Alternative Strategies for Imputing Premiums and Predicting Expenditures Under Health Care Reform

Pat Doyle and Dean Farley, Agency for Health Care Policy and Research

President Clinton's proposed Health Security Act (HSA) would guarantee all U.S. citizens a standard set of insurance benefits to cover the cost of health care [1]. Because the cost of the HSA depends directly on the premiums associated with this benefit package, simulating the cost and distributional impact of the HSA and similar health care reform proposals requires estimates of these premiums. Premium estimates, in turn, require estimates of health expenditures under universal coverage, particularly covered benefits.

The Agency for Health Care Policy and Research (AHCPR) developed the AHSIM microsimulation model to estimate (among other things) the cost of various reform proposals. The AHSIM Model predicts expenditures among the nonelderly population under universal health insurance coverage; calculates premiums associated with an insurance benefit package for this population group; and simulates both direct and indirect costs in total and to the Federal government, employers, and households. Because predicted expenditures include two stochastic components, which are determined by assigning random numbers, estimates of total expenditure, premiums, and financing burdens depend upon random draws in the model.

In this paper we test how sensitive the results of simulating a health care reform proposal are to the stochastic elements in the expenditure model. We examine estimated variations in aggregate costs, premiums, and distributional impacts as the random draws are replicated with different seeds. Our reform proposal resembles, but does not exactly replicate, the HSA; the results should not be interpreted as estimates of the President's plan (H.R. 3600/S. 1757) [2].

■ Background

The Expenditure Model

AHSIM relies on a two-part econometric model estimated from the 1987 National Medical Expenditure Survey (NMES) to predict annual expenditures for each nonelderly person in the NMES data set. The expenditure model predicts the likelihood of incurring expenditures for each of ten different types of services in a year and an average annual expenditure for each service type among consumers of that service type. There are ten pairs of equations in this model, each consisting of a probit for the probability of use and a log-linear OLS regression for level of expenditures, conditioned on some use. Insurance status is parameterized as a set of binary variables indicating full- and part-year coverage under employer-sponsored insurance (ESI), other private insurance, Medicare, Medicaid, and other public medical programs. Other explanatory variables, which are identical in all 20 equations, include demographic characteristics, income, geographic location, and an extensive set of health status measures from the health questionnaires administered in NMES.

Application of the Expenditure Model

In general, the expected impact of a particular reform plan depends upon (1) the level of spending expected in the absence of reform, given existing insurance coverage and other characteristics (referred to as "baseline expenditures"), and (2) expenditures predicted on the basis of insurance coverage as modified by the reform plan. The AHSIM Model simulates both baseline and reform expenditures, so that the only differences between the two are attributable directly to the effects of reform. Simulations are performed in real, 1987 dollars, which are then aged to the point in time of particular analytic interest.

The model simulates expenditures for insured people in two stages. The first stage presumes that all individuals are enrolled in a plan that is typical of ESI. (This is equivalent to using a dummy variable for ESI in the expenditure model.) The second stage adjusts individual expenditures for the relative generosity of his or her "held plan." Adjustments to the probability and level of use are based on findings from the Rand Health Insurance Experiment (Manning et al., 1987) [3]. In this analysis, people are assumed to enroll in either the Health Maintenance Organization or the fee-for-service plan prescribed by the HSA.

The expenditure model uses two stochastic elements to simulate expenditures for each of the 10 service types: a random draw from a uniform (0,1) distribution to establish which persons incur expenditures of a given type and a random draw from a multivariate normal distribution to determine the error terms for the 10 expenditure equations, simulating the error component of the annual expenditure equations. Once assigned, the random variables are constant for both baseline and reform expenditures.

In the aggregate, the baseline predictions of the AHSIM expenditure model differ from actual NMES expenditure data because of these stochastic elements. We control for these differences by calibrating the imputed baseline expenditures to actual NMES totals. The baseline calibration factors are applied in the reform simulation, as well as the baseline simulation.

Deriving Premium Estimates

The AHSIM Model approaches the estimation of insurance premiums differently from other models, in that it builds premiums up from the expected claims of the new insured population. Unlike actuarial methods which rely on previous claims experiences, this approach explicitly takes into account the health status and demographic characteristics of people who obtain insurance as a result of reform. It also ensures that the premiums used in the model are consistent with the assumed expenditures of the insured population.

In particular, the model simulates benefits paid by applying a claims processing module to estimated expenditures. Averaging benefits paid over units and adding an administrative load yields estimates of community-rated premiums for the package. In this analysis, premiums are calculated separately for each of 12 groups, defined by Census region and rating pool (adults with no dependents, single adults with dependents, and couples with dependents).

The model uses estimated premiums to simulate the costs to households and employers of purchasing insurance and the costs to the Federal government of subsidizing insurance purchases and out-of-pocket expenses. Second order effects -- i.e., the change in net expenses resulting from the change in insurance costs (including wage effects and the opportunities to purchase supplemental insurance) -- are simulated as well. The model displays total costs incurred by households, employers and the Federal government and summarizes relative changes in expenditures among individual units.

■ Methods

For this study, we executed the AHSIM Model 50 times, using different random draws each time. Each replication estimated a set of baseline and reform expenditures, premiums derived from these simulated expenditures, household and employer contributions for mandatory insurance coverage, and Federal subsidies of household and employer payments. The financing scheme was patterned after the following proposal:

- ❑ All noninstitutionalized civilians under age 65 who are not recipients of Medicare, Aid to Families with Dependent Children, or Supplemental Security Income enroll in the system.

- ❑ Health insurance units purchase coverage through a regional alliance, unless the policyholder is employed by a firm with at least 5,000 employees.

- Employers pay 80% of the premiums for their workers. Firms in regional alliances are subsidized to ensure that their financial obligations do not exceed 7.9% of payroll; small, low wage firms may receive additional subsidies. Employers of part-time workers pay a pro rata share of the 80% obligation.

- Employers pay for employees' insurance based on the Weighted Average Premium (WAP), which is the average community-rated premium prevailing in the relevant alliance. Employers of persons in units with two adults pay 80% of the WAP divided by the average number of full-time equivalent workers in such units. Other employers pay 80% of the WAP.

- Health insurance units obligations include: 20% of the WAP, the difference between the chosen plan premium and the WAP, and any portion of the 80% employer share that is unpaid because the unit lacks a full-time worker. Unit contributions toward the 20% share are capped at 3.9% of adjusted gross income (AGI). Households below 250% of poverty receive additional subsidies for both the 20% share and the unpaid employer balance.

- Nonelderly early retirees with AGI below a cut-off receive a subsidy to cover the full employer share of premium costs.

- Contributions of low wage workers in units covered through the corporate alliance are limited to 5% of the premium. Subsidies are paid by the employer.

Results

In a typical AHSIM application, AHCPR estimates the impact of program reform on the model population once, using one set of random numbers calibrated to NMES baseline data. The results clearly involve sampling error of unknown magnitude. The purpose of this study is to quantify the information lost by reporting one estimate, rather than a range of estimates, from the model.

The analysis focuses on how the process of imputing expenditures affects three types of measures: (1) total expenditures under baseline and reform, (2) premiums under the proposed minimum benefit plan, and (3) the distribution of the financing burden across the household, employer and Federal sectors. The analysis abstracts from other sources of error, including NMES sampling error, the random assignment of HMO and FFS plans, and specification errors in the model itself.

Expenditures

Cohen and Sommers (in Doyle et al., forthcoming) estimate a lower bound of 3.3% on the relative standard error of predicted baseline mean expenditures from NMES sampling error alone. This implies that a 95% confidence interval would be at least ±6.5% of the estimate. Our analysis increases this interval, based on the stochastic elements of the expenditure imputation, assuming that these two sources of random error are independent.

Table 1 shows that calibrated baseline expenditures are not very sensitive to the stochastic elements in the expenditure model [4]. They range only a few percentage points, with a relative standard error of less than 1% (i.e., the standard deviation is less than 1% of the mean). Assuming independence, these results enable us to calculate a relative standard error for mean expenditures that includes three error components: NMES sampling error, errors in determin-

Table 1: Variation in Total Expenditures Across Replicates (billions)

Scenario	Mean	Minimum	Maximum	Relative Standard Error (percent)
Baseline	380	369	389	1.1
Reform	384	370	396	1.5
Change	4	-6	15	107.2

ing who has expenditures, and errors in forecasting levels of expenditures. This estimate is 3.4%.

The lack of variation in expenditures across replicates is not surprising given that we calibrate baseline expenditures to NMES-reported means. However, we expect more variation in reform expenditures, since they depend upon plan choice and assumed behavioral responses, in addition to the stochastic elements of the baseline expenditure imputation. These other factors increase reform standard errors in ways that are not directly quantified in this analysis. Nonetheless, we observe minimal variation in reform expenditures across random draws. The minimum and maximum values fall within 4% of the mean and the relative standard error of the mean is less than 2%.

On the other hand, the magnitude of the difference between baseline and reform expenditures varies significantly across the replicates. We cannot even infer the direction of the impact of reform on total expenditures, since the difference between total expenditures under reform and total expenditures under baseline ranges from a negative amount (-$6 billion) to a positive number ($15 billion). On average, the proposed plan will cause expenditures to rise by a small amount -- $4 billion dollars -- with a 95% confidence interval of -4.8 to 13.6 billion dollars.

Premiums

AHCPR premiums represent predicted benefits paid under a specified insurance plan, inflated by an appropriate administrative load and averaged over health insurance units, region, and type of unit. Adults (married or not) who have no dependents, single adults with dependents, and married adults with dependents are treated as three separate rate pools. Premiums should vary more than predicted per capita expenditures, because they are averaged over small groups. We also expect more variation in the premiums, because the model does not calibrate average benefits paid in each of the replicates to a known independent estimate.

Table 2 confirms these expectations. The relative standard errors of fee-for-service premiums range from 5% to 23% [5]. Variances are highest for one-adult families, particularly in the northeast and midwest. Larger variances are at least partly due to sample size; the number of one adult family units is roughly half the size of the next largest group, married families with children. Table 2 shows that

Table 2: Variation in Fee-for-Service Premiums Across Replicates

Rating Pool	Region	Mean	Minimum	Maximum	Relative Standard Error (percent)
Individual	Northeast	1520	1306	1761	7.5
	Midwest	1329	1079	1735	10.0
	South	1529	1256	1727	6.9
	West	1478	1073	1926	9.9
One Adult with Children	Northeast	2817	1676	5272	23.2
	Midwest	4725	2990	7089	22.0
	South	2888	2250	3621	10.7
	West	2320	1613	3940	16.5
Two Adults with Children	Northeast	4157	3308	5130	8.0
	Midwest	3565	3173	4157	6.6
	South	3485	3182	3789	4.8
	West	3892	3354	4776	7.9

other relative standard errors range from 5% to 10%, without appreciable differences between individuals and two-adult families. The similarity of these standard errors may seem puzzling at first, since individual units are roughly three times as common as two-adult families. However, the average two-adult family has 3.8 people, roughly offsetting the decrease in variation in premiums from the 3-fold increase in sample size.

Financing

The model estimates both the level and the change in spending by sector (household, employer, Federal government) and type of expense (premiums, out-of-pocket costs, and subsidies). In general, most estimates of important reform levels are fairly insensitive to the random elements of the expenditure model; estimated impacts of reform relative to baseline levels are not.

Household Obligations

Households pay 20% of their premium plus any unpaid portion of the employer obligation. Both components are bounded so that low-income health insurance units are not overburdened with high premium and out-of-pocket costs. We also assume that households bear most of their employers' costs under reform, as higher insurance premiums are passed on to employees in the form of lower wages. On the flip side, when employers pay lower premiums under reform, households benefit in one of three ways. Employers may pay some of the households 20% share, they may purchase supplemental insurance, or they may increase wages.

Total household obligations are not very sensitive to the random draws (Table 3). The 95% confidence interval is \$218 billion ± \$7.6 billion. But, households do not actually pay their full obligations, because of premium subsidies for low-income households and out-of-pocket subsidies for low-income enrollees who do not have access to HMOs. It turns out that household premium discounts are not very sensitive to the random draw, either, averaging \$39 billion dollars with a relative standard error of 3%. Out-of-pocket costs are low on average and moderately sensitive to the random draw (averaging \$6 billion with a relative standard error of 8%). Note that out-of-pocket costs depend upon plan choice, as well as the level of imputed expenditures.

While total households obligations do not vary significantly, estimates of the relative impact on net house-

Table 3: Variation in Financing Burden Across Replicates (billions)

Financing Element	Mean	Minimum	Maximum	Relative Standard Error (percent)
Total Obligation for Household	218	209	226	1.7
Household Premium Discount	39	36	42	3.1
Subsidy of OOP Expenses	6	5	8	8.1
Change in Household Payments	24	16	31	14.5
Total Obligation for Employer	227	223	233	.8
Net Employer Payments (Reform)	195	195	196	.1
Change in Employer Payments	13	12	13	1.7
Employer Premium Discount	31	28	37	5.3
Change in Federal Government Costs	82	76	92	3.5

Source: Agency for Health Care Policy and Research, Center for Intramural Research, AHSIM simulation model.

hold payments do. The model estimates that net household payments (total obligations less premium and out-of-pocket subsidies) increase $24 billion (or 16%), on average, over baseline expenditures for health care and premiums. Estimates of this relative change vary from $16 to $31 billion. The relative standard error of the mean relative change is 14.5%, yielding a fairly large 95% confidence interval ($24 billion ± $6.9 billion). While we can infer that household payments will increase, we cannot be particularly precise about the magnitude of the change.

Employer Obligations

In theory, employers pay 80% of their employees' premium costs under the HSA. However, the effect of this policy is mitigated in two important ways. First, employer contributions for each worker are capped as a percent (7.9%) of average payroll; employers of roughly half of all workers are eligible for reduced payments as a result. In these instances, employer payments do not depend directly upon premiums. Furthermore, employers' contributions are determined by a fairly complex averaging process that pools premiums across types of plans chosen (FFS versus HMO) and family units. The net effect, again, reduces the relationship between employer payments and variations in expenditures.

It is, therefore, not surprising that both the total employer obligations and net employer payments are fairly constant across replicates; the relative standard error is less than 1%. The relative change in employer obligations from baseline to reform is also fairly insensitive to the random draw, with a relative standard error of under 2%.

With a relative standard error of 5%, the size of the discount afforded employers (the difference between the capped employer payments and the employer share of the WAP) is more sensitive to the random draw than net employer payments. Fortunately, because it is paid by the government, this discount is small in relation to total employer obligations. The sensitivity of this output measure is due, in part, to its relatively small level.

Government

The Federal government must pay the employer portion of insurance costs for its employees. It also finances household and employer subsidies. Total Federal obligations estimated directly by AHSIM average $82 billion, but this figure is relatively volatile, ranging from a low of $76 billion to a high of $92 billion. The 95% confidence interval is ± 7% [6]. Such variation is satisfactory for some analytic purposes, it is unfortunately large for the purpose of evaluating alternative health reform proposals.

■ Conclusions

Until the publication of a study of micro-simulation models by the National Academy of Sciences (Citro and Hanushek, 1991), microsimulation models were used heavily to analyze the relative impact of proposed reforms to the welfare and nutrition programs, without much information on the nature of the errors of the model estimates. Cohen et al. (1991) show how alternative models "...that were thought a priori to have similar success in modeling [the Aid to Families with Dependent Children program]..." can produce conflicting predictions as to the size and direction of the program caseload and costs. Doyle and Trippe (1989) illustrated the effectiveness of calibration techniques in reducing the bias in some, but not all, model estimates. The analysis presented here contributes to understanding the effects of uncertainty in model estimates, by illustrating the sensitivity of relative impact measures in one microsimulation model to one of its several stochastic elements.

Relative impact measures are especially influential in debates over public policy. The government often needs to implement programs in years other than those represented by the model output. They may also want to develop model estimates under varying assumptions about the macroeconomic conditions or other interrelated programs. Conventional wisdom has often held that the relative impact measures may, in fact, be more accurate than absolute measures. In fact, this belief underlies the typical modeling approach of simulating, rather than observing, a baseline scenario.

Unfortunately, conventional wisdom is not always correct. This analysis demonstrates that relative estimates can be subject to substantial random variations. In contrast, most of aggregate outcome measures in the model are not very sensitive to the expenditure imputation process, even though such expenditures drive virtually all of the basic cost estimates. The sensitivity of the model estimates depends on the relationship among the structure of the reform and the sources of error in the model, as well as on the size of the relative impact measure. On balance, however, our results indicate that microsimulation modelers cannot continue to ignore the uncertainty that surrounds model estimates, especially when simulation methods rely on additional stochastic processes.

■ Footnotes

[1] Farley and Short (1994) describe the comprehensive benefit package. A summary of the plan appears later in the report.

[2] For estimates of the costs of the President's proposal see Fuchs and Merlis (1993).

[3] The expenditure model and adjustments for behavioral response are described more fully in Doyle et al. (forthcoming).

[4] Baseline expenditures are imputed to a subset of the full AHSIM sample, to take advantage of a fuller set of information on employers collected for this subsample. To correct for the sampling error in selecting the subsample, we calibrate the values imputed to the subsample to those imputed to the full sample. We also calibrate the imputation of expenditures to the reported data upon which the equations are estimated. The calibration raises the mean expenditures (averaged over the 50 replicates) by 5% and lowers the variance over the 50 replicates by over 90%.

[5] Results for HMO plans are available from the authors. The patterns observed in Table 2 for FFS plans holds for HMO plans, as well, with some slight variation in the magnitude of the numbers.

[6] This is not an estimate of total Federal obligations. The analysis excludes, inter alia, changes in tax revenue due to wage changes, as well as changes in programs like Medicaid, CHAMPVA and CHAMPUS.

■ References

Citro, C.F. and Hanushek, E. A. (1991), *Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling.* Washington, D.C.: National Academy Press.

Cohen, M; Billard, L.; Betson, D. M.; and Ericksen, E. P. (1991), "A Validation Experiment with TRIM2," in Citro, C.F. and Hanushek , E. A. (eds), *Improving Information for Social Policy Decisions: The Uses of Microsimulation Modeling*, Washington, D.C.: National Academy Press.

Doyle, P.; Farley, D.; Short, P.; Cohen, S.; Carlson, B.; Beauregard, K.; and Harvey, H. (Forthcoming), The AHSIM Model. Rockville Md: Center for Intramural Research, Agency for Health Care Policy and Research.

Doyle, P. and Trippe, C. (1989), Validation of the Food Stamp Microsimulation Model. Final Report to the Food and Nutrition Service, U.S. Department of Agriculture. Washington, D.C.: Mathematica Policy Research, Inc.

Farley, D.E. and Short, P. F. (1994), Plan Selection in a Microsimulation Model of Health Care Reform through Managed Competition. Rockville Md: Center for Intramural Research, Agency for Health Care Policy and Research.

Fuchs, B. and Merlis, M. (1993), Health Care Reform: President Clinton's Health Security Act, Congressional Research Service Report for Congress, November 22.

Manning, W.; Newhouse, J.; Duan , N.; Keeler, E.; Leibowitz, A.; M.S.; and Marquis (1987), "Health Insurance and the Demand for Medical Care," *American Economic Review*, 77, 3, 251-277, June. ■