

Estimating Heterogeneous Take-up and Crowd-Out

Responses to Existing Medicaid Income Limits and Their Nonmarginal Expansions

John C. Ham, University of Maryland, IFAU, IFS, IRP (UW-Madison) and IZA

Serkan Ozbeklik, Claremont McKenna College

Lara D. Shore-Sheppard, Williams College and NBER

Revised June 2013

We are grateful to Steven Cosslett, Curtis Eberwein, Geert Ridder, Yuxin Su, Sergio Urzúa, Frank Vella and seminar participants at Georgetown, the Institute for Research on Poverty, Ohio State and Princeton for comments. Weijia Dai provided excellent research assistance. This research has been supported by grants from the National Science Foundation (SES-0136928 and SES-0627968), the National Institute of Child Health and Human Development (R01 HD39369-01A1), and the Robert Wood Johnson Foundation. Any opinions, findings, and conclusions or recommendations in this paper are those of the authors and do not necessarily reflect the views of any government agency. Contact information: *John Ham*, ham@econ.umd.edu, Department of Economics, 3105 Tydings Hall, University of Maryland, College Park, MD 20742, phone: (202) 380-8806, fax: (301) 405-3542; *Serkan Ozbeklik*, serkan.ozbeklik@cmc.edu, The Robert Day School of Economics and Finance, Claremont McKenna College, 500 E. Ninth Street, Claremont, CA 91711, phone: (909) 607-0721, fax: (909) 621-8249; *Lara Shore-Sheppard*, Lara.D.Shore-Sheppard@williams.edu, Department of Economics, Williams College, 24 Hopkins Hall Drive, Williamstown, MA 01267, phone: (413) 597-2226, fax: (413) 597-4045.

Abstract

We use a switching probit model and the income-limit-based structure of Medicaid eligibility for children to estimate treatment effects of non-marginal Medicaid expansions on Medicaid take-up, private insurance coverage, and crowd-out, as well as crowd-out for those eligible for Medicaid under rules already in place. Many of these estimates are not found in existing work on public insurance and cannot be calculated with the linear probability model used by previous work in this literature. We provide an estimation approach that is straightforward to implement yet yields precise treatment effects.

Keywords: Medicaid expansions, take-up, crowd-out, treatment effects, switching probit model, counterfactual policy analysis, minimum distance estimation.

1. Introduction

In recent years, eligibility for public health insurance has expanded substantially, leading to a burgeoning research literature on the implications of such expansions for public insurance participation, private insurance coverage and crowd-out, as well as for overall levels of health insurance coverage. A common approach to these questions is to estimate a linear probability model (LPM) of participation (or private or overall insurance coverage) where a dummy variable for eligibility for the program is an endogenous explanatory variable and exogenous variation in eligibility is used to generate an instrumental variable.¹ This model permits the estimation of local average treatment effects (LATEs) on take-up, private insurance coverage, and crowd-out—the average effect of eligibility on insurance coverage among the individuals who are responsive to small changes in the instrument used for identification, the *marginally eligible*.

This approach, as generally implemented, has several drawbacks. The desirability of Medicaid and the availability of outside insurance options are likely to differ across families with different observable and unobservable differences, suggesting that these groups will have different responses to eligibility. Although one can easily allow treatment effects to differ by observable characteristics by interacting group dummy variables with the eligibility variable while treating these interactions as endogenous in a linear probability model, the linear probability model with interactions (LPMI) cannot be used to calculate levels of crowd-out among *currently eligible* individuals (defined as individuals eligible under the rules in place at any particular time) or take-up and crowd out for those made eligible by a *nonmarginal*

¹ See Cutler and Gruber (1996), Currie and Gruber (1996a, 1996b), LoSasso and Buchmueller (2004), Hudson, Selden, and Banthin (2005), Ham and Shore-Sheppard (2005), and Shore-Sheppard (2008), among others.

expansion of eligibility (a group we call the *newly eligible*), either for the whole sample or different population subgroups.² Further, there is also the issue that in general one cannot determine the composition of the group to which the LATE estimates apply, although this issue can be addressed in the case of a binary treatment and a binary instrument (Abadie 2003, Almond and Doyle 2011).³

In this paper we show how to address these drawbacks using a simple economic model, a switching probit model suggested by the economic model, and the fact that Medicaid eligibility is based on explicitly stated criteria. We use this framework to allow both observable and unobservable heterogeneity to affect take-up, private coverage, and crowd-out for currently eligible children and those made eligible by a non-marginal increase in Medicaid eligibility; none of these effects are currently available in the literature.⁴ Our approach is similar to the approach of Aakvik, Heckman, and Vytlacil (2005—hereafter AHV), who examine the effect of job training on employment, although our approach differs in three important ways. First, we

² A partial list of studies allowing for heterogeneous treatment effects in other contexts includes Aakvik, Heckman, and Vytlacil (2005), Angrist (2004) Blundell, Dearden, and Sianesi (2005), Heckman, Smith, and Clements (1997), and Moffitt (2009).

³ Ham, Ozbeklik, and Shore-Sheppard (2013) use an LPMI to allow observable variables to affect Medicaid and private insurance coverage; since they have a continuous instrument, they cannot apply this approach.

⁴ The current literature's LPM approach yields take-up, private coverage, and crowd-out estimates averaged over all of the marginally eligible, and the LPMI used in Ham, Ozbeklik, and Shore-Sheppard (2013) provides estimates for the marginally eligible in different demographic groups.

have two outcomes (public coverage, private coverage) while they only have one (employment). Second, we calculate different policy effects than they do, and calculate our policy effects for different groups in the population. Third, we can distinguish which individuals are affected by a non-marginal increase in Medicaid income limits. We are aided in the latter two contributions by the fact that conditional on family income, eligibility for a given child is observable.

Our approach can thus be used for any existing program (that is, one that already has participants) where eligibility is determined by observable variables available to researchers.⁵ Somewhat surprisingly, the switching probit model has been used relatively infrequently in the literature. We suspect that this has occurred at least in part because these models can be difficult to estimate – for example maximum likelihood estimation of our model requires numerical maximization of a function that involves carrying out trivariate integration for many individuals (for each set of parameters considered in the estimation). Below we show how one can use minimum distance to obtain precise estimates of treatment effects with little computational burden. Specifically, we show that our model can be estimated by linear manipulation of parameters that can be estimated using existing Stata commands. We believe that our approach will make it much easier for researchers to estimate the effects of non-marginal policy changes.

We apply our approach to the most commonly studied Medicaid policy period: the expansions that occurred prior to welfare reform in 1996. There are three reasons why this is an appealing period to examine using our approach. First, this was a period of substantial change in Medicaid eligibility policy. Second, the changes during this period were the most plausibly exogenous, resulting in many cases from states being compelled to raise their income limits.

⁵ Strictly speaking we could analyze a new program as long as the relevant parameters governing behavior can be estimated from existing data; this will be true for any structural model.

Finally, the welfare reform in 1996 that replaced the entitlement program Aid to Families with Dependent Children with the block grant Temporary Assistance to Needy Families was widely believed to have the unintended side effect of affecting Medicaid participation among families eligible for cash assistance because Medicaid and cash assistance were no longer tightly linked. We answer several important questions: (1) How much crowding out occurred among the entire population of children eligible under the original (pre-1996) Medicaid expansions? (2) How would take-up and crowd-out under further expansions compare to existing (1996) levels of take-up and crowd-out? (3) Within the currently eligible or the counterfactually newly eligible, which population subgroups respond more to non-marginal expansions of eligibility by participating in Medicaid or by dropping private coverage, and how large are those differences between groups? Our measure of crowd-out is defined as the difference between the fraction of eligible children who would have private coverage if they were (counterfactually) not eligible and the fraction of those children who actually have private coverage. This measure takes into account differences in observables and unobservables among the newly eligible and the currently eligible.⁶ Our results indicate that taking both of these differences into account is quite important.

⁶ While this approach cannot be used to predict the effects of expansions of Medicaid to populations currently entirely ineligible for the program (such as the able-bodied low-income adults that are the intended beneficiaries of the Medicaid expansion in the Affordable Care Act) since data are not currently available to estimate the parameters of interest for this subgroup, our approach can be easily applied to this setting once some of the intended beneficiaries have been allowed to respond to this expansion (i.e., once we have data on these individuals participating in Medicaid).

Our approach produces sensible and relatively precise effects for Medicaid take-up, private insurance coverage, and crowd-out among the currently eligible and the newly eligible. Further, we find that our approach does remarkably well in terms of fitting take-up and private coverage among the currently eligible when we allow for both observable and unobservable differences within subgroups. We find a wide disparity in our treatment effects across different demographic groups for both the currently eligible and those made newly eligible by a counterfactual nonmarginal Medicaid expansion. This is particularly the case for take-up rates: among both the currently eligible and the newly eligible, take-up rates vary by more than 50 percentage points across groups. As a result, our estimates should prove very helpful to policy makers concerned with outreach to underserved groups. We find that Medicaid take-up rates for the currently eligible are larger than the take-up rates of those made eligible by our policy experiment (50.3 percent as compared to 26.5 percent), while the opposite is true for private insurance coverage. The range of crowd-out estimates is much smaller. We estimate that on average fewer than 5 percent of all children eligible in 1995 had private coverage crowded out, with the largest estimate of crowd-out only 8.6 percent (for children in families where the highest earner has some college education). We find that crowd-out effects among each demographic group are larger for the newly eligible by between 2 and 5 percentage points. While this result may seem counterintuitive, it is plausible within the context of our model, since while one would expect the newly eligible to value Medicaid coverage less than the currently eligible do, private insurance coverage is much more feasible among the newly eligible, e.g. they face a lower cost of private insurance since they are more likely to be able to obtain employer sponsored insurance.

Our paper proceeds as follows. In Section 2 we consider the previous literature on Medicaid expansions. In Section 3 we use an economic model to derive our switching probit model. Section 4 contains our econometric approach while Section 5 discusses the policy effects we consider. Our empirical results are discussed in Section 6 and Section 7 concludes.

2. Medicaid Expansions and Previous Literature

Medicaid was first established as a public health insurance program for welfare recipients and low-income aged and disabled individuals. This focus largely remained until the late 1980s, when expansions in eligibility first permitted, and then required, states to cover pregnant women and children with family incomes that made them ineligible for cash welfare. Following the federally mandated eligibility expansions of 1989 and 1990, states were required to cover children age 6 or younger with family incomes up to 133 percent of the poverty line, and children born after September 30, 1983 with family incomes up to 100 percent of the poverty line. States were also given the option to increase their eligibility income limits up to 185 percent of the poverty line. As these eligibility limits were far more generous than the eligibility limits applying to cash welfare (at the time, Aid to Families with Dependent Children, or AFDC), the link between Medicaid eligibility and AFDC eligibility greatly diminished for young, low-income children. By 1996, of the approximately 30 percent of children age 19 and younger who were eligible for Medicaid, only about half came from typically welfare-enrolled families (Selden, Banthin, and Cohen 1998). While families who enrolled in cash welfare programs were also automatically enrolled in Medicaid, children newly eligible for the program were not. Consequently the establishment of a new route to Medicaid eligibility raised two important policy questions. First, to what extent did expanded eligibility lead to increased health insurance coverage for the targeted population of children? Second, did expanded eligibility lead

to “crowding out” of private health insurance by public insurance availability (and if so, to what extent), since newly eligible children were less poor than previously eligible children and hence more likely to have access to private insurance?

There has been a substantial amount of research on these initial Medicaid expansions, and a nonexhaustive list includes Currie and Gruber (1996a, 1996b), Cutler and Gruber (1996), Dubay and Kenney (1996), Thorpe and Florence (1998), Yazici and Kaestner (2000), Shore-Sheppard (2000), Blumberg, Dubay, and Norton (2000), Card and Shore-Sheppard (2004), Ham and Shore-Sheppard (2005), and Shore-Sheppard (2008). There is also research on the related question of how the further public health insurance expansions of the State Children’s Health Insurance Program (SCHIP) affected coverage and crowd-out (see LoSasso and Buchmueller (2004), Hudson, Selden, and Banthin (2005), and Gruber and Simon (2008)). These papers have provided estimates of a variety of behavioral parameters related to the responsiveness of children’s insurance coverage to expanded eligibility.

The most common approach used for estimating effects of expanded eligibility is an instrumental variable linear probability model (LPM) that we describe in more detail below. This approach produces LATE estimates of take-up and private coverage responses for an unknown (to the researcher) group of individuals who are responsive to small changes in the instrument used for identification. Of course, the LATE estimates are not the same as the average take-up rate or private coverage loss resulting from Medicaid eligibility among all eligible children (that is, including those children who were eligible prior to the expansion being studied). Moreover, they generally do not reflect the change that would occur in take-up or private coverage from a medium-sized or large change in eligibility.⁷ Definitions of “crowd-out”—loosely, the effect of

⁷ These distinctions have often been missed in the literature.

public insurance availability on the propensity to have private coverage—are particularly diverse in the literature. For expositional purposes we will focus on one definition, but our methodology is easily generalized to other definitions. In this paper, we estimate a variety of heterogeneous treatment effects on take-up, private coverage, and crowd-out for the currently eligible and for those made eligible by a nonmarginal policy change.

Since our aim in this paper is to extend previous work rather than summarizing the literature, here we focus on two of the studies that use this now standard approach. An important study using this approach is the seminal paper of Cutler and Gruber (1996—CG hereafter). CG use a linear probability model (LPM) and data on children from the March Current Population Survey (CPS) from 1988 to 1993 to estimate the effect of imputed Medicaid eligibility on insurance status, controlling for demographics and state and year effects. They use an IV version of the LPM since eligibility is likely to be endogenous. This potential endogeneity arises for at least two reasons. First, unobservable factors affecting eligibility are likely to be correlated with unobservable factors affecting health insurance choices, including tastes for private insurance or public insurance. Second, parental wages, which in turn determine eligibility, are likely to be correlated with fringe benefits (including private health insurance) of the parent. Since these benefits are unobserved, they are part of the error term.

To address the endogeneity of the eligibility variable, CG suggest an instrument (which they call *SIMELIG*) that is the fraction of a random sample of 300 children of each age imputed to be eligible according to the rules in each state in each year. This instrument, which is essentially an index of the expansiveness of Medicaid eligibility for each age group in each state and year, is correlated with individual eligibility for Medicaid but not otherwise correlated with the demand for insurance, assuming that changes in a state's Medicaid provisions are not

correlated with changes in the state’s availability of private insurance, which are unobservable to the researcher.⁸

Specifically, their LPM determining public insurance eligibility is

$$elig_i = Z_i\delta + e_i, \tag{1}$$

where $Z_i = (X_i, SIMELIG_i)$, X_i is a vector of demographic variables and e_i is an error term. The

LPM for participation in a public insurance program is given by

$$pub_i = X_i\beta_1 + \gamma_1 elig_i + u_{1i}, \tag{2}$$

where $pub_i = 1$ if child i participates in a public insurance program, $pub_i = 0$ otherwise and

u_{1i} is an error term. The LPM for private insurance coverage is given by

$$priv_i = X_i\beta_2 + \gamma_2 elig_i + u_{2i}, \tag{3}$$

where $priv_i = 1$ if child i has private insurance coverage, $priv_i = 0$ otherwise, and u_{2i} is an

error term. From the discussion in CG, it is clear that they interpret the coefficients γ_1 and γ_2 as

LATEs, i.e. treatment effects for individuals whose eligibility is affected by marginal changes in

SIMELIG, averaged across the different marginal changes in the data. CG estimate LATEs for

take-up and private coverage of 23.5 percent and 7 percent respectively. As noted above these IV

estimates are for those children whose eligibility is sensitive to a small change in *SIMELIG*.

Unfortunately, as constructed CG’s use of *SIMELIG* will not provide consistent

parameter estimates, as there will be correlation between *SIMELIG* and the error terms in (1-3)

for the members of the sample used to create *SIMELIG*. Ham and Shore-Sheppard (2005–HS

hereafter) propose a “jackknife” version of *SIMELIG* that will produce consistent parameter

⁸ One attractive feature of this approach is that *SIMELIG* is an extremely strong instrument.

estimates and in what follows we refer to this instrument as *FRACELIG*. Specifically, they use all sample observations of children of a given age in a SIPP wave *except* those from the state for which the instrument is being calculated. HS use data from the SIPP covering the period from October 1985 to August 1995 and *FRACELIG* to replicate CG's analysis. As with CG, they estimate LATEs, and thus the results of the two papers are comparable. HS find smaller LATEs for take-up rate and crowd-out than CG. They attribute some of the differences between their results and CG's to different samples and recall periods in SIPP and the CPS.⁹

As noted above, the LPM approach can be quite useful but has several limitations. First, the composition of the group to which the LATE estimates apply is unobserved by the researcher. Second, treatment effects are likely to differ for observably and unobservably different groups. (LATEs can be allowed to differ by observable characteristics by using a linear probability model with interactions.) Third, the LATE estimates do not measure treatment effects, such as crowd-out, among the entire currently eligible group. Finally, LATE estimates are not informative about the effects of nonmarginal changes in Medicaid eligibility. We now consider an approach that addresses all of these issues, starting first with a simple economic model of insurance coverage.

⁹ In their specific application they find little actual bias from using *SIMELIG* instead of *FRACELIG*.

3. Economic Model

For simplicity, we consider a family with one child, and when the child is eligible for Medicaid, we describe the family as “eligible.”¹⁰ We emphasize that our work differs from previous theoretical work on this issue since our goal is to guide our empirical work rather than to obtain theoretical results *per se*.

We begin with an ineligible family with fixed income I_i , whose decision focuses on whether to purchase private insurance at a cost $C_{pr,i}$. We assume that the family’s utility is given by

$$U_i(D_{pr,i}) = I_i - C_{pr,i}D_{pr,i} + B_{pr,i}D_{pr,i}, \quad (4)$$

where $D_{pr,i} = 1$ if they purchase private insurance and $D_{pr,i} = 0$ otherwise; hence the direct gross utility produced from having private insurance is $B_{pr,i}$. Utility maximization implies that the family will purchase private insurance if the utility from having this insurance is greater than the utility from not having it, or

$$B_{pr,i} - C_{pr,i} > 0. \quad (5)$$

Now consider an eligible family and assume that participating in Medicaid implies stigma and fixed costs that are equivalent to a dollar cost of $C_{pub,i}$. We first assume that this family’s utility is given by

¹⁰ We believe that the one child assumption greatly simplifies the analysis without obscuring the basic message. In our empirical work we allow the observations of children from the same family to be correlated.

$$U_i(D_{pr,i}, D_{pub,i}) = I_i - C_{pr,i}D_{pr,i} - C_{pub,i}D_{pub,i} + B_{pr,i}D_{pr,i} + B_{pub,i}D_{pub,i}, \quad (6)$$

where I_i , $D_{pr,i}$, $C_{pr,i}$ and $B_{pr,i}$ are defined above, $B_{pub,i}$ is the direct gross utility produced from having public insurance, $D_{pub,i} = 1$ if the family participates in Medicaid and $D_{pub,i} = 0$ otherwise.¹¹ Utility maximization implies that the family participates in Medicaid if

$$B_{pub,i} - C_{pub,i} > 0, \quad (7)$$

while the decision rule for purchasing private insurance is still given by (5).

However, one may reasonably argue that our specification of preferences in (5) is too simple since there is no role for crowd-out; eligible and ineligible families have the same rule for purchasing private insurance, although the crowd-out literature has emphasized the substitution possibilities between public and private insurance. To begin an analysis that allows for crowd-out, we specify family preferences when eligible for Medicaid as

$$U_i(D_{pr,i}, D_{pub,i}) = I_i - C_{pr,i}D_{pr,i} - C_{pub,i}D_{pub,i} + B_{pr,i}D_{pr,i} + B_{pub,i}D_{pub,i} + INT_i\{D_{pr,i} * D_{pub,i}\}, \quad (8)$$

where INT_i represents the interaction effect on utility of having both types of insurance. If there is crowd-out, then INT_i will be negative. However, we allow for the possibility that public and private insurance are complements (such as is the case with private insurance and Medicare) for some families, and thus INT_i can be positive. We draw two conclusions when we analyze this model in Appendix A. First, this simple adjustment to preferences does indeed lead to different decision rules concerning private insurance for eligible and ineligible families. Second, even if we use linear index functions for $B_{pr,i} - C_{pr,i}$, $B_{pub,i} - C_{pub,i}$, and INT_i , the econometric model becomes very complicated because there are now three equations that determine which of the

¹¹ We assume $I_i > C_{pr,i}$ for both eligible and ineligible families.

four insurance states the family occupies (no insurance, private insurance only, public insurance only, or both.) In our empirical work we compromise by requiring the decision rule for private insurance to have the same functional form for eligible and ineligible families, but with different parameter values. The resulting econometric model (described in the next section) is still much more general than those currently used in the literature.¹²

4. Econometric Methodology

4.1 Model Specification

We allow the costs and benefits of private insurance for a family ineligible for public insurance to depend on the family's characteristics X_i :

$$\beta_{pr,i}^{ne} = X_i \gamma_{1ne} + u_{1ne,i} \quad \text{and}$$

$$C_{pr,i}^{ne} = X_i \gamma_{2ne} + u_{2ne,i}.$$

Thus a family randomly made ineligible for Medicaid obtains private insurance coverage for its child if

$$Priv_nelig_i^* = B_{pr,i}^{ne} - C_{pr,i}^{ne} = X_i \gamma_{1ne} - X_i \gamma_{2ne} + u_{1ne,i} - u_{2ne,i} = X_i \gamma_{ne} + u_{ne,i} > 0, \quad (9)$$

where $\gamma_{ne} = \gamma_{1ne} - \gamma_{2ne}$ and $u_{ne,i} = u_{1ne,i} - u_{2ne,i}$. Since we can only observe a function of the difference between $B_{pr,i}^{ne}$ and $C_{pr,i}^{ne}$, we can only identify the effect of an explanatory variable on the (normalized) difference between the costs and benefits of having private insurance.

We write the costs and benefits of private insurance for a family eligible for public insurance as

¹² Another implication of the analysis in Appendix A is that for our economic model the standard LPM model is not consistent with the existence of crowd-out.

$$B_{pr,i}^e = X_i \gamma_{1e} + u_{1e,i} \quad \text{and}$$

$$C_{pr,i}^e = X_i \gamma_{2e} + u_{2e,i}.$$

Thus a family randomly made eligible for Medicaid obtains private insurance coverage for its child if

$$Priv_elig_i^* = B_{pr,i}^e - C_{pr,i}^e = X_i \gamma_{1e} - X_i \gamma_{2e} + u_{1e,i} - u_{2e,i} = X_i \gamma_e + u_{e,i} > 0, \quad (10)$$

where $\gamma_e = \gamma_{1e} - \gamma_{2e}$ and $u_{e,i} = u_{1e,i} - u_{2e,i}$. Similarly, we write the costs and benefits of public insurance for an eligible family as

$$B_{pub,i} = X_i \mu_1 + \varepsilon_{1,i} \quad \text{and}$$

$$C_{pub,i} = X_i \mu_2 + \varepsilon_{2,i}.$$

Thus a family randomly made eligible for Medicaid obtains public insurance coverage for its child if

$$Pub_i^* = B_{pub,i} - C_{pub,i} = X_i \mu_1 - X_i \mu_2 + \varepsilon_{1,i} - \varepsilon_{2,i} = X_i \mu + \varepsilon_i > 0, \quad (11)$$

where $\mu = \mu_1 - \mu_2$ and $\varepsilon_i = \varepsilon_{1,i} - \varepsilon_{2,i}$. To complete our econometric model we need to define an index function determining eligibility. A child of age a in state s in year t will qualify for Medicaid if the family income I_i is below the Medicaid income limit L_{ast} or if¹³

$$L_{ast} - I_i > 0. \quad (12)$$

While we will make use of (12) in our policy experiments, for estimation we follow the existing literature and consider a reduced form version of the rule determining if a child is eligible for Medicaid ($elig_i = 1$),

¹³ Of course it would make no difference to the analysis if we assumed that a family with $L_{ast} = I_i$ was also eligible.

$$elig_i^* = L_{ast} - I_i = Z_i\delta + e_i > 0, \quad (13)$$

where as above $Z_i = (X_i, FRACELIG_i)$.¹⁴ The model will be identified by including $FRACELIG_i$ in the eligibility index function but not in the insurance coverage index functions.

Finally, we assume that

$$(u_{e,i}, u_{ne,i}, \varepsilon_i, e_i) \sim iid N(0, V) \quad \text{and} \quad V = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 \end{bmatrix},$$

noting that ρ_{12} and ρ_{23} will not be identified.

While this switching probit model has many advantages over the LPM as noted in Section 2, the cost would seem to be the need to make a normality assumption. Interestingly, this is not what available Monte Carlo evidence suggests. Angrist (2001) conducts Monte Carlo experiments for a linear probability model with a discrete endogenous regressor estimated by instrumental variables. He finds estimated treatment effects that are very close to those from a properly specified simultaneous equation bivariate probit model. In other words, the simultaneous equation LPM does well with normally distributed errors. However, Bhattacharya, Goldman, and McCaffrey (2006) show that the simultaneous equation LPM does not do well when the error terms are *not* normally distributed. In fact, based on their Monte Carlo evidence, Bhattacharya, Goldman, and McCaffrey (2006) argue in favor of using a simultaneous equation bivariate probit model, rather than a linear probability model, when the error terms are *not* normally distributed. Further, as noted below, AHV find that they obtain similar results across

¹⁴ Note that we could have used the income limits directly, but chose to follow the literature and use $FRACELIG_i$ instead.

different assumptions about the distribution of the error terms, including normality. Thus it appears that the LPM estimates are actually more dependent on the normality assumption than are estimates from the simultaneous equation bivariate probit model.

Of course it would be ideal to eliminate the need to make any distributional assumptions. Shaikh and Vytlacil (2011) show that bounds on the average treatment effect for the overall population as well as for subgroups can be identified in the case of a triangular system of equations where both dependent variables are binary such as we have here. However, given our emphasis on point identification with respect to nonmarginal changes we did not pursue their approach. Further, AHV stress the need for making a distributional assumption for treatment effects such as those considered below, and argue that the normality assumption produces results in their model comparable to those that they obtain with alternative distributional assumptions.

4.2 Likelihood Functions

One way to proceed is to estimate a multivariate model with participation in public insurance, participation in private insurance in the absence of Medicaid eligibility, and participation in private insurance in the presence of Medicaid eligibility as the outcomes of interest using the full quasi-log-likelihood

$$\begin{aligned}
L = & \sum_{\substack{elig=1, \\ pub=0 \\ priv=1}} \log[\Pr(elig_i = 1, priv_i = 1, pub = 0)] + \sum_{\substack{elig=1, \\ pub=1 \\ priv=1}} \log[\Pr(elig_i = 1, priv_i = 1, pub = 1)] + \\
& \sum_{\substack{elig=1, \\ pub=0 \\ priv=0}} \log[\Pr(elig_i = 1, priv_i = 0, pub = 0)] + \sum_{\substack{elig=1, \\ pub=1 \\ priv=0}} \log[\Pr(elig_i = 1, priv_i = 0, pub = 1)] + \quad (14) \\
& \sum_{\substack{elig=0, \\ priv=1}} \log[\Pr(elig_i = 0, priv_i = 1)] \quad + \sum_{\substack{elig=0, \\ priv=0}} \log[\Pr(elig_i = 0, priv_i = 0)].
\end{aligned}$$

We describe this as a quasi-log-likelihood since it does not take into account correlation between observations for a child over time and across children in the same family and over time; while

not efficient, the parameters based on maximizing (14) will be consistent, and one can of course allow for such correlations in constructing standard errors. However, even after simplifying the estimation problem by ignoring these correlations, maximization of (14) involves trivariate integration and serious numerical issues that will be time-consuming even for researchers experienced in nonlinear estimation. Instead, we use a minimum distance procedure that is much easier to implement and feasible for all applied researchers, while producing estimated policy effects with small confidence intervals.

This procedure combines parameter estimates from existing well-behaved Stata routines that maximize three different likelihood functions, all of which produce consistent estimates of the relevant parameters. The first of these likelihood functions is based on the take-up of private insurance by those *currently eligible* for Medicaid

$$\begin{aligned}
L^{(1)} &= \sum_{\substack{elig=1, \\ priv=1}} \log [\Pr(elig_i = 1, priv_i = 1)] + \sum_{\substack{elig=1, \\ priv=0}} \log [\Pr(elig_i = 1, priv_i = 0)] + \sum_{elig=0} \log [\Pr(elig_i = 0)] \\
&= \sum_{\substack{elig=1, \\ priv=1}} \log \Phi_2[(X_i\gamma_e, Z_i\delta, \rho_{13})] + \sum_{\substack{elig=1, \\ priv=0}} \log \Phi_2[-X_i\gamma_e, Z_i\delta, -\rho_{13}] + \sum_{elig=0} \log \Phi_1[-Z_i\delta]. \quad (15)
\end{aligned}$$

The second likelihood function is based on the take-up of public insurance by those *currently eligible* for Medicaid

$$\begin{aligned}
L^{(2)} &= \sum_{\substack{elig=1, \\ pub=1}} \log [\Pr(elig_i = 1, pub_i = 1)] + \sum_{\substack{elig=1, \\ pub=0}} \log [\Pr(elig_i = 1, pub_i = 0)] + \sum_{elig=0} \log [\Pr(elig_i = 0)] \\
&= \sum_{\substack{elig=1, \\ pub=1}} \log \Phi_2[(X_i\mu, Z_i\delta, \rho_{34})] + \sum_{\substack{elig=1, \\ pub=0}} \log \Phi_2[-X_i\mu, Z_i\delta, -\rho_{34}] + \sum_{elig=0} \log \Phi_1[-Z_i\delta]. \quad (16)
\end{aligned}$$

The third likelihood function is based on the take-up of private insurance by those *currently ineligible* for Medicaid

$$\begin{aligned}
L^{(3)} &= \sum_{\substack{elig=0, \\ priv=1}} \log[\Pr(elig_i = 0, priv_i = 1)] + \sum_{\substack{elig=0, \\ priv=0}} \log[\Pr(elig_i = 0, priv_i = 0)] + \sum_{elig=1} \log[\Pr(elig_i = 1)] \\
&= \sum_{\substack{elig=0, \\ priv=1}} \log \Phi_2[(X_i \gamma_{ne}, -Z_i \delta, -\rho_{23})] + \sum_{\substack{elig=0, \\ priv=0}} \log \Phi_2[-X_i \gamma_{ne}, -Z_i \delta, \rho_{23}] + \sum_{elig=1} \log \Phi_1[Z_i \delta]. \quad (17)
\end{aligned}$$

Maximizing (15), (16) and (17) yields three consistent estimates of the coefficients determining eligibility, δ , denoted by $\delta^{(1)}$, $\delta^{(2)}$, and $\delta^{(3)}$, as well as one consistent estimate each of the remaining parameters: γ_{ne} , γ_e , μ , ρ_{13} , ρ_{23} , and ρ_{34} . Note that all parameter estimates are based on the same data and are correlated. The minimum distance estimator (see e.g. Hsiao 2003) uses these estimates efficiently to produce estimates of all parameter values and the covariance matrix for all of the resulting parameters; the latter is necessary for calculating standard errors for the policy effects that we consider below. Defining $\lambda^t = (\delta^t, \delta^t, \delta^t, \gamma_{ne}^t, \gamma_e^t, \mu^t, \rho_{13}, \rho_{23}, \rho_{34})$, $\hat{\lambda}^t = (\hat{\delta}^{(1)t}, \hat{\delta}^{(2)t}, \hat{\delta}^{(3)t}, \hat{\gamma}_{ne}^t, \hat{\gamma}_e^t, \hat{\mu}^t, \hat{\rho}_{13}, \hat{\rho}_{23}, \hat{\rho}_{34})$ (where t denotes transpose), and $\Omega = V(\hat{\lambda})$, the minimum distance estimator minimizes $(\lambda - \hat{\lambda})^t \Omega^{-1} (\lambda - \hat{\lambda})$ with respect to λ . By adopting the approach from Amemiya (1978, 1979) and Ham and Hsiao (1984), we can obtain greater intuition, a closed form solution for the parameter estimates, and a simple expression for their variance-covariance matrix. Let $\theta^t = (\delta^t, \gamma_{ne}^t, \gamma_e^t, \mu^t, \rho_{13}, \rho_{23}, \rho_{34})$, and note that estimating θ is equivalent to estimating λ . Then it is straightforward to show that the minimum distance estimates $\hat{\theta}$ will be the generalized least squares estimates from the regression equation

$$\hat{\lambda} = W\theta + (\lambda - \hat{\lambda}), \text{ where } W = \begin{bmatrix} I_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ I_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ I_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \text{or} \quad (18)$$

$$\hat{\theta} = (W'\Omega^{-1}W)^{-1}W'\Omega^{-1}\hat{\lambda} \quad \text{with } \text{Var}(\hat{\theta}) = \Sigma = (W'\Omega^{-1}W)^{-1}, \quad (19)$$

where I_1 and I_2 are appropriately dimensioned identity matrices. Thus parameter estimation is simple given an estimate of $\text{Var}(\hat{\lambda}) = \Omega$, and we show how to estimate Ω in Appendix B.¹⁵

Using the estimated parameters, we can calculate policy effects for the entire sample and different demographic groups for the currently eligible and for those who become eligible due to a nonmarginal change (for example, raising all income limits by 10 percent).¹⁶ In calculating these effects we use the fact that since the Medicaid income limits determining eligibility are observable we can see exactly which children in the sample are affected. Since the estimation of treatment effects arising from a nonmarginal change is very similar to calculating these effects

¹⁵ MATLAB code for carrying out the minimum distance estimation is available at

http://www.williams.edu/Economics/shoresheppard/MDE_matlab.zip.

¹⁶ Our approach can also be used for those made eligible by a marginal change in the income limits, but since these estimates are also available for the linear probability model with interactions, we do not focus on them here. The equations shown here to calculate treatment effects for a nonmarginal change can be used to examine a marginal change by making the relevant changes “small.”

for all of the currently eligible, we describe the former in the text and the latter in the online Appendix.¹⁷ For expositional ease and better intuition, we begin by assuming that there is no correlation between the unobservable in the eligibility equation and the unobservables in the insurance coverage equations; in what follows we describe this as calculating effects *ignoring selection*. We account for selection in subsection 5.2 and find that it is indeed important to account for both observables and unobservables in estimating our policy effects.

5. Estimated Treatment Effects

5.1 Calculating Take-up, Private Coverage, and Crowd-out Among Those Made Eligible by a Counterfactual Expansion in the Income Limits Ignoring Selection

Consider a nonmarginal change (e.g. a 10 percent increase) in the Medicaid income limits. Applying the new limits to the individuals in the data, we can observe who would become eligible and their characteristics. This assumes that families with incomes above the new limits do not adjust their incomes to qualify for Medicaid. Thus when we ignore selection, the Medicaid take-up rate among all of the newly eligible (denoted by $new=1$) is given by

$$\begin{aligned} NEWATRE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1) \\ &= \frac{1}{N_{new}} \sum_{new=1} \Pr(X_i \hat{\mu} + \varepsilon_i > 0) = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\mu})], \end{aligned} \quad (20a)$$

where N_{new} refers to the number of newly eligible individuals.¹⁸ Similarly, let $new^k = 1$ denote

¹⁷ The online Appendix is available at http://www.williams.edu/Economics/shoresheppard/HOS-S_online_appendices_June2013.pdf.

¹⁸ The conditioning on $new = 1$ in the first equality refers to possible selection pertaining to the newly eligible, and the simplification in the second equality follows because we are ignoring selection for now.

individuals in a particular demographic group k who are made newly eligible, and let N_{new}^k denote the number of these individuals. Then we can estimate the take-up rate for the newly eligible in group k when ignoring selection by calculating an analogous equation to (20a) for the subset of the newly eligible in group k :

$$NEWATRE^k = \frac{1}{N_{new}^k} \sum_{new^k=1} [1 - \Phi_1(-X_i \hat{\mu})]. \quad (20b)$$

The private insurance coverage rate among the newly eligible when we ignore selection is estimated by

$$\begin{aligned} NEWPRE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1) \\ &= \frac{1}{N_{new}} \sum_{new=1} [\Pr(X_i \hat{\gamma}_e + u_{ei} > 0)] = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\gamma}_e)]. \end{aligned} \quad (21)$$

It is straightforward to estimate private insurance coverage for those newly eligible in group k as we did in moving from (20a) to (20b), and we omit this equation to save space. Next, for the newly eligible, the fraction that would have private insurance in the absence of eligibility (when we ignore selection) is given by

$$\begin{aligned} NEWPRNE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1) \\ &= \frac{1}{N_{new}} \sum_{new=1} \Pr(X_i \hat{\gamma}_{ne} + u_{nei} > 0) = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\gamma}_{ne})]. \end{aligned} \quad (22)$$

The private coverage that group k would have in the absence of becoming newly eligible is again a simple extension of (22), and we omit the equation given space constraints. Finally, given the expression for the variance-covariance matrix of the parameter estimates in (19), it is straightforward to calculate standard errors for all treatment effects in this paper using the delta method and we show an example of this in the online Appendix.

Since they are counterfactuals, *NEWATRE* (estimated average take-up rate among the newly eligible) and *NEWPRE* (estimated private insurance coverage for the newly eligible) cannot be directly observed in the data, while *NEWPRNE* (estimated private insurance coverage among the newly eligible in the absence of eligibility) can be directly measured in the data because children who would become eligible under our nonmarginal change are in fact currently observed to be ineligible. We define crowd-out among the newly eligible when we ignore selection as the private coverage rate the newly eligible would be predicted to have minus the private coverage rate they would have if ineligible—*NEWPRE* minus *NEWPRNE*, or (21) minus (22). This is a well-defined crowd-out measure, in contrast to many in the literature where crowd-out is backed out from relative responses to eligibility for public and private coverage. Note that crowd-out among the newly eligible will depend on two factors: the extent to which the benefits of private insurance exceed the costs of private insurance in the absence of eligibility, and the difference in the benefits and costs of Medicaid insurance for these families. We would conjecture that newly eligible families are likely to value Medicaid insurance less than those currently eligible, and may dislike stigma more than the currently eligible; these factors will lead to crowd-out effects being smaller among the newly eligible than the currently eligible. However, the newly eligible are more likely to be able to obtain a significantly lower price for private insurance via employer sponsored insurance. Hence it is not clear which effect will dominate, and whether crowd-out will be higher or lower among newly eligible families than among the currently eligible.¹⁹

¹⁹ For example, consider a very poor family that effectively has a very low probability of purchasing private insurance in the absence of Medicaid. For this family there is essentially no private insurance for Medicaid to crowd out.

The expressions for take-up, private coverage, and crowd-out for all currently eligible (or the currently eligible in specific demographic groups), i.e. average treatment effects for the treated, are analogous to (20a)-(22) when we ignore selection, and as noted above we provide them in the online Appendix. However, we note that these estimates not only provide important policy parameters but also allow us to conduct a specification check on our model by comparing the take-up and private coverage for eligible children and private coverage for ineligible children that are predicted from the model to the actual values observed in the sample. We can also do this for a demographic group k .

5.2 Calculating Take-up and Crowd-out Among Those Made Eligible by a Counterfactual Expansion in the Income Limits while Accounting for Selection

To calculate the policy effects while allowing for selection, we must differentiate between the values of the explanatory variables under the old and new income limits. For each individual let Z_{0i} denote the value of the previous explanatory variables and Z_{1i} denote the value of the explanatory variables under the new income limits; note that Z_{0i} and Z_{1i} differ only in terms of Z_{1i} having a higher value of $FRACELIG_i$. To begin, note that the probability that a child is newly eligible is given by

$$\Pr(new_i = 1) = \Pr(-Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta}) = \Phi_1(-Z_{0i}\hat{\delta}) - \Phi_1(-Z_{1i}\hat{\delta}). \quad (23)$$

Now Medicaid take-up among the newly eligible is given by

$$\begin{aligned} NEWATRESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1 | newelig_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(pub_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\ &= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i\hat{\mu} + \varepsilon_i > 0, -Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta})}{\Pr(-Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta})} \\ &= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i\hat{\mu}, -Z_{0i}\hat{\delta}, \hat{\rho}_{34}) - \Phi_2(X_i\hat{\mu}, -Z_{1i}\hat{\delta}, \hat{\rho}_{34})}{\Phi_1(-Z_{0i}\hat{\delta}) - \Phi_1(-Z_{1i}\hat{\delta})} \right]. \end{aligned} \quad (24a)$$

We predict Medicaid take-up among the newly eligible in demographic group k as

$$NEWATRESEL_k = \frac{1}{N_{new}^k} \sum_{new^k=1} \left[\frac{\Phi_2(X_i \hat{\mu}, -Z_{0i} \hat{\delta}, \hat{\rho}_{\varepsilon.e}) - \Phi_2(X_i \hat{\mu}, -Z_{1i} \hat{\delta}, \hat{\rho}_{34})}{\Phi_1(-Z_{0i} \hat{\delta}) - \Phi_1(-Z_{1i} \hat{\delta})} \right]. \quad (24b)$$

Similarly, private insurance coverage for the newly eligible is given by

$$\begin{aligned} NEWPRESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1 | newelig_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(priv_elig_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\ &= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i \hat{\gamma}_e + u_{ei} > 0, -Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})}{\Pr(-Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})} \\ &= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i \hat{\gamma}_e, -Z_{0i} \hat{\delta}, \hat{\rho}_{13}) - \Phi_2(X_i \hat{\gamma}_e, -Z_{1i} \hat{\delta}, \hat{\rho}_{14})}{\Phi_1(-Z_{0i} \hat{\delta}) - \Phi_1(-Z_{1i} \hat{\delta})} \right]. \end{aligned} \quad (25)$$

In the absence of becoming eligible, these individuals would have private coverage given by

$$\begin{aligned} NEWPRNESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1 | newelig_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(priv_nelig_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\ &= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i \hat{\gamma}_{ne} + u_{nei} > 0, -Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})}{\Pr(-Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})} \\ &= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i \hat{\gamma}_{ne}, -Z_{0i} \hat{\delta}, \hat{\rho}_{23}) - \Phi_2(X_i \hat{\gamma}_{ne}, -Z_{1i} \hat{\delta}, \hat{\rho}_{24})}{\Phi_1(-Z_{0i} \hat{\delta}) - \Phi_1(-Z_{1i} \hat{\delta})} \right]. \end{aligned} \quad (26)$$

(Again note that one could also estimate $NEWPRNESEL$ from the data.) Thus crowd-out for the newly eligible while accounting for selection is given by the difference $NEWPRESEL - NEWPRNESEL$, or (25) – (26). As before, we can calculate the analogous effects (including crowd-out) for the currently eligible in the whole sample or in different demographic groups when we allow for selection; the expressions for these calculations are available in the online Appendix.

6. Comparison to Aakvik, Heckman and Vytlačil (2005)

In their paper evaluating the efficacy of a Norwegian vocational training program, AHV introduce an econometric model for estimating different policy parameters (the average treatment effect, the marginal treatment effect, and the average treatment effect for the treated) when the outcome is discrete and the treatment effect is heterogeneous. Their equation for whether an individual enters training is analogous to our Medicaid eligibility equation (13), while their employment outcome is analogous to our private insurance take-up equations (9) and (10). Specifically, their likelihood function consists of employed trainees contributing a term for the joint probability of taking training and being employed, non-employed trainees contributing a term for the joint probability of taking training and not being employed, employed non-trainees contributing a term for the joint probability of not participating in training and being employed, and non-employed non-trainees contributing a term for the joint probability of not participating in training and not being employed. They estimate the model by maximum likelihood assuming normally distributed error terms, but one could also estimate their model by first using Stata to estimate one equation determining trainee status and the employment outcome for trainees and another equation determining trainee status and the employment outcome for nontrainees, and then optimally combining the parameter estimates using minimum distance.

As noted above, our approach differs from theirs in three key ways. First, we have two outcomes (public coverage, private coverage) while they only have one (employment). Second, we calculate different policy effects than they do, and calculate our policy effects for different groups in the population. Third, we can distinguish which individuals are affected by a non-marginal increase in Medicaid income limits. We are aided in the latter two contributions by the fact that conditional on family income, eligibility for a given child is observable. AHV cannot do

this since the explicit rule for being assigned to training, to the extent that one exists, is unavailable to them.

7. Data

To implement our approach, we use data from the SIPP 1986, 1987, 1988, 1990, 1991, 1992, and 1993 panels, which cover the period 1986-1995. The SIPP is a nationally representative longitudinal household survey specifically designed to collect detailed income and program participation information. The recall period between each interview is four months for every individual, and in our data the panel length ranges from 24 months for the 1988 panel to 40 months for the 1992 panel. The sample universe is the entire U.S., but the Census Bureau did not separately identify state of residence for residents of the nine lowest population states in those panels. Since state of residence information is critical for us to impute Medicaid eligibility, we drop all individuals whose state of residence is not identified. We also restrict our sample to children living in households that are part of the original sample and who are younger than 16 years old at the first time they are observed. Finally, for comparability with earlier studies we drop children who are observed only once, children who leave the sample and then return, and children who move between states during the sample period. In total, these omitted observations constitute less than 8 percent of the sample.

Although the four-month recall period increases the probability of accurate reporting relative to the fifteen-month recall period of the March Current Population Survey (Bennefield 1996), the SIPP suffers from the problem of “seam bias.” Specifically, Census Bureau researchers have shown that there are a disproportionate number of transitions between the last month of the current wave and the first month of the next wave (see, e.g., Young 1989, Marquis and Moore 1990). We use data for all interview months and account for seam bias by including a

dummy variable for the fourth month of each interview wave. When we calculate predicted take-up and crowd-out probabilities, we follow Blank and Ruggles (1995) and Ham, Li, and Shore-Sheppard (2013) and adjust our parameter estimates by dropping the coefficient on the fourth month dummy and adding one-quarter of this coefficient to the intercept. In their study of accounting for seam bias in a multi-state, multi-spell duration model, Ham, Li and Shore-Sheppard (2013) find that this is preferable to using data only from the fourth month of each wave.

We need to impute Medicaid eligibility and use four steps to do so. First, we construct the family unit relevant for Medicaid program participation and determine family income. Second, we assign family-specific poverty thresholds based on the size of the family and the year. Since AFDC eligibility implied Medicaid eligibility over this period, we then use information on the family income and family structure, along with the AFDC parameters in effect in the state and year, to impute eligibility for AFDC. Finally, we assign Medicaid eligibility if any of the following conditions hold: the child is in an AFDC-eligible family; the child is income eligible for AFDC and either lives in a state without a family structure requirement or lives in a state with an AFDC-unemployed parent program and has an unemployed parent; or the child's family income as a percent of the relevant poverty line is below the Medicaid expansion income eligibility cutoff in effect for that age child in his or her state of residence at that time.

In all models we include demographic variables as well as state, year and age dummies for each child to control for state-specific, age-specific and year-specific unobservables. In addition, we cluster the standard errors to account for dependence across person-specific observations. In Table 1 we present the (unweighted) sample means for the variables used in our regressions. Both Medicaid participation and Medicaid eligibility rose over the course of the

sample, while private insurance coverage fell. The rise in eligibility was particularly dramatic between the 1988 and 1990 panels, when the federally mandated expansions took effect. Compared with the changes in insurance eligibility and coverage, the demographic variables are fairly stable across panels.

8. Empirical Results

In section 8.1 we present the standard LATE for the marginal (unknown) currently eligible using the LPM in order to provide a benchmark for the treatment effects from the switching probit model (SPM). Here we also consider the issue of weak instruments, which has not been previously investigated in this literature. In section 8.2 we use the SPM estimates to calculate treatment effects for the currently eligible, while in section 8.3 we do this for the newly eligible. We use all panels of the SIPP data to estimate the SPM parameters, but for ease of comparison of treatment effects for the currently eligible and the newly eligible, we use only the last year of data (1995) to calculate the treatment effects. The parameter estimates are reported in our online Appendix.

8.1 Baseline LATE for Medicaid Take-Up

Using the standard LPM, we estimate that the take-up rate of Medicaid averaged across marginal individuals is a very statistically significant 0.127. Since these are IV estimates, the question of weak IVs is relevant, but to the best of our knowledge this issue has not been explored previously in this context. We cannot investigate whether our instrument is weak using the rule of thumb for the F-test being greater than 10 suggested by Staiger and Stock (1997), or the refinements of their rule in Stock and Yogo (2005), since the F-test is not appropriate if, as we assume, observations from the same family are correlated. Since the first stage equation is a LPM estimated on panel data, the resulting heteroskedasticity will also invalidate the use of an F-

test. Instead we use the rule of thumb from Hansen, Hausman, and Newey (2008) that the Wald statistic for the coefficient on the excluded instrument *FRACELIG* in the first stage equation should be greater than 33 (for one excluded instrument). We find that the Wald statistic for the coefficient on *FRACELIG* is approximately 3,600, so it is safe to conclude that weak instruments are not an issue here.

8.2 Crowd-out, Predicted and Actual Medicaid Take-up and Private Insurance Coverage among the Currently Eligible

Using the approach outlined above, we estimate Medicaid take-up rates and private insurance coverage rates for all of the currently eligible and the currently eligible by demographic group using the 1995 data. We then compare these estimates to the 1995 actual values. This comparison provides a model specification test that exploits the fact that we know the eligibility rules and, to the best of our knowledge, is new to the program evaluation literature. In column (1) of Table 2 we present the 1995 Medicaid take-up rates estimated ignoring selection (using equation (1) in the online Appendix), while in column (2) we present the estimated effects allowing for selection (equation (5) in the online Appendix).²⁰ All of the estimates are highly statistically significant and have relatively narrow confidence intervals. The estimated treatment effects in both columns are qualitatively similar, although the estimates accounting for selection in column (2) are 2 to 5 percentage points higher; this occurs because they take into account the fact (as indicated by our parameter estimates) that those who are eligible for Medicaid have unobservable characteristics that make them more likely to take up Medicaid. In column (3) of Table 2 we present the *actual* take-up rates for the whole sample and the different demographic groups, again using only the 1995 data. Comparing the predicted take-

²⁰ All estimates in the paper are estimates of the average treatment effect on the treated.

up rates in the first two columns with the sample take-up rates in column (3), we see that the estimates that account for selection match the data remarkably well, and much better than the estimates that ignore selection. Given this we focus on the effects in column (2) when discussing the take-up rates across different demographic groups.

The average take-up rate among all currently eligible children is 51 percent, substantially higher than the take-up rate among the marginally eligible estimated from the LPM (12.7 percent). This is not surprising, since the currently eligible include many very low-income children who are also eligible for cash assistance. There is a wide disparity in take-up rates across demographic groups among the currently eligible.²¹ The average take-up rates range from 0.12 for children from families with more than two earners to 0.79 for children from families with no earners. The estimates show a clear pattern: eligible children from traditionally disadvantaged groups take up Medicaid at a higher rate than eligible children from typically less disadvantaged groups. For example, eligible white children have a take-up rate of 0.44 while the take-up rate for nonwhite children is fifty percent larger.²² The estimated take-up rate for children in families in which the family head has less than a high school degree is 0.63, while it is 0.16 for children in families in which the family head has a college degree or more. If the family head has a college degree, their time is likely to be more valuable, making time spent

²¹ Interestingly, we find a similar disparity across groups among the marginally eligible in the LATEs that we estimate using the LPMI (Ham, Ozbeklik, and Shore-Sheppard 2013).

²² This reflects both the fact that nonwhites are more likely to take up Medicaid conditional on the other explanatory variables (see the parameter estimates in the online Appendix table), and that their values for the other explanatory variables make them more likely to participate in Medicaid.

applying for Medicaid more expensive, and as noted above, they may particularly dislike any stigma associated with Medicaid. Moreover, the estimated take-up rate for an eligible child from a family in which a female is a single head is 0.71, while it is only 0.30 for a child from a two-parent family. Thus traditionally welfare-ineligible populations have dramatically lower responses to Medicaid eligibility than do the traditionally welfare-eligible. While this has been suspected in the literature previously, ours are the first quantitative estimates of the differences in take-up across groups, and they suggest that improving take-up among the eligible requires efforts to reduce the transactions costs associated with applying for Medicaid, perhaps by reducing the cost to the family of determining if they are eligible for Medicaid through greater outreach. Also, if it is possible to do so, reducing the stigma associated with Medicaid participation would likely increase participation among the traditionally ineligible groups.

Next we consider predicted private insurance coverage rates for all eligible children and for eligible children in different demographic groups in columns (4) and (5) of Table 2 when we do not, and do, account for selection, using the 1995 data. (We calculate these effects using equations (2) and (6) in the online Appendix, respectively.) These predicted private insurance coverage rates are precisely estimated and again vary widely across groups. For example, the demographic group with the lowest private insurance coverage for all the model specifications is children from families without any earners; the private insurance coverage rate is 0.12 when accounting for selection and 0.15 with selection ignored. This very low rate is consistent with our model in that they are very unlikely to have access to employer-sponsored insurance. The results in column (5) are smaller than those in column (4) because of the (estimated) negative correlation between the unobservables in the eligibility equation and the private insurance coverage equation. The largest difference in the estimated coverage rates between columns (4)

and (5) is for the children in less disadvantaged families: those with two or more earners and those in which the highest earner is a college graduate. Column (6) reports the actual private insurance coverage rates in the 1995 data. Again, we see that the treatment effects in column (5) when we allow for selection mirror the data remarkably well, and much better than when we ignore selection.

Our crowd-out estimates for the currently eligible for the cases where we do not, and do, account for selection are in columns (7) and (8) of Table 2, respectively. They represent the first crowd-out effects estimated for the currently eligible in the population as a whole and in the different demographic groups. (We calculate these effects using equations (4) and (8) in the online Appendix, respectively.) Given that the estimates allowing for selection do a much better job of fitting the data, it clearly seems worth the extra effort that the models with selection require, and we focus on the crowd-out estimates based on them in column (8). The vast majority of the effects are statistically distinguishable from zero and negative, indicating that private and public insurance are indeed substitutes, although the degree of substitution is quite small. Overall, we estimate that fewer than 5 percent of all children eligible in 1995 had private coverage crowded out. Looking separately by demographic group, the estimated effects that are significant at the 5% level range from -0.026 to -0.086. Within this range the size of crowd-out for a group appears unrelated to the private insurance coverage of the group, although it is interesting to note that we find no significant crowd-out from one group that has a very high level of Medicaid take-up—families headed by a high school dropout—and one group that has a very low level of Medicaid take-up—families with more than two earners.

8.3 Treatment Effects on Medicaid Take-up, Private Insurance Coverage, and Crowd-out among Those Made Eligible by a Hypothetical Expansion of Medicaid Eligibility.

We next consider the predicted effects of our estimated counterfactual increase in the Medicaid income limits by 10 percent; a key advantage offered by our proposed approach is the ability to undertake such analysis. In columns (1) and (2) of Table 3 we show the predicted take-up rates for children in the 1995 data made newly eligible by our hypothetical policy experiment when we do not, and do, allow for selection; we focus here on the estimates that allow for selection given our results in Table 2. (We calculate these effects using equations (20a-b) and (24a-b) in the text, respectively.) All of the estimates are statistically significant and precisely estimated. The estimates in column (2) are again larger than those in column (1) because of the positive estimated correlation between the unobservables in eligibility and the unobservables in take-up. The overall levels of take-up in Table 3 for the newly eligible are lower than those in Table 2 for the actually eligible, with an average estimated take-up rate of 26.5 percent among the newly eligible as compared to 50.3 percent among the currently eligible. This result is consistent with our model since the transactions costs of applying for Medicaid are likely to be higher for the newly eligible group, for example because conveying information about eligibility to potential participants is more difficult for potential participants who do not have social networks for obtaining such information (see Bertrand, Luttmer, and Mullainathan 2000) or because of higher costs of time. In addition, this group is much more likely to have access to employer-sponsored insurance and therefore they would face a lower price for private coverage.

As with the case of take-up among the currently eligible, again there are considerable differences across groups of the newly eligible in their take-up of Medicaid, and again the observably less disadvantaged children have substantially lower estimated rates of enrolling in

the Medicaid program for which they are eligible, even though such children were the intended beneficiaries of the expansions. The variation across groups in response to a nonmarginal expansion is still considerable, although it is less than the variation among the currently eligible. For example, the variation in take-up rates by family structure now ranges only from 0.20 (for children in two-parent families) to 0.46 (for children in female-headed families).

We present the predicted private insurance coverage rates when we do not, and do, correct for selection among those made newly eligible by our counterfactual policy change in columns (3) and (4) of Table 3; both sets of estimates are highly significant and precisely estimated. (We calculate these effects using equations (21) and (25) in the text, respectively.) The private coverage rates are higher than the corresponding estimates for the currently eligible in Table 2. This is not surprising as raising the income limits makes higher income children eligible, and our model would predict that such children would be more likely to have private coverage since they are more likely than the currently eligible to have access to private coverage through their parents' employment and thus face a lower cost of private insurance. The predicted coverage rates that take selection into account are smaller than those ignoring selection because the SPM estimates indicate a negative correlation between unobservables in the eligibility index and the private coverage index. Focusing on the estimates allowing for selection in column (4), we see that estimated private coverage rates range from 0.30 for a household headed by a high school dropout to 0.79 for a household in which the head is a college graduate.²³

Finally, columns (5) and (6) of Table 3 contain the crowd out estimates for the newly eligible when we do not, and do, allow for selection. (We calculate these effects using equations

²³ Here we are ignoring the estimated coverage rate for a family with no earners of 0.15 as there will be very few of these among the newly eligible.

(21-22) and (25-26) in the text respectively.) Focusing on the estimated effects that allow for selection (column 6), most are statistically significant and have small confidence intervals. Among the groups with crowd-out effects statistically distinguishable from 0, estimated crowd-out rates range from 0.03 to 0.11. For each demographic group, the crowd-out effects are larger for the newly eligible than the currently eligible; as we noted above the predicted effect here is ambiguous since these families are likely to face a lower cost of private insurance (through employer insurance) which can offset the likely lower valuation of Medicaid services by these families. On average, we estimate that private insurance coverage is about 8 percentage points lower among children made newly eligible by a counterfactual policy change than it would be in the absence of such an expansion of Medicaid eligibility; recall that the corresponding crowd-out effect for the currently eligible is 5 percentage points.

7. Conclusions

In this paper we use an economic model and a switching probit model implied by it to investigate private and public insurance coverage across different demographic groups. Employing the fact that Medicaid eligibility is observable given child characteristics and the Medicaid income limits, and using the estimates from the switching probit model, we estimate the effects of Medicaid eligibility on take-up, private coverage, and crowd-out for currently eligible children, as well as for those made newly eligible by a counterfactual nonmarginal increase in the Medicaid income limits. None of these effects are currently available in the literature, and they cannot be estimated using the standard linear probability model approach and the extension to the linear probability model with interactions of Ham, Ozbeklik and Shore-Sheppard (2013). Moreover, since a full maximum likelihood implementation of the switching probit model will be overly computationally demanding for many researchers, we suggest a

minimum distance approach that simply involves linear transformations of parameters obtained using existing, well-behaved, Stata commands.

Our approach produces sensible and relatively precise effects for Medicaid take-up, private insurance coverage, and crowd-out among the currently and newly eligible. Moreover, our economic model helps us interpret and understand our results. Our treatment effect estimates for the currently eligible that allow for selection do a very good job of fitting actual Medicaid take-up and private insurance coverage rates. We find a wide disparity in estimated take-up rates across different demographic groups for both the currently and newly eligible; in general, we estimate that more disadvantaged groups have greater Medicaid participation rates among both the currently and newly eligible. While this is perhaps not surprising, the magnitude of the differences across groups is substantial, with some groups taking up the coverage for which they are eligible at two times or more the rate of other groups. The pattern of take-up effects suggests a potentially important role for information provision and information flows in determining insurance coverage following an expansion in order to reduce the cost of Medicaid participation among newly eligible families. As expected based on our model, we find that Medicaid take-up rates for the currently eligible are larger than the take-up rates of those made eligible by our policy experiment, while the newly eligible have higher private insurance coverage than the currently eligible. Crowd-out, which we have defined as the difference between the fraction of eligible children who would have private coverage if they were (counterfactually) not eligible and the fraction of those children who actually have private coverage, is statistically distinguishable from 0 for most groups among the currently and newly eligible. Perhaps more surprising, we find that crowd-out tends to be higher among the newly eligible families, who are relatively better off, than among the currently eligible families.

Going forward, it will be possible to use our approach to examine the impact of Medicaid expansions under the Affordable Care Act (ACA). Under the ACA, states may expand Medicaid coverage to low-income able-bodied adults who are not single parents. Because this population was heretofore ineligible for Medicaid, little is known about the effect of extending eligibility to them. Once the ACA has been in effect for a period of time and data for estimation are available, our model is well suited to examining the situation where some, but not all, states expand their Medicaid coverage for adults. In this situation, able-bodied adults will be eligible at much higher income limits in some states than others. Thus we could focus on, for example, low-skill single men. The men in the states not introducing Medicaid expansions would be analogous to our ineligible families, while the men in the states introducing Medicaid expansions would be analogous to our eligible families. Once data are available, our approach can be used to estimate the increase in coverage and extent of crowd-out if the states not expanding Medicaid were to do so later.

References

- Aakvik, Arild, James J. Heckman, and Edward J. Vytlacil. 2005. "Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs." *Journal of Econometrics* 125(1-2): 15-51.
- Abadie, Alberto. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113(2): 231-263.
- Almond, Douglas and Joseph J. Doyle. 2011. "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays." *American Economic Journal: Economic Policy* 3(3): 1-34.
- Amemiya, Takeshi. 1978. "The Estimation of a Simultaneous Equation Generalized Probit Model." *Econometrica* 46(5):1193-1205.
- _____. 1979. "The Estimation of a Simultaneous-Equation Tobit Model." *International Economic Review* 20(1):169-181.
- Angrist, Joshua D. 2001. "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice." *Journal of Business and Economic Statistics* 19(1):2-16.
- _____. 2004. "Treatment Effect Heterogeneity in Theory and Practice." *Economic Journal* 114(494):C52-C83.
- Bennefield, Robert L. 1996. "A Comparative Analysis of Health Insurance Coverage Estimates: Data from CPS and SIPP." U.S. Bureau of the Census SIPP Working Paper No. 218.
- Bertrand, Marianne, Erzo F. P. Luttmer and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics* 115(3):1019-1055.

- Bhattacharya, Jay, Dana Goldman and Daniel McCaffrey. 2006. "Estimating Probit Models with Self-Selected Treatments." *Statistics in Medicine* 25(3): 389-413.
- Blank, Rebecca M. and Patricia Ruggles. 1996. "When Do Women Use Aid to Families with Dependent Children and Food Stamps? The Dynamics of Eligibility Versus Participation." *Journal of Human Resources* 31(1):57-89.
- Blumberg, Linda J., Lisa Dubay, and Steven A. Norton. 2000. "Did the Medicaid Expansions for Children Displace Private Insurance? An Analysis Using the SIPP." *Journal of Health Economics* 19(1):33-60.
- Blundell, Richard, Lorraine Dearden and Barbara Sianesi. 2005. "Evaluating the Effect of Education on Earnings: Models, Methods and Results from the National Child Development Survey." *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 168(3):473-512.
- Card, David and Lara D. Shore-Sheppard. 2004. "Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low Income Children." *Review of Economics and Statistics* 86(3):752-766.
- Carneiro, Pedro, James J. Heckman and Edward J. Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101(6):2754-2781.
- Currie, Janet and Jonathan Gruber. 1996a. "Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women." *Journal of Political Economy* 104(6):1263-1296.
- _____. 1996b. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." *Quarterly Journal of Economics* 111(2):431-466.
- Cutler, David M. and Jonathan Gruber. 1996. "Does Public Insurance Crowd Out Private Insurance?" *Quarterly Journal of Economics* 111(2):391-430.

- Dubay, Lisa C. Genevieve M. and Kenney. 1996. "The Effects of Medicaid Expansions on Insurance Coverage of Children." *The Future of Children* 6(1):152-161.
- Gruber, Jonathan and Kosali Simon, K. 2008. "Crowd-out 10 Years Later: Have Recent Public Insurance Expansions Crowded Out Private Health Insurance?" *Journal of Health Economics* 27(2):201-217.
- Ham, John C. and Cheng Hsiao. 1984. "Two-Stage Estimation of Structural Labor Supply Parameters Using Interval Data from the Canadian Census," *Journal of Econometrics*, 24, 138-158
- Ham, John C., Xianghong Li, and Lara D. Shore-Sheppard. 2013. "Seam Bias, Multiple-State, Multiple-Spell Duration Models and the Employment Dynamics of Disadvantaged Women." University of Maryland Working Paper.
- Ham, John C., Serkan Ozbeklik, and Lara D. Shore-Sheppard. 2013. "Estimating Heterogeneous Medicaid Take-up and Private Insurance Crowd-out Rates: Evidence from a Linear Probability Model." University of Maryland Working Paper.
- Ham, John C. and Lara D. Shore-Sheppard. 2005. "The Effect of Medicaid Expansions for Low-Income Children on Medicaid Participation and Private Insurance Coverage: Evidence from the SIPP." *Journal of Public Economics* 89(1):57-83.
- Hansen, Christian, Jerry Hausman and Whitney Newey. 2008. "Estimation with Many Instrumental Variables." *Journal of Business and Economic Statistics* 26(4):398-422.
- Heckman, James J., Jeffrey Smith, and Nancy Clements. 1997. "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts." *The Review of Economic Studies* 64(4):487-535.

- Hsiao, Cheng. 2003. *Analysis of Panel Data, 2nd Edition*. Cambridge: Cambridge University Press.
- Hudson, Julie L., Thomas M. Selden, and Jessica S. Banthin. 2005. "The Impact of SCHIP on Insurance Coverage of Children." *Inquiry* 42(3):232-54.
- LoSasso, Anthony T. and Thomas C. Buchmueller. 2004. "The Effect of the State Children's Health Insurance Program on Health Insurance Coverage." *Journal of Health Economics* 23(5):1059-1082.
- Marquis, Kent H. and Jeffrey C. Moore. 1990. "Measurement Errors in the Survey of Income and Program Participation Program Reports." *1990 Annual Research Conference Proceedings*. Washington, D.C.: U. S. Bureau of the Census.
- Moffitt, Robert. 2009. "Estimating Marginal Treatment Effects in Heterogeneous Populations." *Annales d'Economie et de Statistique*. Special Issue on Econometrics of Evaluation, Issue 91-92:239-261.
- Selden, Thomas M., Jessica S. Banthin, and Joel W. Cohen. 1998. "Medicaid's Problem Children: Eligible but not Enrolled." *Health Affairs* 17(3):192-200.
- Shaikh, Azeem M. and Edward J. Vytlacil. 2011. "Partial Identification in Triangular Systems of Equations with Binary Dependent Variables." *Econometrica* 79(3):949-955.
- Shore-Sheppard, Lara D. 2000. "The Effect of Expanding Medicaid Eligibility on the Distribution of Children's Health Insurance Coverage." *Industrial and Labor Relations Review* 54(1):59-71.
- _____. 2008. "Stemming the Tide? The Effect of Expanding Medicaid Eligibility on Health Insurance Coverage." *The B.E. Journal of Economic Analysis and Policy* 8 (Advances), Article 6.

Staiger, Douglas and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65(3):557-286.

Stock, James H. and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*, Donald W.K. Andrews and James H. Stock, eds. Cambridge University Press, Cambridge, MA: 80–108.

Thorpe, Kenneth E. and Curtis S. Florence. 1998. "Health Insurance among Children: The Role of Expanded Medicaid Coverage." *Inquiry* 35(4):369-379.

Yazici, Esel and Robert Kaestner. 2000. "Medicaid Expansions and the Crowding Out of Private Health Insurance Among Children." *Inquiry* 37(1):23-32.

Young, Nathan. 1989. "Wave-seam Effects in the SIPP." *Proceedings of the Section on Survey Research Methods*. Alexandria, VA: American Statistical Association Annual Meeting: 393-398.

Table 1: Means of the Variables Used in Estimation

| SIPP Panel | 1986 | 1987 | 1988 | 1990 | 1991 | 1992 | 1993 |
|-------------------------------|--------|--------|--------|--------|--------|--------|--------|
| Medicaid | 0.12 | 0.12 | 0.11 | 0.16 | 0.17 | 0.18 | 0.20 |
| Private Coverage | 0.73 | 0.75 | 0.75 | 0.71 | 0.73 | 0.71 | 0.69 |
| Imputed Eligibility | 0.19 | 0.18 | 0.18 | 0.27 | 0.29 | 0.31 | 0.34 |
| Size of Health Insurance Unit | 4.22 | 4.16 | 4.18 | 4.16 | 4.22 | 4.17 | 4.22 |
| White | 0.83 | 0.83 | 0.82 | 0.78 | 0.82 | 0.80 | 0.81 |
| Male | 0.51 | 0.51 | 0.51 | 0.51 | 0.51 | 0.52 | 0.51 |
| Two Parents | 0.76 | 0.77 | 0.77 | 0.71 | 0.74 | 0.73 | 0.73 |
| Male Head Only | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 | 0.03 | 0.02 |
| No Earners | 0.14 | 0.13 | 0.13 | 0.16 | 0.15 | 0.16 | 0.16 |
| One Earner | 0.41 | 0.42 | 0.42 | 0.43 | 0.42 | 0.41 | 0.41 |
| Two Earners | 0.38 | 0.39 | 0.40 | 0.37 | 0.38 | 0.38 | 0.38 |
| Age of Highest Earner | 36.49 | 36.47 | 36.49 | 36.67 | 36.90 | 36.89 | 37.10 |
| High School Drop-out | 0.22 | 0.21 | 0.20 | 0.21 | 0.19 | 0.18 | 0.19 |
| High School Graduate | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.37 | 0.35 |
| Some College | 0.20 | 0.19 | 0.20 | 0.20 | 0.20 | 0.21 | 0.22 |
| HSFRAC | 0.20 | 0.19 | 0.19 | 0.28 | 0.30 | 0.32 | 0.34 |
| Years Covered | 86-88 | 87-89 | 88-89 | 90-92 | 91-93 | 92-95 | 93-95 |
| Number of Observations | 175709 | 182307 | 163165 | 397187 | 267589 | 433545 | 406833 |

Notes: Shown are unweighted means from the respective SIPP panels noted above. See the text for a description of the sample construction.

Table 2: Estimated Medicaid Average Take-up Rates, Private Insurance Coverage and Crowd-out Rate for All Eligible Children

| Group | <u>Medicaid</u> | | | <u>Private Insurance</u> | | | <u>Crowd-out</u> | |
|--------------------------|-----------------------|-----------------------|---------------|--------------------------|-----------------------|---------------|------------------------|------------------------|
| | <i>ATRE</i> (1) | <i>ATRESEL</i> (2) | Actual (3) | <i>PRE</i> (4) | <i>PRESEL</i> (5) | Actual (6) | No Selection (7) | Selection (8) |
| All | 0.4702*** (0.0028) | 0.5025*** (0.0034) | 0.51 | 0.4369*** (0.0049) | 0.3436*** (0.0040) | 0.34 | -0.0232** (0.0097) | -0.0486*** (0.0127) |
| <i>Race</i> | | | | | | | | |
| White | 0.3973*** (0.0032) | 0.4322*** (0.0039) | 0.44 | 0.4960*** (0.0055) | 0.3918*** (0.0046) | 0.38 | -0.0247** (0.0102) | -0.0535*** (0.0136) |
| Non-White | 0.6417*** (0.0042) | 0.6679*** (0.0045) | 0.67 | 0.2980** (0.0053) | 0.2304*** (0.0048) | 0.22 | -0.0198* (0.0106) | -0.0369*** (0.0125) |
| <i>Education</i> | | | | | | | | |
| High School Drop-out | 0.6041*** (0.0039) | 0.6265*** (0.0042) | 0.63 | 0.2445*** (0.0053) | 0.1809*** (0.0044) | 0.18 | 0.0037 (0.0098) | -0.0120 (0.0099) |
| High School Graduate | 0.4463*** (0.0039) | 0.4830*** (0.0045) | 0.49 | 0.4785*** (0.0061) | 0.3737*** (0.0054) | 0.36 | -0.0434*** (0.0156) | -0.0731*** (0.0135) |
| Some College | 0.3859*** (0.0060) | 0.4277*** (0.0069) | 0.44 | 0.5764*** (0.0075) | 0.4586*** (0.0080) | 0.46 | -0.0500*** (0.0109) | -0.0862*** (0.0140) |
| College Graduate | 0.1298*** (0.0062) | 0.1644*** (0.0080) | 0.16 | 0.8385*** (0.0063) | 0.7200*** (0.0090) | 0.69 | 0.0004 (0.0098) | 0.0220* (0.0132) |
| <i>Family Structure</i> | | | | | | | | |
| Female Head | 0.6901*** (0.0036) | 0.7141*** (0.0038) | 0.71 | 0.2626*** (0.0045) | 0.2060*** (0.0039) | 0.21 | -0.0364*** (0.0104) | -0.0518*** (0.0115) |
| Male Head | 0.2771*** (0.0120) | 0.3108*** (0.0133) | 0.30 | 0.4851*** (0.0154) | 0.3645*** (0.0159) | 0.37 | -0.0308 (0.0220) | -0.0641*** (0.0213) |
| Two parents | 0.2592*** (0.0036) | 0.2998*** (0.0046) | 0.31 | 0.6103*** (0.0064) | 0.4815*** (0.0057) | 0.46 | -0.0095 (0.0104) | -0.0444*** (0.0134) |
| <i>Number of Earners</i> | | | | | | | | |
| No earner | 0.7739*** (0.0035) | 0.7870*** (0.0035) | 0.78 | 0.1541*** (0.0036) | 0.1247*** (0.0031) | 0.13 | -0.0183* (0.0105) | -0.0255*** (0.0102) |
| One earner | 0.2707*** (0.0037) | 0.3161*** (0.0047) | 0.33 | 0.6098*** (0.0071) | 0.4777*** (0.0060) | 0.46 | -0.0342*** (0.0116) | -0.0708*** (0.0145) |
| Two earners | 0.1153*** (0.0041) | 0.1698*** (0.0062) | 0.17 | 0.8162*** (0.0077) | 0.6328*** (0.0085) | 0.61 | -0.0010 (0.0096) | -0.0521*** (0.0149) |
| More than two earners | 0.0812*** (0.0090) | 0.1199*** (0.0124) | 0.14 | 0.8626*** (0.0143) | 0.7212*** (0.0224) | 0.63 | 0.0048 (0.0169) | 0.0290 (0.0228) |

Notes: All index functions include demographic main effects, year, age, and state dummies and are clustered by child when estimated. Estimates are based on parameters estimated for the entire sample by minimum distance and the characteristics of the children in the 1995 SIPP data. Columns (1), (4), and (7) ignore selection while columns (2), (5), and (8) account for selection*** significantly different from zero at 1%; ** significantly different from zero at 5%; * significantly different from zero at 10%.

Table 3: Estimated Medicaid Average Take-up Rates, Private Insurance Coverage and Crowd-out Rate for Children Newly Eligible after Raising the 1995 Income Limits by 10%

| Group | <u>Medicaid</u> | | <u>Private Insurance</u> | | <u>Crowd-out</u> | |
|--------------------------|-----------------------|--------------------------|--------------------------|-------------------------|------------------------|------------------------|
| | <i>NEWATRE</i> (1) | <i>NEWATRESEL</i> (2) | <i>NEWPRE</i> (3) | <i>NEWPRESEL</i> (4) | No Selection (5) | Selection (6) |
| All | 0.2035*** (0.0048) | 0.2652*** (0.0067) | 0.7019*** (0.0094) | 0.5058*** (0.0092) | -0.0236** (0.0114) | -0.0799** (0.0176) |
| <i>Race</i> | | | | | | |
| White | 0.1762*** (0.0047) | 0.2361*** (0.0067) | 0.7254*** (0.0094) | 0.5310*** (0.0095) | -0.0259** (0.0112) | -0.0834*** (0.0175) |
| Non-White | 0.3132*** (0.0070) | 0.3818*** (0.0089) | 0.6075*** (0.0113) | 0.4047*** (0.0108) | -0.0143 (0.0143) | -0.0656*** (0.0199) |
| <i>Education</i> | | | | | | |
| High School Drop-out | 0.2914*** (0.0065) | 0.3519*** (0.0082) | 0.4957*** (0.0125) | 0.2965*** (0.0099) | 0.0210 (0.0138) | -0.0274 (0.0178) |
| High School Graduate | 0.2080*** (0.0059) | 0.2733*** (0.0080) | 0.7145*** (0.0107) | 0.5106*** (0.0109) | -0.0494*** (0.0126) | -0.1122*** (0.0191) |
| Some College | 0.1938*** (0.0068) | 0.2625*** (0.0092) | 0.7610*** (0.0105) | 0.5557*** (0.0124) | -0.0439*** (0.0114) | -0.1104*** (0.0182) |
| College Graduate | 0.0492*** (0.0040) | 0.0909*** (0.0071) | 0.9375*** (0.0051) | 0.7876*** (0.0108) | 0.0074 (0.0064) | -0.0274* (0.0148) |
| <i>Family Structure</i> | | | | | | |
| Female Head | 0.3808*** (0.0077) | 0.4559*** (0.0092) | 0.5571*** (0.0118) | 0.3676*** (0.0103) | -0.0520** (0.0152) | -0.1057*** (0.0196) |
| Male Head | 0.1467*** (0.0106) | 0.2063*** (0.0140) | 0.6611*** (0.0192) | 0.4223*** (0.0211) | -0.0412* (0.0215) | -0.1125*** (0.0286) |
| Two parents | 0.1444*** (0.0045) | 0.2015*** (0.0066) | 0.7550*** (0.0090) | 0.5590*** (0.0096) | -0.0113 (0.0105) | -0.0690*** (0.0173) |
| <i>Number of Earners</i> | | | | | | |
| No earner | 0.6346*** (0.0085) | 0.6626*** (0.0086) | 0.2060*** (0.0080) | 0.1461*** (0.0065) | 0.0013 (0.0123) | -0.0154 (0.0124) |
| One earner | 0.2107*** (0.0055) | 0.2756*** (0.0073) | 0.6842*** (0.0105) | 0.4856*** (0.0098) | -0.0335*** (0.0130) | -0.0900*** (0.0187) |
| Two earners | 0.1015*** (0.0049) | 0.1643*** (0.0078) | 0.8370*** (0.0092) | 0.6141*** (0.0115) | -0.0098 (0.0099) | -0.0771*** (0.0184) |
| More than two earners | 0.0676*** (0.0090) | 0.1173*** (0.0140) | 0.9074*** (0.0130) | 0.7308*** (0.0253) | 0.0210** (0.0100) | -0.0113 (0.0248) |

Notes: See notes to Table 2. Columns (1), (3), and (5) ignore selection while columns (2), (4), and (6) account for selection.

Appendix A: Econometric Model When There are Interaction Effects for Eligible Families

A.1 Contribution of Medicaid Ineligible Families

The analysis for ineligible families is not affected by allowing for interaction effects and their contribution to the likelihood function will not change. Combining equations (5) and (9) in the text, an ineligible family that does not purchase private insurance will contribute the following term to the likelihood function for insurance choice

$$\Pr(\text{elig}^* \leq 0, B_{pr,i} - C_{pr,i} \leq 0) = \Pr(Z_i\delta + e_i \leq 0, X_i\gamma_{ne} + u_{ne,i} \leq 0) \quad (\text{A1})$$

while an ineligible family that purchases private insurance will contribute

$$\Pr(\text{elig}^* \leq 0, B_{pr,i} - C_{pr,i} > 0) = \Pr(Z_i\delta + e_i \leq 0, X_i\gamma_{ne} + u_{ne,i} > 0). \quad (\text{A2})$$

A.2 Contribution of Medicaid Eligible Families

Recall that preferences take the form

$$U_i(D_{pr,i}, D_{pub,i}) = I_i - C_{pr,i}D_{pr,i} - C_{pub,i}D_{pub,i} + B_{pr,i}D_{pr,i} + B_{pub,i}D_{pub,i} + INT_i\{D_{pr,i} * D_{pub,i}\}.$$

In this section we will assume

$$INT_i = X_i\tau + v_i.$$

where X_i is a vector of demographic variables and v_i is an error term.

A.2.1 Contribution of Medicaid Eligible Families Who Have No Insurance

The family will choose no insurance if utility is highest in this state, which implies:

i) No insurance dominates only having private insurance

$$I_i \geq I_i + B_{pr,i}D_{pr,i} - C_{pr,i}D_{pr,i} \text{ or } B_{pr,i} - C_{pr,i} \leq 0;$$

ii) No insurance dominates only public insurance

$$I_i \geq I_i + B_{pub,i}D_{pub,i} - C_{pub,i}D_{pub,i} \text{ or } B_{pub,i} - C_{pub,i} \leq 0;$$

iii) No insurance dominates having both public and private insurance

$$I_i \geq I_i + B_{pub,i}D_{pub,i} - C_{pub,i}D_{pub,i} + B_{pr,i}D_{pr,i} - C_{pr,i}D_{pr,i} + INT_i$$

or

$$B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \leq 0.$$

Using i) - iii) these families will make the following contribution to the likelihood function

$$\begin{aligned} & \Pr(\text{elig}^* > 0, B_{pub,i} - C_{pub,i} \leq 0, B_{pr,i} - C_{pr,i} \leq 0, B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \leq 0) \\ & = \Pr(Z_i\delta + e_i > 0, X_i\mu + \varepsilon_i \leq 0, X_i\gamma_e + u_{e,i} \leq 0, X_i\mu + \varepsilon_i + X_i\gamma_e + u_{e,i} + X_i\tau + v_i \leq 0). \end{aligned} \quad (A3)$$

A.2.2 Contribution of Medicaid Eligible Families Who Have Only Private Insurance

The family will choose only private insurance if utility is highest in this state, which implies:

i) Private insurance dominates no insurance

$$I_i < I_i + B_{pr,i} - C_{pr,i} \text{ or } B_{pr,i} - C_{pr,i} > 0;$$

ii) Private insurance dominates only public insurance

$$I_i + B_{pr,i} - C_{pr,i} \geq I_i + B_{pub,i} - C_{pub,i} \text{ or } B_{pr,i} - C_{pr,i} \geq B_{pub,i} - C_{pub,i};$$

iii) Private insurance only dominates having both public insurance and private insurance

$$I_i + B_{pr,i} - C_{pr,i} \geq I_i + B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \text{ or } 0 \geq B_{pub,i} - C_{pub,i} + INT_i;$$

Using i) - iii) these families will contribute

$$\Pr(\text{elig}^* > 0, B_{pub,i} - C_{pub,i} \leq B_{pr,i} - C_{pr,i}, B_{pr,i} - C_{pr,i} > 0, B_{pub,i} - C_{pub,i} + INT_i \leq 0) \text{ or}$$

$$\Pr(Z_i\delta + e_i > 0, X_i\mu + \varepsilon_i \leq X_i\gamma_e + u_{e,i}, X_i\gamma_e + u_{e,i} > 0, X_i\mu + \varepsilon_i + X_i\tau + v_i \leq 0). \quad (A4)$$

A.2.3 Contribution of Medicaid Eligible Families Who Have Only Public Insurance

The family will choose only public insurance if utility is highest in this state, which implies:

i) Only public insurance dominates no insurance

$$I_i < I_i + B_{pub,i} - C_{pub,i} \text{ or } B_{pub,i} - C_{pub,i} > 0;$$

ii) Only public insurance dominates only private

$$\text{insurance } I_i + B_{pr,i} - C_{pr,i} \leq I_i + B_{pub,i} - C_{pub,i} \text{ or } B_{pr,i} - C_{pr,i} \leq B_{pub,i} - C_{pub,i};$$

iii) Only public insurance dominates having both public insurance and private insurance

$$I_i + B_{pub,i} - C_{pub,i} \geq I_i + B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \quad \text{or} \quad 0 \geq B_{pr,i} - C_{pr,i} + INT_i.$$

Using i) - iii) these families will contribute

$$\begin{aligned} & \Pr(elig^* > 0, B_{pub,i} - C_{pub,i} \geq B_{pr,i} - C_{pr,i}, B_{pub,i} - C_{pub,i} > 0, B_{pr,i} - C_{pr,i} + INT_i \leq 0) \\ & = \Pr(Z_i\delta + e_i > 0, X_i\mu + \varepsilon_i \geq X_i\gamma_e + u_{e,i}, X_i\mu + \varepsilon_i > 0, X_i\gamma + u_i + X_i\tau + v_i \leq 0). \end{aligned} \quad (A5)$$

A.2.4 Contribution of Medicaid Eligible Families Who Have Both Private and Public Insurance

The family will choose both private and public insurance if utility is highest in this state, which implies:

i) Having both public insurance and private insurance dominates no insurance

$$I_i < I_i + B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \quad \text{or} \quad B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i > 0;$$

ii) Having both public insurance and private insurance dominates only private insurance

$$I_i + B_{pr,i} - C_{pr,i} < I_i + B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \quad \text{or} \quad 0 < B_{pub,i} - C_{pub,i} + INT_i;$$

ii) Having both public insurance and private insurance dominates only public insurance

$$I_i + B_{pub,i} - C_{pub,i} < I_i + B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i \quad \text{or} \quad 0 < B_{pr,i} - C_{pr,i} + INT_i;$$

Using i) - iii) these families will contribute

$$\begin{aligned} & \Pr(elig^* > 0, B_{pub,i} - C_{pub,i} + B_{pr,i} - C_{pr,i} + INT_i > 0, B_{pub,i} - C_{pub,i} + INT_i > 0, B_{pr,i} - C_{pr,i} + INT_i > 0) \\ & = \Pr(Z_i\delta + e_i > 0, X_i\mu + \varepsilon_i + X_i\gamma_e + u_{e,i} + X_i\tau + v_i > 0, X_i\mu + \varepsilon_i + X_i\tau + v_i > 0, X_i\gamma_e + u_{e,i} + X_i\tau + v_i > 0). \end{aligned} \quad (A6)$$

This appendix illustrates two points. First, the equations determining private insurance coverage will now be quite different between the eligible and ineligible families. Second, a likelihood function for the Medicaid eligible based on the product of the contributions (A1)-(A6) is much

more complicated than the one in the text and we do not pursue maximizing this likelihood function here. Instead we estimate separate equations determining the private insurance decision for the eligible and ineligible families.

Appendix B: Covariance Matrix of All First Stage Parameter Estimates

To implement the minimum distance estimator we need an estimate of Ω . The diagonal will consist of the variance-covariance matrices for $(\hat{\mu}^t, \hat{\delta}^{(1)t})$, $(\hat{\gamma}_e^t, \hat{\delta}^{(2)t})$, and $(\hat{\gamma}_{ne}^t, \hat{\delta}^{(3)t})$, respectively, all of which will be consistently estimated by Stata as long as one clusters by family. However we need to calculate the off-diagonal sub-matrices $\text{cov}[(\hat{\gamma}_e^t, \hat{\delta}^{(2)t}), \hat{\gamma}_{ne}^t, \hat{\delta}^{(3)t}]$, $\text{cov}[(\hat{\gamma}_e^t, \hat{\delta}^{(2)t}), \hat{\mu}^t, \hat{\delta}^{(1)t}]$ and $\text{cov}[(\hat{\gamma}_{ne}^t, \hat{\delta}^{(3)t}), \hat{\mu}^t, \hat{\delta}^{(1)t}]$. To do this, let $\varphi_1^t = (\mu_e^t, \delta^{(1)t})$, $\varphi_2^t = (\gamma_e^t, \delta^{(2)t})$ and $\varphi_3^t = (\gamma_{ne}^t, \delta^{(3)t})$.

We know that asymptotically

$$\hat{\varphi}_1 - \varphi_1 = \left(\frac{\partial^2 L_1(\hat{\varphi}_1)}{\partial \varphi_1 \partial \varphi_1^t} \right)^{-1} \left(\frac{\partial L_1(\hat{\varphi}_1)}{\partial \varphi_1} \right) = \left(\frac{\partial^2 L_1(\hat{\varphi}_1)}{\partial \varphi_1 \partial \varphi_1^t} \right)^{-1} \left(\sum_i \left[\sum_t \frac{\partial L_{1it}(\hat{\varphi}_1)}{\partial \varphi_1} \right] \right), \quad (\text{B1})$$

where an estimate of the inverse of the second derivative matrix in (B1) is available from Stata when one maximizes the first likelihood function – it is the variance covariance matrix of the parameters if one does not cluster the standard errors.

Since similar expressions hold for $\varphi_2 - \hat{\varphi}_2$ and $\varphi_3 - \hat{\varphi}_3$ we can use them to obtain the off-diagonal elements of Ω . For example, again allowing observations across children from the same family to be correlated, we have

$$\begin{aligned}
\text{Cov}(\hat{\varphi}_1, \hat{\varphi}_2) &= E\left([\hat{\varphi}_1 - \varphi_1][\hat{\varphi}_2 - \varphi_2]'\right) \\
&= \left(\frac{\partial^2 L_1(\hat{\varphi}_1)}{\partial \varphi_1 \partial \varphi_1'}\right)^{-1} E\left(\sum_{i=1}^N \left[\sum_t \frac{\partial L_{1it}(\hat{\varphi}_1)}{\partial \varphi_1}\right]\right) \left(\sum_{j=1}^N \left[\sum_t \frac{\partial L_{2jt}(\hat{\varphi}_2)}{\partial \varphi_2}\right]\right) \left(\frac{\partial^2 L_2(\hat{\varphi}_2)}{\partial \varphi_2 \partial \varphi_2'}\right)^{-1} \\
&= \left(\frac{\partial^2 L_1(\hat{\varphi}_1)}{\partial \varphi_1 \partial \varphi_1'}\right)^{-1} E\left\{\sum_{f=1}^F \left(\sum_{i \in f} \left[\sum_t \frac{\partial L_{1it}(\hat{\varphi}_1)}{\partial \varphi_1}\right]\right) \left(\sum_{j \in f} \left[\sum_t \frac{\partial L_{2it}(\hat{\varphi}_2)}{\partial \varphi_2}\right]\right)\right\} \left(\frac{\partial^2 L_2(\hat{\varphi}_2)}{\partial \varphi_2 \partial \varphi_2'}\right)^{-1}.
\end{aligned} \tag{B2}$$

The other covariance terms can be calculated in a similar fashion.