



Abt Associates Inc.

Cambridge, MA
Lexington, MA
Hadley, MA
Bethesda, MD
Washington, DC
Chicago, IL
Cairo, Egypt
Johannesburg, South Africa

Abt Associates Inc.
55 Wheeler Street
Cambridge, MA 02138

**The AHCA
Long-Term Care
Microsimulation Model:
A Technical Brief**

**Health Services
Research and
Evaluation (HSRE)**

**Working Paper
No. 4**

March 29, 2001

Prepared by
Austin B. Frakt, Ph.D.
Steven D. Pizer, Ph.D.

1. Model Overview

This document is intended to provide a technical overview of the American Health Care Association's (AHCA's) long-term care microsimulation model. The purpose of the microsimulation model is to provide a consistent framework for computing and analyzing the implications of various approaches to financing long-term care for elderly and non-elderly persons with disabilities through the middle of the 21st century. The model has been designed with two main goals in mind: (1) It is extremely flexible in that it is capable of simulating a wide variety of policy proposals including, but not limited to, the status quo, the expansion of publicly-financed Medicare benefits to include long-term care, tax credits and deductions toward long-term care insurance premiums, changes to the Medicaid program, and public/private partnership programs; (2) The model provides detailed output in an easy-to-read format so that the implications of the simulated policy proposal are clear. Output statistics include, but are not limited to, long-term care insurance market penetration, state and federal government healthcare and long-term care spending, insurance company spending on long-term care, and out-of-pocket spending on long-term care.

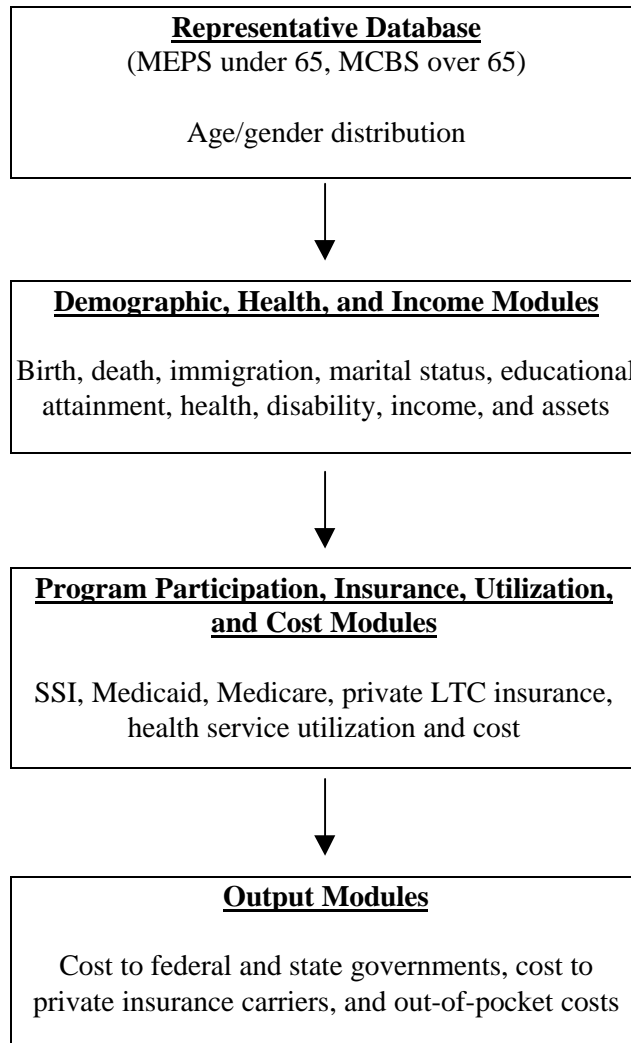
The main steps in developing the microsimulation model are described in subsequent sections of this document. The remainder of Section 1 describes the structure of the model and how it works in general terms. Section 2 discusses the estimation of the statistical relationships used in model components (called "modules"). Section 3 discusses the verification of model output (called "benchmarking") through comparisons with other standard projections (e.g., those of the Board of Trustees of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (OASDI)).

1.1. Model Structure and Dynamics

The microsimulation model is created as a series of independent modules, each performing calculations and simulations related to a specific variable (e.g., income), then passing the results to the next module. These calculations and simulations are based on statistical relationships among individual-level variables estimated from survey data. Each such module is designed to predict a new value of one individual-level variable (e.g., income) from current values of other variables (e.g., employment status, sex), and the previous year's value (e.g., lagged income).

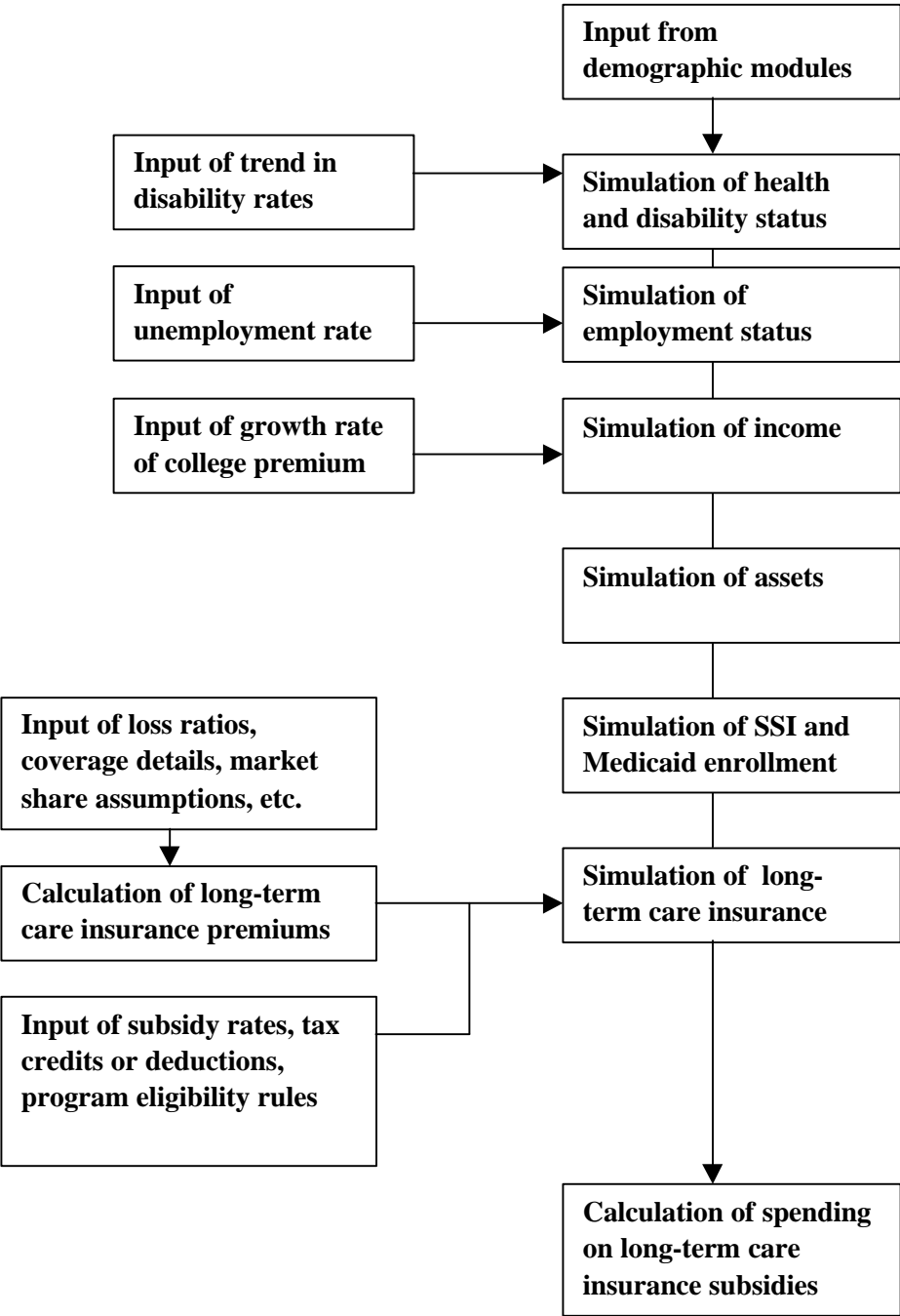
Figure 1 illustrates the hierarchical structure of the model. The model starts with a database of observations selected to be representative of the entire U.S. population in the base year (1996). The Medical Expenditure Panel Survey (MEPS) and the Medicare Current Beneficiary Survey (MCBS) are used to form this initializing database. Econometric equations are applied to update the data in numerous dimensions each year until 2050. For example, individuals represented by an observation might die or continue to live; if they live, they will age one year, their marital status might change, their income and assets will change, their health status and health care utilization will change, and so on. A separate module of computer code simulates each of these changes for each year.

Figure 1 Overview of Simulation Model



Each module builds on results generated by the modules that execute ahead of it, so the final results depend on the entire series of simulations, done in order. Figure 2 illustrates how a particular final result (total spending on long-term care insurance subsidies) depends on a series of preceding simulations. As indicated in the figure, subsidy spending depends on who purchases long-term care insurance, as simulated by the preceding module. This module in turn reads insurance premium rates from a preceding module. However, before insurance premiums can be calculated the user must input loss ratios, coverage details, and market share assumptions. In addition, the user inputs subsidy rates, tax credits or deductions. Finally, eligibility for Medicaid and SSI are considered along with assets, income, employment, health, and disability status, all determined in order by preceding modules that also rely on inputs from the user. Assets depend on income; income depends on employment; and employment depends on health and disability status.

Figure 2 Example: how spending on long-term care insurance subsidies is simulated



1.2. Inputs and Outputs

The long-term care microsimulation model offers a large degree of user control through a variety of inputs. It also provides many outputs to permit policy comparison and analysis. The inputs fall into five main categories:

1. Demographic parameters such as the fertility rate, trend in mortality rate, trend in educational attainment, and trend in disability rate;
2. Economic parameters such as the unemployment rate, the college premium (together, these two parameters provide control over the growth rate of the economy), and real rate of growth in costs for medical goods and services;
3. LTC insurance market parameters such as the target market penetration, loss ratio, and dollar value of benefits;
4. Consumer behavior parameters such as the moral hazard rate, and the rate of substitution of LTC for acute and primary care; and
5. Policy parameters such as the minimum age for non-disabled Medicare eligibility, rates of tax deductibility or credit for LTC insurance premiums (by income band), year of Medicare expansion to more fully cover LTC (if any), federal Medicaid match rate, amount and type of assets protected from spend down to Medicaid eligibility, and parameters that indicate which segments of the Medicaid population are the responsibility of the federal government, the state governments, or shared.

The model outputs data for analysis for each year in the simulation. The outputs fall into five main categories:

1. Demographic trends such as the total population, number of workers per retiree, and the proportion of the population 80 or more years old;
2. Economic trends such as the personal income, gross domestic product (GDP), and consumer price index (CPI);
3. LTC insurance market trends such as collected premiums, achieved market penetration, and losses;
4. Cost to individuals such as out-of-pocket direct LTC spending and out-of-pocket spending on LTC insurance premiums; and
5. Cost to governments and programs such as total federal and state health care and LTC costs, Medicaid acute/primary and long-term care spending, and federal LTC insurance subsidy spending or lost tax revenue.

In addition, the model outputs each year's microdataset. Any questions not addressed by pre-specified model outputs can be investigated using these microdata.

1.3. Modules

The general categories of modules are described in Figure 1. There are modules that update demographic variables (simulate births, deaths, marital status, and the like), modules that update health and disability status (e.g., whether or not the individual has impairment in 2 of 5 activities of

daily living), modules that update income and assets, modules that update program participation (simulate SSI and Medicaid eligibility and enrollment), modules that update LTC insurance purchase, modules that simulate utilization (e.g., doctor's visits, hospitalizations, nursing facility use, home health use, assisted living use), and modules that compute the cost associated with utilization. Finally, there are modules that divide up these costs among payors and produce the previously described output.

2. Statistical Estimation Methodology

Each module consists of several statistical equations governing behavior. The estimation of these equations is described in this section.

2.1. Data Sources for Estimation

Three survey data sources were used for the estimation of the statistical equations that underlie model modules: the 1995 and 1996 Medicare Current Beneficiary Survey (MCBS), the 1996 Medical Expenditure Panel Survey (MEPS), and the 1994 and 1996 Health and Retirement Survey (HRS). The last of these was used primarily for estimating the dynamics of assets since neither MCBS nor MEPS have asset information. These three datasets are described in more detail in the following subsections.

2.1.1. Medicare Current Beneficiary Survey (MCBS)

The Medicare Current Beneficiary Survey (MCBS) is a continuous, multipurpose survey of a representative sample of the Medicare population conducted under the auspices of HCFA. Sample persons are interviewed three times a year over several years to form a continuous profile of their health care experience. Field work for Round 1 began in September of 1991 and was completed in December. New rounds, which involve re-interviewing the same sample persons (or other appropriate respondents), begin every four months. Interviews are conducted regardless of whether the sample person resides at home or in a long-term care facility, using the questionnaire version appropriate to the setting. Full-year Medicare utilization and other administrative data for the sample persons are included in this data collection.

Respondents were sampled from the Medicare enrollment file to be representative of the Medicare population as a whole and by age group: under 45, 45 to 64, 65 to 69, 70 to 74, 75 to 79, 80 to 84, and 85 and over. Because of interest in their special health care needs, the oldest old (85 and over) and the disabled (64 and under) were oversampled to permit detailed analysis of these subpopulations. The sample is replenished annually for attrition, as well as for newly eligible persons. Useful variables include demographic characteristics, health and functional status, insurance coverage, financial, resources, family support, Medicare utilization for hospital, SNF, hospice, home health, outpatient and physician services.

2.1.2. Medical Expenditure Panel Survey (MEPS)

The Medical Expenditure Panel Survey (MEPS) is a nationally representative survey of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population, as well as a national survey of nursing homes and their residents.

MEPS is co-sponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS). The survey is designed to yield comprehensive data that allow researchers to estimate the level and distribution of health care use and expenditures, monitor the dynamics of the health care delivery and insurance systems, and assess health care policy implications.

2.1.3. Health and Retirement Survey (HRS)

The Health and Retirement Study (HRS) is an on-going survey by the University of Michigan that examines economic and health issues in conjunction with the process of retiring from the labor force. The survey has a sample of approximately 8,000 people from age 53 to 63. This survey additionally covers family structure, retirement plans, job demands and details about health insurance and pension plans. This survey furnishes information on how people prepare for retirement and how they expect to meet their health needs as they grow older. Useful variables include age, gender, marital status, health and functional status, program participation, service utilization, income, and assets.

Other data sources used in developing module dynamics include the 1998 Annual Report of the OASDI Board of Trustees, and U.S. Census Bureau data.

2.2. Estimation Example: Income

As a specific example, in this section we discuss the estimation of an equation for income. In the model, there are separate income equations for different age categories (16-24, 25-54, 55-64, 65+). Within these age categories, there are separate income equations for males, females, those in the labor force, and those not in the labor force. Each equation is log-linear, i.e., the log of income was modeled using ordinary least-squares regression. The general form of the income equation for 55-64 year olds in the labor force is

$$\log(\text{income}) = f(\text{lagged log income, age, age}^2, \text{male crossed with marital status, female crossed with marital status, work disability, educational attainment}).$$

While other independent variables (e.g., number of dependents) and other interactions were considered, they were dropped from the model due to lack of statistical significance.

3. Model Benchmarking

Survey and published data were used to validate and calibrate the model. As each module was completed we checked the model's output against published sources. This "benchmarking," as it is called, serves two purposes. First, it provides confidence that the microsimulation model is performing as intended. Second, it suggests the reasonable range of a variety of user-controlled inputs. Through benchmarking we were able to adjust the inputs so the outputs would match results published elsewhere (e.g., to match the population projections provided by the U.S. Census Bureau and the OASDI Board of Trustees).

3.1. Data Sources for Benchmarking

Data sources for benchmarking include those described in Section 2.1 as well as the Current Population Survey (CPS) and the National Health Interview Survey (NHIS). These are described in the next two subsections.

3.1.1. Current Population Survey (CPS)

The Current Population Survey (CPS) is a monthly survey of about 50,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics. The survey has been conducted for more than 50 years. The CPS is the primary source of information on the labor force characteristics of the U.S. population. The sample is scientifically selected to represent the civilian noninstitutional population. Respondents are interviewed to obtain information about the employment status of each member of the household 15 years of age and older. However, published data focus on those ages 16 and over. The sample provides estimates for the nation as a whole and serves as part of model-based estimates for individual states and other geographic areas. Estimates obtained from the CPS include employment, unemployment, earnings, hours of work, and other indicators. They are available by a variety of demographic characteristics including age, sex, race, marital status, and educational attainment. They are also available by occupation, industry, and class of worker. Supplemental questions to produce estimates on a variety of topics including school enrollment, income, previous work experience, health, employee benefits, and work schedules are also often added to the regular CPS questionnaire.

3.1.2. National Health Interview Survey (NHIS)

The National Health Interview Survey (NHIS) is a continuing nationwide survey of noninstitutionalized U.S. residents by household conducted by the National Center for Health Statistics. The survey is comprised of approximately 36,000 to 47,000 households (92,000 to 125,000 people) who are surveyed annually, with the number of households depending upon the year. Information is collected on the health status of individuals in the household, including number of school or work days lost, number of physician visits and medical services utilization and acute or chronic conditions responsible for these events. Additionally, data regarding chronic illness and resultant disability are collected. This survey uses a modified version of the ICD-9 classification system to record health conditions for respondents. Useful variables include, demographics, health status, medical service utilization, acute or chronic conditions responsible for utilization, and disability status.

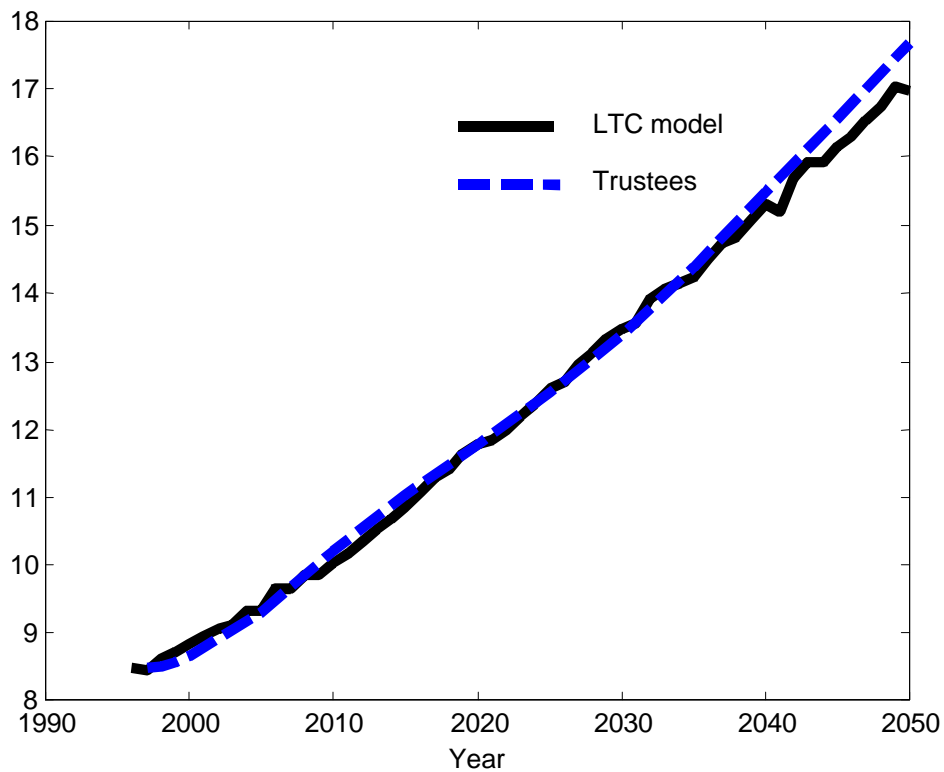
In addition to the aforementioned survey datasets, we used the Health Care Financing Administration's National Health Accounts to benchmark health care spending.

3.2. Benchmarking Example: GDP

Model outputs were benchmarked by comparing them to other established projections. In this section we illustrate just one example, GDP. Personal income has historically been proportional to GDP, so increasing the sum of all simulated personal income by a fixed factor can provide an accurate simulation of GDP.¹ Figure 3 illustrates a GDP benchmark where the microsimulation GDP projections have been obtained from income as just described. In this figure we have plotted the microsimulation projection of GDP and the OASDI Board of Trustees medium projection (all values in 1996 dollars).

Figure 3

GDP Projections in 1996 Dollars



Sources: AHCA microsimulation model, OASDI Board of Trustees, 1998 Annual Report

¹ Simulated personal income does not include categories of income like retained earnings by corporations, the value of employee benefits, social insurance, capital investments, and depreciation.