expenditures for goods and services used in day-to-day living. Data are also collected on the amount and sources of family income, changes in savings and debts, and demographic and economic characteristics of family members The current survey actually consists of two separate surveys the Interview Survey, where each consumer unit in the sample is interviewed every three months over five calendar years, and the Diary Survey, which is completed at home by participating families over two consecutive one-week periods	CES are a source of excise tax information and source of information on individuals' health premiums and out-of-pocket health care spending under the CBO model	CBO

(continued)

TABLE C-1: Databases Used to Develop Health Reform Estimates (Con'td.)

Database	Description	Use in models	Models used
County Business Patterns (CBP)	Sponsored by BC. All business establishments with one or more paid employees (except agriculture production, railroad, most government, and household employment) in all U.S. counties are represented. Data collected on employment, payroll, and number of establishments by employment-size classes; data tabulated by detailed industry based on 1987 edition of the Standard Industrial Classification (SIC) manual. At the three-digit SIC code level, 450,000 aggregate industry records from over 3,000 counties in the U.S. exist.	Within the CBO model, CBP is used as a source of wage information. Within AHCPR, CBP is used to impute average wages.	AHCPR AHSIM model CBO
National Health Accounts (NHA)	Aggregate time series database developed by HCFA. NHA are statistics representing total national health expenditures for all goods and services relating to health care, and the sources of payments for these goods and services.	NHA are used within the CBO, AHCPR, HCFA, and Lewin-VHI models for benchmarking health care spending by category	HCFA's SPAM Model AHCPR AHSIM model Lewin-VHI's HBSM model CBO
Statistics of Income (SOI) Individual income Tax Returns	Developed by the IRS. Samples of tax returns and supporting schedules are abstracted each year from approximately 100 million tax returns; sample sizes are about 80,000 returns in even years and 120,000 returns in odd years. Data collected on taxpayers' income, exemptions, deductions, credits, and taxes owed.	SOI serve as a source for income tax information within the CBO and AHCPR models.	AHCPR AHSIM model CBO
Health insurance Association of America (HIAA) survey	Sponsored by Health Insurance Association of America The HIAA Survey was discontinued in 1993. The HIAA survey database includes data on employee health benefits coverage and health costs. The HIAA database originally presented data in terms of employee-level health benefits. In recent years the database has been weighted by national population counts of both firms and employees, by size of firm, public versus private firm categorization, and region.	Used in the AHCPR model to project market shares for FFS, HMO, and PPO plans by region. Used by Lewin-VHI for premium data. Lewin-VHI also uses HIAA for information on coverage of workers, average payroll, share of individuals with family coverage, eligibility of workers, and income level. The HIAA Survey is a source for information on employment-based insurance premiums within the CBO model.	Lewin-VHI's HBSM model AHCPR AHSIM model CBO

(continued)

TABLE C-1: Databases Used to Develop Health Reform Estimates (Con'td.)

Database	Description	Use in models	Models used
American Hospital association (AHA) annual survey of hospitals	Sponsored by the American Hospital Association's Hospital Data Center Mailed to all hospitals, both AHA registered and nonregistered in the United States and its associated areas, U S government hospitals outside United States not Included Data collected Include a profile of services, utilization, personnel, and finances of hospitals in U.S. and associated areas	In the AHCPR model, AHA data are used to AHCPR AHSIM model segregate overall rates of growth in hospital spending into components attributable to Inpatient and outpatient services CBO uses AHA Surveys in Medicare cost estimates	CBO

KEY AHA - American Hospital Association, AHCPR = U S Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research, AHISM = Agency Health Simulation Model BLS - U S Department of Labor, Bureau of Labor Statistics; CBO = U S Congress, Congressional Budget Office, CBP = County Business Patterns, CES = Consumer Expenditure Survey, CPS = Current Population Survey, FFS - fee-for-service, HBSM - Health Benefits Simulation Model, HCFA = U S Department of Health and Human Services, Health Care Financing Administration; HIAA - Health Insurance Association of America, HIPS = Health Insurance Plan Survey, HMO = health maintenance organization; IRS = Internal Revenue Survey, NHA = National Health Accounts, NMES - National Medical Expenditure Survey, PPO = preferred provider organization; SOI = Statistics of Income, SPAM = Special Policy Analysts Model

SOURCE Lewin-VHI, Inc., *The Financial Impact of the Health Security Act* (Fairfax, VA Dec 9, 1993), Rivlin, A M , Cutler, D M , and Nichols, L M "Financing, Estimation, and Economic Effects," *Health Affairs* 13(1) 30-49, 1994; U S Congress, Congressional Budget Office, *An Analysis of the Administration's Health Proposal* (Washington, DC U S Government Printing Office, February 1994), U S Congress, Office of Technology Assessment, *Understanding Estimates of the National Health Expenditures Under Health Reform*, OTA-H-594 (Washington, DC U S Government Printing Office, May 1994)

54 | Understanding Estimates of the Impact of Health Reform on the Federal Budget

Each model has a host database that serves as the central source of data. The CPS serves as the host database for the Health Care Financing Administration's (HCFA) and the Congressional Budget Office's (CBO) model. NMES serves as a host for the Agency for Health Care Policy and Research's (AHCPR) and Lewin-VHI's analyses.

No one database regularly provides the majority of the variables needed to evaluate alternative health care policies. In addition to the host database, several other databases are used piecemeal to fill in various gaps left by the host database.

Analysts use various types of procedures to link different databases. This in turn may lead to differences in the estimates of different analysts. Statistical matching is one such technique that is used to add or impute information to one database by using information from a different database. Records in the two files are classified by elements they share in common (such as age, sex, or income). Then other data elements from the records in the "donor file" can be attached to the host file records drawn from the same matching group.

For example, CPS has information on demographic and economic characteristics of persons and families (including health insurance coverage and employment characteristics), but it does not have information on health care expenditures. In the Health Care Financing Administration's (HCFA) SPAM Model, each CPS record is

matched to the record of a similar person or family in the NMES Household Survey database (in terms of age, sex, income, and insurance coverage), and the health expenditures data for that person or family are added to the CPS record.

All survey databases correspond to a past year when the data were collected. (The NMES data were collected in 1987. The most recent March CPS data used by many analysts were collected in 1993 and provide information for 1992.) To provide estimates pertinent to the population in current or future years, the existing survey data must be adjusted (or "aged") to correspond to the changes in population characteristics and other relevant information (e.g., expenditures, income). This is usually done in two stages. First, each record is reweighted to correspond to projections of the population by specific characteristics relevant to the analysis, such as age, sex, race, and insurance coverage. Most analysts adjust population weights in their models to match population projections by the U.S. Bureau of the Census or by the Office of the Actuary of the Social Security Administration. Second, individual income and health expenditure data on each record are adjusted so that, in the aggregate, the totals either reflect the income growth over time or match certain aggregate projections, such as the National Health Accounts produced by the Office of the Actuary of the Health Care Financing Administration.