

**WHAT DO WE REALLY KNOW ABOUT
TRENDS IN OUTPATIENT MEDICAL EXPENDITURES FOR CHILDREN,
1977 TO 1987?**

Carol Rapaport
Federal Reserve Bank of New York
Domestic Research Function
(212)720-5657
carol.rapaport@ny.frb.org

Christopher A. Trenholm
Mathematica Policy Research
Research Division
(609) 936-2796
ctrenholm@mathematica-mpr.com

January 2000

The views expressed in this article are ours and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System. We thank Andrea Biddle, Tom Mroz, Frank Sloan, Sally Stearns, and seminar participants at the Institute for Research on Poverty, the Sixth Annual Health Economics Meetings, and the Triangle Health Economics Workshop for comments. Partial support for this project came from The Commonwealth Fund grant 93-32 and NIH grant R01 HD29551-03.

Abstract

This analysis of outpatient medical expenditures for children identifies which children experience a relative decline in medical expenditures between 1977 and 1987. The paper also evaluates some standard methodologies used in medical demand estimation. Our semiparametric approach models expenditures simultaneously with the choice of insurance plan and the decision to incur any expenditures. Children in poor families and Hispanic children witness a decline in expenditures relative to other children. Children on Medicaid and black children experience stable expenditures over time. These results imply that the recent Medicaid expansions may not help all children attain good health. The results are sensitive to assumptions of insurance exogeneity, but are insensitive to the assumptions governing the decision to incur any expenditures. The general emphasis in the health economics literature on sample selection instead of insurance endogeneity therefore seems misguided.

I. Introduction

This paper examines outpatient medical expenditures for children in 1977 and 1987. The analysis first identifies which socio-economic groups of children experience a relative decline in expenditures over the decade. It is these children who should be the focus of public policy. The analysis then evaluates some of the standard methodologies used to estimate the demand for medical services. Our econometric framework nests much of the medical care demand literature, and we evaluate the adequacy of conventional estimation schemes.

Improving the health status of children has become an increasingly important public policy goal. Legislation often focuses on Medicaid, the federal-state entitlement program that provides health insurance to needy children. For example, Congress expanded Medicaid eligibility to healthy children in 1986, 1987, and 1989. These eligibility expansions are intended to reduce the disparities in medical services between low income and higher income children. This research examines pre-expansion data to determine the degree to which medical expenditures are indeed a function of Medicaid use, and to quantify the interrelationships between household demographics and health outcomes.

We analyze outpatient expenditures using an econometric model of discrete and continuous choice. The paper models the simultaneous selection between three discrete insurance states, the discrete decision of whether to have an initial contact with an outpatient provider (i.e. to participate in the outpatient system), and the continuous level of outpatient medical expenditures for the participants. The choice of insurance is between private insurance, Medicaid, and no insurance. The analysis uses two single-year panel data sets, the 1977 National Medical

Care Expenditure Survey (NMCES) and the 1987 National Medical Expenditure Survey (NMES).

We feel this simultaneous framework captures some of the best aspects of the existing literature while avoiding some of its shortcomings. Partial analyses of the relationships between medical utilization, insurance, and demographic characteristics include Cunningham and Monheit (1990), Newacheck (1992), and Spillman (1992). These papers do provide information on the correlations across some variables. However, the absence of a behavioral model and the failure to account for unobservable factors common to the choice of insurance and expenditures make the results difficult to interpret. At the other extreme, Gilleskie (1995) and Schone, Selden and Zabinski (1995) estimate behavioral models of the demand for medical services. These fully-structural models incorporate the influence of unobservable variables and yield parameters with clear interpretations. Nevertheless, behavioral stochastic dynamic models cannot include more than a handful of observed variables and are heavily driven by distributional and functional form assumptions. Our work, like that of Currie and Thomas (1995), Goldman (1995), and Currie and Gruber (1996), is intended to strike a balance between these two approaches, and provide structural estimates of expenditures that incorporate a large number of observed variables. In addition, we can evaluate the extent to which the estimation technique matters in a fully-simultaneous selection model with economically justified exclusion restrictions.

The methodological literature concerning self-selection into positive health expenditures is already quite large. The standard methodology is the "two-part" model popularized by the Rand Health Insurance Experiment (RHIE) and extensively discussed in Duan *et. al.* (1983), Duan *et. al.* (1984), and Manning, Newhouse *et. al.* (1987). On the other hand, Hay and Olsen (1984) and

Maddala (1985) argue for a sample selection (Heckit) approach. Despite these theoretical disagreements, Manning, Duan *et. al.* (1987) concludes through Monte Carlo simulations that the two-part and sample selection approaches yield very similar results.

The potential endogeneity of insurance has received less attention in the medical expenditure literature than has selection bias. Although Cameron, Trivedi, Milne, and Piggott (1988) estimate the joint choice of health care and health insurance in Australia, much research relies on the RHIE data, where insurance is exogenous by construction. Given that the RHIE data date from the mid 1970s, an investigation into the possible endogeneity of insurance in expenditure estimation seems especially timely.

In our model, the endogeneity of participation and insurance status comes about because common unobservable factors influence all choices. For example, parental attitudes toward physicians can influence the decision to purchase insurance, to use any outpatient medical services, and the observed expenditure level. Our semiparametric technique uses a discrete approximation to control for the common factor. The econometric model is in the spirit of Heckman and Singer (1984), and is an application of the Mroz (1999) estimator to health economics.

The results indicate that both economic and demographic variables have grown in importance between 1977 and 1987. Children in poor families are increasingly likely to have lower medical expenditures than their wealthier peers. Hispanic children experience a decline in expenditures relative to white children. Black children have lower expenditures than white children, but the differential does not widen over time. The relationship between Medicaid and

expenditures is also stable. Like Currie and Thomas (1995), we conclude that public policies that equalize budget constraints will not equalize care across different groups of children.

Our methodological comparisons suggest that methodology matters, and that insurance is endogenous to expenditure estimation. The importance of common shocks to Medicaid enrollment and medical expenditures, such as would occur if a medical clinic enrolled an eligible child at the first visit, seems to have grown over time. Research that does not allow for the simultaneous determination of Medicaid and medical expenditures will produce inconsistent estimates of the key policy parameters. On the other hand, the paper confirms that selection into positive expenditures is empirically unimportant.

This paper proceeds as follows. Section II first presents an econometric model of medical expenditures and insurance choice, and then describes the data. Section III presents estimation results, and Section IV concludes.

II. An Econometric Model of Outpatient Medical Expenditures

Maddala (1985) characterizes the estimation of medical expenditures by writing "it is time that more thought be given to the exact sources of selectivity bias and a 'structural' formulation be attempted. Otherwise all we have is lip service to the basic problems. The health care area is one where policy questions are very important, and these can be analyzed only if a structural formulation is adopted. Otherwise, there is no need to depart from mother OLS." (p. 16).

Although over a decade old, we consider this point still valid, and our econometric model is in its spirit. We evaluate outpatient medical expenditures using an econometric model of the simultaneous choice of insurance status, participation into positive expenditures, and the level of

expenditures. The model is structural because participation and insurance are explicitly endogenous.

This section first characterizes the underlying optimization problem, then presents the empirical model, and finally describes the data. Although our ultimate empirical work deals with household heads making choices on behalf of their children, we refer here to the utility-maximizing individual for ease of exposition. In order to minimize notation, we do not use an individual-level subscript.

A. The Economic Model

This subsection outlines the simplest possible economic model needed to generate endogenous selection and insurance. The model is a two stage model under uncertainty. The consumer ex ante chooses the insurance plan, observes the health outcome, and ex post chooses health services and consumption. This model is a direct application of Cameron, Trivedi, Milne, and Piggott (1988) to a world with two types of individuals. Some individuals already have knowledge of the health care system (including how to acquire health care), and some do not. Participating in the health care system requires a minimum threshold of knowledge, and uninformed individuals must acquire knowledge in order to participate. In short, an individual must know where the doctor is in order to see the doctor. The acquisition of this knowledge requires "work" and generates disutility.

In the first period, the individual maximizes expected utility by choosing among $i=1$ to I mutually exclusive and exhaustive insurance plans. The choice of insurance plan influences the subsequent price of medical services and income. In the second period, the individual maximizes

utility given the budget constraint and the realization on health. As is usually the case with dynamic problems, the model is solved recursively.

In the second period, the individual picks consumption and health services to solve

$$\begin{aligned}
 & \text{Max } U(C, H, d_i, W, s, X, \theta_i) \text{ subject to } P_C C + P_{Hi} H + Y_i \\
 & \text{where } W = W_1 \text{ if uninformed and participates} \\
 & \quad \quad \quad W = 0 \text{ if informed} \\
 & \quad \quad \quad W = 0 \text{ if uninformed and does not participate.}
 \end{aligned} \tag{1}$$

where C is consumption of a composite commodity; H is health services; d_i is a health quality transformation factor; W is the additional knowledge required to participate in the health care system; s is the realized value from the distribution, $F(S)$, of health outcomes; X and θ_i are vectors of characteristics of the child and household; P_C and P_{Hi} are the respective unit prices for the composite commodity and health services; and Y_i is exogenous income net of the insurance premium. This specification abstracts away from deductibles.

The exogeneity of income implies that families cannot change their hours of work to alter their welfare eligibility. Marital status is similarly an exogenous element of X . By assumption, $U()$ is strongly separable in W , and $U_1 > 0$, $U_2 > 0$, $U_3 > 0$, $U_4 < 0$, and $U_5 > 0$. The transformation variable d_i includes factors such as access to facilities and pediatrician quality and is most likely a function of insurance status. For example, children on Medicaid may face a restricted choice of pediatricians and/or types of treatment.

Informed individuals are not required to participate in the health care system. If an uninformed individual chooses to participate in the health care system, however, then that person must become informed. The "cost" of acquiring knowledge of the health care system is measured

in terms of the decrease in utility associated with moving from $W=0$ to $W=W_1$. The participation rule for positive expenditures then becomes

$$H^i > 0 \text{ iff } U(C^i(P_C, P_{Hi}, Y_i, d_i, S, X_i), H^i(P_C, P_{Hi}, Y_i, d_i, S, X_i), d_i, W, S, X_i) > U(C^i(P_C, P_{Hi}, Y_i, d_i, S, X_i), 0, d_i, 0, S, X_i) \quad (2)$$

where H^* and C^* are the optimal demands for health services and consumption, respectively. H^* does not depend on W because of the assumption of strong separability.

In the first period, the consumer makes the discrete choice of insurance policy knowing only the cumulative distribution of S , $F(S)$. The consumer selects the policy i to solve

$$\text{Max}_{i,S} EU^i = \int U[C^i(P_C, P_{Hi}, Y_i, d_i, S, X_i), H^i(P_C, P_{Hi}, Y_i, d_i, S, X_i), d_i, W, S, X_i] dF(S) \quad (3)$$

Given the precise form for $U()$, the solution to equation (1) yields the demand for health services. Equation (2) yields a specification for the participation decision, and equation (3) yields a specification for the choice of insurance. Cameron, Trivedi, Milne, and Piggott (1988) specify a functional form and present a detailed version of this model. Indeed, their specifications are particularly appealing because the resulting optimal health expenditure equation is linear, and therefore easy to estimate (although not estimated in their paper).

Our model is also somewhat related to Kenkel's (1990) empirical analysis of medical information and the demand for medical services. Kenkel measures the individual's level of medical information by the score on a test of general medical knowledge. He finds that more informed consumers are more likely to use medical care, but that information does not affect the level of use, conditional on any use.

B. The Empirical Implementation

The empirical implementation of the above model consists of five equations. Three equations model latent variables measuring the expected utility associated with each of the three insurance plans evaluated in equation (3). The fourth equation models the latent equation (2) variable measuring the utility associated with positive health expenditures after choosing an insurance plan. The fifth equation models the level of log total outpatient expenditures observed for those with positive expenditures. This expenditure equation is analogous to the optimal quantity of health care from equation (1) multiplied by the gross price of care. In other words, our dependent variable, total outpatient expenditures, is a true measure of medical utilization.

The empirical model is as follows. The individual solves equation (3), evaluates the expected utility in each insurance state, and selects the state with the highest expected utility:

$$\begin{aligned}
 I_i^* &= \arg \max_{i=1,2,3} \{ \beta_0 + \beta_1 X_{1i} + \beta_2 Z_{2i} + \epsilon_{1i} \} \\
 I_i &= 1 \text{ iff } I_i^* = \text{Max}(I_1^*, I_2^*, I_3^*) \\
 I_i &= 0 \text{ otherwise}
 \end{aligned} \tag{4}$$

where I_i^* is a latent variable that determines whether the discrete observed outcome $I_i=1$ occurs, X is a vector of economic and demographic characteristics, Z is a vector of variables distinct from X used to help identify insurance choice, the ϵ_{1i} represent unobservables influencing the expected utility from the i th insurance choice, and the β are parameters to be estimated. If $I_1=1$, the individual is uninsured, if $I_2=1$, the individual has private insurance, and if $I_3=1$, the individual has Medicaid.

Because government policy uses the federal poverty status as a benchmark, we replace income, Y , with X variables indicating that the child's household falls into one of five poverty

classifications. Premiums and deductibles are ignored, and nine Census division dummy variables help control for regional variation in the price of medical care. The data contain a crude measure of health status, s_i : three parentally-reported health classifications. Finer measures of health are incorporated into ϵ_i . The health quality transformation factor, d_i , is unobservable and therefore becomes a part of ϵ_i .

This empirical model assumes that all children may choose any of the three insurance plans with a nonzero, but perhaps very small, probability. This assumption is theoretically justified by the presence of d_i and necessary given the available data. In particular, although healthy children in families with a current income greater than four times the federal poverty line are generally not eligible for Medicaid, children in recently less-well-off families would be eligible. In fact, the data do not distinguish between children who are eligible for Medicaid, and those who are not. For example, we observe some children in households above four times the federal poverty line on Medicaid in 1977 and 1987.¹

The individual also solves equation (1) and equation (2) and compares the utility associated with participating in the health care system with the utility associated with not

¹Because Medicaid eligibility is not known with certainty, equation (4) cannot be estimated on the full set of three choices for the Medicaid-eligible and on a restricted set of two choices for the Medicaid-ineligible. Equation (4) could, however, be specified such that eligibility is probabilistic. Under this estimation scheme, observations on Medicaid are by definition eligible. Observations in the other two insurance states are Medicaid-eligible with probability P , and Medicaid-ineligible with probability $(1-P)$. We reject this scheme for two reasons. First, the absence of data on state of residence, assets, and other key variables implies that P is identified by functional form only. Results from this more-complicated estimation scheme therefore provide no real information beyond the current equation (4). Second, the focus of this paper is on trends in medical expenditures, and estimation involving probabilistic eligibility is not needed to control for the potential endogeneity of Medicaid in expenditure estimation.

participating. The individual chooses whether to participate in the health care system and, for participants, the level of outpatient expenditures. Formally,

$$E^* = \alpha_0 + \alpha_1 X + \alpha_2 W + \alpha_3 I_2 + \alpha_4 I_3 + \epsilon_2 \quad (5)$$

$$Exp = 1 \text{ if } E^* \geq 0$$

$$Exp = 0 \text{ otherwise}$$

where E^* is a latent variable that determines whether the discrete outcome $EXP=1$ occurs, W is a vector of variables distinct from X used to help identify the selection into positive expenditures, ϵ_2 represents unobservables influencing the participation decision, and the α are parameters to be estimated. If $EXP=1$ the individual has positive outpatient medical expenditures and if $EXP=0$ if the individual has no medical expenditures. The expenditure equation is

$$E = \beta_0 + \beta_1 X + \beta_2 I_2 + \beta_3 I_3 + \epsilon_3 \quad (6)$$

$$\text{observe } E \text{ iff } EXP = 1$$

where E is the log of outpatient medical expenditures, ϵ_3 represents unobservables influencing the expenditures, and the β are parameters to be estimated.

A fundamental feature of this model is that equations (4) through (6) are simultaneous, and need to be estimated jointly. Unobservable factors such as parental attitudes toward physicians and health shocks to the child are likely to influence all outcomes. For example, a needy child may be uninsured and healthy, become ill, and then use medical services. If the provider enrolls this child in Medicaid at the time of treatment, then all the dependent variables are influenced by the shock of becoming ill. The estimated effects of private insurance and Medicaid are therefore inconsistent because the insurance variables are endogenous in the log

expenditure regression. In addition, failure to account for common unobservables will reduce the efficiency of all the estimates.

We use the "discrete factor approximation" estimator of Mroz (1999) to estimate equations (4) through (6) simultaneously. The unknown joint distribution of the unobserved variables is approximated using a random effects, components-of-variance structure. Each stochastic term includes a common factor, and this factor allows for nonzero covariances across the disturbances. The distribution of the factor is modeled as a discrete step function. This semiparametric specification of the error structure is an application of the Heckman and Singer (1984) control for unobservable heterogeneity in duration models. Mroz (1999) provides Monte Carlo evidence that discrete factor estimators perform better than maximum likelihood estimators that assume joint normality when the true distribution of the disturbances is not normal. Their Monte Carlo results also suggest that discrete factor estimators perform well when the true distribution is normal. Goldman (1995) uses this estimator to analyze utilization patterns within the military-health services system, and Card and Sullivan (1988), Gritz (1993), and others use the method to analyze labor market outcomes.

The stochastic structure of each of the five equations is the sum of an equation-specific error and a common error. In other words, a common linear factor captures the dependence of all health outcomes on the unobserved variables. We first derive the likelihood function for an individual's observed outcomes conditional upon the value of the unobserved factor, and then integrate out over the distribution of the unobserved factor.

Formally,

$$f_j(\mu_j) \propto \mathbf{D}_j v \quad j = 1, 2, 3 \quad (7)$$

where μ_{11} , μ_{12} , μ_{13} , μ_2 , and μ_3 are mutually independent. The joint density of the equation-specific component of the disturbances is therefore

$$f_j(\mu_{11}, \mu_{12}, \mu_{13}, \mu_2, \mu_3) = f(\mu_{11})f(\mu_{12})f(\mu_{13})f_L(\mu_2) \frac{1}{\mathbf{F}} \mathbf{N}\left(\frac{\mu_3}{\mathbf{F}}\right) \quad (8)$$

where $f(\cdot)$ denotes the Type I extreme value density function, $f_L(\cdot)$ denotes the logistic density function, and $\mathbf{N}(\cdot)$ denotes the standard normal density function. If there are no unobservable common to all equations, the model reduces to a multinomial logit estimating insurance choice, a logit estimating participation, and a log expenditures OLS regression.

Nevertheless, the common unobserved variable v is present. V may include the individual's risk aversion, detailed health status, and attitudes toward medical care. The joint density function for the five simultaneous disturbances conditional on v is

$$f_j(\mu_{11}, \mu_{12}, \mu_{13}, \mu_2, \mu_3 | v) = f(\mu_{11} \& \mathbf{D}_{11} v) f(\mu_{12} \& \mathbf{D}_{12} v) f(\mu_{13} \& \mathbf{D}_{13} v) f_L(\mu_2 \& \mathbf{D}_2 v) \frac{1}{\mathbf{F}} \mathbf{N}\left(\frac{\mu_3 \& \mathbf{D}_3 v}{\mathbf{F}}\right) \quad (9)$$

The variance components structure yields a joint conditional density function that is the product of univariate densities, and is therefore computationally feasible.

The estimation is by maximum likelihood. For each individual, there are six possible outcomes for the discrete variables: the individual could be in one of three insurance states, and one of two expenditure states. In other words, there are six possible regions in $(\mu_{11}, \mu_{12}, \mu_{13}, \mu_2)$

space, together with the continuous outcome characterized by y_i . The likelihood function for individual i , conditional on v_i , is

$$\begin{aligned}
 R(v) = & [prob(I_1^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(<0^*v))]^{1_{\{I_1^{(i)}=1, Exp=0\}}} \times \\
 & [prob(I_2^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(<0^*v))]^{1_{\{I_2^{(i)}=1, Exp=0\}}} \times \\
 & [prob(I_3^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(<0^*v))]^{1_{\{I_3^{(i)}=1, Exp=0\}}} \times \\
 & [prob(I_1^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(\$0^*\mu_3, v) \times (N(\mu_3)^*v))]^{1_{\{I_1^{(i)}=1, Exp=1\}}} \times \\
 & [prob(I_2^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(\$0^*\mu_3, v) \times (N(\mu_3)^*v))]^{1_{\{I_2^{(i)}=1, Exp=1\}}} \times \\
 & [prob(I_3^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)}), E(\$0^*\mu_3, v) \times (N(\mu_3)^*v))]^{1_{\{I_3^{(i)}=1, Exp=1\}}}
 \end{aligned} \tag{10}$$

where $1_{\{ \cdot \}}$ is an indicator function. The first three branches represent the likelihood that the individual is uninsured and does not participate, is covered by private insurance and does not participate, and is on Medicaid and does not participate, respectively. The fourth branch represents the likelihood that the individual is insured, participates, and has expenditures equal to the observed level, and the fifth and sixth branch are analogous.

This likelihood function simplifies because each joint density conditional on the value of the common factor is the product of the component univariate densities. The simplified function for individual i , conditional on v_i , is

$$\begin{aligned}
 R(v) = & [prob(I_1^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)*v})]^{1_{\{I_1^{(i)}=1\}}} \times \\
 & [prob(I_2^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)*v})]^{1_{\{I_2^{(i)}=1\}}} \times \\
 & [prob(I_3^{(i)} | Max(I_1^{(i)}, I_2^{(i)}, I_3^{(i)*v})]^{1_{\{I_3^{(i)}=1\}}} \times \\
 & [prob(E(<0^*v))]^{1_{\{Exp=0\}}} \times \\
 & [prob(E(\$0^*v) | (N(\mu_3)^*v))]^{1_{\{Exp=1\}}}
 \end{aligned} \tag{11}$$

where the first three branches are the conditional probabilities of being observed in each of the insurance states; the fourth branch is the conditional probability of being observed with no expenditures; and the fifth branch is the observed level of expenditures conditional on positive expenditures (and v).

The unconditional likelihood is found by integrating out over the distribution of v . We approximate the distribution of v by a step function where

$$Prob(v \in \mathbf{O}_k) = P_k \quad k = 1, \dots, K, \quad \text{where } P_k \geq 0 \text{ and } \sum_{k=1}^K P_k = 1 \quad (12)$$

That is, the step function has K points of support, the \mathbf{O}_k , and P_k is the probability of being at each support. The unconditional likelihood function for a single child is therefore a weighted sum of the conditional likelihood functions

$$L = \sum_{k=1}^K P_k R(\mathbf{O}_k) \quad (13)$$

The final likelihood function is the product of L over all individuals in each sample. As discussed below, the data are stratified random samples. We therefore maximize the following weighted function

$$\text{Log } \ell = \sum_{n=1}^N w_n (\text{Log}[L_n]) \quad (14)$$

where there are N individuals in each sample, each individual has sample weight w_n , and L_n is the unconditional likelihood function for a single child-- L in equation (13).

As in all polychotomous choice models, the coefficients on β_{11} , β_{21} and D_{11} are normalized to zero. We additionally need at least one normalization for \mathbf{O} because all equations involve the

product of D_j and O_k . In our estimation, we normalize O_1 to -0.5 and O_K to 0.5. These mass point normalizations imply no substantive restrictions because the product $D_j O_k$ remains unrestricted. We set $K=2$. Maximization of equation (14) yields estimates of the impact of all observed variables from equations (4) through (6), the D_j from equation (7), and the P_k from equation (12).

This empirical framework nests much of the existing literature. For example, the expenditure estimates from the RHIE two-part approach set $D_j=0$. The expenditure equation is estimated independently of participation under the assumption of exogenous insurance status. The traditional Heckit sample selection approach sets the $D_{1i}=0$, and proceeds in two stages. This procedure is consistent but typically produces very imprecise estimates. Instrumental variable procedures set $D_2=0$ and $\text{cov}(D_{1i}, D_3) \neq 0$, and estimate the effects of endogenous insurance on the expenditures in samples restricted to those with positive expenditures. These two-step procedures again typically produce imprecise estimates. Recent research supplements the instrumental variable approach by assuming D_{1i} is constant for either the individual over time, or for all individuals in a family. These fixed effects models have a number of potentially serious problems. Identification of the key parameters comes about from data on a nonrandom sample; in the current context, the effects of insurance on expenditures are identified solely through the behavior of children in the same family with different insurance statuses. Measurement error problems are magnified, and the calculation of out-of-sample predictions is problematic. This paper evaluates the efficacy of the RHIE, Heckit-type, instrumental variable, and family fixed effects approaches by comparing results from these methods with the baseline results obtained from the maximization of equation (14).

Although equations (4) through (6) let exogenous variables such as poverty status and race influence medical outcomes, the model has not drawn the distinction between differences in tastes and differences in constraints. For example, black children might receive lower expenditures than white children either because their parents face the same constraints but have different preferences, or because their parents have similar preferences but face different constraints over access to care. This distinction would have been easy to incorporate theoretically, but essentially impossible to operationalize empirically without untestable and unreasonable assumptions.

Government policy has indeed tried to change both tastes and constraints since the mid 1970s. For example, recent federal public health policies include national campaigns for stopping smoking, exercising, proper nutrition, medical care in the first trimester of pregnancy, and safe sex. At least some of these policies might have shifted the family's preferences toward the children's health care over time. At the same time, policies regulating the required acceptance of patients by providers and reducing residential segregation may have improved access to care. The relationships between the exogenous X variables and outcomes can therefore be expected to have shifted over time in such a way that incorporates both tastes and constraints.

C. Data and Identification

We perform identical analyses on two single-year panel survey data sets, the 1977 NMCES and the 1987 NMES. The endogenous variables are whether the child has Medicaid, private insurance, no insurance, positive outpatient medical expenditures, and the level of log expenditures (conditional on positive expenditures). Key exogenous X variables are the child's

age and race, and the family's poverty status and family type (e.g., a never-married female head). The specifications contain numerous other explanatory variables, including measures of health status of the child, age and education of the head, region and SMSA status of residence, number of children in the family, and whether the child is the eldest in the family.

The two single-year surveys are ideal for intertemporal comparisons. Both NMCES and NMES are administered by the Agency for Health Care Policy and Research (AHCPR), and are random stratified samples of the civilian, non-institutionalized population. The surveys have similar sample designs. All analysis uses the full-year sample weights provided in the data. The child's insurance status is observed at five points spread throughout the survey year. The insurance variable is determined through a hierarchical screening where Medicaid takes precedence over private insurance. We exclude from the analysis a small number of children ever observed on Medicare, CHAMPUS, or with an unknown insurance status. The data sets contain 10,334 and 7,913 single children between ages 2 and 17 inclusive in 1977 and 1987, respectively. All monetary values are in 1987 dollars. Further details on variable creation are in the Data Appendix.

The model is identifiable through functional form. Nevertheless, we use economically-meaningful exclusion restrictions to improve the efficiency of estimation. The Z vector contains variables influencing the insurance status of the child, but not influencing the presence and level of medical expenditures. Newhouse, Phelps, and Marquis (1980) observe that employment characteristics are theoretically appropriate instruments for insurance. The Z vector includes dummy variables indicating the family head's employment status, industry, and occupation.

The W variable represents knowledge of the health care system, and helps identify the participation decision. We use the presence of a new baby (born in the survey year) in the family as a proxy for parental knowledge of health services. Recent parents are exposed to health providers during the birth (or adoption) process, and therefore must have some minimum information about health provision. Because otherwise-identical parents may not have this minimum level of information, the disutility associated with participation is lower for the older children of recent parents than for other children. At the same time, the presence of a new baby is unlikely to influence the level of expenditures for the older siblings because the expenditure equations already control for the number of children in the family, the child's age, and other relevant variables. The presence of a new baby is also included in the Z vector.

III. Results

Table 1 presents descriptive statistics for the two samples of children. The fraction of children with private insurance falls by 5 percentage points between 1977 and 1987, while the fraction of children on Medicaid increases by less than 3 percentage points. Just under 70% of children have positive expenditures in both years. The changes in the level of expenditures are more dramatic, with average real expenditures increasing by about 46%. The exogenous family characteristics display the expected patterns, with increases in the number of never-married female heads, formerly-married female heads, and heads younger than 45. The family income distribution widens over time as more families are either poor or at greater than four times the poverty level. Changes in the exogenous characteristics of the children include increases in the number of black and Hispanic children, and a decrease in the average age of the children.

Table 2 presents partial results from the estimation of the full model, equation (14). Although the insurance choice coefficients are not presented, Table 3 provides a summary of the intertemporal trends in insurance choice. The summary decomposition holds the population constant and focuses on structural changes as measured by changes in the estimated coefficients. The share of children on Medicaid increases from 13.8% in 1977 to 16.6% in 1987. However, if the 1977 children received Medicaid under the 1987 structure, then 14.2% would have received Medicaid. The structural changes therefore lead to an estimated 0.4 percentage point increase in the number of children on Medicaid, while changes in the demographic composition of the children led to a 2.4 percentage point increase in the number of children on Medicaid. In short, the increase in the share of children on Medicaid is largely caused by demographics and not by any change in the structure of the Medicaid program. Perhaps this result is unsurprising because there were few Federal legislative changes in Medicaid over the period. On the other hand, it does suggest that there was little change in the enrollment behavior of the Medicaid-eligible population.

The first and third columns of Table 2 present key coefficients describing the participation decision. The economic variables are significant in both years. Children with private and Medicaid coverage, who face a lower price of care, experience an increased probability of any outpatient expenditures. The participation probability for a child with the median characteristics increases from 70% to 71% if the child has private coverage, decreases from 78% to 76% if the child has Medicaid, and decreases from 65% to 56% if the child is uninsured. Children in families below four times the federal poverty line are less likely than wealthier children to have positive expenditures. However the relationship between poverty and participation generally weakens over time.

Black children are substantially less likely than white children to have any expenditures, although the magnitude of this difference seems stable. On the other hand, Hispanic children experience a worsening over time in the probability of any expenditures. Despite a popular concern over the quality of the lives of children of never-married mothers, these children are indistinguishable from the children in conventionally married families. In fact, it is the children of unmarried male headed families who experience a smaller probability of participation in 1987. In both years, the probability of expenditures is U shaped in age, with the minimum near age 12. The presence of a new baby is a significant determinant of participation in 1977, but not in 1987. Although it is desirable to use an exclusion restriction that is statistically significant in both years, it is important to note that the model is identified without any exclusion restrictions.

The second and fourth columns of Table 2 present key coefficients describing log expenditures. The economic variables display strong intertemporal changes, with the effects of family income becoming more pronounced and the effects of having private insurance becoming less pronounced. Poor children have about 7% lower expenditures than those at greater than four times the poverty level in 1977, but 19% lower expenditures in 1987. Indeed the coefficients on all the included poverty status variables increase from small and insignificant to larger and/or significant. As expected, the coefficients on the insurance variables are positive. Having insurance lowers the price of care. The substitution effect toward increased health services should overwhelm any offsetting income effect induced by (not modeled) premiums in private insurance. The coefficient on private insurance decreases from .39 to .22, while the coefficient on Medicaid is stable at about .33.

Demographic variables also influence expenditures. The black/white expenditure differential is roughly constant over time at about 23%. On the other hand, the Hispanic/white differential increases from .11 to .31. There are no expenditure differentials between children in a family headed by a never-married woman and children in a two-parent family. In fact, in a result consistent with Currie and Gruber (1996), it is the children of unmarried men who experience lower expenditures than the children of married couples in 1987. Expenditures are more strongly influenced by the age of the family head in 1987 than in 1977, but the relationship between the child's own age and expenditures is largely unchanged. The expenditure premium associated with being the eldest child falls over time, and there are no strong relationships between expenditures and the number of children in the family.

Table 4 presents partial results from the estimation of equation (14) under three alternative assumptions. The first column for each year duplicates the results from Table 2. The second column presents results under the assumption that $D_j=0$. This OLS estimation implies exogenous insurance and no selection into positive expenditures. The third column presents results under the assumption that the D_j are constant across families. This family fixed effects estimator implies no selection into positive expenditures, but permits family-specific unobservables such as attitudes toward doctors and risk to influence expenditures.

Comparisons between the first and second columns in each year show that the estimated insurance effects differ dramatically across estimation schemes. In particular, estimation under independent errors suggests that the expenditure differential between Medicaid and uninsured nearly doubles, while the results from the joint estimation show no change. The difference between the two estimation schemes is driven largely by changes in the coefficient for 1987,

which is substantially larger under independence. One explanation for the difference is that uninsured Medicaid-eligible children may become enrolled once they have contact with a health-care provider. Therefore, the results under independence are picking up not only that Medicaid increases expenditures, but also that having larger expenditures increases the probability of being enrolled in Medicaid.

The coefficients on the common factor are also consistent with this explanation. The coefficient on D_{12} , not shown, is positive (but insignificant), and the coefficient on D_3 is also positive. The correlation across equations is therefore positive. A common shock, such as an illness, influences both the probability of having Medicaid and the level of expenditures. The correlation in the stochastic terms across equations is negative in 1977, although the differences in the estimated Medicaid differential are smaller. The estimated Medicaid effects on the probability of participation are stable across estimation schemes.

The privately insured/uninsured differential is similarly sensitive to estimation scheme. Once again, the results under the assumption of exogenous insurance and no selection understate the differential in 1977 (relative to the baseline), but overstate it in 1987. In addition, the effect of private insurance on participation is overstated in 1977.

The family fixed effects technique is appealing conceptually because the estimator accommodates unobservables that are correlated with the regressors. On the other hand, the insurance coefficients are identified from only those children with a different insurance plan from their siblings--approximately 660 children in each year. The estimates presented in the third column in each year have extremely large standard errors. Estimates from a variety of two-stage instrumental variable procedures, not shown, are broadly consistent with the fixed effects results,

but are even more imprecise. Although the family fixed effects and instrumental variable estimates might be consistent with the baseline results in the first column, they are just too imprecise to permit useful inferences.

The estimation setting only $D_{1i}=0$ assumes that insurance is exogenous, but allows for selection effects in the expenditure equations. The results, not shown, are virtually identical to those for exogenous selection and insurance. Controlling for sample composition has no effect on the estimated impacts of insurance, poverty status, or other variables. We therefore confirm the fundamental conclusion from the RHIE/selection debate. Controls for sample selection do not alter inferences from medical demand estimation, even though our estimation scheme is fully simultaneous.

IV. Conclusions

What do we really know about trends in outpatient medical expenditures for children? The results of the joint estimation of insurance status, participation into medical expenditures, and the level of expenditures suggest that both economic and demographic factors influence expenditures in 1977 and 1987. Children in poor families, Hispanic children, and children in families headed by a single male witness a decline in expenditures relative to other children. Although the black/white expenditure differential has not grown over time, this racial differential remains substantial. The relationship between Medicaid and expenditures is stable, while private insurance has a somewhat weaker influence on expenditures in the later year.

We also know that methodology matters. The health economics literature generally focuses on issues of selection, and concludes that inferences are robust across the RHIE two part

and Heckit models. We confirm that selection is unimportant empirically. The literature's focus on selection, however, is somewhat misplaced because it masks the importance of the endogeneity of insurance. In 1987, for example, treating Medicaid as exogenous substantially overstates the Medicaid/uninsured expenditure differential. As the RHIE data becomes increasingly dated, it is increasingly important to improve the estimation techniques used with nonexperimental data. Although the discrete factor framework for simultaneous equations has not yet been widely used in health economics, the methodology appears promising. Instrumental variables and family fixed effects are often too imprecise to be useful.

Finally, we know that the recent Medicaid eligibility expansions should partly achieve the goals of increased care for children. Children on Medicaid do have significantly higher expenditures than uninsured children. At the same time, the effectiveness of the expansions will vary with the race and family composition of the child. We conclude, as did Currie and Thomas (1995), that simply equalizing budget constraints by expanding Medicaid (or the Children's Health Insurance Program) will not equalize care across different groups of children.

What don't we know about trends in outpatient medical expenditures? It is impossible to tell whether the observed differences are caused by tastes or constraints. At the same time, the observed racial differences clearly indicate a public policy problem. For example, the American Academy of Pediatrics recommends that a child have at least one physician visit at age four. We use the estimates in Table 2 to calculate the probability that this recommendation is followed for representative four year olds. Even among those children on Medicaid, about 16% of whites, 23% of Hispanics, and 28% of blacks have no predicted medical expenditures. The Medicaid (or CHIP) expansions will not help these children meet the suggested guidelines for good health.

Data Appendix

This Appendix describes the creation of the key variables. Most variables are measured as of the final survey round. The insurance status of each child is created using the following hierarchy. The child has Medicaid if, at any of five points over the year, the child's parent or guardian reports the child covered by Medicaid. Each remaining child has private insurance if, at any of the five points, the child's parent or guardian reports the child has private coverage. All other children are uninsured.

Outpatient expenditures are an annual total of all medical expenditures accrued by the child for ambulatory services. Outpatient expenditures include visits to a physician's office, clinic, or emergency room; home health visits; and hospital stays under one night. The level of outpatient expenditures generally represents the total charges for medical services. There are two exceptions imposed by the way AHCPR releases the data. First, in cases where the actual payment is reduced through the use of a third-party payer (such as Medicaid or private insurance), the data are the amount paid rather than the actual charge. Second, in cases where no charge is specified (such as for an HMO visit) or where the stated charge is otherwise unreliable, expenditure data are imputed. All monetary values are expressed in 1987 dollars, with the 1977 data inflated using the Consumer Price Index for health services.

References

- Cameron, A.C., P.K. Trivedi, Frank Milne, and J. Piggott. 1988. "A Microeconometric Model of the Demand for Health Care and Health Insurance in Australia." Review of Economic Studies 55:85-106.
- Card, David and Daniel Sullivan. 1988. "Measuring the Effect of Subsidized Training Programs on Movements In and Out of Employment." Econometrica 56:497-530.
- Cunningham, Peter J. and Alan C. Monheit. 1990. "Insuring the Children: A Decade of Change." Health Affairs 9:76-90.
- Currie, Janet and Jonathan Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." Quarterly Journal of Economics 111:431-466.
- Currie, Janet and Duncan Thomas. 1995. "Medical Care for Children: Public Insurance, Private Insurance, and Racial Differences in Utilization." Journal of Human Resources 30:135-162.
- Duan, Naihua, Willard G. Manning, Jr., Carl N. Morris, and Joseph P. Newhouse. 1983. "A Comparison of Alternative Models for the Demand for Medical Care." Journal of Business and Economic Statistics 1:115-126.
- , "Choosing Between the Sample-Selection Model and the Multi-Part Model." Journal of Business and Economic Statistics 2:283-289.
- Gilleskie, Donna. 1995. "A Dynamic Stochastic Model of Medical Care Use and Work Absence." The University of North Carolina, mimeo.
- Goldman, Dana. 1995. "Managed Care as a Public Cost-Containment Mechanism." Rand Journal of Economics 26:277-295.
- Gritz, R. Mark. 1993. "The Impact of Training on the Frequency and Duration of Employment." Journal of Econometrics 57:21-51.
- Hay, Joel W. and Randall J. Olsen. 1984. "Let Them Eat Cake: A Note on Comparing Alternative Models of the Demand for Medical Care." Journal of Business and Economic Statistics 2:279-282.
- Heckman, J. and B. Singer. 1984. "A Method For Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." Econometrica 52:271-320.
- Kenkel, Don. 1990. "Consumer Health Information and the Demand for Medical Care." Review of Economics and Statistics 587-595.

Maddala, G.S. 1985. "A Survey of the Literature on Selectivity Bias as it Pertains to Health Care Markets." Advances in Health Economics and Health Services Research. JAI Press Inc. 6:3-18.

Manning, W.G., N. Duan, and W.H. Rogers. 1987. "Monte Carlo Evidence on the Choice Between Sample Selection and Two-Part Models." Journal of Econometrics. 35:59-82.

Manning, Willard G, Joseph P. Newhouse, Naihua Duan, Emmet B. Keeler, Arleen Leibowitz, and M. Susan Marquis. 1987. "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment." American Economic Review. 77:251-277.

Mroz, Thomas A. 1999. "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome." Journal of Econometrics, forthcoming.

Newacheck, Paul W. 1992. "Characteristics of Children with High and Low Usage of Physician Services." Medical Care. 30:30-41.

Newhouse, Joseph, Charles Phelps, and Susan Marquis. 1980. "On Having Your Cake and Eating It Too: Econometric Problems in Estimating the Demand for Health Services." Journal of Econometrics 13:365-390.

Schone, Barbara Steinberg, Thomas Selden, and Daniel Zabinski. 1995. "Insurance Choice and Medical Care Utilization: An Empirical Analysis Using Restrictions from Self-Selection Equilibrium." Agency for Health Care Policy and Research, mimeo.

Spillman, Brenda. 1992. "The Impact of Being Uninsured on Utilization of Basic Health Care Services." Inquiry. 29:457-466.

Table 1
Variable Definitions and Descriptive Statistics

		1977 means	1987 means
Endogenous Variables			
posexp	=1 if positive expenditures	0.696	0.696
expall	level of expenditures, all cases	137.9	201.7
expifpos	level of expenditures, posexp=1	198.1	289.8
uninsure	=1 if uninsured throughout the year	0.089	0.111
ins_priv	=1 if Medicaid=0 and privately insured at least once during the year	0.773	0.723
Medicaid	=1 if on Medicaid at least once during the year	0.138	0.166
Characteristics of the Child			
male	=1 if male	0.512	0.515
female	=1 if female	0.488	0.485
white	=1 if white	0.654	0.701
black	=1 if black	0.117	0.154
hispanic	=1 if Hispanic	0.056	0.103
othrace	=1 if other race	0.015	0.042
dkrace	=1 if race unknown	0.158	--
age	age in years	9.98	9.42
eldest	=1 if eldest child in family	0.400	0.488
noteldest	=1 if child not eldest in the family	0.600	0.512
healthex	=1 if parentally-reported health excellent	0.545	0.471
healthgd	=1 if parentally-reported health good	0.354	0.313
healthfp	=1 if parentally-reported health fair/poor	0.060	0.052
dkhealth	=1 if health not reported	0.041	0.163
Characteristics of the Family			
newb_yes	=1 if child in family born during survey year	0.035	0.046
newb_no	=1 if no child in family born during survey year	0.965	0.954
poor	=1 if residing in family below the poverty line	0.151	0.189
poor1_125	=1 if residing in family from 100% to 125% of the poverty line	0.048	0.050
poor125_2	=1 if residing in family between 125% and 200% of the poverty line	0.195	0.159
poor2_4	=1 if residing in family from 200% to 400% of the poverty line	0.398	0.376
poor4up	=1 if residing in family above 400% of the poverty line	0.208	0.226

Table 1
Variable Definitions and Descriptive Statistics

		1977 means	1987 means
smsa	=1 if residing in an smsa	0.691	0.722
non_smsa	=1 if not residing in an smsa	0.309	0.278
mhead_cmar	=1 if family head male and currently married, spouse present	0.807	0.726
mhead_oth	=1 if household head male and not currently married, spouse present	0.026	0.032
fhead_fmar	=1 if family head female and widowed, divorced, or separated	0.147	0.187
fhead_nmar	=1 if family head female and never married	0.020	0.055
hda17_24	=1 if family head age from 17 to 24	0.028	0.034
hda25_34	=1 if family head age from 25 and 34	0.297	0.334
hda35_44	=1 if family head age from 35 to 44	0.401	0.441
hda45_54	=1 if family head age from 45 to 54	0.201	0.127
hda55_64	=1 if family head age from 55 to 64	0.056	0.044
hda65up	=1 if family head age over 64	0.017	0.019
hdedlt12	=1 if family head has less than 12 years of education	0.335	0.253
hded12	=1 if family head has 12 years of education	0.336	0.345
hded1315	=1 if family head has 13 to 15 years of education	0.125	0.196
hded16pl	=1 if family head has 16 or more years of education	0.163	0.195
hddked	=1 if education of family head unknown	0.041	0.010
totkids	total children under 18 in family	2.76	2.49
# cases		10334	7913

Notes: All expenditures are outpatient expenditures, and are expressed in 1987 dollars using the CPI-health. All data except the number of cases are weighted. The standard deviations of unconditional expenditures for the two series are 282.4 and 561.1 respectively. The standard deviations of conditional expenditures are 320.4 and 653.3 respectively.

Table 2
Estimates of the Probability of Positive Expenditures and the (Log) Level of Outpatient Medical Expenditures

	1977		1987	
	Probability of Expenditures	Level of Expenditures	Probability of Expenditures	Level of Expenditures
female	-0.0167 (.043)	-.1317* (.024)	0.0385 (.048)	-0.0406 (.030)
black	-0.7529* (.081)	-0.2242* (.053)	-0.7549* (.095)	-0.1992* (.071)
hispanic	-0.2149* (.106)	-0.1091 (.058)	-0.4694* (.099)	-0.3106* (.071)
othrace	-0.2860 (.196)	-0.3079* (.112)	-0.7409* (.106)	-0.1796* (.074)
dkrace	-1.0390* (.062)	0.1830* (.040)	--	--
age	-0.2579* (.025)	-0.1143* (.013)	-0.1953* (.027)	-0.1100* (.016)
age2	0.0103* (.001)	0.0059* (.001)	0.0082* (.001)	0.0057* (.001)
eldest	0.2116* (.049)	0.1308* (.027)	0.2133* (.053)	0.0532 (.033)
healthgd	0.1349* (.048)	0.1811* (.027)	0.2045* (.057)	0.2262* (.036)
healthfp	0.9438* (.114)	0.5257* (.051)	0.9264* (.134)	0.6791* (.063)
dkhealth	0.1715 (.136)	0.0999 (.070)	0.0763 (.071)	0.1043* (.044)
ins_priv	0.2330 (.156)	0.3886* (.067)	0.6934* (.096)	0.2212* (.090)
Medicaid	0.6191* (.114)	0.3355* (.072)	0.9574* (.118)	0.3246* (.098)
newb_yes	0.2981* (.135)		0.1519 (.126)	
poor	-0.3871* (.099)	-0.0683* (.052)	-0.3339* (.102)	-0.1742* (.073)
poorl1_125	-0.5356* (.132)	0.0525 (.073)	-0.3918* (.143)	-0.0892 (.086)
poorl25_2	-0.3708* (.076)	0.0416 (.042)	-0.3007* (.096)	-0.1745* (.057)
poor2_4	-0.0869 (.062)	-0.0581 (.032)	-0.1947* (.073)	-0.1043* (.041)
smsa	0.0754 (.050)	0.1831* (.029)	0.0633 (.057)	0.2131* (.037)

Table 2
Estimates of the Probability of Positive Expenditures and the (Log) Level of Outpatient Medical Expenditures

	1977		1987	
	Probability of Expenditures	Level of Expenditures	Probability of Expenditures	Level of Expenditures
mhead_oth	-0.1002 (.134)	0.0438 (.081)	-0.2690* (.127)	-0.2364* (.093)
fhead_fmar	-0.0154* (.073)	0.1400* (.042)	0.2303* (.072)	0.0199 (.045)
fhead_nmar	0.2675 (.205)	0.1639 (.115)	0.2063 (.142)	-0.0229 (.088)
hda25_34	0.2716 (.158)	0.0737 (.080)	0.1344 (.145)	0.0864 (.085)
hda35_44	0.3235* (.163)	0.1775* (.085)	0.3316* (.153)	0.1860* (.091)
hda45_54	0.1550 (.168)	0.1161 (.088)	0.1151 (.166)	0.2353* (.101)
hda55_64	0.0419 (.183)	0.0661 (.097)	-0.1456 (.183)	0.1765 (.119)
hda65up	0.3102 (.227)	0.1633 (.133)	0.3945 (.255)	0.3557* (.166)
hded12	0.2167* (.059)	0.0533 (.033)	0.1030 (.067)	0.0675 (.044)
hded1315	0.1334 (.076)	0.1155* (.043)	0.4021* (.081)	0.1570* (.052)
hded16pl	0.3748* (.077)	0.2543* (.040)	0.6001* (.089)	0.0530 (.055)
hddked	0.0274* (.135)	0.1308 (.080)	0.1327 (.248)	-0.2355 (.141)
totkids	-0.1167* (.045)	-0.0193 (.028)	-0.2356* (.066)	-0.0766 (.049)
totkids2	-0.0071 (.005)	-0.0005 (.004)	0.0118 (.009)	0.0031 (.008)
constant	2.4802* (.422)	4.1373* (.152)	1.3013* (.280)	5.3895* (.180)
rho	-0.4637 (.652)	0.6716* (.167)	-0.6635 (.416)	1.7812* (.054)
probsup1	-1.6005* (16.8%)		1.1474* (75.9%)	
MSE		1.0351* (.017)		1.2426* (.017)

Table 2
Estimates of the Probability of Positive Expenditures and the (Log) Level of Outpatient Medical Expenditures

	1977		1987	
	Probability of Expenditures	Level of Expenditures	Probability of Expenditures	Level of Expenditures
Mean of Dependent Variable	0.696	4.67	0.696	4.76

Notes: For probsup1, the values in parentheses are the probability weights on the first support point. For all other coefficients, these values are standard errors. The probability coefficient for support one is highly significant in both years. The first and second supports are normalized to -0.5 and 0.5, respectively. The omitted dummy categories are male, white, noteldest, healthex, uninsured, newb_no, poor4up, non_smsa, mhead_cmar, hda17_24, and hdedlt12. Each specification also includes dummy variables for Census division. The value of the log likelihood is -21,044.7 in 1977 and -17,028.4 in 1987.

Table 3
Actual and Predicted Insurance Choice Probabilities

	Percent of Population on Medicaid	Percent of Population with Private Insurance	Percent of Population that Is Uninsured
1977	13.8%	77.3%	8.9%
1987	16.6	72.3	11.1
Decomposition ¹	14.2	74.4	11.4

¹ The decomposition assumes that the 1977 population chooses insurance based on the structure estimated in the 1987 model.

Table 4
Estimates of (Log) Level of Outpatient Medical Expenditures Under Alternative Specifications

	1977			1987		
	Baseline Model	Ordinary Least Squares	Family Fixed Effects	Baseline Model	Ordinary Least Squares	Family Fixed Effects
female	-0.1317* (0.024)	-0.1337* (.024)	-0.1410* (.038)	-0.0406 (.030)	-0.0634* (.030)	0.0014 (.053)
age	-0.1143* (.013)	-0.1166* (.013)	-0.1428* (.021)	-0.1100* (.016)	-0.1194* (.016)	-0.1080* (.031)
age2	0.0059* (.001)	0.0060* (.001)	0.0080* (.001)	0.0057* (.001)	0.0063* (.001)	0.0061* (.001)
eldest	0.1308* (.027)	0.1274* (.027)	0.0117 (.042)	0.0532 (.033)	0.0554 (.033)	-0.0283 (.061)
healthgd	0.1811* (.027)	0.1798* (.027)	0.4324* (.068)	0.2262* (.036)	0.2556* (.036)	0.3171* (.079)
healthfp	0.5257* (0.051)	0.5191* (.050)	0.7934* (.098)	0.6791* (.063)	0.7538* (.061)	0.5975* (.141)
dkhealth	0.0999 (.070)	0.0806 (.069)	0.0896 (.204)	0.1043* (.044)	0.1138* (.044)	0.3345* (.103)
ins_priv	0.3886* (.067)	0.1928* (.052)	-0.0601 (.290)	0.2212* (.090)	0.3159* (.069)	-0.5094 (.448)
Medicaid	0.3355* (.072)	0.2734* (.064)	0.2096 (.294)	0.3246* (.098)	0.4769* (.074)	0.1035 (.392)

Notes: Standard errors are in parentheses.