How Managed Care Affects Medicaid Utilization
A Synthetic Difference-in-Differences Zero-Inflated Count Model

Deborah A. Freund[1,3]
Thomas J. Kniesner[2,3]
Anthony T. LoSasso[2,3]

April 18, 1996

We have benefited from the helpful comments of Doug Staiger and audience members at the Sixth Annual Health Economics Conference, the 1995 Annual Meetings of the Association for Public Policy Analysis and Management, and the 1995 International Conference on Health Policy Research. We gratefully acknowledge the financial support of Tilburg University's CentER for Economic Research and the Office of the Vice-President for Research and the University Graduate School of Indiana University. The research presented here has been facilitated by funding from the Agency for Health Care Policy Research, AHCPR grant
#R03 HS08944; the agency bears no responsibility for the research and policy conclusions.

Abstract: We develop a synthetic difference-in-differences statistical design to apply to experimental data for adult women living in Hennepin County, Minnesota, to estimate the impact of Medicaid managed care on various modes of medical care use. Because the outcomes of interest are utilization counts with many persons using none of a particular mode of care we use count regression models that are adjusted for excessive zeros. We find no reductions in physician visits or hospital inpatient and emergency department care use, but reductions in hospital outpatient care. Simulations designed to judge the economic significance of our results suggest a program effect that is a savings of about ten percent.

Key Words: Medicaid, managed care, difference-in-differences, count data models, zero inflated Poisson

JEL Classification: I18, C25, C23

[1]School of Public and Environmental Affairs, Indiana University
[2]Department of Economics, Indiana University, Bloomington
[3]Bowen Research Center, Indiana University
Communications: Thomas J. Kniesner
Department of Economics, Ballantine Hall (901),
Indiana University, Bloomington, IN 47405-6620 USA.
Phone: (812) 855–1021, Fax: (812) 855–3736,
Internet: Kniesne@Ophelia.UCS.Indiana.Edu

1. Introduction

With national health reform on the backburner and Congress increasingly favoring policy decentralization the locus of activity for health reform is returning with a vengeance to the states. States have long been the nation's laboratories for experimentation with different forms of health care delivery systems. At the heart of most ongoing state health reforms is overhauling Medicaid through the strategy of moving Medicaid eligibles out of fee-for-service (FFS) coverage and into managed care.[1] Managed Medicaid is likely to be increasingly popular under block grant programs. Over 40 states now use Medicaid managed care as a cost saving device or as a platform for increasing insurance coverage (Hurley, Freund, and Paul 1993). By the year 2000, over 30 percent of Medicaid eligibles in all aid categories will be enrolled in managed care programs; within the next three years enrollment of Medicaid eligibles in managed care will likely reach nearly 100 percent in the largest Medicaid states, New York and California (Freund and Hurley 1995). We use a synthetic difference-in-differences specification of a count regression model to examine the effect on utilization and program cost of randomly switching Medicaid adults out of fee-for-service into managed care.

Estimation of the program effect (of managed care on use) is the central aim of our work. There are different aspects of medical care utilization, and each may be distinctly affected by managed care. We examine use of four modes of care: doctor office visits, hospital outpatient department visits, emergency room visits, and hospital inpatient days.

---

[1] As background for our international readers Medicaid provides hospital, physician, and nursing home care for the U.S. low income population. Medicaid programs are administered by the states and differ in terms of eligibility and generosity of coverage across states subject to federal guidelines and financial assistance. Patient co-payments for Medicaid services are minimal and the health care providers who participate in the program must accept all Medicaid patients who seek care. Currently Medicaid enrolls about 30 million people (20 million of whom are receiving income maintenance payments under Aid to Families with Dependent Children) with total expenditures of about $90 billion.

Specifically, we explore the answers to two questions. What is the impact of managed care on the utilization of various modes of health care by Medicaid enrollees? How do the estimated utilization changes affect Medicaid program cost?

As is realistic, the Medicaid experiment we examine has multiple provider settings. Enrollees in the managed care program got care from one of six different managed care plans, and we estimate a set of program effects across managed care provider characteristics. In describing the use and cost implications of a switch to managed care we are careful to report the robustness of our conclusions to alternative estimating equations and to note how the program effects inferred from our synthetic difference-in-differences design differ from the program effects one would infer from simpler designs such as either a one group before and after treatment design or a post treatment cross-sectional comparison of use by the experimental versus control group.[2]

To summarize our results, we find evidence of substitution among adult Medicaid users from more expensive forms of ambulatory care to cheaper modes of care. We also find that the particular health maintenance organization (HMO) in which Medicaid users enroll does not matter much in terms of the estimated program effect. The substitution among care modes due to a switch to managed care leads to an estimated annual decrease in Medicaid program cost of about ten percent.

2. Brief Background on Medicaid Experiments

Previous evaluations of the impact of Medicaid managed care on use patterns for single mothers receiving welfare benefits differ from our research in three ways: the type of evaluation, the type of data, and the extent of generalizability.

---

[2] As we will soon make clear, we coin the term synthetic difference-in-differences to note that we use two distinct, statistically equivalent, comparison groups to infer any change in use by Medicaid recipients continuously under fee-for-service medical care during the experiment.

With one exception (Leibowitz, Buchanan, and Mann 1992) previous studies have used quasi-experimental designs. Hurley, Freund, and Paul (1993) examined the 25 best executed Medicaid managed care evaluations for the AFDC population; only one evaluation was designed as an experiment. Quasi-experimental studies suffer some obvious difficulties in inferring program effects by comparing cross-sectional data from pre and post intervention periods. Only some of the cross sectional comparisons use findings from an area employing Medicaid managed care and findings from a sufficiently similar area retaining fee-for-service Medicaid, which means that events other than the switch to managed care may have caused the observed differences in outcomes between areas. Similarly, quasi-experimental evaluation of programs in areas where Medicaid managed care is not mandatory may be contaminated by researchers' inability to correct for preferred or adverse selection.

Another difficulty in previous research lies in the two data sources generally used: either (1) claims/encounter data reported by Medicaid programs or managed care providers or (2) telephone or personal surveys of beneficiaries enrolled in Medicaid in managed care or fee-for-service areas. Although claims data may suffer less than survey data from inaccurate reports of the existence of health care events, such as doctor visits or hospital admissions, and claims data are cheaper to collect than survey data, use other than hospital care often is underreported particularly when the claims come from capitated providers who have no distinct fee for the service provided. Neither claims data nor encounter data are collected with research in mind, and frequently events are not comparably defined. Accurate measures of health status needed to control for individual differences that confound the assessment of managed care's effect on use are also missing from claims/encounters data.

The third difficulty with previous research is that it has generally evaluated the success of Medicaid managed care plans that were newly created solely to service the Medicaid population. The demonstration that managed care plans studied are not ones likely to proliferate through the rest of the 1990s. Rather, Medicaid recipients are beginning to be enrolled in HMOs and other types of managed care entities that have already been marketed successfully in the private sector. In the best previous evaluation using experimental data, Leibowitz, Buchanan, and Mann (1992) were limited not only to using data from a Medicaid only managed care plan but also were limited to studying a plan so small in enrollment that the data are unlikely to generalize to other situations.

Although our research will not correct every problem with the existing empirical research, we believe that what we do is informative because of quality of the data and the productive marriage of samples. Also notable is that our data cover a group of health plans operating in a mature managed care market with commercial enrollment. We contend that we are able to learn more from our data concerning how the market will operate when most, if not all, Medicaid recipients are in regularly functioning managed care arrangements.

3. Econometric Framework

There are two special features of the econometric framework we use to estimate the cost change of a switch to managed care from fee-for-service under Medicaid. One distinctive feature of our work is that we develop a synthetic difference-in-differences statistical design because of the idiosyncratic sampling feature of the Medicaid experiment we evaluate. A second distinctive feature of our research is to use count data regressions that have been purged of two complicating forms of heteroskedasticity: (1) overdispersion in the conditional mean and variance and (2) excess zeros because many people use no medical care (Lambert 1992, Greene 1994, Grootendorst 1995, Gurmu and Trivedi 1995, Mullahy 1995b). We now lay out the algebra of the conventional difference-in-differences regression setup, which aids in explaining how the sampling design in the Hennepin County experiment necessitates developing a synthetic difference-in-differences specification.

A. The Standard Difference-in-Differences Research Design

An obvious way to infer the behavioral effects of an experimental policy intervention is to compare the difference in outcomes over time for a control group to the difference in outcomes during the same period for a randomly selected treatment group. The difference-in-differences data organization is prevalent in social science research including psychology where it has been termed the untreated control group design with pretest and post test (Meyer 1995). The advantage of the difference-in-differences design is that the behavior change for the control group picks up any naturally occurring changes in behavior while the experimental group's behavior change reflects both the (same) naturally occurring change in behavior plus the impact of the intervention. A comparison of the changes in behavior for the two otherwise demographically and economically homogeneous groups reveals the behavioral impact of the experimental intervention. Expressed algebraically in a linear in the parameters regression with control covariates x_{it} the difference-in-differences research design is

$$y_{it}^{j} = \alpha_0 + \alpha_1 d_t + \alpha^{1} d^{j} + \beta d_t^{j} + \gamma x_{it}^{j} + \varepsilon_{it}^{j} \qquad \textbf{(1)}$$

where y_{it} is use of a medical care mode by person i in either the pretest period ($t = 0$) or the post-test period ($t = 1$), and the superscript indicates either membership in the control group ($j = 0$) or the treatment group ($j = 1$). The ds are 0-1 binary variables indicating pre and post treatment periods for the control and treatment groups such that

$$\begin{Bmatrix} d_t = 1 \forall t = 1 \\ d^j = 1 \forall j = 1 \\ d_t^j = 1 \forall t, j = 1 \end{Bmatrix}. \quad \textbf{(2)}$$

Notice that the interaction binary variable d_t^j switches on to indicate the experimental group's medical care use after it is (randomly) switched into Medicaid managed care from fee-for-service. With exogenous random assignment to the managed care group so that $E(\varepsilon_{it}^j | d_t^j) = 0$, β would be zero in the absence of treatment.[3] Under truly random assignment of treatment an unbiased estimate of the average treatment effect is the estimated value of β, which is also (conditional on x_{it}) the average difference between the change in medical care use by the treatment group pre and post policy intervention and the change in the medical care use by the control group pre and post treatment period.

B. Synthetic Difference-in-Differences

The data we use for our examination of the effect of managed care on Medicaid health care utilization were collected as part of the Nationwide Evaluation of Medicaid Competition Demonstrations sponsored by the Health Care Financing Administration. The six state project (California, Florida, Minnesota, Missouri, New Jersey, and New York) involved implementation and assessment of alternative strategies for delivering and

---

[3] Heckman and Smith (1995) emphasize that the researcher should be careful to check that endogenous self selection does not invalidate the exogeneity of treatment group membership. Note that the estimated value of α_1 captures a common time trend in use by both the treatment and control groups. Similarly, the estimated value of α^1 captures any overall time-invariant difference in the use by the treatment and control groups not due to the covariates in x_{it}.

financing Medicaid benefits. We use data from Hennepin County (Minneapolis), Minnesota. Our data are unique because, unlike the other demonstration sites, Hennepin County's featured random assignment to prepaid plans and a panel structure for a portion of the data set in addition to cross-sections of controls and experimental Medicaid users, which led us to develop a synthetic difference-in-differences regression setup.[4] Our research represents the first time the Hennepin data have been used to estimate econometric models that take maximum advantage of the data the experiment generated.

The Three Samples. The core of our data is a stratified random sample of persons qualifying for Aid to Families with Dependent Children (AFDC) in 1986; the 86 percent response rate yielded a sample size of 307 people. All 307 AFDC eligibles were enrolled in FFS Medicaid in 1986 when they were surveyed concerning health and other personal attributes. The intention of the demonstration was to return 12–18 months later, after the 307 people had been enrolled in prepaid plans, and re-administer the survey. Due to various administrative problems the second survey was not administered for 30 months so that only 152 of the original 307 interviewees could be followed.

Because of the attrition the project staff selected an additional random sample in 1988, which included complete information for 194 people in FFS Medicaid.[5] The 194 persons added to the demonstration data in FFS Medicaid in 1988 plus the 155 usable people who were part of the 1986 survey but had attrited by 1988 combine to make up the control/comparison group for our synthetic difference-in-differences comparison.[6]

---

[4] The main drawback of the Hennepin County data is a lack of cost information; there also are no data regarding specialist referrals or prescription drug use.

[5] The second sample had a response rate of 85 percent. Overall the second sample had 430 people, 222 of which were randomly selected to have a prepaid plan in 1988. Section 4 discusses how we used the 1988 cross section of persons in Medicaid managed care in a robustness check of our synthetic difference-in-differences results.

[6] For all survey respondents one of their children was randomly selected and detailed information obtained for the randomly selected child. We only consider AFDC eligible women; there was only one male in the

In the synthetic difference-in-differences approach we calculate the average difference between the medical use in 1986 and 1988, before and after the random assignment of managed care, for a treatment group and a synthetic control group using the three samples from the Hennepin County experiment. Using regressions of the form of equation (1) we compare (i) the difference between the actual post and pre-intervention medical care use by the panel members and (ii) the difference between the 1988 medical care use by the supplemental cross section of FFS Medicaid users and the 1986 medical care use by the baseline cross section of FFS Medicaid users in 1986 who had attrited before the resurvey in 1988. Conditional on the covariates x_{it} our synthetic difference-in-differences estimate of the effect of Medicaid managed care on use is algebraically summarized as

$$\hat{\beta} = (\bar{y}_1^1 - \bar{y}_0^1) - (\bar{y}_1^{\tilde{0}} - \bar{y}_0^0), \qquad \textbf{(3)}$$

where the tilde over the superscript 0 indicates that two different cross sections were used in the baseline of how medical care use changed during the treatment period for persons not receiving the policy intervention. We provide a detailed look at the underlying assumptions behind synthetic difference-in-differences in the appendix.

Definitions for Hennepin variables and descriptive statistics appear in Table 1. We judge our synthetic difference-in-differences to be an informative regression specification based on the variable-by-variable t-test results displayed in Table 1. First, the descriptive statistics generally do not reject the null hypothesis that the 1986 (pre-managed care intervention) means for the panel (treatment group) are identical to the 1986 means for the baseline cross section (control group). Second, there is little statistical difference between

---

panel subset of the data set who we dropped for the sake of homogeneity. We also dropped seven men in the 1986 attrition group and 14 men in the 1988 nonenrollee cross section.

the means of the 1988 supplemental control group cross section and the 1986 control group cross section (of attriters), which leads us not to worry about possible attrition bias. Because all three samples we use are similar in terms of the mean values of the independent variables, and because the pre-intervention (1986) medical care use is statistically indistinguishable for the 1986 cross section and the panel group who were later switched to managed care in 1988, we believe that comparing the two separate cross sections to the panel sample in a synthetic difference-in-differences regression is informative. [Insert Table 1 here.]

Other Notable Aspects of the Survey. The recall period for physician office visits, OPD visits, and ER visits is the previous three to four months, and the recall period for inpatient use is the previous 12 months, which are short enough for reasonable recall accuracy with medical care use data. Our data values also are in the range of other surveys and therefore have face validity (Hurley, Freund, and Paul 1993).

Persons who enrolled in managed care in 1988 could select from six prepaid plans. Three plans were large well-established HMOs in the Minneapolis market at the time, and three others were smaller local HMOs. The managed care plans were capitated by 95 percent of the predicted FFS expenditures. Of persons randomly assigned managed care, about 60 percent chose one of the three large well-established HMOs, and about 40 percent of the persons in managed care chose one of the three smaller local HMOs. A statistical advantage of the Hennepin County experiment is that differing HMOs create heterogeneity in the economic treatment administered to the enrollees. We estimate econometric models parameterizing heterogeneous treatment to improve the accuracy of program intervention effect estimates, specifically whether the particular HMO membership mattered. An instrumental variables approach to self-selected treatment site is called for if our single equation count models suggest significant differences in cost containment by HMO type (Heckman 1995, Mullahy 1995a).

In 1992 there were about 31 million persons on Medicaid with about seven million or 22 percent AFDC adults, and about 15 million or 49 percent AFDC children. Of the $92 billion total vendor payments under Medicaid about $12 billion or 14 percent were for AFDC adults and about $15 billion or 16 percent were for children; the bulk of the remainder of AFDC vendor payments went to the aged, blind, and disabled (*Medicaid Statistics 1993*). In the aggregate, AFDC Medicaid enrollees, the group we examine with the Hennepin County experiment, are disproportionately costly relative to children but not costly relative to the expenditures for the aged, blind, and disabled.

A final point to note is that our data cover persons continuously eligible for AFDC during the 12 months preceding the 1986 (pre-intervention) survey and continuously eligible for AFDC during the 12 months preceding the 1988 (post-intervention) survey. How important are the relatively long-term AFDC eligibles among the Medicaid population? The Survey of Income and Program Participation (SIPP), a large random sample of the United States focusing on work history and use of the social safety net, provides insight into the intersection of the AFDC and Medicaid populations. The 1984 SIPP had over 229,000 observations of which over 25,000, or approximately 11 percent of the people, had ever enrolled in Medicaid during the 32 month observation period of the survey. About 9,000, or 37 percent of the people, enrolled in Medicaid in 1984 were also receiving AFDC. Virtually all persons receiving AFDC in 1984 had been enrolled in Medicaid at some time during the previous 32 months, and two-thirds of the AFDC recipients in 1984 had been enrolled in Medicaid during the entire 32 month survey period (Short, Cantor, and Monheit 1988).[7] From the joint use of AFDC and Medicaid as reported in the SIPP we conclude that most AFDC recipients who receive Medicaid are long-term welfare recipients so that our experimental group is representative of the typical AFDC Medicaid enrollee.

C. Count Data Regression Models

We estimate five regression models for each of the four modes of medical care studied: doctor visits, emergency room visits, hospital clinic and outpatient visits, and hospital inpatient visits. We now describe how each modeling approach offers successively more safeguards against misspecification of the conditional mean function and error structure that if present and ignored lead to inconsistent parameter estimates in limited dependent variable models.

Poisson and Negative Binomial Models. The well-known Poisson specification is a basic econometric model able to incorporate the discrete non-negative integer aspect of medical care utilization data. The distribution of medical care use under the Poisson is

$$f(y_i) = \frac{\lambda_i^{y_i} \exp(\lambda_i)}{y_i!}; \; i = 1, ..., \mathrm{n}, \qquad \textbf{(4)}$$

where y_i represents a discrete dependent variable, such as the number of doctor office visits. The typical parameterization of the Poisson sets

$$\lambda_i = \exp(x_i'b), \qquad \textbf{(5)}$$

where x_i is a collection of independent variables and b is the corresponding vector of parameters the investigator wants to estimate, including the program effect. A Poisson

---

[7] In addition, claims data from all sites in the Medicaid Competition Demonstrations indicate that 90 percent of all Medicaid eligibles were eligible for at least 11 months in a typical year (Paul, Sherrill, and Iannacchoine 1989).

distribution of use implies the restriction that the conditional mean equals the conditional variance, or

$$E(y_i|x_i) = \text{var}(y_i|x_i) = \lambda_i\,, \tag{6}$$

which is known as equidispersion.

Another popular count regression specification is the negative binomial model. The negative binomial loosens the deterministic specification by including a stochastic term in the parameter λ_i so that

$$\lambda_i = \exp(x_i'b + \varepsilon_i)\,, \tag{7}$$

where ε_i follows a gamma distribution. The distribution of medical care use under the negative binomial is

$$f(y_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(\alpha)\Gamma(y_i + 1)} \left(\frac{\alpha}{\phi_i + \alpha}\right)^{\alpha} \left(\frac{\phi_i}{\phi_i + \alpha}\right)^{y_i}, \tag{8}$$

where $\alpha > 0$ is known as the precision parameter, $\Gamma(\cdot)$ is the gamma function, and the conditional mean, λ_i, can be parameterized as the linear exponential in (5). The negative binomial distribution has the properties

$$E(y_i|x_i) = \phi_i \text{ and} \tag{9}$$

$$\text{var}(y_i|x_i) = \phi_i(1 + \delta\phi_i)\,, \tag{10}$$

where $\delta \equiv 1/\alpha$. The term δ is known as the overdispersion parameter, which permits the form of heteroskedasticity where the conditional variance exceeds the conditional mean, which is prevalent in count data. The negative binomial model collapses into the Poisson specification as δ approaches zero, which makes the negative binomial model a robust generalization of the Poisson specification.

Excess Zeros. As emphasized earlier, it is common for persons to use none of a narrowly defined medical care mode during a year. Specifically, during the two observation years 56 percent of the treatment group had no doctor office visits, 79 percent no hospitalizations, 82 percent no emergency room visits, and 86 percent of the treatment group had no outpatient department visits. Statistically modeling the outcome of many people using no care, known as the problem of excess zeros in count regression, necessitates hypothesizing a specific form for the overdispersion created by the so-called excess zeros. The particular form of the conditional variance we use is

$$\text{var}(y_i) = \lambda_i(1 - q_i)(1 + \lambda_i q_i), \qquad \textbf{(11)}$$

where λ_i is the mean of the Poisson process and q_i is the probability of observing a structural zero, which is a choice not to consume any of a mode of care. As q_i approaches zero in (11) the variance collapses into the well-known Poisson variance, λ_i. The general economic theory of demand says the q_i term should be parameterized with the same covariates as is λ_i. When using a logit specification for the binary choice probability

$$q_i = \frac{\exp(x_i'(\tau\beta))}{1+\exp(x_i'(\tau\beta))} = \frac{\lambda_i^\tau}{1+\lambda_i^\tau}, \qquad \textbf{(12)}$$

where τ is the parameter reflecting the proportionate difference between the parameters of the count model and the logit parameters. Correctly including excess zeros into the estimation process is necessary for consistent and efficient estimates of the effect of a switch to managed care on care use across modes.

It is possible that the Poisson portion of the model also has the standard form of overdispersion, gamma heterogeneity, leading to the negative binomial count regression specification. In the case of excess zeros coupled with gamma heterogeneity the error variance is

$$\mathrm{var}(y_i) = \lambda_i(1 - q_i)[1 + \lambda_i(q_i + \delta)], \qquad \textbf{(13)}$$

where δ is the familiar overdispersion parameter in the negative binomial model. Once again the count data regression model is robust in the sense that as q_i and δ approach zero in (13) the specification collapses into the simple Poisson.[8]

In zero altered models, the conditional mean is

$$\mathrm{E}(y_i \mid x_i) = \mathrm{E}(y_i \mid x_i, z_i = 0)\ \mathrm{Prob}(z_i = 0) + \mathrm{E}(y_i \mid x_i, z_i = 1)\ \mathrm{Prob}(z_i = 1), \qquad \textbf{(14)}$$

where z_i represents an indicator of whether or not a structural zero is observed. Continuing,

$$\mathrm{E}(y_i \mid x_i) = 0\ q_i + \mathrm{E}(y_i \mid x_i, z_i = 1)\ (1 - q_i), \qquad \textbf{(15)}$$

because y_i is always zero in the case of a structural zero, and the probability of observing a structural zero is defined as q_i. Conditional on not observing a structural zero the data are assumed to follow a Poisson or negative binomial distribution so that

$$\mathrm{E}(y_i \mid x_i) = \lambda_i\ (1 - q_i). \qquad \textbf{(16)}$$

Zero altered models provide the opportunity to decompose the program effect of managed care on Medicaid utilization into (1) the effect of the program on the probability

---

[8] It is possible that the zero inflated count models also have additional (extrinsic) heteroskedasticity, which extends beyond the intrinsic heteroskedasticity in the variance functions specified in (11)–(13). We checked for possible extrinsic heteroskedasticity by generalizing the overdispersion parameter, δ, in the ZINB model by making it a linear exponential function of personal attributes — age, household income, and the number of chronic conditions. We selected the arguments in the skedasticity function to capture the most likely remaining sources of subtle nonlinearity in the regression model (Grootendoorst 1995). In general, none of the estimated coefficients in the skedasticity function appended to the ZINB models were significantly different from zero.

of observing a structural zero, and (2) the effect on the level of utilization conditional on not observing a structural zero. Formally,

$$\begin{aligned} \frac{\partial E(y_i|x_i)}{\partial x_i} &= (1-q_i)\frac{\partial E(y_i|x_i, z_i = 1)}{\partial x_i} + E(y_i|x_i, z_i = 1)\frac{\partial(1-q_i)}{\partial x_i} \\ &= (1-q_i)\lambda_i\beta' - \lambda_i q_i(1-q_i)\tau\beta' \\ &= (1-\tau q_i)(1-q_i)\lambda_i\beta'. \end{aligned} \qquad \textbf{(17)}$$

We will use the decomposition in (17) to present results displaying the impact of managed care on the decision to use any of a particular mode of care as well as the impact on quantity of care used once the mode is chosen.

4. Econometric Estimates

Our four outcomes of interest are the numbers of doctor visits, hospital outpatient department visits, hospital emergency room visits, and hospital inpatient days. We are concerned whether $\hat{\beta}$, the estimated coefficient of the enrollment period-treatment group dummy variable, d_t^j, in the specification of the conditional mean in (1), which measures the effect of a random switch into Medicaid managed care out of fee-for-service, differs statistically and economically from zero. Because the count regression models we examine are largely nested within one another our organization scheme is to begin with the simplest count model, the stand-alone Poisson, and then proceed through the results for increasingly more elaborate specifications of the conditional mean and error structure. We conclude by discussing how much difference the complex count models make to the answer to our research question and whether the estimated treatment effect differs significantly by HMO site.

A. Program Effect Estimates from the Basic Count Models

Tables 2 and 3 display results from the two basic count data models — Poisson and negative binomial. The first column of Table 2 shows no significant change in doctor visits following a switch to managed care. Other results of interest for their conformity with prior expectations, which is an informal model specification check, include that schooling positively affects doctor visits, African-Americans have fewer office visits, children reduce the mother's office visits, and greater functional limitation leads to additional doctor office visits. Regarding outpatient department services, a summary of results for the statistically significant control variables are that African-Americans use more OPD services, and additional chronic conditions and functional limitation raise OPD use. Most importantly, Table 2 illustrates how the introduction of managed care results in a statistically significant drop in outpatient department use. The fall in OPD visits is sizable, 80–90 percent of the 1986 (pre-intervention) average, and indicates that managed care has the intended effect of restraining utilization of a relatively costly mode of health care. [Insert Table 2 here.]

Results for emergency room (ER) use appear in the third column of Table 2. No significant program effect on emergency room use appears as was the case for doctor visits. Indeed, the only statistically significant coefficients in the Poisson specification show age negatively associated with ER visits and additional acute conditions positively associated with ER visits. The finding of no managed care effect on emergency room use runs counter to many previous findings in the literature (Hurley, Freund, and Taylor 1989a, 1989b). Our final outcome of interest, Poisson results for inpatient days, appears in the last column of Table 2. The simple Poisson specification suggests a marginally significant program effect that is an increase in inpatient use. From Table 2 alone it is unclear how potential overdispersion will affect the parameter estimate.

Table 3 presents results from negative binomial regressions for all four modes of care. The negative binomial results are similar to the Poisson results presented in Table 2. In the first column of Table 3 the overdispersion parameter, δ, is significantly different from zero, which rejects the Poisson model for doctor visits. The results for doctor visits are robust across the basic count specifications in Tables 2 and 3; from a policy perspective the focal result is the lack of any change in doctor visits following a switch to Medicaid managed care. The estimates in Table 3 also locate overdispersion that rejects the Poisson model for outpatient department (OPD) visits. As in Table 2 there is a statistically significant drop in utilization after a switch to managed care under the negative binomial specification; the magnitude of the program effect in Table 3 is nearly identical to that in Table 2. [Insert Table 3 here.]

In contrast to the previous two cases of doctor and outpatient department visits emergency room use regressions do not locate overdispersion rejecting the simple Poisson count model against the negative binomial alternative. As before, no change in ER visits is observed under managed care. For inpatient days, overdispersion in the data rejects the Poisson model. Additionally, in the negative binomial model presented in the last column of Table 3 there is no significant change in inpatient use after a switch to managed care. The result that inpatient days do not change due to managed care is likely due to the experiment's enrollment of only persons qualifying for AFDC (Freund, et al. 1989). Most hospital utilization for the mothers and children who qualify for AFDC is birth related, which is unlikely to change if the number of pregnancies does not change.

As noted in Section 2 a substantial proportion of people (56–86 percent depending on mode) use no medical care. Because the problem of excessive zeros is so obvious in the Hennepin County experiment we estimate a variety of zero inflated count models.

B. Program Effect Estimates from the Zero Inflated Count Models

Tables 4 and 5 each display results for the four modes of care from two two-part models — zero inflated Poisson (ZIP) and zero inflated negative binomial (ZINB). Unlike Grootendorst (1995), our model specification includes all independent variables in both the first-part logit portion of the model accounting for the binary choice between using a particular health care mode or not and the second-part count potion of the model accounting for the level of utilization. Based on the economic theory of demand there is no a priori reason to assume that any of the independent variables would affect one part of the use outcome and not the other. Note that the first-stage logit coefficient estimates can be derived by multiplying the tabulated coefficients by the estimated value of τ. Also note that the derived logit coefficients indicate the probability of a person using none of a particular mode of care.

The Vuong (1989) statistic at the bottom of Table 4 tests the non-nested hypothesis whether the zero inflated Poisson model provides a better fit than the simple Poisson model. Values of the Vuong statistic exceeding 1.96 reject the Poisson model. Note that with the exception of ER visits all Vuong statistics reject the Poisson in favor of the zero inflated Poisson. Estimates of the two-part zero inflated Poisson models in Table 4 show outpatient department visits falling significantly under managed care but by a fraction of the amount estimated in the one-part regression functions of the parallel count models in Tables 2 and 3. According to Table 4 the extent of the decline in outpatient department visits is also more than proportionally offset by a significant increase in the amount of inpatient use under managed care. Doctor visits and emergency room use do not change significantly due to a switch to managed care in the Poisson zero inflated count model estimates displayed in Table 4. [Insert Table 4 here.]

Comparing Table 4 to Table 5 underscores that accounting for excess zeros in a Poisson specification insufficiently controls for overdispersion in three of the four modes of medical care use; only emergency room visits do not display overdispersion. Table 5 also shows that when controlling for overdispersion beyond that resulting from excess zeros with a negative binomial in the second stage the significant drop in OPD visits again appears under managed care but at a marginal level of significance; no other mode changes significantly under managed care. The Vuong statistic rejects the basic negative binomial specification in favor of the zero inflated negative binomial model for all modes of care. [Insert Table 5 here.]

In Table 6 we present the decomposition of the program effect developed in equation (17) into the contribution from the binary portion versus the count part of the zero inflated specification. We do program effect decompositions for both the ZIP and ZINB models for all four modes of care. In both the zero inflated Poisson and negative binomial specifications doctor visits show little program effect from the binary portion of the model, but a small drop in the number of visits conditional on not observing a structural zero. Across models Table 6 there is an increase in non-use of OPD services leading to a drop of between 0.18 and 0.23 visits. The ZIP model indicates an increase in utilization given that the person uses some OPD services, while the ZINB model indicates a decrease in utilization given that the person uses some OPD services. Because the ZINB model dominates the ZIP model due to overdispersion, we place more weight on the ZINB results. As expected from the simpler models ER use shows little change through either avenue of the two-part models. Across specifications IP use shows a marked increase in the incidence of non-use, resulting in an estimated per-person decrease in utilization of between 0.23 and 0.28 days annually from the increase in structural zeros. Curiously, when conditioning on not observing a structural zero with both the ZIP and ZINB specifications there is a marked rise in the estimated length of stay under managed care ranging from 0.33 to 0.42 additional hospital days annually. Although without greater data detail at the episode level such a conclusion cannot be confirmed the longer average hospital stays for hospital inpatient service users under managed care might be partly explained by greater use of outpatient surgery, such that the average amount of inpatient use rises because some below average users of hospital inpatient care are induced to shift to outpatient care.

Summary. Results from the two-part models suggest that significant changes in utilization come in the form of lower levels of hospital outpatient department usage under managed care.

C. Additional Model Specification Results

The treatment effects we estimate and present in Tables 2–6 do not disaggregate by HMO. As noted earlier, a person could select one of six different HMOs if picked to switch into managed care out of fee-for-service. It is important for the policymaker with scarce program dollars to allocate to know whether some kinds of HMOs are more effective than others in limiting the use of relatively high-cost modes of care.

Additionally, the program effect estimates we present come from specialized count data regression models that incorporate multiple avenues for heteroskedasticity that if present and ignored cause inconsistent parameter estimates. Because of the complexity of our estimating equations it seems informative to determine ex post how much practical difference the specialized count models made to the results.

Third, we develop a synthetic difference-in-differences statistical design. Cross-sectional comparisons are the focus of much of the econometric literature on treatment effects (Manski 1995, Heckman and Smith 1995). It would be interesting to know how estimates that use the ability of longitudinal data to condition out person specific latent heterogeneity differ from program effect estimates based on point-in-time comparisons of outcomes for a treatment versus a control group.

We now conclude our presentation of parameter estimates with a discussion of the consequences of the three model specification decisions that ex ante matter most to the estimated program effect — degree of aggregation of the treatment effect variable, the use of specialized count regressions models, and the difference-in-differences design.

HMO Heterogeneity. Possible differential effectiveness across HMOs in cost containment success admits the possibility for policymakers to target enrollments into the plan(s) with the greatest cost savings. Moreover, because persons chosen (randomly) for managed care could in turn choose their HMOs, any differential plan effectiveness detected signals the need for model respecification where the selection of membership in the treatment group is exogenous but the selection of specific HMO for treatment is not exogenous.

In models not tabulated we disaggregated the enrollment (treatment effect) variable according to which of the six possible HMOs serviced the treatment group member. We found insignificant variability of program effects across managed care providers. We conclude that for Hennepin County treatment heterogeneity is unimportant; no HMO was significantly more effective than the others in reducing utilization under Medicaid.[9]

Count Regressions Versus Ordinary Least Squares. On theoretical grounds the specialized count data regression models we use are econometrically preferable. Ex post it may not be the case that the more complex count regression models estimates lead to different conclusions than simple ordinary least squares coefficients. As a check on how much difference it makes in our case to use specialized count models we ran OLS regressions for the four outcomes using the synthetic difference-in-differences statistical design of Tables 2–5. Similar to the specialized count models the estimated program effects from OLS are insignificant for the numbers of doctor visits, emergency room visits, and inpatient stays. Also similar to the count models the estimated treatment effect for outpatient department visits is significantly negative; the estimated marginal effect of managed care on OPD use is about –0.43, which is slightly larger than the estimated program effect from the preferred count model for OPD visits, the zero inflated negative binomial in Table 5. For the Hennepin county experiment there is little difference between OLS and the results from the substantially more complex count regressions in terms of statistical significance of the estimated program effects, but the magnitudes of the

[9] As additional verification of the lack of differential cost savings across HMOs we aggregated the six HMOs into their two naturally occurring groups, larger well established HMOs and smaller newer HMOs. Again we found no difference between the estimated program effects for the two different broad HMO types.

statistically significant estimated program savings from OLS are slightly larger than from the specialized count models.

Difference-in-Differences Versus Cross-Sectional Comparisons. As a final check of our results before proceeding to simulations examining their economic significance we produced program effect estimates from cross section regressions comparing the 1988 use by the randomly selected treatment group to the control group. Count regression models estimated with the 1988 (post-intervention) cross section show no significant program effects for all four modes of medical care.

What matters most to our results is the difference-in-differences design rather than the specialized count regression models or the degree of aggregation of the HMO variable. The result that using the available panel dimension is key to our finding is of general importance. Much of the current debate over whether experimental data are a necessary improvement over survey data takes the use of a post treatment cross section as given (Heckman and Smith 1995, Manski 1995). Our results point to the value of experiments with longitudinal surveys of the treatment and control groups, which allow the researcher to condition out latent individual specific heterogeneity.

4. Simulations Examining Program Effect Economic Significance

Economists have recently criticized themselves for devoting too much attention to statistical significance and too little attention to the more important policy issue of the economic significance of empirical results (McCloskey and Ziliak 1996). Having discussed the statistical significance and magnitudes of the estimated program effects for four modes of care use we now examine the economic significance of our results, which we judge in the context of simulated cost savings of switching AFDC Medicaid recipients from fee-for-service to managed care.

A. Background

The focal policy issue of our research is whether managed care saves an economically significant amount when applied to Medicaid. The Medicaid competition demonstration evaluation collected no cost or payment data specific to Hennepin county. As an approximation we use estimates of cost of care by mode for people continuously enrolled in AFDC during 1987 in Maryland's Medicaid system (Stuart, et al. 1990). The AFDC status and year for Maryland fit nicely with our Hennepin data. We acknowledge up front that the Maryland Medicaid experience is not identical to Minnesota's.[10] Because we examine proportionate payment (expenditure) changes it is only necessary that the relative prices be similar in Maryland and Minnesota for our calculations to be an informative estimate of the percentage savings possible from a switch to managed care.[11]

Stuart, et al. (1990) also found that the average payment for an office-based physician visit was $34, the average payment for an OPD visit was $77, and the average payment for an ER visit was $47.[12] The relative payment values for Maryland confirm previous findings that average payments to outpatient departments exceed average

---

[10] Using a 50 percent sample of the population of AFDC enrolled Medicaid recipients, Stuart, et al. (1990) found that ambulatory care users are about two and a half more times likely to use an office based physician than outpatient department based care, and about three times more likely to use an office based physician than emergency room based care. In the Hennepin County data for 1986, doctor office visits were about three times more likely than OPD visits, and about eight times more likely than ER visits. Although ER usage is not similar in proportion in Maryland and Minnesota, OPD usage is similar in both states, which is most important for our purposes because doctor office visits and OPD visits are where we find the most evidence of reductions under managed care.

[11] The savings we estimate is limited to Medicaid payments for adults, which are roughly 14 percent of total Medicaid payments. We do not consider overhead and other administrative payments in our simulation exercises.

[12] Because average ER payments seem unbelievably low in Stuart, et al. (1990) we do sensitivity checks by first setting the average ER payment to $47 then doubling it to $94.

payments to office-based physicians (see also McDevitt and Dutton (1989) for Michigan's Medicaid system). Additionally, we use the 1985 average Medicaid inpatient per diem reimbursement rate for hospitals in Minnesota, $311, as a measure of the payments for providing IP services to Medicaid users (*Medicaid Source Book, 1988*, p. 463). To summarize, we take Maryland's average payment figures for ambulatory services and the Medicaid inpatient per diem rate for Minnesota hospitals as representative of the proportionate payment differences among physician office visits, OPD visits, ER visits, and IP visits to form proportionate payment estimates under fee-for-service versus managed care.

In policy regime change simulations examining the economic significance of a switch from fee-for-service to managed care under Medicaid we consider four alternative econometric count model results — (i) the simplest of all models, ordinary least squares, (ii) the simplest of all specialized count models, Poisson, (iii) the most highly parameterized model we estimated, zero inflated negative binomial, and (iv) a hybrid simulation model, which uses the statistically most preferred model's estimated program effect for each of the four modes of care.[13] To construct the hybrid model we used the Vuong statistics and estimated overdispersion parameters in Tables 2–5, permitting the data for each mode of care to select the most preferred of the different specialized count models. It is reasonable to use each of the four modes statistically dominant program effect estimates in an economic simulation to see what difference it makes. Based on the results for the estimated variance function the statistically dominant model for doctor office visits is the ZINB, the dominant model for outpatient department visits is also the ZINB, the dominant model for emergency room visits is the Poisson, and the dominant

[13] Program effect estimates are annualized by multiplying the estimate by 12 divided by the average number of months of the ambulatory observation period, or about 3.5 months.

model for inpatient visits is the ZINB. We also use simple ordinary least squares for a baseline comparison. Table 7 presents the cost savings estimates from our simulation exercises. [Insert Table 7 here.]

B. Simulation Results

The estimates in Table 7 are the average percentage change in expenditures relative to the 1986 expenditures for the treatment group prior to their switch to managed care. The range of estimated program effects is a 5–12 percent savings from a switch from fee-for-service to managed care under Medicaid.

The Poisson results in the first row of Table 7 give a baseline for the cost savings from managed care. The Poisson model is the least econometrically complex of the specialized count models because the Poisson admits the fewest potential problems in the error structure. In the Poisson regressions outpatient department visits drop by about 1.2 visits per year, which is the most of any of the three specialized count models in Table 7 we use to simulate a switch to managed care; except for OPD all other use modes showed virtually no change in utilization. Using the Poisson coefficients the baseline expected effect of a switch to managed care is an annual payment decrease of about five percent.

The second simulation in Table 7 uses the most highly parameterized of the specialized count models, the zero inflated negative binomial. In the negative binomial regressions taking account of non-use again only the program effect for outpatient visits is statistically significant. OPD use in the ZINB based simulations drop by about 1.3 visits per year, which is about the effect of a switch to managed care in the simple Poisson model in row 1. The ZINB model is statistically more conservative than the simple Poisson because the ZINB in row 2 controls for overdispersion and excess zeros. The estimated economic impact of a switch to managed care is a cost savings of 12 percent when computed with parameter estimates from the highly parameterized ZINB for all for medical care use modes.

The third row of Table 7 gives simulation results using the statistically dominant count regression specification for each of the four separate medical care modes we study. The only statistically significant coefficient in the hybrid simulation model in row 3 is the program effect for outpatient visits, which comes from the ZINB specification. The simulated OPD effect from the ZINB is slightly smaller than the simulated OPD program effect from the OLS estimate, but slightly larger than the Poisson estimate. Because the estimated ER use effect is somewhat larger in its econometrically preferred specification, the Poisson, the hybrid model's estimated cost savings are slightly smaller than in the ZINB, about 10–11 percent.

Finally, we apply the results yielding the largest estimated program impact, ordinary least squares. OLS estimates show a fairly large, albeit insignificant, reduction in doctor office visits as well as a reduction in emergency room use to go with the large estimated reduction in outpatient department use. The slightly larger estimated marginal effects under OLS make the simulated program effect in row 4 in the range 11–12 percent. The OLS estimates in general tend to be larger in absolute value than the specialized count data models because OLS imposes a simple linear relationship between the covariates and the dependent variable. We emphasize that the cost savings from OLS is presented only as an additional benchmark, and that the model selection procedures generally point to the value of a count model with a specification giving special attention to so-called excess zeros.

Discussion. The extent of cost savings hinges crucially on any movement out of relatively expensive hospital inpatient use and outpatient department based care into relatively cheaper physician office based care. The value of average payment for an ER visit matters only slightly in the predicted proportionate total payment change because there are so few emergency visits for the Hennepin sample members. Based on model selection tests and the simulations in Table 7 we conclude that cost savings of about 10 percent are possible given a switch from fee-for-service to managed care under Medicaid.

Federal and sate governments together spend about $90 billion annually in the U.S. on the Medicaid program. Because Medicaid is the fastest growing item in the typical state budget and nationally consumes a large fraction of discretionary expenditures, legislators are anxious to find cost savings. Although it is common to hear policymakers long for cost savings of about a third, the 10 percent savings we find is an economically significant sum. If every state had a well designed program of managed care as we studied for Minnesota then the U.S. could save about $9 billion per year. Other programs would generally be able to save such an amount only by eliminating major components or beneficiaries.

5. Conclusion

We find that as Medicaid programs move from fee-for-service to managed care the expected maximum program payment savings is about 10 percent. Although our estimates are similar to some in the literature (Freund et al. 1989) they are much smaller than the cost savings found in studies of waiver applications performed by the states for the Health Care Financing Administration (Hurley, Freund, and Paul 1993).

Our modeling procedures take advantage of a unique data set that longitudinally surveyed a treatment group before and after a switch from fee-for-service into managed care under Medicaid. The data set we used also contains several homogeneous cross section control groups, which we used to develop a synthetic difference-in-differences statistical design. Of note is that the signs and statistical significance of estimated program effects on medical care use across the four modes studied are similar when estimated by ordinary least squares and complex count data models correcting for multiple sources of heteroskedasticity. Our preferred estimate, 10 percent savings from the hybrid simulation model, is roughly double the basic Poisson and slightly lower than the cost savings estimated by simulations using either OLS or the most complex count model, the ZINB, results.

More important than the regression estimator to our results is the statistical design. Using longitudinal data with latent individual heterogeneity conditioned out yields an estimated program effect of about a 5–12 percent cost savings from universal adoption of Medicaid managed care. Using the difference in use from the post-intervention cross section of the (randomly selected) treatment versus control group indicates no statistically significant effect on utilization of medical care services under managed care. We believe the difference in program effects across statistical designs comes from necessarily unmodeled latent person-specific heterogeneity that produces biased estimates of the program effect from a switch to managed care when estimated with the 1988 (post-intervention) cross section data alone.

Because hospital inpatient use is the largest component of Medicaid outlays substantial cost reductions are impossible without control of inpatient use. We do not find evidence of decreased use of inpatient services under managed care. We also emphasize that our preferred 10 percent estimated cost savings is for the initial effect of switching to managed care. Cost changes can be difficult to pin down in a one-time experimental study because the capitation rate may be set inappropriately yielding an apparently higher cost to switching from fee-for-service to prepaid plans (Freund, et al., 1989, p. 86). A broader picture of the managed care effect would come from viewing the utilization of a Medicaid population over a period of several years, which would include the adjustment in capitation rates as HMOs become accustomed to the Medicaid population and permit estimation of managed care's ability to make people healthier or change health seeking behavior, in turn maintaining lower utilization over time.

Appendix

The age-old problem in social experimentation is the inability to observe an outcome of interest, such as utilization of medical care services, for the same set of persons simultaneously with and without the policy intervention. That is, the researcher would most like to know

$$E(Y_1) - E(Y_0) = \Delta, \tag{A1}$$

where Δ is the pure treatment effect, the subscript 1 indicates the presence of the policy intervention, and the subscript 0 indicates the absence of the policy intervention.

Instead, the usual cross-sectional experiments measure

$$E(Y_1|z = 1) - E(Y_0|z = 0) = \Delta + \text{error}, \tag{A2}$$

where z indicates membership in either the treatment group (z = 1) or membership in the other (untreated) group (z = 0).

Notice that in (A2)

$$E(Y_1) = E(Y_1|z = 1)\ Pr(z = 1) + E(Y_1|z = 0)\ Pr(z = 0) \quad \textbf{(A3)}$$

and

$$E(Y_0) = E(Y_0|z = 1)\ Pr(z = 1) + E(Y_0|z = 0)\ Pr(z = 0). \quad \textbf{(A4)}$$

When persons are randomly assigned to treatment and control groups, error = 0 in (A2) because

$$E(Y|z = 1) = E(Y_1|z = 1) = E(Y_1|z = 0) \quad \textbf{(A5)}$$

and

$$E(Y|z = 0) = E(Y_0|z = 1) = E(Y_0|z = 0). \quad \textbf{(A6)}$$

So altogether,

$$E(Y_1) = E(Y_1|z = 1) \quad \textbf{(A7)}$$

and

$$E(Y_0) = E(Y_0|z = 0). \quad \textbf{(A8)}$$

An alternative approach recognizes that we can observe the same people both with and without treatment when over time. That is, we can construct a pre-/post-treatment experiment designed to measure,

$$E(Y_1) – E(Y_0) = \Delta, \quad \textbf{(A9)}$$

with

$$E(Y_1|t = 1) – E(Y_0|t = 0). \quad \textbf{(A10)}$$

A difficulty is a possible secular trend in the outcome measure that will be confound measuring the true intertemporal program effect. That is, it may more properly be the case that

$$E(Y_1| t = 1) - E(Y_0| t = 0) = \Delta + k, \qquad \textbf{(A11)}$$

where k is an unknown constant additive secular trend. Purifying the common trend from (A11) requires estimating k. First, the researcher must make a formal assumption about the nature of the secular trend such as

$$E(Y|t = 1) - E(Y|t = 0) = k. \qquad \textbf{(A12)}$$

If another independent untreated sample (denoted by a prime) is available one can form an estimate of the intervention effect that is

$$[E(Y_1|t = 1) - E(Y_0|t = 0)] - [E(Y_{0'}|t = 1) - E(Y_{0'}|t = 0)] = \Delta. \qquad \textbf{(A13)}$$

Note the similarity of (A13) to a difference-in-differences estimator. One needs no new apparatus to cope with the added time dimension, and the results follow with only notation changes from cross-section model in (A1)–(A8).

Consider the first difference in equation (A13):

$$E(Y_1|t = 1) = E(Y_1|t = 1,z = 1) \Pr(z = 1|t = 1) + E(Y_1|t = 1,z = 0) \Pr(z = 0|t = 1) \qquad \textbf{(A14)}$$

and

$$E(Y_0|t = 0) = E(Y_0|t = 0,z = 1) \Pr(z = 1|t = 0) + E(Y_0|t = 0, z = 0) \Pr(z = 0|t = 0). \qquad \textbf{(A15)}$$

As before, under random selection

$$E(Y|t = 1, z = 1) = E(Y_1|t = 1, z = 1) = E(Y_1|t = 1, z = 0) \qquad \textbf{(A16)}$$

and

$$E(Y|t = 0, z = 0) = E(Y_0|t = 0, z = 1) = E(Y_0|t = 0, z = 0). \qquad \textbf{(A17)}$$

Altogether

$$E(Y_1|t = 1) = E(Y_1|t = 1, z = 1) \quad \textbf{(A18)}$$

and

$$E(Y_0|t = 0) = E(Y_0|t = 0, z = 0). \quad \textbf{(A19)}$$

Note that we observe $E(Y_1|t = 1)$ and $E(Y_0|t = 0)$. When the same people are observed at both points in time the researcher does not need to worry about the potential for selection biases. What the researcher cannot observe using a single difference in outcomes is an underlying trend, which make us consider estimation of the second difference in equation (A13).

The data issues concerning a control group over time are directly analogous to the treatment group issues we have just discussed. The key to a viable measure of the time trend, k, is a random sample at the two points in time. It is less important that the survey involve the same people over time because no treatment is ever administered to the control group. We can rewrite the difference-in-differences estimator to reflect the fact that the second difference (secular trend estimator) need not involve the same people over time

$$[E(Y_1|t = 1) – E(Y_0|t = 0)] – [E(Y_{0''}|t = 1) – E(Y_{0'}|t = 0)] = \Delta, \quad \textbf{(A20)}$$

which we call synthetic difference-in-differences because the double prime denotes a second control group.

The manipulations of the synthetic difference-in-differences design follow closely what we have above.

$$E(Y_{0''}|t = 1) = E(Y_{0''}|t = 1,z = 1) \Pr(z = 1|t = 1) + E(Y_{0''}|t = 1,z = 0) \Pr(z = 0|t = 1) \quad \textbf{(A21)}$$

and

$$E(Y_{0'}|t = 0) = E(Y_{0'}|t = 0,z = 1)\ Pr(z = 1|t = 0) + E(Y_{0'}|t = 0, z = 0)\ Pr(z = 0|t = 0). \quad \textbf{(A22)}$$

As before, under random selection,

$$E(Y|t = 1, z = 0) = E(Y_{0''}|t = 1, z = 1) = E(Y_{0''}|t = 1, z = 0) \quad \textbf{(A23)}$$

and

$$E(Y|t = 0, z = 0) = E(Y_{0'}|t = 0, z = 1) = E(Y_{0'}|t = 0, z = 0). \quad \textbf{(A24)}$$

Altogether,

$$E(Y_{0''}|t = 1) = E(Y_{0''}|t = 1, z = 0) \quad \textbf{(A25)}$$

and

$$E(Y_{0'}|t = 0) = E(Y_{0'}|t = 0, z = 0). \quad \textbf{(A26)}$$

Putting our observables together we have

$$E(Y_{0''}|t = 1) – E(Y_{0'}|t = 0) = k. \quad \textbf{(A27)}$$

Subtracting (A27) from (A11) gives (A20), our synthetic difference-in-differences estimator.[14] In the Hennepin County data set the sample denoted by 0' is a potential sore spot because it is the group of attriters from the original pre-/post-survey. As such, we note the additional requirement that the attrition could reasonably be viewed as random or ignorable.

[14] The synthetic difference-in-differences design is perhaps better called semi-synthetic. A completely synthetic difference-in-differences design would involve four separate cross-section samples, two per pre- and post intervention period, where one group in the second period gets the policy intervention.

References

Freund, Deborah A. and Robert E. Hurley, “Medicaid Managed Care: The Contribution to Issues of Health Reform,” *Annual Review of Public Health*, 1995, 16, pp. 473–496.

Freund, Deborah A., Louis F. Rossiter, Peter D. Fox, Jack A. Meyer, Robert E. Hurley, Timothy S. Carey, and John E. Paul, “Evaluation of the Medicaid Competition Demonstrations,” *Health Care Financing Review,* Winter 1989, 11:2, pp. 81–97.

Greene, William H., “Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models,” Working Paper EC-94-10, Department of Economics, New York University, March 1994.

Grootendorst, Paul V., “A Comparison of Alternative Models of Prescription Drug Utilization,” *Health Economics*, 1995, 4, pp. 183–198.

Gurmu, Shiferaw and Pravin K. Trivedi, “Recent Developments in Models of Event Counts,” Working Papers in Economics, Indiana University, Bloomington, 1995.

Heckman, James J., “Instrumental Variables: A Cautionary Tale,” NBER Technical Working Paper No. 185, September 1995.

Heckman, James J. and Jeffrey A. Smith, “Assessing the Case for Social Experiments,” *Journal of Economic Perspectives*,” Spring 1995, 9:2, pp. 85–110.

Hurley, Robert E., Deborah A. Freund, and John E. Paul. *Managed Care in Medicaid: Lessons for Policy and Program Design.* Ann Arbor: Health Administration Press, 1993.

Hurley, Robert E., Deborah A. Freund, and Donald E. Taylor, Gatekeeping the Emergency Room Department: Impact of a Medicaid Primary Care Case Management Program,” *Health Care Management Review*, Spring 1989a, 14:2, pp. 63–72.

Hurley, Robert E., Deborah A. Freund, and Donald E. Taylor, “Emergency Room Use and Primary Care Case Management: Evidence from Four Medicaid Demonstration Programs,” *American Journal of Public Health*, July 1989b, 70:7, pp. 843–847.

Lambert, Diane, “Zero-Inflated Poisson Regression With an Application to Defects in Manufacturing,” *Technometrics*, 1992, 34:1, pp. 1–14.

Leibowitz, Arleen, Joan L. Buchanan, and Joyce Mann, “A Randomized Trial to Evaluate the Effectiveness of a Medicaid HMO,” *Journal of Health Economics*, 1992, 11, pp. 235–257.

Manski, Charles F. *Identification Problems in the Social Sciences*. Cambridge, MA: Harvard University Press, 1995.

McCloskey, Deirdre N. and Stephen T. Ziliak, "The Standard Error of Regressions," *Journal of Economic Literature*, March 1996, XXXIV:1, pp. 97–114

McDevitt, R. D., and B. Dutton, "Expenditures for Ambulatory Episodes of Care: The Michigan Medicaid Experience," *Health Care Financing Review*, Winter 1989, 11:2, pp. 43–55.

Medicaid Source Book: Background Data and Analysis. Washington D.C.: Government Printing Office, 1988.

Medicaid Statistics: Program and Financial Statistics, Fiscal Year 1992. Washington, D.C.: U.S. Department of Health and Human Services, Health Care Financing Administration Medical Bureau, HCFA Publication #10129, October 1993.

Meyer, Bruce D., "Natural and Quasi-Natural Experiments in Economics," *Journal of Business and Economic Statistics*, 1995, 13:2, pp. 151–162.

Mullahy, John, "Instrumental Variable Estimation of Poisson Regression Models: Applications to Models of Cigarette Smoking Behavior," Working Paper, Department of Economics, Trinity College, February 1995a.

Mullahy, John, "Heterogeneity and the Structure of Count Data Models," Working Paper, Department of Economics, Trinity College, August 1995b.

Paul, John E., Terri A. Sherrill, and Vincent G. Iannacchoine, "Nationwide Evaluation of Medicaid Competition Demonstrations. Medicaid Consumer Survey Results, Minnesota Prepaid Demonstration Project, 1986–1988," Research Triangle Institute, May 1989.

Short, Pamela F., Joel C. Cantor, and Alan C. Monheit, "The Dynamics of Medicaid Enrollment," *Inquiry*, 1988, 25:Winter, pp. 504–516.

Stuart, Mary, Donald Steinwachs, Jennifer Harlow, and Michael Fox, "Ambulatory Practice Variation in Maryland: Implication for Medicaid Cost Management," *Health Care Financing Review*, 1990 Annual Supplement, pp. 57–67.

Vuong, Quang H., "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 1989, 57:2, pp. 307–333.

Table 1
Descriptive Statistics for Hennepin County Medicaid Data Set

	1986 Attriters (n = 155)	1988 Nonenrollees (n = 194)	1986 (Pre-Enrollment) (n = 152)	1988 (Post-Enrollment) (n = 152)
Variable: Definition	*Mean (Std. Dev.)*	*Mean (Std. Dev)*	*Mean (Std. Dev.)*	*Mean (Std. Dev.)*
DOCTOR: Doctor office and health center visits.	1.400 (2.462)	1.392 (2.871)	1.632 (4.162)	1.645 (3.546)
ER: Emergency room visits.	0.187 (0.453)	0.278 (0.671)	0.204 (0.493)	0.237 (0.561)
OPD: Hospital clinic and outpatient visits.	0.368 (1.455)	0.577 (1.895)	0.539 (1.783)	0.283 (1.136)
HOSPSTAY: Indicator of inpatient hospital admission.	0.213 (0.411)	0.222 (0.416)	0.197 (0.399)	0.184 (0.389)
LOS: Length of hospital stay (days).	1.529 (7.113)	1.126 (3.312)	0.717 (1.693)	0.809 (2.742)
SCHOOL: Number of years completed.	11.632 (2.344)	11.948 (1.587)	11.237 (2.716)	11.408 (2.448)
AGE: Age of survey member.	30.213 (8.478)	28.861 (6.928)	28.395† (7.393)	30.730 (7.386)
HHINC: Household income level.	6841.9 (3318.5)	7164.9 (2362.5)	6578.9 (2524.8)	7006.6 (2203.2)
CHRONIC: Number of chronic conditions.	1.761 (1.802)	1.866 (1.836)	1.704 (1.826)	1.934 (2.397)
ACUTE: Number of acute conditions.	4.181 (3.362)	3.784 (3.675)	4.132 (4.339)	3.382‡ (3.711)
PARTNER: Indicator if woman is married.	0.071 (0.258)	0.057 (0.232)	0.053 (0.224)	0.066 (0.249)
TOTCHIL: Total number of children in household.	2.239 (1.358)	2.196 (1.121)	2.237 (1.301)	2.382‡ (1.162)
HSTATUSI: Measure of functional ability, 0–7.	0.529 (1.191)	0.392 (1.143)	0.500 (1.297)	0.612 (1.447)
BLACK: Race indicator.	0.226 (0.419)	0.407* (0.493)	0.289 (0.455)	---------
EXPOSURE: Recall period for ambulatory care (mos.).	3.563 (0.307)	3.441 (0.325)	3.516 (0.300)	3.322‡ (0.469)
GROUP: Indicator if person was in experimental group.	0.0	0.0	1.0	1.0
TIME: Indicator if year was 1988.	0.0	1.0	0.0	1.0
ENROLL: Interaction term of GROUP and TIME.	0.0	0.0	0.0	1.0

* Indicates a statistically significant difference in means at the 5% level between the 1986 attriters and the 1988 nonenrollees.
† Indicates a statistically significant difference in means at the 5% level between 1986 attriters and nonattriters.
‡ Indicates a statistically significant difference in means at the 5% level between the pre- and post-enrollment groups.

Table 2
Poisson Regression Models — All Modes
(absolute values of t-statistics) [marginal effects]

	Doctor Visits	Outpatient Dept. Visits	Emergency Rm. Visits	Inpatient Days
Regressors	n = 653	n = 653	n = 653	n = 652
Constant	–0.352 (0.947)	–1.768 (2.303)	0.189 (0.163)	1.556 (4.794)
SCHOOL	0.037 (2.495) [0.045]	0.046 (1.407) [0.014]	0.058 (1.243) [0.011]	–0.032 (1.555) [–0.024]
AGE	–0.002 (0.482) [–0.003]	–0.054 (5.763) [–0.016]	–0.058 (4.092) [–0.011]	–0.078 (10.542) [–0.059]
BLACK	–0.774 (8.939) [–0.904]	0.550 (4.259) [0.165]	0.132 (0.752) [0.025]	0.231 (2.607) [0.173]
HHINC	–0.000007 (0.487) [–0.000008]	0.00006 (2.517) [0.00002]	–0.00002 (0.470) [–0.000003]	0.00005 (3.373) [0.00004]
CHRONIC	–0.023 (1.117) [–0.027]	0.204 (8.337) [0.061]	0.050 (1.096) [0.009]	–0.032 (1.177) [–0.024]
ACUTE	0.119 (12.588) [0.139]	–0.012 (0.704) [–0.004]	0.070 (2.835) [0.013]	0.053 (4.495) [0.044]
PARTNER	–0.093 (0.690) [–0.109]	1.234 (7.887) [0.371]	0.510 (1.866) [0.096]	0.495 (3.598) [0.371]
TOTCHIL	–0.212 (6.619) [–0.248]	–0.106 (2.043) [–0.032]	0.071 (1.016) [0.013]	0.133 (3.952) [0.100]
HSTATUSI	0.127 (5.094) [0.149]	0.233 (5.464) [0.070]	0.091 (1.394) [0.017]	0.026 (0.670) [0.019]
EXPOSURE	0.090 (1.024) [0.106]	0.254 (1.399) [0.076]	–0.484 (1.642) [–0.091]	--------
GROUP	0.142 (1.507) [0.167]	0.301 (1.718) [0.090]	0.005 (0.019) [0.001]	–0.434 (3.390) [–0.326]
TIME	0.123 (1.322) [0.143]	0.278 (1.660) [0.083]	0.326 (1.391) [0.062]	0.077 (0.724) [0.058]
ENROLL	0.009 (0.065) [0.010]	–1.167 (4.379) [–0.350]	0.073 (0.212) [0.014]	0.257 (1.497) [0.193]

Table 3
Negative Binomial Regression Models — All Modes
(absolute values of t-statistics) [marginal effects]

	Doctor Visits	Outpatient Dept. Visits	Emergency Rm. Visits	Inpatient Days
Regressors	n = 653	n = 653	n = 653	n = 652
Constant	–0.681 (0.795)	–2.392 (1.264)	0.090 (0.073)	2.065 (1.917)
SCHOOL	0.065 (1.978) [0.076]	0.005 (0.067) [0.002]	0.063 (1.248) [0.012]	–0.022 (0.391) [–0.016]
AGE	0.003 (0.255) [0.003]	–0.029 (1.336) [–0.010]	–0.056 (3.679) [–0.011]	–0.100 (4.452) [–0.073]
BLACK	–0.779 (4.623) [–0.914]	0.443 (1.337) [0.147]	0.143 (0.750) [0.027]	0.350 (1.212) [0.258]
HHINC	–0.00001 (0.524) [–0.00001]	0.00004 (0.665) [0.00001]	–0.00001 (0.360) [–0.000003]	0.00005 (0.953) [0.00003]
CHRONIC	0.033 (0.679) [0.039]	0.154 (1.939) [0.051]	0.043 (0.815) [0.008]	0.044 (0.484) [0.033]
ACUTE	0.116 (4.901) [0.136]	–0.004 (0.079) [–0.001]	0.068 (2.482) [0.013]	0.014 (0.354) [0.011]
PARTNER	0.071 (0.239) [0.083]	0.779 (1.371) [0.258]	0.415 (1.317) [0.079]	0.375 (0.670) [0.277]
TOTCHIL	–0.114 (1.717) [–0.134]	0.002 (0.014) [0.001]	0.063 (0.819) [0.012]	0.151 (1.349) [0.111]
HSTATUSI	0.118 (1.987) [0.139]	0.186 (1.344) [0.062]	0.086 (1.201) [0.016]	0.129 (0.998) [0.095]
EXPOSURE	0.007 (0.032) [0.008]	0.454 (0.986) [0.150]	–0.484 (1.563) [–0.092]	--------
GROUP	0.005 (0.024) [0.006]	0.152 (0.368) [0.050]	0.010 (0.037) [0.002]	–0.447 (1.198) [–0.330]
TIME	0.079 (0.392) [0.093]	0.051 (0.130) [0.017]	0.304 (1.203) [0.058]	0.068 (0.189) [0.050]
ENROLL	–0.098 (0.331) [–0.115]	–1.073 (1.703) [–0.355]	0.095 (0.256) [0.018]	0.051 (0.094) [0.038]
δ	0.767 (8.108)	2.192 (14.042)	–0.663 (1.133)	2.127 (17.811)

Table 4
Zero Inflated Poisson Regression Models — All Modes
(absolute values of t-statistics) [overall marginal effects]

	Doctor Visits	Outpatient Dept. Visits	Emergency Rm. Visits	Inpatient Days
Regressors	n = 653	n = 653	n = 653	n = 652
Constant	0.534 (1.270)	–0.233 (0.473)	0.643 (0.456)	0.449 (1.641)
SCHOOL	0.044 (0.235) [0.059]	0.741 (2.813) [–0.106]	1.228 (1.570) [0.191]	0.644 (2.899) [0.177]
AGE	0.066 (0.495) [0.089]	–0.238 (1.263) [0.034]	1.954 (3.437) [–0.303]	0.317 (1.950) [0.087]
BLACK	–0.650 (6.710) [–0.877]	–0.075 (0.917) [0.011]	0.218 (0.853) [0.034]	0.324 (6.030) [0.089]
HHINC	–0.055 (0.531) [–0.074]	0.247 (2.148) [–0.035]	–0.051 (0.172) [–0.008]	0.121 (2.105) [0.033]
PARTNER	–0.122 (0.862) [–0.164]	0.264 (2.747) [–0.038]	0.463 (1.258) [0.072]	–0.014 (0.152) [–0.004]
CHRONIC	–0.030 (0.723) [–0.040]	0.028 (0.570) [–0.004]	0.085 (0.624) [0.013]	–0.038 (1.138) [–0.010]
ACUTE	0.301 (7.779) [0.406]	–0.078 (1.499) [0.011]	0.291 (1.934) [0.045]	0.172 (6.315) [0.047]
TOTCHIL	–0.451 (5.634) [–0.608]	–0.053 (0.565) [0.008]	0.068 (0.346) [0.011]	–0.168 (4.101) [–0.046]
HSTATUSI	0.129 (5.065) [0.174]	0.101 (3.384) [–0.014]	0.085 (1.007) [0.013]	–0.088 (3.078) [–0.024]
EXPOSURE	0.427 (1.356) [0.576]	0.578 (1.529) [–0.083]	–1.936 (1.721) [–0.300]	--------
GROUP	0.171 (1.725) [0.230]	–0.073 (0.684) [0.010]	0.107 (0.346) [0.067]	–0.281 (2.710) [–0.077]
TIME	0.182 (1.846) [0.246]	–0.029 (0.291) [0.004]	0.320 (1.194) [0.050]	0.007 (0.098) [0.002]
ENROLL	–0.130 (0.965) [–0.176]	0.296 (1.875) [–0.042]	–0.083 (0.216) [–0.013]	0.354 (2.713) [0.097]
τ	–0.185 (2.104)	1.560 (13.070)	0.660 (1.530)	0.891 (13.105)
Vuong Statistic	6.127	3.300	0.496	6.619

Table 5
Zero Inflated Negative Binomial Regression Models — All Modes
(absolute values of t-statistics) [overall marginal effects]

	Doctor Visits	Outpatient Dept. Visits	Emergency Rm. Visits	Inpatient Days
Regressors	n = 653	n = 653	n = 653	n = 652
Constant	–0.034 (0.096)	0.158 (0.191)	0.729 (0.384)	–0.672 (0.814)
SCHOOL	0.251 (1.534) [0.391]	–0.129 (0.375) [–0.104]	1.212 (1.094) [0.188]	1.046 (1.672) [0.340]
AGE	0.179 (1.615) [0.278]	–0.533 (1.482) [–0.431]	–1.954 (2.698) [–0.304]	1.101 (2.236) [0.358]
BLACK	–0.260 (3.507) [–0.405]	0.270 (1.382) [0.218]	0.215 (0.821) [0.033]	0.308 (1.793) [0.100]
HHINC	–0.104 (1.573) [–0.161]	0.040 (0.252) [0.033]	–0.053 (0.173) [–0.008]	0.059 (0.345) [0.019]
PARTNER	–0.123 (0.841) [–0.191]	0.449 (1.300) [0.363]	0.461 (1.035) [0.072]	–0.050 (0.196) [–0.016]
CHRONIC	0.098 (1.945) [0.152]	0.159 (1.511) [0.128]	0.084 (0.592) [0.013]	–0.055 (0.536) [–0.018]
ACUTE	0.217 (3.780) [0.338]	0.016 (0.194) [0.013]	0.291 (1.707) [0.045]	0.202 (2.167) [0.066]
TOTCHIL	–0.024 (0.469) [–0.037]	–0.024 (0.171) [–0.019]	0.069 (0.228) [0.011]	–0.216 (1.545) [–0.070]
HSTATUSI	0.114 (1.878) [0.178]	0.075 (1.104) [0.061]	0.086 (0.831) [0.013]	–0.127 (1.566) [–0.041]
EXPOSURE	–0.110 (0.428) [–0.172]	0.039 (0.063) [0.032]	–2.002 (1.418) [–0.311]	--------
GROUP	0.000 (0.001) [0.000]	0.099 (0.590) [0.080]	0.110 (0.284) [0.017]	–0.283 (1.008) [–0.092]
TIME	–0.012 (0.161) [–0.019]	0.039 (0.253) [0.032]	0.321 (0.925) [0.050]	–0.008 (0.036) [–0.002]
ENROLL	–0.050 (0.448) [–0.078]	–0.470 (1.407) [–0.380]	–0.090 (0.184) [–0.014]	0.441 (1.230) [0.143]
τ	–15.775 (2.266)	–2.264 (1.251)	0.663 (0.858)	0.853 (9.514)
δ	0.744 (7.038)	1.267 (5.567)	–9.047 (0.003)	0.483 (1.582)
Vuong Statistic	5.994	4.475	5.192	5.940

Table 6
Decomposition of Program Effects for ZIP and ZINB Models

Zero-Inflated Poisson

Care Mode	Program Effect from Zero Portion	Program Effect from Count Portion
Doctor Visits	–0.014	–0.162
Outpatient Department Visits	–0.181	0.138
Emergency Room Visits	0.003	–0.016
Inpatient Days	–0.234	0.331

Zero-Inflated Negative Binomial

Care Mode	Program Effect from Zero Portion	Program Effect from Count Portion
Doctor Visits	–0.002	–0.076
Outpatient Department Visits	–0.228	–0.162
Emergency Room Visits	0.003	–0.017
Inpatient Days	–0.281	0.424

Table 7
Summary of Results of Payment Savings for Medicaid
Table Entries Represent Annual Changes in Utilization Per Person
Alternate Cost Values Used for Emergency Room Visits

ER Visit Cost = $47

Scenario	Doctor Office Visits ($34)	OPD Visits ($77)	ER Visits (**$47**)	Hospital Days ($311)	Percent Savings
1: Poisson	0.035	–1.228*	0.049	0.193	–5.20%
2: ZINB	–0.274	–1.334*	–0.049	0.143	–11.70%
3: Hybrid**	–0.274	–1.334*	0.049	0.143	–10.92%
4: OLS	–0.395	–1.508*	0.055	0.185	–11.63%

* Coefficient estimate is significantly less than zero at the 5% level.
** Uses the ZINB for Doctor Visits, OPD Visits, and Hospital Days; uses the Poisson for ER Visits.

ER Visit Cost = $94

Scenario	Doctor Office Visits ($34)	OPD Visits ($77)	ER Visits (**$94**)	Hospital Days ($311)	Percent Savings
1: Poisson	0.035	–1.228*	0.049	0.193	–4.56%
2: ZINB	–0.274	–1.334*	–0.049	0.143	–11.44%
3: Hybrid**	–0.274	–1.334*	0.049	0.143	–9.98%
4: OLS	–0.395	–1.508*	0.055	0.185	–10.60%

* Coefficient estimate is significantly less than zero at the 5% level.
** Uses the ZINB for Doctor Visits, OPD Visits, and Hospital Days; uses the Poisson for ER Visits.