Volume Title: The Role of Health Insurance in the Health Services Sector

Volume Author/Editor: Richard N. Rosett

Volume Publisher: NBER

Volume ISBN: 0-87014-272-0

Volume URL: http://www.nber.org/books/rose76-1

Publication Date: 1976

Chapter Title: The Effect of National Health Insurance on the Price and Quantity of Medical Care

Chapter Author: Martin Feldstein, Bernard Friedman

Chapter URL: http://www.nber.org/chapters/c3826

Chapter pages in book: (p. 505 - 541)

12

MARTIN FELDSTEIN
Harvard University
and
BERNARD FRIEDMAN
Northwestern University

The Effect of National Health Insurance on the Price and Quantity of Medical Care

The study reported in this paper uses a microsimulation model to estimate the effects of alternative national health insurance policies. Unlike previous microsimulation studies, the current analysis uses an explicit model of the supply and price response in the markets for hospital care and physicians' services. Indeed, two quite different models of supply and price response are examined and their implications are contrasted.

A *microsimulation* model of household demand is necessary if the analysis is to provide useful results in the comparison of specific health insurance proposals. The effects of alternative national health insurance policies depend crucially on the stochastic character of health care demand. More specifically, the effects of

The methods and programs used in this paper were developed in a project supported by the Department of Health, Education, and Welfare. The model in Section 1 was described in an unpublished technical report, "National Health Insurance Simulation Model" (August 1972), and the methods in Sections 2 and 4 were described in "Supply and Price Response in National Health Insurance Analysis" (September 1972). We are grateful to the Department of Health, Education, and Welfare for its support of this research and to B. Mitchell, G. Moyer, and D. Schenker for useful discussions.

different sets of deductibles, coinsurance rates, and other parameters of insurance policies depend on particular stochastic distributions of health expenditures. *Aggregate* specification of demand behavior cannot capture the subtle differences in the response of demand to different types of insurance policies. The current investigation uses a stochastic simulation model of demand based on the actual experience of more than 300,000 families. The basic demand model, described in detail in Section 1, is an extension of the simple aggregate health care demand model used in Feldstein, Friedman, and Luft (1972).

A serious weakness of all previous microsimulation studies of national health insurance, including our own (1972), has been the neglect of the supply and price response to national health insurance. The current analysis shows how an aggregate model of supply and price response can be combined with a microsimulation model of demand. Although there are no explicit *aggregate* demand equations, the complete model is solved for prices that equate supply and demand. The supply model and the method of finding the equilibrium are described in Section 2. Some illustrative results are then presented in Section 3.

The supply and price response of sections 2 and 3 is based on the simplest model of aggregate supply and the assumption of market clearing equilibrium. The markets for hospital care and for physicians' services may not behave in this way. Hospital prices may rise in response to increases in demand because hospitals change the nature of their product and not because it is more expensive to produce a larger quantity of the old product. Physicians may increase prices in response to greater demand or increased insurance without setting a market clearing price. Section 4 develops a model with these characteristics, describes the simultaneous interaction of demand with this supply behavior, and presents some illustrative results.

Most of the debate about the effects of national health insurance has focused on the uncertainty about the responsiveness of household demand. The current study shows that our uncertainty about supply response may be even more important.

1. A MICROSIMULATION MODEL OF DEMAND

The annual health care expenditures of a group of families with the same demographic composition, income, and insurance coverage can be described by a joint frequency distribution of expenditures on hospital services and medical services. Each such distribution is conditional on the gross prices charged for hospital and physician services. Let $F^i(E_h, E_m | P_h, P_m)$ be such a distribution for insurance coverage i with:

E_h = family's total expenditure for hospital services
E_m = family's total expenditure for medical services
P_h = gross price per unit of hospital services
P_m = gross price per unit of medical services

There is an associated distribution of net out-of-pocket expense $G^i(N | P_h, P_m)$ that is related to F^i by the insurance reimbursement formula.

The expenditure distribution associated with any particular insurance structure and prices is derived from a "baseline" *quantity* distribution that would prevail in the absence of any insurance and with prices equal to unity: $F^0(X_h, X_m | P_h = 1, P_m = 1)$. There is, of course, a different baseline expenditure distribution for each family type. A specific national health insurance proposal can be described in terms of the deductibles, coinsurance rates, and maximum net out-of-pocket expenditure for each type of family. D_h and D_m will be used to denote the deductibles, C_h and C_m the coinsurance rates, and MAX the maximum net out-of-pocket expenditure.

The equations relating expenditure in the presence of insurance to the baseline distribution and the prevailing gross prices is an extension of a traditional constant elasticity demand model. The most appropriate way to extend a constant elasticity specification to deal with deductibles and a maximum net spending limit is uncertain. One approach, offered as a tentative specification until better empirical evidence is available, assumes a constant elasticity of the quantity demanded with respect to the *net price* paid for expenditures over the deductible. The net price of an additional unit of hospital care depends on the family's current level of expenditure. More generally, the net price paid by the family depends on (1) the gross price charged by the hospital (P_h), (2) the effective coinsurance rate (1 for expenditures below the deductible, (C_h) between the deductible and the maximum net expenditure limit, and 0 above that limit) and (3) a parameter, λ, representing the nonmonetary costs (Acton, 1972; Phelps and Newhouse, 1972) to the consumer of health services.

If there were no deductibles ($D_h = D_m = 0$) and no maximum net expenditure (MAX $= \infty$), the two expenditure equations would be:

$$(1) \qquad E_h = P_h \cdot X_h \left[\frac{P_h(C_h + \lambda)}{1 + \lambda}\right]^{\alpha_h} \left[\frac{P_m(C_m + \lambda)}{1 + \lambda}\right]^{\alpha_m}$$

$$(2) \qquad E_m = P_m \cdot X_m \left[\frac{P_h(C_h + \lambda)}{1 + \lambda}\right]^{\beta_h} \left[\frac{P_m(C_m + \lambda)}{1 + \lambda}\right]^{\beta_m}$$

Notice that if there is no insurance, $C_h = C_m = 1$ and the equation is the usual constant elasticity demand equation. With complete insurance, $C_h = C_m = 0$ but demand remains finite because $\lambda > 0$ implies a positive nonmonetary cost.

To allow for deductibles and for the maximum net expenditure limit, it is necessary to distinguish four separate cases. Let $\hat{E}_h = X_h P_h^{1+\alpha_h} P_m^{\alpha_m}$ and $\hat{E}_m = X_m P_h^{\beta_h} P_m^{1+\beta_m}$, the expenditures that would occur at prices P_h, P_m if there were no insurance.

Case i If $\hat{E}_h < D_h$ and $\hat{E}_m < D_m$, then $E_h = \hat{E}_h$ and $E_m = \hat{E}_m$. Here the insurance is irrelevant because the deductibles exceed the expenditure that would be made in the absence of insurance. The total net out-of-pocket expenditure is $N = E_h + E_m$.

Case ii If $\hat{E}_h > D_h$ and $\hat{E}_m > D_m$, then

$$(a) \qquad E_h = D_h + P_h(X_h - D_h P_h^{-1}) \left[\frac{P_h(C_h + \lambda)}{1 + \lambda}\right]^{\alpha_h} \left[\frac{P_m(C_m + \lambda}{1 + \lambda}\right]^{\alpha_m}$$

$$E_m = D_m + P_m(X_m - D_m P_m^{-1}) \left[\frac{P_h(C_h + \lambda)}{1 + \lambda}\right]^{\beta_h} \left[\frac{P_m(C_m + \lambda)}{1 + \lambda}\right]^{\beta_m}$$

and

$$N = D_h + D_m + C_h(E_h - D_h) + C_m(E_m - D_m) < \text{MAX}$$

or

$$(b) \qquad E_h = D_h + P_h(X_h - D_h P_h^{-1}) \left[\frac{P_h \lambda}{1 + \lambda}\right]^{\alpha_h} \left[\frac{P_m \lambda}{1 + \lambda}\right]^{\alpha_m}$$

$$E_m = D_m + P_m(X_m - D_m P_m^{-1}) \left[\frac{P_h \lambda}{1 + \lambda}\right]^{\beta_h} \left[\frac{P_m \lambda}{1 + \lambda}\right]^{\beta_m}$$

and

$$N = \text{MAX}$$

Case iii If $E_h > D_h$ and $\hat{E}_m < D_m$ then either:[1]

(a) $$E_h = D_h + P_h(X_h - D_h P_h^{-1}) \left[\frac{P_h(C_h + \lambda)}{1 + \lambda}\right]^{\alpha_h} \left[\frac{P_m(C_m + \lambda)}{1 + \lambda}\right]^{\alpha_m}$$

$$E_m = \hat{E}_m$$

$$N = D_h + C_h(E_h - D_h) + \hat{E}_m < \text{MAX}$$

or

(b) $$E_h = D_h + P_h(X_h - D_h P_h^{-1}) \left[\frac{P_h \lambda}{1 + \lambda}\right]^{\alpha_h} \left[\frac{P_m \lambda}{1 + \lambda}\right]^{\alpha_m}$$

$$E_m = P_m X_m \left[\frac{P_h \lambda}{1 + \lambda}\right]^{\beta_h} \left[\frac{P_m \lambda}{1 + \lambda}\right]^{\beta_m}$$

$$N = \text{MAX}.$$

Case iv If $\hat{E}_h < D_h$ and $\hat{E}_m > D_m$, the results bear an obvious analogy to Case iii.

These four sets of demand equations can be used to generate the distribution F^i corresponding to any gross prices, insurance characteristics, and demand parameters. More specifically, given a baseline distribution $F^0(X_h, X_m)$ we can draw values (X_h, X_m) with the appropriate probability and calculate the corresponding E_h, E_m. Average gross and net expenditures ($\overline{E}_h$, $\overline{E}_m$, and $\overline{N}$) are then readily computed. This procedure is done separately for each family type. These calculations and the aggregates produced by combining the averages for different family types could be used to assess the effects of alternative national health insurance proposals if supplies were infinitely elastic at fixed values of P_h and P_m. This was essentially the procedure used in Feldstein, Friedman, and Luft (1972) with a simpler model that did not distinguish hospital and medical services. The more general use of the demand simulation model when prices are endogenous is discussed in the next section. The remainder of this section describes our data sources, the derivation of the family baseline distributions, and the calibration of the model to 1970 aggregate experience.

The primary data are the individual insurance claims for more than 300,000 federal government employees and their dependents in 1970. All of these persons had the very comprehensive Aetna "High Option" coverage. A tabulation of the joint frequency distribution of hospital and medical services (a 24-by-24 matrix of relative frequencies with an associated matrix of cell means) was derived for all male employees. Similar tabulations were derived separately for female employees, dependent spouses, and children. The Aetna coverage uses a \$50 deductible per individual (except that the first \$1,000 of hospital room and board is fully reimbursed) and a coinsurance rate of 20 per cent for both hospital and medical services. For each category of individual, the observed bivariate frequency distribution was used to infer the baseline distribution corresponding to each of several alternative sets of demand parameters (α's, β's, and λ). This procedure uses the inverse of the demand function with $D_h = 0$, $D_m = 50$, $C_h = C_m = 0.20$, and MAX $= \infty$. The prices were both normalized to be 1.

Family baseline distributions for sixteen different family compositions (e.g., husband and wife; husband, wife, and three children; etc.) were produced by convoluting the baseline distributions for individuals.[2] Persons over 65 were assumed to be excluded from the basic national health insurance plan and were therefore ignored in calculating family distributions. For each of the sixteen family compositions, twelve income categories were defined in order to implement national health insurance specifications that are income related and to assess the tax burdens that balance the new government expenditures.

The demand simulation model provides average gross and net expenditures for each of the 192 family types. These are also aggregated to national averages and subaverages by using population counts computed from the Current Population Survey of 1971.

Although the experience of the federal employees with Aetna High Option coverage is an extremely rich and valuable source of data, these employees are not a representative sample of all U.S. families, particularly in regard to geographic distribution, employment status, and occupation. For each set of demand parameter assumptions (α_n, α_m, β_n, β_m, λ) the following additional steps are taken to calibrate the baseline distributions to known national aggregates for 1970. The typical family insurance coverage in 1970 was assumed to be $D_h = 100$, $D_m = 100$, $C_h = 0.25$, $C_m = 0.40$, and MAX $= \infty$.[3] This typical coverage is used for a preliminary simulation to estimate the national aggregate expenditures corresponding to each set of demand parameters. Suppose that the estimated expenditures for some set of demand parameters are E_h^* and E_m^* and that the actual expenditures are E_h^0 and E_m^0. The ratios E_h^*/E_h^0 and E_m^*/E_m^0 are then used to deflate the expenditure units of the corresponding baseline distributions. With this calibration completed, the implications of various national health insurance plans may be compared for the given set of demand parameters.

2. A SUPPLY RESPONSE MODEL AND THE EQUILIBRIUM SOLUTION

In the supply response model, the usual interaction of supply and demand provides a market clearing reaction of quantities and prices to a change in insurance coverage. The computational problems and novelty of the model occur because there are no aggregate demand equations but only individual demand equations and a microsimulation model.

The basic idea and computational procedure for the supply response model can be described most easily by ignoring the distinction between hospital and medical services. The solution for the more general case will be discussed below. Let Q be the aggregate quantity of health services consumed by all households, P the price level, and E the expenditure ($E = Q \cdot P$). Since there are no natural units in which to measure Q, we take the current price prevailing in the absence of national health insurance (P_0) to have the value 1 and thus define the quantity in the absence of national health insurance (Q_0) to be equal to the expenditure (E_0).

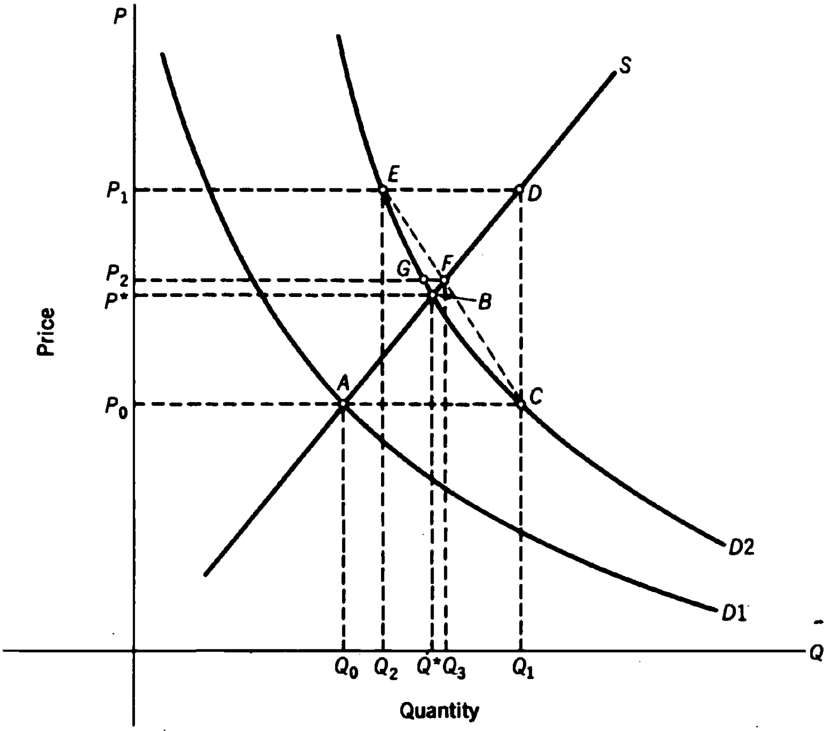
An aggregate supply function with constant price elasticity can be written:

(3) $$\ln Q = \ln Q_0 + \gamma \ln P$$

If an *aggregate* market demand function could be written in terms of the gross market price and the features of the national health insurance program, the two equations could be solved for the changes in P and Q that would accompany alternative NHI plans. However, such an aggregate demand function is the outcome of a very complex and stochastic set of individual demand functions that cannot be given an aggregate parametric summary. Only by operating the demand simulation model described in Section 1 can points on the demand curve be calculated. The market equilibrium is found by combining the aggregate supply function of Equation (3) with demand generated by the simulation model.

The process of convergence to an equilibrium solution of this iterative simulation process is best described with the aid of a diagram. Figure 1 shows the change in the price-quality equilibrium that results from the introduction of national health insurance. More specifically, in the absence of NHI the market is in equilibrium at point A. The demand curve ($D1$) relates the quantity demanded to the *gross* price for the structure of private health insurance prevailing before NHI. The $D1$ curve is not actually known but points on it can be found by using the demand simulation model. The supply function S corresponds to Equation

FIGURE 1



(3). The introduction of NHI shifts the demand function to $D2$ and the equilibrium to B. The computational problem is to determine the coordinates of the point B even though the two demand curves are not directly observable.

The following feasible and efficient procedure is used. First, the demand simulation model is used to find the aggregate quantity that would be purchased under NHI if price were unchanged—i.e., it locates the point C and the quantity Q_1. Second, the supply function is used to solve for the price P_1 at which the quantity Q_1 would be supplied. Third, the demand simulation model is used again to find the aggregate quantity demanded in the presence of NHI but with the gross market price P_1; this is Q_2 at point E. It is clear from Figure 1 that (if the aggregate demand function is well behaved) the equilibrium price after NHI (P^*) lies between P_0 and P_1. Similarly, the new equilibrium quantity (Q^*) lies between Q_1 and Q_2. The fourth step in the analysis is to approximate the unknown demand curve ($D2$) in the relevant range by the straight line connecting points C and E. This is shown as a broken line in the figure. For computational purposes it is defined by the equation

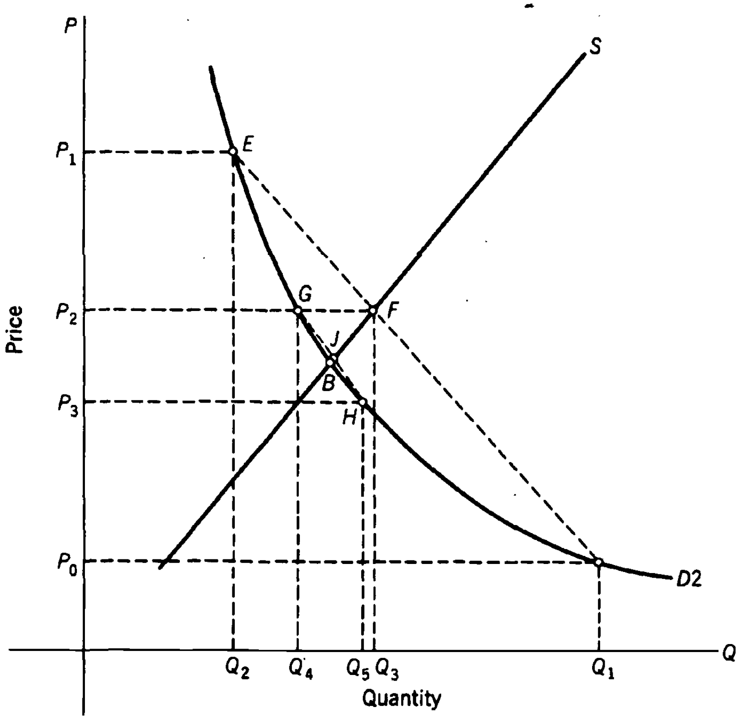
$$Q = Q_2 - \frac{Q_1 - Q_2}{P_1 - P_0}(P_1 - P) \tag{4}$$

Equations (3) and (4) may now be combined and solved for equilibrium price quantity point. This corresponds to point F and thus to P_2 and Q_3.

The next step checks on the closeness of the approximations of F to the true new equilibrium B. Since, by construction, F is on the supply function, the test of closeness depends on the gap between the trial solution (point F) and the demand curve. To assess this, the demand simulation model is again used. Simulating with gross price P_2 yields the point G. If the quantities at G and F are sufficiently close, the analysis is complete. If they differ by more than some prespecified amount, the iterative procedure can be continued in order to achieve greater accuracy.

The method of increasing accuracy is illustrated in Figure 2, with

FIGURE 2



notation carried over from Figure 1. If Q_3 is greater than Q_4, the curve $D2$ is convex toward the origin in the relevant range; this is the case that will now be described. If the inequality is reversed, the curve is concave and the calculations are modified accordingly. Since the equation for the supply curve is given, it is possible to solve for the price P_3 at which the quantity Q_4 would be supplied. Demand simulation then identifies the point H and the quantity Q_5 that would be demanded at price P_3. The points G and H are analogous to the old points E and C. The equation of the line joining G and H is found and used as a local approximation to the demand curve. The intersection of this approximation and the supply curve identifies the point J, which is much closer to B than F was. This is verified and the accuracy evaluated by a further simulation at the price corresponding to J.

This iterative simulation method of finding the post-NHI market equilibrium is easily extended to separate markets for hospital care and medical services. The procedure begins by using the demand simulation model described in Section 1 to calculate the hospital and medical expenditures, E_h and E_m, that would prevail with NHI if the prices remained unchanged.[4] Since the pre-NHI prices are normalized at 1, these expenditures are also equivalent to quantities. This yields quantities that may be denoted QH_1 and QM_1, corresponding to point C of Figure 1.

The two aggregate constant elasticity supply functions are:

(5)
$$\ln QH = \ln QH_0 + \gamma_H \ln PH$$
$$\ln QM = \ln QM_0 + \gamma_M \ln PM$$

where QH_0 and QM_0 are the aggregate quantities before NHI, PH and PM are prices of hospital and medical care ($PH_0 = PM_0 = 1$), and γ_H and γ_M are supply elasticities. Substituting the values QH_1 and QM_1 into QH and QM yields the prices PH_1 and PM_1 at which these quantities would be supplied; these prices correspond to point D of Figure 1. The demand simulation model is then repeated using PH_1 and PM_1 and the deductibles, coinsurance rates, and values of MAX provided for by the NHI plan. The aggregate quantities demanded at these prices, QH_2 and QM_2, correspond to point E of Figure 2.

Although the coordinates of two points like C and E were sufficient to define a linear demand equation when hospital and medical services were not distinguished, a third point is now needed if the cross-price effects are to be taken into account. A new price corresponding to the average of PH_1 and the pre-NHI price ($PH_0 = 1$) is selected for hospital care. A similar value is selected for medical care. At these prices (PH_2 and PM_2), a new set of quantities

(QH_3 and QM_3) is obtained by simulation. The coordinates of the three quantity-price points ($QH_1, QM_1, 1, 1$), (QH_2, QM_2, PH_1, PM_1), and (QH_3, QM_3, PH_2, PM_2) can be used to evaluate the parameters of the two linear equations that approximate the demand function in the presence of NHI. More specifically, we take the aggregate demand equations to be:

(6) $$QH = a_1 + a_2 PH + a_3 PM$$

and

$$QM = b_1 + b_2 PH + b_3 PM$$

The three parameters of each equation can be obtained by substituting the values of the variables for each of the three price quantity points. The two supply equations (5) and the two demand equations (6) are then solved simultaneously to obtain trial equilibrium values (QH_4, QM_4, PH_3, PM_3) that correspond to point F of Figure 1. This set of values corresponds to a point on the supply functions. To check the accuracy of the approximation to the demand, the demand simulation is recomputed for prices PH_3 and PM_3. This yields a point that is analogous to G of Figure 1. If this is not sufficiently accurate, a further iteration is computed as described above for Figure 2.

3. AN ANALYSIS OF TWO NHI PLANS

This section calculates the equilibrium quantities and prices associated with two alternative NHI options. The analysis emphasizes the substantial sensitivity of the results to the elasticities of supply as well as to the demand parameters.

The first NHI plan (NHI-1) has low annual deductibles of \$50 for hospital care and \$50 for medical services and very low coinsurance rates of 10 per cent for both types of services. There is no limit, however, to the family's maximum net expenditure (MAX = ∞). These characteristics of the NHI plan are the same for all income levels and all family demographic compositions. The second NHI plan (NHI-2) has the same \$50 deductibles but a coinsurance rate of 20 per cent for both types of expenditure. Each family's net expenditure is limited to 10 per cent of family income. This is a substantial reduction in risk, especially for lower- and middle-income families.

The individual demand equations are simplified by assuming no cross-price elasticities (i.e., $\alpha_m = \beta_h = 0$ in the demand equations of Section 1). Although hospital services and other services are substitutes for some purposes, they are also complements in other contexts. The assumption of zero cross-elasticity may therefore be a reasonable starting point for a preliminary analysis.[5] Two different sets of price elasticities have been used. The "moderate" price elasticities assume that the own-price of hospital care (α_h) is 0.5 and the own-price elasticity of medical care (β_m) is 0.4. The "low" elasticity pair assumes that $\alpha_h = 0.25$ and $\beta_m = 0.20$. The relative nonmonetary cost parameter, λ, is assumed to be 0.10 in all the calculations. The same baseline distributions are used in all income classes; this implicitly assumes that there is zero income elasticity of demand.

Table 1 analyzes the first NHI plan. Column 1 presents baseline figures for 1970 with no NHI program. Total expenditure on covered services for the population under age 65 is \$23.3 billion. This corresponds to hospital services of \$12.9 billion and other medical care expenditure of \$10.4 billion.[6] Because the pre-NHI prices are normalized to be unity, these expenditures can also serve as measures of quantities for comparison with the post-NHI quantities.

Column 2 shows the impact that NHI-1 would have if prices remained unchanged. Total national expenditures on the covered health services would increase 38 per cent to \$32.2 billion. The quantity of hospital services rises 38 per cent to 17.8, and the quantity of other medical services rises 38 per cent to 14.4. The effective coinsurance rates are 17 per cent for hospital care and 25 per cent for other medical care. The cost to the government is therefore \$25.6 billion.

An analysis with supply elasticities of 0.8 for both hospital care and other services is presented in column 3. The results are substantially different from the "pure demand case" of column 2. Prices rise by approximately 30 per cent. Total expenditure of \$37 billion is 15 per cent higher than the estimate that ignored the endogenous price increase. The higher gross prices also imply a smaller increase in the quantities of care. The quantity of hospital services is 15.7, indicating that the rise from the pre-NHI value of 12.9 is only 22 per cent, or about half of the estimated increase when the price response was ignored. The comparison is similar for medical services. It is interesting that national spending increases \$13.7 billion, but the extra volume of services is worth only \$5.3 billion at the pre-NHI prices. Column 4 shows comparable

TABLE 1 Effects of NHI Plan 1

	Baseline Simulation (no NHI) (1)	"Moderate" Demand Elasticities: Supply ∞ (2)	Elasticities 0.8 (3)	Elasticities 0.2 (4)	Price Response Model (5)	"Low" Demand Elasticities: Supply ∞ (6)	Elasticities 0.8 (7)	Elasticities 0.2 (8)	Price Response Model (9)
National health expenditure[a]	23.3	32.2	37.0	42.3	38.1	27.2	30.5	35.9	35.0
Cost to government[a]	N.A.	25.7	29.9	34.5	30.8	21.0	23.9	28.6	27.8
Quantity of hospital care[a]	12.9	17.8	15.7	14.1	14.3	15.0	14.5	13.8	13.7
Quantity of medical care[a]	10.4	14.4	12.9	11.6	10.4	12.2	11.8	11.3	10.4
Price of hospital care	1.00	1.00	1.28	1.58	1.55	1.00	1.15	1.40	1.44
Price of medical care	1.00	1.00	1.31	1.72	1.54	1.00	1.17	1.48	1.47
Effective coinsurance rate, hospital care	N.A.	0.17	0.16	0.16	0.16	0.19	0.19	0.18	0.18
Effective coinsurance rate, medical care	N.A.	0.25	0.23	0.21	0.23	0.27	0.25	0.23	0.24

NOTE: The "moderate" demand elasticities are $\alpha_h = 0.5$ and $\beta_m = 0.4$; the "low" elasticities are $\alpha_h = 0.25$ and $\beta_m = 0.20$.
[a]Billions of 1970 dollars.

results for lower supply elasticities of $\gamma_H = \gamma_M = 0.2$. The total spending increases are greater but the quantity increases are smaller.

Columns 6, 7, and 8 present the same comparison of supply elasticities but with a lower pair of demand elasticities.[7] Although the effects of the NHI plan are now smaller, the implications of different supply elasticities are still very important. It is clear, moreover, that plausible differences in supply elasticities are at least as important as a source of uncertainty in total expenditure and in the cost to the government as the plausible differences in demand elasticities.

Table 2 presents a corresponding analysis for the second national health insurance option. The higher coinsurance rates ($C_H = C_M = 0.2$) decrease total expenditure but the maximum out-of-pocket expenditure of 10 per cent of income increases total expenditure. If prices remained constant, the net effect of these two changes in the NHI program would be a small reduction in total cost; with the moderate demand elasticities, total expenditure is \$31.9 billion under plan NHI-2 in contrast to \$32.2 billion under plan NHI-1. The effect of less elastic supply is to increase prices under both plans. The higher price substantially increases the probability that each family's expenditure will exceed 10 per cent of income. At this point, the MAX limit becomes effective and the coinsurance rate ends. Although demand is limited by the non-monetary price $[\lambda/(1 + \lambda)]$, there is a substantial increase in demand. The result is that NHI-2 becomes more expensive than NHI-1. The government's cost is generally lower under the second plan because the higher coinsurance rate on relatively small expenditures more than outweighs the extra cost of providing complete protection for expenditures above 10 per cent. Only if the supply elasticities are very low would prices rise enough to reverse this situation and make NHI-2 more expensive; this happens here with $\gamma_H = \gamma_M = 0.2$. Notice that with the lower demand elasticities the rise in expenditure is always sufficiently small so that NHI-2 entails lower total expenditure and lower cost to the government.

The supply elasticity can also affect the distributional impact of a national health insurance plan. Because the current simulations assume a zero income elasticity of demand for health services, the first NHI plan provides the same expected benefits at all income levels to families with any fixed demographic composition. NHI-2, on the other hand, limits each family's net expenditure to no more than 10 per cent of family income. For low-income families, this is a substantial reduction in the net price of health services, whereas for

TABLE 2 Effects of NHI Plan 2

		"Moderate" Demand Elasticities				"Low" Demand Elasticities			
	Baseline Simulation (no NHI)	Supply ∞	Elasticities 0.8	Elasticities 0.2	Price Response Model	Supply ∞	Elasticities 0.8	Elasticities 0.2	Price Response Model
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
National health expenditure[a]	23.3	31.9	37.3	43.8	36.5	26.5	29.5	34.9	31.7
Cost to government[a]	N.A.	24.2	28.9	34.6	28.3	19.0	21.4	25.9	23.2
Quantity of hospital care[a]	12.9	18.4	16.3	14.5	15.4	14.8	14.4	13.8	14.1
Quantity of medical care[a]	10.4	13.4	12.4	11.4	10.4	11.7	11.5	11.1	10.4
Price of hospital care	1.00	1.00	1.34	1.77	1.45	1.00	1.15	1.42	1.28
Price of medical care	1.00	1.00	1.25	1.59	1.37	1.00	1.13	1.37	1.31
Effective coinsurance rate, hospital care	N.A.	0.19	0.18	0.17	0.18	0.24	0.23	0.22	0.23
Effective coinsurance rate, medical care	N.A.	0.31	0.29	0.27	0.30	0.34	0.33	0.31	0.32

NOTE: The "moderate" demand elasticities are $\alpha_n = 0.50$ and $\beta_m = 0.40$; the "low" demand elasticities are $\alpha_n = 0.25$ and $\beta_m = 0.20$.
[a]Billions of 1970 dollars.

higher-income families the effect on price is much smaller. The result is a more substantial increase in spending at lower incomes.

These distributional effects are presented in Table 3. The analysis refers to the second NHI plan and to the moderate price elasticities. Columns 1 through 3 describe the impact on families of two adults and two children in the case in which prices are unchanged—i.e., infinitely elastic supplies of hospital and medical services at the original prices. Column 1 presents the average net benefit received by families at each income level—i.e., the average cost to the government as insurer. These net benefits fall rapidly for the first few income classes and then fall more slowly, reflecting the highly skewed distribution of health spending. Similarly, the average direct out-of-pocket payments by the family (column 2) increase rapidly for the first few income classes and then more slowly. The total quantity of care received is, with prices fixed at unity, the sum of the net benefits and direct payments; these quantities are shown in column 3. The quantity of care consumed also falls rather rapidly at first and then more slowly.

All three columns show that substantial progression is introduced by the single feature of a 10 per cent maximum limit on direct payments, even when there is a relatively low 20 per cent coinsurance rate. It is convenient to have a summary measure of the distributional impact and a method of combining the benefits (or payments or quantities) in different income classes into a single measure that reflects a constant value judgement about distributional equity. The "uniformly distributed dollar" (UDD) measure is useful for this purpose. For example, the UDD value of benefits is a weighted sum of the average benefits per family in each income class:

$$B_{UDD} = \frac{\sum_i B_i W_i N_i}{\sum_i W_i N_i} \tag{7}$$

where B_i is the benefit per family in income class i, W_i is the weight given to a marginal dollar of a family in income class i, and N_i is the number of families in class i. Notice that if $B_i = 1$, $B_{UDD} = 1$; thus one unit in the B_{UDD} measure is the social value of \$1.00 given to each family—i.e., the social value of a uniformly distributed dollar. It is convenient to relate the W_i's to income by a simple functional relation. The constant elasticity function $W_i = Y_i^{-\alpha}$ is both familiar and convenient. For $\alpha = 1$ it implies that the weight given to a marginal dollar of income varies inversely with the income of the recipient family. The higher the value of α, the more egalitarian the implied preferences.

TABLE 3 Distributional Aspects of NHI Plan 2: Supply Response Model

Income Class	2 Adults and 2 Children						All Families		
	$\gamma_H = \gamma_M = \infty$			$\gamma_H = \gamma_M = 0.8$			$\gamma_H = \gamma_M = \infty$		
	Net Benefits (1)	Direct Payments (2)	Quantity (3)	Net Benefits (4)	Direct Payments (5)	Quantity (6)	Net Benefits (7)	Direct Payments (8)	Quantity (9)
<$ 2,000	$823	$ 87	910	$968	$ 90	814	$201	$ 28	229
$ 2,000–	700	132	832	830	140	745	202	42	245
$ 3,000–	647	143	790	770	154	709	261	62	323
$ 4,000–	620	151	771	735	163	689	297	77	374
$ 5,000–	589	157	746	709	170	675	346	95	441
$ 6,000–	564	162	726	678	176	656	365	107	473
$ 7,000–	536	165	701	651	180	638	387	120	507
$ 8,000–	513	169	682	617	185	617	409	134	543
$10,000–	487	172	659	586	189	596	424	148	572
$12,000–	465	175	640	551	192	576	429	158	586
$15,000–	434	178	612	520	197	552	424	168	592
$25,000+	413	179	592	484	199	527	395	167	562
UDD Values									
$\alpha = 0.0$	497	169	667	596	186	602	351	112	462
$\alpha = 0.5$	522	165	687	626	180	620	307	86	393
$\alpha = 1.0$	562	156	718	672	170	647	262	62	324
$\alpha = 1.5$	623	141	765	742	153	688	231	45	276
$\alpha = 2.0$	696	122	819	825	131	735	214	36	250

NOTE: All calculations use "moderate" demand elasticities, $\alpha_n = 0.5$ and $\beta_m = 0.4$.

Table 3 shows the B_{UDD} values corresponding to values of α between zero and 2. The value of $\alpha = 0$ corresponds to the simple average of benefits with no weighting for distribution. The average benefit (cost to the government) per family with two adults and two children is thus \$497. For someone whose distributional preferences correspond to $\alpha = 1$, these benefits are equivalent to \$562 distributed uniformly to all such families. With more egalitarian preferences ($\alpha = 2$), the benefits are equivalent to \$696 distributed uniformly. Conversely, the average value of direct out-of-pocket payments is \$169 per family of two adults and two children, but this amount is equivalent to a smaller uniformly distributed payment of \$156 for $\alpha = 1$ and \$122 for $\alpha = 2$. Finally, the average quantity of services is \$667 per family. Applying the same UDD evaluation to these benefits implies that for $\alpha = 1$, they are equivalent to a constant \$718 per family.

The effect of introducing supply elasticities of 0.8 for hospital and medical services is to increase prices and therefore expenditures at all income levels. The equilibrium quantities are now smaller than before. Benefits rise by about 18 per cent in the lower-income classes and about 14 per cent in the higher-income classes. Direct costs rise by about 7 per cent in the lower-income classes and about 11 per cent in the higher-income classes. Thus in both of these ways the NHI-2 plan is slightly more redistributive when the supply response is explicitly recognized. But the relatively greater direct payments by higher-income families just about offset the relatively lower benefits from the insurer and make the proportional change in the quantity of services approximately equal at all income levels; the quantities shown in column 6 are almost exactly 90 per cent of the quantities in column 3.

Although the NHI-2 plan is very progressive with respect to income when attention is focused on families with a single demographic composition, this characteristic is disguised when all family types are combined. Average family size increases with family income; there are fewer single-person families and larger average numbers of children. Columns 7 through 9 show that this has striking effects on the distribution of average benefits, direct payments, and quantities of care. Average benefits rise with income until \$15,000 and then fall only slightly. Average direct payments rise much more sharply with income. The net effect is that quantity increases with family income up to \$25,000, despite the income-related limit on direct payments. It is clear from this comparison that it is important to take demographic structure into account in evaluating the distributional impact of alternative NHI plans.

4. A PRICE RESPONSE MODEL

Neither hospitals nor physicians are like the typical economic agents to which the traditional theory of price and quantity determination applies. These differences—the nonprofit nature of hospitals, the special expertise of physicians, and the physicians' professional interest—may not be enough to vitiate the applicability of the traditional theory in sections 2 and 3. Nevertheless, it seems useful to provide an alternative response model that contains special features of the health care sector. The "price response model" presented in this section incorporates ideas about the markets for physicians' services and hospital care that were previously developed by Feldstein (1970, 1971a, 1971b, 1974). An important characteristic of this alternative model is that the price-quantity equilibrium need not be market clearing; excess demand and nonprice rationing may prevail in equilibrium.

Consider first the model of physicians' behavior. The price response model specifies that the effect of NHI is to raise the gross price of physicians' services by an amount that depends on the increased insurance coverage of physicians' services. More complete insurance raises the physicians' price not only because it increases demand, but also because physicians take into account the financial impact of their fees on their patients. Moreover, physicians may seek to maintain excess demand in order to have the opportunity to select the types of patients and diagnoses that they like to treat; an increase in insurance permits gross prices to be raised without reducing the desired degree of excess demand.[8]

More specifically, the change from the pre-NHI price, PM_0, to the post-NHI price, PM_1, is given by the function:

$$\frac{PM_1}{PM_0} = \left[\frac{NPM_1}{NPM_0}\right]^{-\delta_M} \tag{8}$$

where NPM_0 is the average *net* price of physicians' services prevailing before NHI (i.e., the product of PM_0 and the effective coinsurance rate before NHI) and NPM_1 is the average *net* price that would prevail after NHI if physicians did not alter their gross price.[9] The computational procedure for deriving PM_1 is straightforward. The demand simulation model is used to calculate NPM_0 as the ratio of aggregate *direct* patient expenditure on physicians' services to aggregate *total* expenditure on those services. The insurance coverage is then changed to the NHI plan and the calculation is repeated to obtain NPM_1. Since NPM_1 depends on the gross price PM_0, no iterative procedure is necessary. Applying Equation (8) then yields the new price that prevails under NHI.

The supply function of the physician now indicates the *desired* supply at each price. The same constant elasticity function will be used:

(9) $\ln QM_1 = \ln QM_0 + \gamma_M \ln PM_1$

If the aggregate demand at price PM_1 (and corresponding hospital price PH_1) is less than or equal to the desired supply QM_1, the equilibrium quantity is "demand-determined"; i.e., each family gets the quantity of physicians' services that it wants at the new prevailing price. If, however, as is probably more likely, the aggregate demand exceeds supply, the new equilibrium is "supply-determined." Each family obtains only some fraction of the services that it would like to purchase with the new prices and insurance coverage. In the absence of better information about nonprice rationing, the current model specifies that each family receives the same fraction of the quantity that it demands regardless of income, demographic composition, or desired expenditure. More specifically, the "rationing constant" for physicians' services is defined as:

(10) $$RM = \frac{QM_1}{QMD}$$

where QMD is the aggregate quantity of physicians' services demanded at prices PM_1 and PH_1 under the NHI plan, and QM_1 is the desired aggregate supply defined in Equation (9). Each individual family then obtains RM times the quantity that it demands according to the basic demand equations in Section 1, with $P_h = PH_1$ and $P_m = PM_1$.

Notice that the use of nonprice rationing increases the likelihood that some families would receive less care than in the absence of NHI even if NHI improves everyone's coverage. This will clearly happen when the supply elasticity is zero but the demand elasticity is non-zero. NHI then increases demand and results in a rationing parameter RM less than 1. Unless all families' demands are increased in exactly the same proportion, the NHI would reduce the quantity of care received by some families.

Although the hospital services section of the price response model has the same formal structure as the model of physicians' services, the interpretation of this behavior is quite different. An analysis of hospitals' response to the growth of private insurance in the 1960s and to the introduction of Medicare and Medicaid suggests that hospitals respond to insurance by increasing the cost per patient day through more sophisticated care and higher staff wages. Prices rise in response to additional insurance not because of a greater unit cost of providing more of the same type of care (i.e., a rising supply curve in the traditional sense) but because hospitals produce a different product and choose to pay higher wages. To analyze this as a response to NHI, the price response model with a demand-determined equilibrium would be used. The price of hospital services after the introduction of NHI is given by

$$\frac{PH_1}{PH_0} = \left[\frac{NPH_1}{NPH_0}\right]^{-\delta_H} \tag{11}$$

where NPH_0 is the average *net* price of hospital services prevailing before NHI and NPH_1 is the average *net* price that would prevail after NHI if hospitals did not alter their gross price. The analysis in Feldstein (1971a) suggests this type of behavior with $0 < \delta_H \leq 1$ and with the actual quantity determined by household demand. The model is thus completed by using PH_1 and PM_1 to calculate each family's demand and assuming that hospitals will supply this quantity.

Alternatively, the price response model may be evaluated with a supply-determined equilibrium by using the hospital supply equation

$$\ln QH_1 = \ln QH_0 + \gamma_H \ln PH_1$$

where γ_H is the elasticity of supply. The hospital rationing parameter (RH) is then defined by an equation analogous to (10). The individual demand equations and RH are then combined to determine the allocation of the rationed hospital care.

The price response model in which the quantities are constrained by supply can also be used to examine the case in which price controls are used to limit the price rise. The prices PM_1 and PH_1 are then determined by the price control agency instead of by equations (8) and (11). The corresponding supplies are then calculated with equations (9) and (12). The individual demand simulations yield rationing parameters RM and RH and the rationed allocation of services corresponding to the NHI plan, the controlled prices, and the quantities supplied.

The two alternative NHI options discussed in Section 3 have been reanalyzed with the current price response model. More specifically, the price response parameters δ_H and δ_M are both assigned the value 0.5; a 20 per cent decrease in the effective coinsurance rate (e.g., from 0.40 to 0.32) thus raises the gross prices by approximately 10 per cent and therefore also lowers the net price by about 10 per cent. For medical services, the total supply is assumed fixed—i.e., $\gamma_M = 0$. The equilibrium quantities of medical services are therefore supply-determined. For hospital services, a positive supply elasticity ($\gamma_H = 0.5$) is assumed. The same two alternative demand specifications as in Section 3 are again investigated. With the "moderate" demand elasticities ($\alpha_n = 0.5$, $\beta_m = 0.4$), the increase in demand exceeds the increase in supply and the allocation of hospital services is also supply-determined. Only with the "low" demand elasticities ($\alpha_n = 0.25$, $\beta_m = 0.2$) is there no excess demand and a demand-determined allocation.

The aggregate implications of these price response models are shown in columns 5 and 9 of tables 1 and 2. With the moderate demand elasticities, the total cost implications are quite similar to the previous analysis with supply elasticities of 0.8 and market-clearing prices. Total national spending under NHI-1 is \$38.1 billion, in comparison to the earlier value of \$37 billion; for NHI-2, the figures are \$36.5 billion and \$37.3 billion. Estimated costs to the government are also quite similar. The underlying price and quantity changes are, however, very different; price rises are greater and quantity increases are smaller. For NHI-1, total national spending increases by \$14.8 billion to buy only \$1.4 billion worth of additional services valued at original prices (i.e., quantity increases from 23.3 billion to 24.7 billion). With NHI-2, the extra spending of \$13.2 billion induces only an extra quantity worth \$2.5 billion at original prices.

Although the prices and aggregate quantities are quite different under the two models, for the two sets of assumptions examined here the distributional implications are approximately the same. Table 4 shows the distributions of benefits, direct out-of-pocket payments, and quantities for families with two adults and two children under the price response model and under the supply response model with $\gamma_H = \gamma_M = 0.8$. Although benefits and direct payments are lower under the price response model, the ratios of corresponding values under the two models is approximately constant. Similarly, quantities are some 10 per cent lower at each income level. It should, of course, be stressed that this result depends on the particular assumptions made for this comparison. With a more complex insurance structure (e.g., deductibles related to income) or different elasticities of demand and supply, the two different models of provider behavior may imply quite different distributional patterns.

TABLE 4 Distributional Aspects of NHI Plan 2: Comparison of Price Response and Supply Response Models

Income Class	Price Response Model: Net Benefits (1)	$\gamma_H = 0.5$ Direct Payments (2)	$\gamma_M = 0$ Quantity (3)	Supply Response Model ($\gamma_H = \gamma_M = 0.8$): Net Benefits (4)	Direct Payments (5)	Quantity (6)
<$ 2,000	$945	$ 89	730	$968	$ 90	814
$ 2,000–	812	137	669	830	140	745
$ 3,000–	750	150	634	770	154	709
$ 4,000–	723	159	621	735	163	689
$ 5,000–	696	166	607	709	170	675
$ 6,000–	664	172	589	678	176	656
$ 7,000–	638	176	574	651	180	638
$ 8,000–	604	181	554	617	185	617
$10,000–	572	185	534	586	189	596
$12,000–	543	188	516	551	192	576
$15,000–	506	193	494	520	197	552
$25,000+	470	195	470	484	199	527
UDD Variables						
$\alpha = 0.0$	582	182	539	596	186	602
$\alpha = 0.5$	611	176	556	626	180	620
$\alpha = 1.0$	657	166	581	672	170	647
$\alpha = 1.5$	725	149	617	742	153	688
$\alpha = 2.0$	806	128	659	825	131	735

NOTE: All calculations use moderate demand elasticities and refer to families of two adults and two children.

5. CONCLUSION

The primary purpose of this paper has been to emphasize that any analysis of the effects of alternative national health insurance plans should take into account the effect of insurance on the prices and supplies of health services. An operational method was presented for combining a stochastic microsimulational model of household demand with aggregate supply and price determination equations.

The supply models used in this analysis are preliminary and can only be regarded as illustrative. Neither the traditional supply model in Section 2 nor the price response model in Section 4 can be eliminated as completely inconsistent with the data. More econometric research is therefore required to provide conditional estimates of the parameters of both models. We hope that the current evidence of the importance of these parameters will encourage others to continue work on these empirical issues.

NOTES

1. For simplicity, the actual calculations assume $E_m = \hat{E}_m$ in both subcases.
2. Although the independence assumption seems strong, there are several countervailing forces that may produce such independence. Some preliminary comparisons of convoluted "synthetic" family distributions and actual family distributions supported the assumption of independence.
3. This assumption for "typical" coverage is based on Reed (1969) and information supplied by the Department of Health, Education, and Welfare.
4. Notice that the simulation model in Section 1 permits specifying a separate set of deductibles, coinsurance rates, and MAX value for each of the 192 family demographic types and income classes.
5. Davis and Russel (1972) did provide some evidence of positive cross-price elasticities, but medical services in their study were limited to hospital outpatient care.
6. Actual costs in 1970 were $13.2 billion for hospital services and $10.1 billion for medical services. The calibration method described above does not yield these exact figures because of the nonlinearity of the insurance schedules. These dollar amounts refer to persons under age 65; see Cooper *et al.* (1973).
7. Columns 5 and 9 present results that will be discussed in Section 4.
8. These ideas are developed more fully in Feldstein (1970, 1974).
9. This model of price response is clearly a simplification that is used because no more specific hypothesis seems either theoretically or empirically superior.

REFERENCES

1. Acton, Jan Paul, "Demand for Health Care Among the Urban Poor, with Special Emphasis on the Role of Time," R-1151-OEO, The Rand Corporation, October 1972.
2. Cooper, Barbara *et al.*, *Medical Care Expenditures, Prices and Costs: Background Book* (Washington, D.C.: Department of Health, Education, and Welfare, 1973).
3. Davis, K., and L. Russel, "The Substitution of Hospital Outpatient Care for Inpatient Care," *Review of Economics and Statistics*, 54 (May 1972), pp. 109–120.
4. Feldstein, M. S., "The Rising Price of Physicians' Services," *Review of Economics and Statistics*, 52 (May 1970), pp. 121–133.

5. ______,"Hospital Cost Inflation: A Study of Nonprofit Price Dynamics," *American Economic Review*, 61 (December 1971a), pp. 853–872.
6. ______, *The Rising Cost of Hospital Care*, published for the National Center for Health Services Research and Development, Department of Health, Education, and Welfare (Washington, D.C.: Information Resources Press, 1971b).
7. ______, "Econometric Studies of Health Economics," in M. Intriligator and D. Kendrick (editors), *Frontiers of Quantitative Economics*, II (Amsterdam: North-Holland, 1974).
8. ______, B. Friedman, and H. Luft, "Distributional Aspects of National Health Insurance Benefits and Finance," *National Tax Journal*, 25 (December 1972), pp. 497–510.
9. Phelps, Charles E., and Joseph P. Newhouse, "Effect of Coinsurance: A Multivariate Analysis," *Social Security Bulletin*, 35 (June 1972), pp. 20–44.
10. Reed, Louis, *Private Health Insurance in the United States, 1968; Coverage and Benefits* (Washington, D.C.: Social Security Administration, Office of Research and Statistics, 1969).

12 | COMMENTS

Barry R. Chiswick

Council of Economic Advisers and
National Bureau of Economic Research

Within the last few years there has been considerable interest in the adoption of national health insurance (NHI) in the U.S. Although the three letters NHI may have wide appeal, as with apple pie, there is substantial disagreement over which recipe is best. Legislative proposals range from the Long-Ribicoff mandatory catastrophic insurance to the virtual cradle-to-grave-universal-zero-out-of-pocket cost coverage of the Health Security Bill.

The various proposals, including the one to "do nothing," have elicited much discussion. Everyone seems to agree that NHI would increase the amount of services demanded and the price of a unit of service. Concern for the increased share of GNP devoted to the medical sector and the change in the distribution of income arising from NHI has led to a variety of collateral proposals to control units purchased and prices charged.

Part of the uncertainty about NHI is attributable to the difficulty economists

---

The views expressed in these comments, based on the version of the paper presented at the conference, are solely those of the author.

have in measuring magnitudes. We are quite good at measuring *directions* of change, but quite poor in measuring magnitudes. This is not unique to the health field—for example, it is now well established that a 10 per cent increase in the minimum wage will decrease the employment of teenagers, but by how much is far less certain. However, research on the impact of NHI has encountered additional handicaps. The data on current utilization are inadequate. Also, the economist's tools are designed for the analysis of marginal changes in a partial equilibrium model, whereas NHI would have such profound widespread effects that a model that allows for a variety of long-run interactions may be the best approximation.

These points indicate that measuring the impact of NHI is not an easy research topic. For some of the same reasons, however, the quantification of the impact is of vital concern. We know that NHI will influence the share of GNP devoted to the health sector. The magnitude of the change and what we get for it in terms of improved health will influence our view of the wisdom of NHI. A wrong decision with respect to NHI will be very costly. And, if an NHI is adopted, the political costs of reversing our policy may be substantial as new interest groups develop, even if it turns out that for the country as a whole NHI is economically inefficient.[1] In a world of uncertainty, the larger the cost of enduring a wrong decision and the greater the cost of reversing a decision, the greater the amount of resources that should be devoted to finding the correct decision, and the more we should appear to be risk averse.[2]

It is for these reasons that the Feldstein-Friedman paper, and the work of others, on estimating the impact of NHI is of considerable importance. Feldstein and Friedman correctly recognize that NHI has a direct impact on the demand for medical care and that this change in demand induces a supply response. It is the combined effects of the change in the demand curve and the movement along the supply curve for medical care that generate the change in the quantity and price of the units of medical care provided.

Feldstein and Friedman use a microsimulation model of household demand and an aggregate model of supply to predict the price and quantity of medical care services that would arise from alternative models of NHI. They prefer a microsimulation model because an "aggregate specification of demand behavior cannot capture the subtle differences in the response of demand to different types of insurance policies." There is truth in this statement.

An aggregate specification, however, may be better able to account for the effects, if any, of interactions in demand among individuals.[3] Although interacting individual demand may not be important for marginal changes, it may be important for a large nationwide change in the medical care system brought about by NHI. Also, the microsimulation model as formulated by Feldstein and Friedman—and any microsimulation model—embodies a variety of built-in behavioral assumptions. The authors assume, for example, that each family has the same price elasticity of demand, the income elasticity of demand for medical care is zero, the cross-price elasticity of demand for hospital and out-of-hospital care is zero, individual demand curves are independent of one another, etc. These are probably reasonably good assumptions. Unfortunately, by making these assumptions they cannot capture all the subtle differences in demand.

A priori, it is not clear that an aggregate model of demand would necessarily be a worse predictor than a microsimulation model. It would be interesting to know the extent to which the two procedures yield different estimates. If they generate essentially the same estimates, the aggregate model has the advantage of simplicity but the microsimulation model would provide predicted use rates for specific demographic groups. If the two procedures generate widely different estimates, the cause of the difference should be investigated. Unfortunately, we may not be able to determine which is a better predictor until after we have had some experience with NHI.

A preliminary answer could perhaps come from an analysis of our two current targeted NHI programs—Medicare and Medicaid. Using pre-1966 data on the utilization of health services by the aged, the poor, and all others, which of the two procedures more accurately predicts the current utilization pattern?

In their microsimulation model of household demand, Feldstein and Friedman estimate expenditures for hospital and out-of-hospital medical services by assuming a constant (gross price) elasticity of demand equation in the absence of insurance. By introducing modifications for deductibles and coinsurance, including zero per cent coinsurance beyond some level of expenditure, they are able to estimate the effect on demand of alternative models of NHI.

They explicitly include the parameter λ, which measures the nonmoney cost of medical care. The units for λ are not spelled out and I am troubled by the manner in which it is included. In equations (1) and (2), for example, for a zero per cent coinsurance, the market price, P_h, should have no effect on the demand for hospital care, yet the price variable is $\lambda P_h/(1 + \lambda)$. Since λ is (arbitrarily) assumed to be 0.1 in the computations, the price to the household is $0.09 P_h$, which is a function of P_h. It is not clear why the nonmoney cost to the household would be proportional to the market price.[4]

Unfortunately, we know very little about the role of nonmarket costs in the medical care field, and in others. With the growth of private and public insurance for out-of-pocket expenditures and the rise in the value of time, there has presumably been a substitution toward goods-intensive forms of medical care that economize on time. This suggests that the relative nonmoney cost parameter λ may not be constant but may decline in response to the expansion of insurance coverage for out-of-pocket expenditures. It would be useful if the microsimulation model allowed for this effect. It would also be interesting to see the effect of alternative specifications of the magnitude of nonmoney costs.

From economic theory we know that demand and supply curves are less elastic in the short run than in the long run. Using the very low demand and supply elasticities ($\alpha_h = 0.25$, $\beta_m = 0.20$, $\gamma = 0.2$), the simulation of the two NHI models predicts a modest 7 per cent increase in the quantity of hospital and medical care. However, prices increase by at least 40 per cent. For the long run we can expect a very large supply elasticity. Using an infinite supply elasticity and the Feldstein-Friedman "moderate" demand elasticities ($\alpha_h = 0.5$, $\beta_m = 0.4$), the quantity of care provided increases by approximately 37 per cent in their simulation.

The long-run estimate of the increase in utilization is likely to be biased downward. With the subsidization of direct costs but not of time costs, we would expect a substitution of goods for time, and hence a decline in the parameter λ. By holding λ constant, Feldstein and Friedman ignore what may be an increasingly important source of increased demand for medical services.

Feldstein and Friedman present simulations of the distributional impact among income groups. They assume a zero income elasticity of demand, the same price elasticity of demand for all income groups, that λ equals 0.1 for all income groups, etc. Although these simplifying assumptions may be adequate if we are interested only in an aggregate simulation, they are clearly highly suspect if we focus on different income classes.

For example, the nonmonetary cost of a day of hospitalization is likely to be substantially higher for a professional than for a skilled craft worker, and the latter's time cost is likely to be higher than that of a domestic day worker. Hence, if λ falls with lower-income levels, NHI results in a larger relative decline in price for lower-income than for higher-income groups. Then, by assuming a constant λ, Feldstein and Friedman underestimate the increase in the utilization of services by low-income groups relative to higher-income groups. Thus, *ceteris paribus*, NHI would be more progressive in redistributing medical service to lower-income groups than would be indicated by Feldstein and Friedman's Table 3.

Feldstein and Friedman have made an important contribution in this paper by the explicit incorporation of nonmoney costs into the demand for medical care.

NOTES

1. The oil depletion allowance was but one example of an apparently politically irreversible policy.
2. K. Arrow and A. C. Fisher, "Environmental Preservation, Uncertainty and Irreversibility," *Quarterly Journal of Economics* (May 1974), pp. 312–320.
3. These interactions may be attributable to local styles of medicine, including bandwagon effects, and to the level and spread of information.
4. A positive nonmoney price, however, does play an important role, as indicated by the finite demand with "free" (zero per cent coinsurance) medical care. For the aged and the poor, an important component of the nonmoney price of medical care—the value of time—is very low. Does this explain the very large response to the introduction of Medicaid and Medicare?

Michael D. Intriligator

University of California, Los Angeles and Human Resources Research Center, University of Southern California

In their paper, Feldstein and Friedman (FF hereafter) use a microsimulation model of the health care system in order to study the effects of alternative national health insurance (NHI) proposals. These comments are organized into three sections. The first discusses the use of a microeconometric model as a tool for policy evaluation. The second summarizes the FF microeconometric model and compares it to the Human Resources Research Center (HRRC) microeconometric model of the health care system.[1] The third discusses the major conclusions of the FF and HRRC models with regard to NHI.

1. THE USE OF A MICROECONOMETRIC MODEL FOR POLICY EVALUATION

Figure 1 summarizes the procedure by which a microeconometric model could be used for policy evaluation in a wide variety of areas, including not only health but also education, transportation, housing, economic stabilization, and many others. It also specifies how it could be used in the particular application area treated by FF.

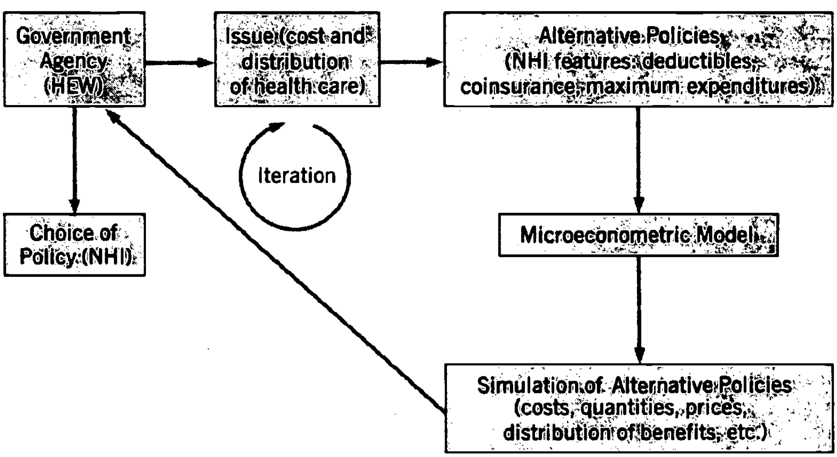


FIGURE 1 Use of a Microeconometric Model for Policy Evaluation

The procedure begins with a government agency responsible for policy in a particular area. For FF the agency is the Office of the Secretary of the Department of Health, Education, and Welfare (HEW). Responsible individuals in the agency identify a particular issue that is both of concern and subject to influence by the agency. For FF the issue is the cost and distribution of health care.

The agency individuals and/or model builders specify alternative policies that could affect the issue to be treated. In this instance, among the alternative policies are the various payment features of a NHI plan, specifically the levels of deductibles, coinsurance rates, and maximum net out-of-pocket expenditure for individuals or families.[2]

The alternative policies are analyzed and evaluated by use of a formal framework, which can take the form, as in FF, of a microeconometric model, involving interacting microsimulation and econometric models. The microsimulation portion of the structure facilitates both treatment of detailed features, such as changes in eligibility or differences in costs or benefits to different individuals, and treatment of distributional effect. The aggregate econometric portion of the structure facilitates treatment of macro interactions on markets, such as price/quantity determination in product markets and wage/employment determination in factor markets. FF use this interactive framework in relating demand, obtained from an aggregation of estimated demands via the microsimulation model, to supply, obtained from aggregate relationships, in order to determine prices.

The methodology used to evaluate the alternative policies is *simulation*. Each set of policies implies an alternative future course of relevant variables determined as a forecast of the estimated model, conditional on the particular set of policies. For FF the alternative payment features imply alternative values of total expenditure, cost to government, and quantity and price of medical and hospital care, as reported in their tables 1 and 2.

The simulations are then reported to the initiating agency, in this case the Office of the Secretary of HEW. If necessary, the entire process can be iterated by selecting related issues or other policies to derive new simulations. Eventually, the policy-makers select a particular simulation they desire and adopt the relevant set of policies implying this simulation. Of course, actually carrying out the adopted set of policies entails a complex process of legislative enactment, organizational development, budgetary appropriations, etc.

This procedure is quite general; it can be applied to a wide variety of policy-making situations. The future of policy evaluation could very well be based on this procedure, with close relationships being formed between policy-makers and model-builders. In effect, through the microeconometric model, the model-builders will be providing the policy-makers with a "wind tunnel" in which alternative policy configurations can be tested before actual use. Much of existing policy has been based on only the sketchiest idea about its immediate and primary effects and virtually no information concerning its delayed and secondary effects. The procedure outlined in Figure 1 could provide government agencies with needed guidance on important policy issues.

2. COMPARISON OF THE FF MICROSIMULATION AND HRRC MICROECONOMETRIC MODELS

The FF and HRRC models are two different microeconometric models of the health care system, both of which have been used for policy evaluation. A comparison of some of their more important structural components is presented in Table 1.

TABLE 1 Comparison of the Feldstein-Friedman (FF) Micro-Simulation Model and the Yett-Drabek-Intriligator-Kimbell (HRRC) Microeconometric Model

	Feldstein-Friedman (FF)	Human Resources Research Center (HRRC)
Population	Families: 192 types 16 family compositions 12 income categories	Individuals 9 age 2 sex 2 race 3 income Nurses 6 age Physicians 11 age 14 specialty 5 professional activity 2 place of medical school training, domestic or foreign
Demand	Constant elasticity: $D = D_0 p^{-\epsilon}$ D_0: base quantity; p = net price; ϵ: elasticity, assumed at "moderate" or "low" levels	Constant elasticity: $D = D_0 p^{-\epsilon}$ D_0: base quantity; p = net price; ϵ: elasticity, estimated for physicians and length of hospital stay
Supply	Constant elasticity: $S = S_0 p^{\gamma}$ S_0: base quantity γ: elasticity, assumed at alternative levels: ∞, 0.8 or 0.2	Quantity supplied determined from 1. Numbers of physicians, nurses, and hospitals 2. Demand for non-physician labor 3. Production functions
Price Response	Either market clearing: P_t) or price response model: $\frac{P_{t+1}}{P_t} = \left(\frac{NP_{t+1}}{Np_t}\right)^{1-\delta}$ where N_p = net price	Adjustment of prices based on trend and excess demand: $P_{t+1} = \alpha P_t + \beta(D_t - S_t)$
Data Base	Over 300,000 federal government employees and dependents enrolled in Aetna "High Option" coverage in 1970	Over 130,000 responses from the Health Interview Survey, conducted by the U.S. National Center for Health Statistics in 1967.

An essential feature of a microsimulation model is a population (or populations) being simulated. In the FF study the population consists of families of 192 types, determined from family composition and income categories. These families determine demand for health services. In the HRRC Microeconometric Model there are three populations being simulated.[3] The first is a population of individuals demanding health services. This population is described by age, sex, race, and income characteristics. The second population is one of physicians supplying outpatient services. This population is described by age, specialty, professional activity, and training (domestic- or foreign-trained medical school graduates). The third population includes nurses assisting in the provision of outpatient and inpatient services. This population is described by age.

The demand specification of both models is of the constant elasticity type in which a demand curve is passed through a particular base quantity-price point by assuming that the price elasticity of demand is constant. In this instance price is net price—price times the coinsurance rate.[4] The base quantities are derived from historical data. In FF the elasticities of demand for physician and hospital services are assumed at two different levels, "moderate" or "low." In the HRRC study the demand elasticities are estimated through regression analysis and they differ, for example, for different specialties for outpatient services. Also, the demand for hospital services is treated somewhat differently in the HRRC study than in the FF study. Instead of estimating a single demand function for hospital services, separate demand functions are estimated for hospital admission and length of hospital stay.

The supply specification is radically different in the two models. In the FF model a supply function analogous to the demand function is postulated and is assumed to exhibit constant elasticity. Alternative assumptions are made about the value of this elasticity: ∞, 0.8, or 0.2. In the HRRC model, on the other hand, the quantities supplied of inpatient or outpatient services are determined from estimated production functions, using factor inputs that are determined both from input demand functions, dependent on wages, and from the simulated population of physicians.[5] Thus, supply in the HRRC model depends on simulated populations—of physicians, nurses and hospitals, factors not treated explicitly in the FF model. Explicit treatment of these factors, however, facilitates analysis of various policy initiatives, such as added training programs for physicians in medical schools, added nurse training programs, and hospital construction, that might be undertaken in conjunction with NHI. It might also be pointed out that such a supply specification must be based on the assumption of perfectly competitive markets for both inpatient and outpatient services, since it is only in such markets that supply functions of the type they treat exist. Such an assumption is both extreme and at considerable variance with previous analyses of these markets, which typically treat the relevant markets as ones of oligopoly, or monopolistic competition.

Because of the nature of their demand/supply specifications, FF do not have independent estimates of demand and supply. They therefore cannot directly treat disequilibrium phenomena in the relevant markets. This is a severe limitation given the existence of waiting lists, discussions of shortages, and rapid price increases in these markets, all of which suggest disequilibrium. FF postulate either a market-clearing model, wherein prices clear the markets, or a price response model, wherein price changes depend on changes in net price. In the HRRC model, on the other hand, wherein demand and supply are determined on the basis of fundamentally different considerations, disequilibrium phenomena can be treated directly. In this model, price in period $t+1$ is the following function of price and excess demand in period t

$$P_{t+1} = \alpha P_t + \beta(D_t - S_t)$$

Thus, prices increase according to the trend rate $\alpha - 1$, with acceleration or deceleration around the trend rate depending on whether excess demand $(D_t - S_t)$ is positive or negative, respectively. If $\alpha = 1$, the price adjustment equation reduces to the usual tatonnement model, whereas if $\beta = 0$ (or $D_t = S_t$) the price adjustment equation reduces to a simple trend. Values of α are estimated from historical time trends. Alternative β's were considered and tested for sensitivity.[6]

Turning to the data, FF utilize data concerning over 300,000 families of federal government employees and their dependents who are enrolled in the Aetna "High Option" coverage. The HRRC study utilizes data concerning over 130,000 individuals from the 1967 Health Interview Survey, conducted by the National Center for Health Statistics. Although the FF data may be more precise, more complete, and available for greater numbers of individuals than the HRRC data, there is some question about its relevance. FF extrapolate from their data to national demands, but the behavior of federal employees who choose a particularly complete package of health insurance may not be representative of the behavior of the entire population. The HIS sample is a more representative sample of the entire population.

3. CONCLUSIONS REGARDING THE TWO MICROSIMULATION MODELS

There are two major conclusions that FF draw from their model: First, they conclude that supply, particularly the price elasticity of supply, is of major importance in calculating effects of NHI on prices, quantities, cost to the government, etc. Second, they conclude that the demographic composition of the population is of major importance in evaluating distributional effects. Their conclusions, in qualitative form, are apparent from a simple supply-demand diagram wherein the effects of a NHI-induced shift in demand depend on the elasticity of the supply schedule and the extent of the shift in demand depends on the family composition of the population. The quantitative counterparts of this diagrammatic observation are summarized in the tables in the FF paper.

There are five major conclusions of the HRRC study. First, demographic information and prices are not nearly so effective in explaining demand for health services as demographic information, prices, and health conditions.[7] The health status of the population, as well as such NHI features as coinsurance, etc., determine demand for health services.

The second conclusion of the HRRC study is the importance of the distribution of the population for the composition of health services provided. Thus, a "bulge" in the number of women of child-bearing age has a significant effect on demands for the services of pediatricians and obstetricians, two of the physician specialties explicitly treated in the HRRC model.

The third conclusion of the HRRC study is the importance of the foreign medical graduates (FMG's), based on a simulation study that exploited the capability of the model to track separately physicians trained domestically and those trained abroad. The FMG's are important components of the health care system, especially for hospital staffs, and the number of FMG's is more susceptible to policy choices, especially in the short run, than the numbers of physicians trained domestically. In fact, the most critical factor influencing the supply of physicians in the short run is the net migration rate of FMG's. A more liberal policy toward FMG's could play a significant role in meeting the demands created by NHI. There are, however, indications that a less liberal policy will be pursued over the next several years. This policy could have profound effects on the supply of physician services.

The fourth conclusion of the HRRC study is the fundamental importance to NHI outcomes of the productivity of physicians and other labor inputs in producing outpatient and inpatient services. Thus, the HRRC study found, as FF did, that supply is of major importance to NHI outcomes, but it treated explicitly the determinants of supply—manpower and productivity.

The fifth major conclusion of the HRRC study is the importance of organizational factors. Changing the mix of practice settings (e.g., a shift from solo to group practices) and changing the mix of institutions (e.g., types of hospitals) can be of considerable significance in evaluating the impact of NHI.

CONCLUSION

Feldstein and Friedman and the Human Resources Research Center have developed microeconomic models of the health care system that exhibit some similarities (e.g., in the specification of demand) but differ in both certain structural components (e.g., population, supply, price adjustment) and in the nature of their conclusions. Both are pathbreaking studies, however, in the application of formal analyses to the study of policy issues, specifically the application of microeconometric models to the study of features of national health insurance.

NOTES

1. The HRRC microeconometric model of the health care system, developed by Donald E. Yett, Leonard J. Drabek, Michael D. Intriligator, and Larry J. Kimbell, was first conceptualized in "The Development of a Microsimulation Model of Health Manpower Demand and Supply." See *Proceedings and Report of Conference on a Health Manpower Simulation Model*, Vol. 1 (Washington, D.C.: Bureau of Health Manpower, Department of Health, Education, and Welfare, December 1970), pp. 9–172.

   Elements of the model and comparisons with the HRRC macroeconometric model were presented in "Health Manpower Planning: An Econometric Approach," *Health Services Research*, 7 (1972), pp. 134–147.

   Components of the model were presented in "A Microsimulation Model of the Health Care System in the United States: The Role of the Physician Services Sector," paper presented at the 5th Conference on Optimization Techniques, Rome, May 1, 1973, published in R. Conti and A. Ruberti (editors), *5th Conference on Optimization Techniques, Part II* (Heidelberg: Springer-Verlag, 1974), and in "A Microsimulation Model of the Health Care System: The Role of the Hospital Sector," paper presented at the TIMS meetings, Tel Aviv, 1973, published in *Applied Mathematics and Computation*, 1 (1975), pp. 105–130.

   An overview of the model was presented in "A Microeconometric Model of the Health Care System in the U.S.," paper presented at the Econometric Society Meetings, New York, December 29, 1973, published in *Annals of Economic and Social Measurement*, 4/3 (1975), pp. 407–433.

   The complete model will be published as *A Forecasting and Policy Simulation Model of the Health Care Sector: The HRRC Prototype Microeconometric Model* (Lexington, Massachusetts: Lexington Books, 1977).

   All publications are by the four authors listed above.
2. For another treatment of NHI, using a macroeconometric model, see D. E. Yett, L. Drabek, M. D. Intriligator, and L. J. Kimbell, "Econometric Forecasts of Health Services and Health Manpower" in M. Perlman (editor), *The Economics of Health and Medical Care* (London: Macmillan, 1974). This paper identifies and treats several different features of alternative NHI plans, including not only payment features (coinsurance, deductibles, etc.) but also training and reorganization features, such as the expansion of medical schools and the development of health maintenance organizations (HMO's), respectively.
3. In addition, there is a population of hospitals in the HRRC microeconometric model.
4. FF also appropriately treat nonmonetary costs in their specification of demand functions.
5. Production function estimates are adopted from Larry J. Kimbell and John H. Lorant, "Production Functions for Physicians' Services," paper presented at the Econometric Society Meetings, December 1972. Demand functions for aides were derived from estimates in Michael D. Intriligator and Barbara H. Kehrer, "A Simultaneous Equations Model of Allied Health Personnel Employed in Physicians' Offices," in M. Perlman (editor), *The Economics of Health and Medical Care* (London: Macmillan, 1974).
6. The β's are the only coefficients of the HRRC microeconometric model that are not explicitly estimated. They constitute fewer than 1/10 of 1 per cent of the parameters in the model.
7. Health conditions are used in the demand functions, but they are not maintained in the population simulation of the HRRC microeconometric model.

12 REPLY

Martin Feldstein and Bernard Friedman

The comments of Chiswick and Intriligator suggest to us that we did not adequately emphasize the role of *stochastic* microsimulation. Although our model and the HRRC model discussed by Intriligator are both described as "microsimulation" studies, ours alone uses a *stochastic* simulation of *individual* behavior. The HRRC model might better be called a "detailed" or "disaggregated" macrosimulation model since no allowance is made for variations in individual experience within demographic groups.

The stochastic simulation method is particularly important for studying the effects of different insurance structures. Changes in deductibles, in upper limits, and in coinsurance rates affect the mean costs and benefits of insurance in complex and nonlinear ways that depend on the *distribution* of health expenditures and not just on the mean of that distribution. We see no adequate method of comparing, say, a \$200 deductible and a \$400 deductible without a model that contains information on the proportion of expenses below \$200 and between \$200 and \$400. We therefore feel that the HRRC model cannot be used to analyze the types of policy alternatives with which we are concerned. Similarly, we do not understand Chiswick's remark that "it is not clear that an aggregate model of demand would necessarily be a worse predictor than a microsimulation model."

The two alternative specifications of supply and price response are very aggregate and very simple. Greater disaggregation would clearly be desirable if reliable parameter estimates were available.[1] We see no reason, however, to provide a stochastic model of supplies behavior.

Although we made no attempt to deal with the time path of the system's response to a change in demand, it is not true that we do not "directly treat disequilibrium phenomena in the relevant markets," as Intriligator says. In Section 4 we explicitly consider the possibilities of excess demand, describe a "supply-determined" equilibrium, and, in Equation (10), posit a model of rationing behavior.

Chiswick correctly points out that our simulations assume no cross-elasticities of demand, no income elasticity, and the same price elasticities for all individuals. Although we obviously agree with Chiswick that these are reasonable assumptions, they should be borne in mind in considering the specific empirical results. These restrictions are, of course, not inherent in the microsimulation method. Equations (1) and (2) explicitly allow for cross-elasticities of demand. The first sentence in Section 1 indicates that separate demand equations can be specified for each income and demographic group. The original simulations prepared as part of a study for the Department of Health, Education, and Welfare did use these features to examine the implications of alternative income elasticities and different sets of demand parameters.

Finally, we recognize that our treatment of the nonmonetary costs of health care was only a first attempt at a difficult problem. We are grateful to Chiswick for suggesting ways to extend and improve our specification.

NOTE

1. We had previously studied the data in the 1967 Health Interview Survey and concluded that although it is very valuable for certain purposes, the information on prices paid is so poor as to invalidate an estimated demand equation.