

**Estimating Heterogeneous Take-up and Crowd-Out
Responses to Current Medicaid Limits and Their Nonmarginal Expansions**

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

I. Serkan Ozbeklik, Claremont McKenna College

Lara Shore-Sheppard, Williams College and NBER

Abstract

We use a switching probit model and the income-limit-based structure of Medicaid eligibility for children to estimate treatment effects of Medicaid expansions not found in existing work on public insurance. In particular, we estimate rates of Medicaid take-up, private insurance coverage, and crowd-out for the currently eligible overall and for different demographic groups, and we estimate corresponding rates for children made eligible by a counterfactual nonmarginal increase in the Medicaid income limits. We find strikingly different rates across demographic groups, with individuals in traditionally less disadvantaged groups having a considerably lower response to the coverage for which they are eligible.

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

1. Introduction

In recent years, eligibility for public health insurance has expanded substantially, leading to a burgeoning of research on the implications of such expansions for public insurance participation, private insurance coverage and crowd-out, and the overall levels of health insurance coverage. A common approach to these questions is to estimate a linear probability model 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, while quite useful, has several drawbacks. Crucially, the desirability of Medicaid and the availability of outside insurance options are likely to differ across families. Families with both observable differences (from different demographic groups, for example) and unobservable differences (with different tastes for public insurance, for example) may respond to eligibility in different ways. As a result, there are treatment effects likely to be of substantial use to policymakers that are not captured by the LATE estimates. These include: i) levels of crowd-out among all *currently eligible* individuals; ii) take-up and private coverage responses and levels of crowd-out among currently eligible individuals in observably different population subgroups; and iii) take-up, private coverage, and crowd-out among individuals made eligible by

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.

a *nonmarginal* expansion of eligibility (both all of those made newly eligible and those made newly eligible in observably different population subgroups). Moreover, the composition of the group to which the LATE estimates apply is unobserved by the researcher.²

In this paper we show how to address these drawbacks using a variant of the switching probit model and the fact that Medicaid eligibility is based on explicitly stated criteria. We use this model to provide estimates of take-up, private coverage, and crowd-out for all currently eligible children and currently eligible children from different demographic 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. Using this model we also calculate the first estimates in the literature of the effect on take-up, private coverage, and crowd-out for children who would be made eligible by a nonmarginal Medicaid expansion, and then examine how these effects vary across different subgroups of these newly eligible children. Importantly, throughout the paper we use existing Stata commands to obtain our estimates.

The estimates that our model generates allow us to 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 levels of take-up and crowd-out? (3) Within the currently eligible or the counterfactually newly eligible, which population subgroups respond more to

² The estimated LATEs for the entire sample also provide no information about how the LATE effects vary across population subgroups; this omission is corrected in Ham, Ozbeklik and Shore-Sheppard (2012) using a very different approach than that considered here.

eligibility by enrolling or by dropping private coverage, and how large are those differences between groups?

Recently, the empirical program evaluation literature has seen a number of papers that explicitly estimate heterogeneous treatment effects. The majority of these effects have been estimated within the education and training literature (see, e.g., Aakvik, Heckman, and Vytlačil (2005—hereafter AHV), Blundell, Dearden, and Sianesi (2005), Card and Payne (2002), Heckman, Smith, and Clements (1997), Moffitt (2007), and Carneiro, Heckman, and Vytlačil (2011)), although there are other applications as well.³ One popular approach in this literature is to estimate a linear probability model where the treatment dummy variable is interacted with demographic variables to estimate LATEs for different demographic groups. Another approach is to use methods derived from the switching probit model (SPM) to predict the effect of changing treatment on the basis of observables and unobservables to estimate marginal effects.⁴

³ For example, Angrist (2004) allows for treatment effect heterogeneity in the effects of childbearing on marital dissolutions, poverty status, and welfare participation while Ham, Ozbeklik and Shore-Sheppard (2012) estimate heterogeneous treatment effects of Medicaid expansions.

⁴ The first approach is used by Blundell, Dearden, and Sianesi (2005), Card and Payne (2002), and Ham, Ozbeklik and Shore-Sheppard (2012), while the second is used by AHV. The approach introduced in Moffitt (2007) can be considered a combination of the two, since he allows treatment effects to depend on both observables and unobservables that are correlated with the error in the outcome of interest within a random coefficients framework. Heckman, Smith, and Clements (1997) use propensity score matching allowing for treatment heterogeneity, while Angrist (2004) uses an IV framework to show the relationship between LATEs and average

We use the SPM and the explicit eligibility criteria for Medicaid to extend the program evaluation literature in several ways. First, we show how to estimate the impact, on any demographic group, of nonmarginal changes in the Medicaid income limits that determine eligibility. This approach is suitable for any social program where eligibility is based on observable criteria, as opposed to programs (e.g. job training) which have less clear-cut criteria. (Moreover, by making the change in the income limits small, our approach can also be used to see the impact of marginal changes on any demographic group.) We then use our model to calculate predictions for Medicaid take-up and private coverage, again across demographic groups, for those currently eligible. We compare the predicted take-up and private coverage for each demographic group to its empirical counterpart to perform an informal investigation of how well our approach is likely to do in practice. Further, we calculate crowd-out for the currently eligible, again across demographic groups. Finally, we estimate Medicaid take-up, private insurance coverage, and crowd-out for those made newly eligible by a counterfactual 10 percent increase (across the board) in the Medicaid income limits. These estimates of the effects of nonmarginal changes in the income limits, and crowd-out among the currently eligible, are currently unavailable in the literature for both the entire sample and different demographic groups.

Our approach produces sensible and relatively precise effects for Medicaid take-up and private insurance coverage among 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

treatment effects under certain assumptions. Finally, Carneiro, Heckman, and Vytlacil (2011) use local instrumental variables estimators to estimate the effect of marginal policy changes, although they also estimate a normal switching regression model for comparison.

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. As a result, our estimates should prove very helpful to policy makers concerned with outreach to underserved groups. Our results suggest that the marginal group for whom the standard LATE estimate from the LPM applies is two earner families. 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 opposite is true for private insurance coverage. Estimating precise crowd-out effects among the newly eligible is more challenging, but our results suggest similar levels of crowd-out among the newly eligible as among the currently eligible given the respective standard errors.

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). Further, 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)). The above 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 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.⁵ Definitions of “crowd-out”—loosely, the effect of 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.

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 (discussed briefly in Section 3.1 below) since eligibility is likely to be endogenous. This potential endogeneity arises for several reasons. First, unobservable factors affecting eligibility are likely to be correlated with unobservable factors driving family characteristics that determine eligibility. Second, eligibility may serve as a proxy for family income if income, which is also likely to be endogenous, is not included as an independent variable. Finally, parental wages, which in turn determine eligibility, are likely to be correlated with fringe benefits

⁵ These distinctions have often been missed in the literature.

(including private health insurance) of the parent. Since these benefits are unobserved, they are part of the error term, thus providing an additional factor necessitating treating eligibility as endogenous.

To address the endogeneity of the eligibility variable, CG suggest an instrument (which we denote *FRACELIG*) 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.⁶ 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 *FRACELIG*.

Ham and Shore-Sheppard (2005–HS hereafter) use data from the SIPP covering the period from October 1985 to August 1995 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. HS also modify the CG instrument by using all sample observations of children of a given age in a SIPP wave except for those from the state for which the instrument is being calculated. Since this instrument is created using a larger sample, it is theoretically superior to the version using a random sample, but in

⁶ One attractive feature of this approach is that *FRACELIG* is an extremely strong instrument.

practice its use makes very little difference to the results. We use the data and instrument of HS in our estimation.

3. Estimating the Effects of Marginal and Nonmarginal (Counterfactual) Medicaid

Expansions across Demographic Groups

3.1 The Conventional Linear Probability Model for Medicaid Expansions

To set the stage for our approach, we first describe the simultaneous equation LPM approach typical in the Medicaid expansion literature. The LPM for participation in a public insurance program is given by

$$pub_i = X_i\beta_1 + \gamma_1 elig_i + u_{1i}, \quad (1a)$$

where X_i is a vector of demographic variables for child i , $elig_i$ is a dummy variable coded one if the child is eligible for public insurance and zero otherwise, u_{1i} is an error term, and $pub_i = 1$ if child i participates in a public insurance program, otherwise $pub_i = 0$. The LPM for private insurance coverage is given by

$$priv_i = X_i\beta_2 + \gamma_2 elig_i + u_{2i}, \quad (1b)$$

where $priv_i = 1$ if child i has private insurance coverage and $priv_i = 0$ otherwise. The LPM determining public insurance eligibility is

$$elig_i = Z_i\delta + e_i, \quad (2)$$

where $Z_i = (X_i, FRACELIG_i)$, and e_i 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 $FRACELIG$, averaged across the different marginal changes in the data. As noted above, this approach is 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. Third, the LATE estimates do not measure treatment effects among the entire currently eligible group. Finally, LATE estimates are not informative about the effects of nonmarginal changes.

3.2 Estimating the Switching Probit Model

To address these limitations, we use the switching probit model (SPM) developed by Quandt (1958, 1960, 1972) and Heckman (1979), which has been applied in the bivariate setting with binary outcomes and selection by van de Ven and van Praag (1981). This model allows us to control for both observables and unobservables, once we make a distributional assumption, when estimating policy effects for the entire currently eligible population, for different demographic groups among the currently eligible, and for individuals eligible under a nonmarginal expansion of eligibility. Our approach builds on the important work of AHV and the papers they cite. 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.

As is clear from the material we present below, our approach differs from that of AHV in several important ways. First, we estimate policy effects for different demographic groups. Second, Medicaid eligibility--the treatment in our analysis--is observable given family income, state of residence, year and child's age, and we exploit this feature in estimation and in calculating treatment effects (just as earlier papers did using the LPM). In particular, it allows us to examine nonmarginal (counterfactual) Medicaid income limit expansions. In other words, when we conduct a counterfactual policy experiment, we can identify whose treatment status

would actually change, and not just the change in the probability of receiving treatment. Knowing who is eligible also allows us to consider how well our models mirror the data.⁷ We do this by comparing noncounterfactual quantities (such as the take-up rate among all currently eligible children) with the same quantities predicted from the model. However, we emphasize that these advantages come at the cost of assuming there is no measurement error in eligibility. This assumption is required for the consistent estimation of treatment effects in our approach, the LPM model of CG, and the linear probability model with interactions (LPMI) model of Ham, Ozbeklik and Shore-Sheppard (2012), but not in AHV.⁸ Third, our estimated policy effects for Medicaid are precisely estimated, while this was not the case for AHV. Fourth, we are able to use preprogrammed commands in Stata to estimate all relevant effects and most standard errors.

The cost of the SPM versus the LPM would seem to be the need for a normality assumption, but 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

⁷ Note that these features of Medicaid apply to other programs, such as the Supplemental Nutrition Assistance Program (formerly Food Stamps).

⁸ CG's IV approach does not eliminate measurement error bias since the measurement error in eligibility is correlated with the true value of eligibility; if true eligibility is 0, the measurement error must be 0 or 1, while if true eligibility is 1, the measurement error must be 0 or -1.

on their Monte Carlo evidence, Bhattacharya, Goldman, and McCaffrey (2006) argue in favor of using simultaneous equation bivariate probit, rather than the linear probability model, when the error terms are *not* normally distributed. 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, but unfortunately Shaikh and Vytlacil (2005) show that, in general, only bounds on the average effect of treatment can be identified in the case of a triangular system of equations where both dependent variables are binary such as we have here. Further, AHV stress the need for making a distributional assumption for treatment effects such as those considered below. Of course, researchers are free to estimate the model with a different distributional assumption, e.g. assume that the error terms are drawn from a mixture of normals (see e.g., Carneiro, Hansen, and Heckman (2003)). We do not adopt this approach since it would force us to violate our commitment to using prewritten Stata modules.

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. However, this is quite a complicated model to estimate since it requires trivariate integration, and there is no way to use preprogrammed routines in Stata to estimate this model. Fortunately, it is straightforward to show that all parameters can be estimated consistently using preprogrammed routines in Stata by separately maximizing two likelihood functions: one containing the parameters for eligibility and for public insurance participation and one containing the parameters for eligibility and for private insurance participation. Using separate likelihood

functions has the advantage that misspecification in the public take-up equation will not cause bias in the private participation equation or vice-versa, and the disadvantage that one gives up some efficiency in estimation; these issues are essentially identical to those that come up in the trade-off between two-stage and three-stage least squares.

We first construct the likelihood function containing the parameters for eligibility and for public insurance participation. We assume that the index function for eligibility is

$$Elig_i^* = Z_i\delta + e_i, \quad (3)$$

where Z_i is defined as in (2). Next, we assume that the index function for participation in public insurance, once being made exogenously eligible for Medicaid, is

$$PPub_i^* = X_i\mu + \varepsilon_i, \quad (4)$$

where $(\varepsilon_i, e_i) \sim iid N(0, V_{pub})$ and $V_{pub} = \begin{bmatrix} 1 & \rho_{\varepsilon, e} \\ \rho_{\varepsilon, e} & 1 \end{bmatrix}$. Following the econometrics literature, we

consider the index function for take-up for a randomly chosen (in terms of ε_i) child exogenously made eligible; note that this also allows us to do counterfactual policy analysis. However, it is worth emphasizing that unless $\rho_{\varepsilon, e} = 0$, those actually eligible will not be a randomly chosen subgroup of the population (so estimating the take-up equation using only the eligible would result in selection bias). The resulting log likelihood for eligibility and participation in public insurance is

$$\begin{aligned} L &= \sum_{elig=1, pub=1} \log[\Pr(elig_i = 1, pub_i = 1)] + \sum_{elig=1, pub=0} \log[\Pr(elig_i = 1, pub_i = 0)] + \sum_{elig=0} \log[\Pr(elig_i = 0)] \\ &= \sum_{elig=1, pub=1} \log \Phi_2[(X_i\mu, Z_i\delta, \rho_{\varepsilon, e})] + \sum_{elig=1, pub=0} \log \Phi_2[-X_i\mu, Z_i\delta, -\rho_{\varepsilon, e}] + \sum_{elig=0} \log \Phi_1[-Z_i\delta], \end{aligned} \quad (5)$$

where $\Phi_2[\cdot, \cdot, \rho_{\varepsilon, e}]$ is the bivariate standard normal distribution function and $\Phi_1(\cdot)$ is the univariate standard normal distribution function. For Medicaid take-up calculations we will only need the parameters estimated by maximizing (5).

For private coverage and crowd-out calculations, we need to incorporate participation in private insurance. We assume that for a randomly chosen individual, the index function for participation in private insurance, given exogenous eligibility for public insurance, is

$$Priv_elig_i^* = X_i \gamma_e + u_{ei}. \quad (6a)$$

We define the observed outcome variable $priv_elig_i = 1$ if $Priv_elig_i^* > 0$ and zero otherwise.

Further, we assume that for a randomly chosen individual the index function for participation in private insurance given exogenous ineligibility for public insurance is

$$Priv_nelig_i^* = X_i \gamma_{ne} + u_{nei}, \quad (6b)$$

and define $priv_nelig_i = 1$ if $Priv_nelig_i^* > 0$ and zero otherwise. Finally we assume⁹

$$(u_{ei}, u_{nei}, e_i) \sim iidN(0, V_{priv}) \text{ and } V_{priv} = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix}. \quad (6c)$$

The appropriate log likelihood function in this case is

$$L = \sum_{\substack{elig=1, \\ priv=1}} \log[\Pr(elig_i = 1, priv_i = 1)] + \sum_{\substack{elig=1, \\ priv=0}} \log[\Pr(elig_i = 1, priv_i = 0)] \\ + \sum_{\substack{elig=0, \\ priv=1}} \log[\Pr(elig_i = 0, priv_i = 1)] + \sum_{\substack{elig=0, \\ priv=0}} \log[\Pr(elig_i = 0, priv_i = 0)] \quad (7)$$

⁹ Note that $\rho_{12} = corr(u_{ei}, u_{nei})$ is not identified (because no one is simultaneously eligible and ineligible for Medicaid) but also is not needed for any calculations below.

$$\begin{aligned}
&= \sum_{\substack{elig=1, \\ priv=1}} \log \Phi_2[(X_i \gamma_e, Z_i \delta, \rho_{13})] + \sum_{\substack{elig=1, \\ priv=0}} \log \Phi_2[-X_i \gamma_e, Z_i \delta, -\rho_{13}] + \\
&\quad \sum_{\substack{elig=0, \\ priv=1}} \log \Phi_2[(X_i \gamma_{ne}, -Z_i \delta, -\rho_{23})] + \sum_{\substack{elig=0, \\ priv=0}} \log \Phi_2[-X_i \gamma_{ne}, -Z_i \delta, \rho_{23}].
\end{aligned}$$

Unfortunately, there is no preprogrammed routine in Stata to maximize this function. Instead, to obtain an approach that can be implemented with preprogrammed commands in Stata, we randomly divide the sample into two subsamples, each containing half the sample. On the first subsample we focus on the take-up of private insurance by those *currently eligible* for Medicaid and maximize

$$\begin{aligned}
L &= \sum_{\substack{elig=1, \\ priv=1}} \log [\Pr(elig_i = 1, priv_i = 1)] + \sum_{\substack{elig=1, \\ priv=0}} \log [\Pr(elig_i = 1, priv_i = 0)] + \sum_{elig=0} \log [\Pr(elig_i = 0)] \\
&= \sum_{\substack{elig=1, \\ priv=1}} \log \Phi_2[(X_i \gamma_e, Z_i \delta, \rho_{13})] + \sum_{\substack{elig=1, \\ priv=0}} \log \Phi_2[-X_i \gamma_e, Z_i \delta, -\rho_{13}] + \sum_{elig=0} \log \Phi_1[-Z_i \delta].
\end{aligned} \tag{8}$$

Next, on the other subsample we focus on the take-up of private insurance by those *currently ineligible* for Medicaid and maximize

$$\begin{aligned}
L &= \sum_{\substack{elig=0, \\ priv=1}} \log [\Pr(elig_i = 0, priv_i = 1)] + \sum_{\substack{elig=0, \\ priv=0}} \log [\Pr(elig_i = 0, priv_i = 0)] + \sum_{elig=1} \log [\Pr(elig_i = 1)] \\
&= \sum_{\substack{elig=0, \\ priv=1}} \log \Phi_2[(X_i \gamma_{ne}, -Z_i \delta, -\rho_{23})] + \sum_{\substack{elig=0, \\ priv=0}} \log \Phi_2[-X_i \gamma_{ne}, -Z_i \delta, \rho_{23}] + \sum_{elig=1} \log \Phi_1[Z_i \delta].
\end{aligned} \tag{9}$$

Note that since we randomly chose the subsamples, the parameter estimates from (8) and (9) will be independent.

Using the estimated coefficient vectors $\hat{\mu}$, $\hat{\gamma}_e$, and $\hat{\gamma}_{ne}$ and the correlation coefficients $\hat{\rho}_{\varepsilon,e}$, $\hat{\rho}_{13}$ and $\hat{\rho}_{23}$ we can calculate policy effects for the entire sample and different demographic groups for: i) the currently eligible; ii) those who become eligible due to a marginal change; and

iii) those who become eligible due to a nonmarginal change (for example, raising all income limits by 10 percent). In calculating these effects we use the fact that the Medicaid income limits determining eligibility are observable, and therefore we can calculate counterfactual quantities (like the predicted private coverage rate among children made eligible by a nonmarginal expansion) by combining our estimated parameters with the observed characteristics of the appropriate individuals in our sample. Since the estimation of treatment effects arising from a nonmarginal change is the most novel contribution of this approach, we present the equations necessary to calculate these treatment effects in the text. The equations needed to calculate take-up, private coverage, and crowd-out for the currently eligible are presented in Appendix A, while the equations needed to calculate take-up and crowd-out for a marginal change are straightforward extensions of those presented below. For expositional ease and a 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 3.4.

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

Consider a 10 percent increase (or other nonmarginal change) 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. Thus when we ignore selection, the Medicaid take-up rate among all of the *newly eligible* (denoted by $new=1$) is given by

$$\begin{aligned}
 NEWATRE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1) \\
 &= \frac{1}{N_{new}} \sum_{new=1} \Pr(X_i \hat{\mu} + \varepsilon_i > 0) = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\mu})],
 \end{aligned} \tag{10a}$$

where N_{new} refers to the number of newly eligible individuals.¹⁰ Let $new^k = 1$ denote 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 (10a) for the subset of newly eligible in group k :

$$NEWATRE^k = \frac{1}{N_{new}^k} \sum_{new^k=1} [1 - \Phi_1(-X_i \hat{\mu})]. \quad (10b)$$

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

$$\begin{aligned} NEWPRE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1) \\ &= \frac{1}{N_{new}} \sum_{new=1} [\Pr(X_i \hat{\gamma}_e + u_{ei} > 0)] = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\gamma}_e)]. \end{aligned} \quad (11)$$

It is straightforward to estimate private insurance coverage for those newly eligible in group k as we did in moving from (10a) to (10b), 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

$$\begin{aligned} NEWPRNE &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1 | new_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1) \\ &= \frac{1}{N_{new}} \sum_{new=1} \Pr(X_i \hat{\gamma}_{ne} + u_{nei} > 0) = \frac{1}{N_{new}} \sum_{new=1} [1 - \Phi_1(-X_i \hat{\gamma}_{ne})]. \end{aligned} \quad (12)$$

¹⁰ The conditioning on $new_i = 1$ 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.

The private coverage that group k would have in the absence of becoming newly eligible is again easily calculated and we omit the equation given space constraints. Since they are counterfactuals, $NEWATR$ (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 observed to be ineligible. Further, these measures are easily calculated for a specific demographic group by taking the summation just over those in that group among the newly eligible. 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 (11) minus (12).¹¹ 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. To calculate marginal effects ignoring selection for any group k , one can simply use the above expressions for smaller increases in the income limits.

The expressions for take-up, private coverage, and crowd-out for the *currently eligible* when we ignore selection are analogous to (10a)-(12), and for expositional ease we provide them in the 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 take-up and private coverage for eligible children and private coverage for ineligible children that are

¹¹ Note that calculating the standard errors for the crowd-out estimates is simplified by the fact that the parameter estimates used in (11) and (12) will be independent.

predicted from the model to the values observed in the sample. We can also do this for any demographic group k .

3.4 Calculating Take-up and Crowd-out Among Those Made Eligible by a Counterfactual Expansion in the Income Limit 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

$$\Pr(new_i = 1) = \Pr(-Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta}) = \Phi_1(-Z_{0i}\hat{\delta}) - \Phi_1(-Z_{1i}\hat{\delta}). \quad (13)$$

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

$$\begin{aligned} NEWATRESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(pub_i = 1 \mid newelig_i = 1) = \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(pub_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\ &= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i\hat{\mu} + \varepsilon_i > 0, -Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta})}{\Pr(-Z_{1i}\hat{\delta} < e_i \leq -Z_{0i}\hat{\delta})} \\ &= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i\hat{\mu}, -Z_{0i}\hat{\delta}, \hat{\rho}_{\varepsilon,e}) - \Phi_2(X_i\hat{\mu}, -Z_{1i}\hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\Phi_1(-Z_{0i}\hat{\delta}) - \Phi_1(-Z_{1i}\hat{\delta})} \right]. \end{aligned} \quad (14a)$$

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

$$NEWATRESEL_k = \frac{1}{N_{new}^k} \sum_{new^k=1} \left[\frac{\Phi_2(X_i\hat{\mu}, -Z_{0i}\hat{\delta}, \hat{\rho}_{\varepsilon,e}) - \Phi_2(X_i\hat{\mu}, -Z_{1i}\hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\Phi_1(-Z_{0i}\hat{\delta}) - \Phi_1(-Z_{1i}\hat{\delta})} \right]. \quad (14b)$$

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

$$\begin{aligned}
NEWPRESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_elig_i = 1 | newelig_i = 1) \\
&= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(priv_elig_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\
&= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i \hat{\gamma}_e + u_{ei} > 0, -Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})}{\Pr(-Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})} \\
&= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i \hat{\gamma}_e, -Z_{0i} \hat{\delta}, \hat{\rho}_{13}) - \Phi_2(X_i \hat{\gamma}_e, -Z_{1i} \hat{\delta}, \hat{\rho}_{13})}{\Phi_1(-Z_{0i} \hat{\delta}) - \Phi_1(-Z_{1i} \hat{\delta})} \right].
\end{aligned} \tag{15}$$

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

$$\begin{aligned}
NEWPRNESEL &= \frac{1}{N_{new}} \sum_{new=1} \Pr(priv_nelig_i = 1 | newelig_i = 1) \\
&= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(priv_nelig_i = 1, newelig_i = 1)}{\Pr(newelig_i = 1)} \\
&= \frac{1}{N_{new}} \sum_{new=1} \frac{\Pr(X_i \hat{\gamma}_{ne} + u_{nei} > 0, -Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})}{\Pr(-Z_{1i} \hat{\delta} < e_i \leq -Z_{0i} \hat{\delta})} \\
&= \frac{1}{N_{new}} \sum_{new=1} \left[\frac{\Phi_2(X_i \hat{\gamma}_{ne}, -Z_{0i} \hat{\delta}, \hat{\rho}_{23}) - \Phi_2(X_i \hat{\gamma}_{ne}, -Z_{1i} \hat{\delta}, \hat{\rho}_{23})}{\Phi_1(-Z_{0i} \hat{\delta}) - \Phi_1(-Z_{1i} \hat{\delta})} \right].
\end{aligned} \tag{16}$$

Thus crowd-out for the *newly eligible* while accounting for selection is given by the difference $NEWPRESEL - NEWPRNESEL$, or (15) – (16). Finally, 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; again we have placed the expressions for these effects in the Appendix.

4. 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 period increases the probability of accurate reporting, particularly 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 (2011) 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 (2011) 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.

We include in all models 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.

5. Empirical Results

In section 5.1 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 SPM model. Here we

also consider the issue of weak instruments which has not been previously investigated in this literature. In section 5.2 we use the SPM estimates to calculate treatment effects for the currently eligible, while in section 5.3 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.

5.1 Baseline LATE for Medicaid Take-Up

Using the standard LPM, we estimate that the take-up rate averaged across marginal individuals is a very statistically significant 0.127 with a standard error of (0.01). 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 the observations are correlated or if heteroskedasticity is present. Since the first stage equation is a LPM estimated on panel data, both problems will occur in our application. 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.

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

We next estimate Medicaid take-up rates and private insurance coverage rates for the currently eligible overall and by demographic group using the 1995 data.¹² 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 (A1) in the Appendix), while in column (2) we present the estimated effects allowing for selection (equation (A4)).¹³ All of the estimates are highly statistically significant. The estimated treatment effects in both columns are quantitatively similar, although the estimates accounting for selection in column (2) are 2 to 6 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 sample take-up rates for the whole sample and the different demographic groups, again using only the 1995 data. Comparing the predicted take-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

¹² The current literature's LPM approach provides take-up, private coverage, and crowd-out estimates averaged over all of the marginally eligible, and the LPMI approach of Ham, Ozbeklik and Shore-Sheppard (2012) provides estimates for the marginally eligible in different demographic groups.

¹³ All estimates in the paper are estimates of the Average Treatment Effect on the Treated.

remarkably well and better than the estimates that ignore selection, with the exception of the take-up rate for children from families whose highest earner is a college graduate.

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. To discuss the differences in take-up across groups, we focus on the estimates in column (2). 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.¹⁴ The estimated take-up rate for children in families in which the family head has less than a high school degree is 0.62, while it is 0.27 for children in families in which the family head has a college degree or more. Moreover, the estimated take-up rate for an eligible child from a family in which a female is a single head is 0.72, 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

¹⁴ This reflects both the fact that nonwhites are more likely to take up Medicaid conditional on the other explanatory variables (see Table C1), and that their values for the other explanatory variables make them more likely to participate in Medicaid.

among the eligible requires efforts to promote public coverage primarily among populations that had not previously been eligible for coverage.

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 (A2) and (A5) in Appendix A, respectively.) These predicted private insurance coverage rates are precisely estimated and again vary widely across groups. The estimates in column (5) that account for selection are qualitatively similar to, though smaller than, the estimates in column (4) that do not account for selection. 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 0.16 with selection ignored. 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 but this is less true 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 the respective discussions in the

Appendix under equations (A3) and (A6) respectively.) The results when we do not account for selection are small and have narrow confidence intervals. However, when we account for selection in column (8), the confidence intervals widen, resulting in only one estimate that is statistically distinguishable from zero. There is a relatively small range of crowd-out rates across the different groups, however, ranging from -0.01 to -0.05,¹⁵ and the size of crowd-out for a group appears unrelated to the private insurance coverage of the group.

Given the above results we conclude that it is preferable for researchers to incur the additional computational burden of calculating treatment effects that account for selection when estimating Medicaid take-up rates and private insurance coverage rates.¹⁶ However, when calculating the much smaller crowd-out rates it is better to use the simpler treatment effects that ignore selection.

5.3 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 calculate these effects using equations (10a, 10b) and (14a, 14b) in the text respectively.) The estimates in column (2) are again larger than

¹⁵ This ignores the counterintuitive, but very noisy, positive estimate for two earner families.

¹⁶ Of course, this burden is in terms of computation time and not programming time, since all estimators in the paper can be obtained with preprogrammed Stata commands, and we provide our Stata code at <http://econ.williams.edu/people/lshore>.

those in column (1) because of the positive estimated correlation between the unobservables in eligibility and the unobservables in take-up. As with the actual expansions, again there are considerable differences across groups of the newly eligible in their take-up response to the policy change, 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. However, the variation across groups in response to a nonmarginal expansion is less than the variation among the currently eligible. For example, the variation in take-up rates by family structure ranges from 0.3 (for children in two-parent families) to 0.72 (for children in female-headed families) when considering actual eligibility, but only 0.2 to 0.46 for a counterfactual expansion. In addition, the overall levels of take-up are lower. This is consistent with a story where 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)). However, it is also consistent with newly eligible individuals not enrolling because they are higher in the income distribution and have access to private coverage through parental employment.

Finally, it is worth noting for these predicted values that the average take-up rate for marginal individuals of 0.12 estimated from the standard LPM is relatively close to the average take-up rates for newly eligible children from two-earner or two-parent families, suggesting that these latter groups may each approximate the relevant marginal (unknown) individuals in the standard estimates. This interpretation seems especially plausible given that the expansions focused on extending benefits to children in two-parent 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. (We calculate these effects using equations (11) and (15) in the text, respectively.) The 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 such children are more likely than the currently eligible to have access to private coverage through their parents' employment. The predicted coverage rates that take selection into account are smaller than those ignoring selection because the SPM estimates indicate a negative correlation between the eligibility index and the private coverage index. Focusing on the estimates in column (4) we see that the coverage rates are precisely estimated and range from 0.33 for a household headed by a high school dropout to 0.70 for a household in which the head is a college graduate. (Here we are ignoring the estimated coverage rate for a family with no earners of 0.16 as there will be very few of these among the newly eligible.)

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 the discussions below equations (12) and (16) in the text respectively.) The results when we do not control for selection in column (5) are small and have narrow confidence intervals. However, when we account for selection in column (6), the confidence intervals widen, resulting in only one estimate that is statistically distinguishable from zero. This pattern of statistical significance is similar to that of the currently eligible in columns (7) and (8) of Table 2, and indeed the magnitude of these crowd-out effects is also similar to those presented in Table 2, given the respective standard errors. Although there is some evidence that crowd-out rates are higher for groups that have higher levels of private insurance coverage (children in families where the

highest earner has some college or is a high school graduate exhibit higher levels of crowding out and higher rates of private coverage than children in families where the highest earner does not have a high school degree, for example), this explanation is clearly not the only one, as crowd-out rates are generally highest (though still small) among groups with the largest Medicaid take-up rates. Overall, we estimate that private insurance coverage is on average 2 to 5 percentage points lower among currently or newly eligible children than it would be in the absence of Medicaid eligibility.

6. Conclusions

In this paper we demonstrate how a switching probit model (SPM) can be used to estimate a variety of nonmarginal (as well as marginal) effects of Medicaid eligibility on Medicaid take-up, private insurance coverage, and crowd-out overall and across different demographic groups. Relying on the fact that Medicaid eligibility is observable given child characteristics and the Medicaid income limits, we are able to estimate the effects of Medicaid eligibility on take-up, private coverage, and crowd-out rates for currently eligible children as well as to predict the impact of a counterfactual nonmarginal expansion of eligibility on Medicaid participation and crowd-out. Our ability to carry out counterfactual policy analysis demonstrates a significant advantage of our approach over the standard LPM approach and the extension to the LPMI of Ham, Ozbeklik and Shore-Sheppard (2012). Using those approaches, all one can do is estimate take-up rates, private insurance coverage, and crowd-out for unknown marginal individuals in the sample as a whole and in different demographic groups respectively. Our approach here permits us to use estimated take-up rates among observable groups of children among both the currently eligible and those made eligible by a counterfactual policy change while allowing for differences in observables and unobservables in families.

The approach we present produces sensible and relatively precise effects for Medicaid take-up and private insurance coverage among the newly eligible. Further, our treatment effect estimates that allow for selection do a good job of fitting actual Medicaid take-up and private insurance coverage rates among the currently eligible. We find a wide disparity in estimated take-up rates across different demographic groups for both the currently eligible and those made newly eligible. In general, more disadvantaged groups appear to have larger responses to eligibility. 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. Finally, as expected, we find that Medicaid take-up rates for the currently eligible are smaller than the take-up rates of those made eligible by our policy experiment, while the opposite is true for private insurance coverage.

Estimating crowd-out effects among the newly eligible is more challenging, and we find that only when we do not allow for selection do our estimates produce confidence intervals for crowd-out among the newly eligible which will be useful for policy makers. Notably, this model indicates similar levels of crowd-out among the newly eligible to levels among the currently eligible given the respective standard errors. These levels are fairly small, with point estimates in the range of 2 to 5 percent. Crowd-out does not appear to be higher among eligible children in groups with higher levels of private coverage originally.

7. References

Aakvik, A., Heckman, J. J., and Vytlacil, E. J. (2005), "Estimating Treatment Effects for Discrete Outcomes When Responses to Treatment Vary: An Application to Norwegian Vocational Rehabilitation Programs," *Journal of Econometrics* 125, 15-51.

Angrist, J. D., (2001), "Estimation of Limited Dependent Variable Models with Dummy Endogenous Regressor: Simple Strategies for Empirical Practice," *Journal of Business and Economic Statistics* 19, 2-16.

_____, (2004), "Treatment Effect Heterogeneity in Theory and Practice," *Economic Journal* 114, 52-83.

Bennefield, R.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, M., Luttmer, E. F. P., and Mullainathan, S., (2000), "Network Effects and Welfare Cultures," *Quarterly Journal of Economics* 115, 1019-1055.

Bhattacharya, J., Goldman, D., and McCaffrey, D., (2006), "Estimating Probit Models with Self-Selected Treatments," *Statistics in Medicine* 25, 389-413.

Blank, R., and P. 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, 57-89.

Blumberg, L. J., Dubay, L., and Norton, S. A., (2000), "Did the Medicaid Expansions for Children Displace Private Insurance? An Analysis Using the SIPP," *Journal of Health Economics* 19, 33-60.

Blundell, R., Dearden, L., and Sianesi, B., (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, 473-512.

Card, D. and Payne, A. A., (2002), "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores," *Journal of Public Economics* 83, 49-82.

Card, D. and Shore-Sheppard, L. D. (2004), "Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low Income Children," *Review of Economics and Statistics* 86, 752-766.

Carneiro, P., Heckman J.J. and Vytlacil, E.J, (2011), "Estimating Marginal Returns to Education," *American Economic Review* 101, 2754-2781.

Carneiro, P., Hansen, K., and Heckman J.J. (2003), "Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement," *International Economic Review* 44, 361-422.

Currie, J. and Gruber, J., (1996a), "Saving Babies: The Efficacy and Cost of Recent Expansions of Medicaid Eligibility for Pregnant Women," *Journal of Political Economy* 104, 1263-1296.

_____, (1996b), "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *Quarterly Journal of Economics* 111, 431-466.

Cutler, D. M. and Gruber, J., (1996), "Does Public Insurance Crowd Out Private Insurance?" *Quarterly Journal of Economics* 111, 391-430.

Dubay, L. C. and Kenney, G. M., (1996), "The Effects of Medicaid Expansions on Insurance Coverage of Children," *The Future of Children* 6, 152-161.

Gruber, J. and Simon, K., (2008), "Crowd-out 10 Years Later: Health Insurance?" *Journal of Health Economics* 27, 201-217.

- Ham, J. C., Li, X., and Shore-Sheppard, L. D., (2011), "Seam Bias, Multiple-State, Multiple-Spell Duration Models and the Employment Dynamics of Disadvantaged Women," Mimeo, University of Maryland.
- Ham, J. C., Ozbeklik, S., and Shore-Sheppard, L. D., (2012), "Estimating Heterogeneous Medicaid Take-up and Private Insurance Crowd-out Rates: Evidence from a Linear Probability Model," Mimeo, University of Maryland.
- Ham, J. C. and Shore-Sheppard, L. D., (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, 57-83.
- Hansen, C., Hausman, J. and Newey, W., (2008), "Many Weak Instruments and Microeconomic Practice," *Journal of Business and Economic Statistics* 26, 398-422.
- Heckman, J. J., (1979), "Sample Selection Bias as a Specification Error," *Econometrica* 47, 153-161.
- Heckman, J. J., Smith, J., and Clements, N., (1997), "Making the Most Out of Programme Evaluations and Social Experiments: Accounting for Heterogeneity in Programme Impacts," *The Review of Economic Studies* 64, 487-535
- Hudson, J. L., Selden, T. M., and Banthin, J. S., (2005), "The Impact of SCHIP on Insurance Coverage of Children," *Inquiry* 42, 232-54.
- LoSasso, A. T. and Buchmueller, T. C, (2004), "The Effect of the State Children's Health Insurance Program on Health Insurance Coverage," *Journal of Health Economics* 23, 1059-1082.

Marquis, K. H. and Moore, J. C., (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, R., (2007), "Estimating Marginal Returns to Higher Education in the UK," NBER Working Paper 13534.

Quandt, R. E., (1958), "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association* 53, 873-880.

_____, (1960), "Tests of the Hypothesis That a Linear Regressions System Obeys Two Different Regimes," *Journal of the American Statistical Association* 55, 324-330.

_____, (1972), "A New Approach to Estimating Switching Regression," *Journal of the American Statistical Association* 67, 306-310.

Selden, T. M., Banthin, J. S., and Cohen, J. W., (1998), "Medicaid's Problem Children: Eligible but not Enrolled," *Health Affairs* 17, 192-200.

Shaikh, A. M. and Vytlačil, E. J., (2005), "Threshold Crossing Models and Bounds on Treatment Effects: A Nonparametric Analysis," NBER Technical Working Paper No. 307.

Shore-Sheppard, L. D., (2000), "The Effect of Expanding Medicaid Eligibility on the Distribution of Children's Health Insurance Coverage," *Industrial and Labor Relations Review* 54, 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, available at: <http://www.bepress.com/bejeap/vol8/iss2/art6>.

Staiger, D. and Stock, J., (1997), "Instrumental Variables Regression with Weak Instruments," *Econometrica* 65, 557-286.

- Stock, J. H. and Yogo, M., (2005), "Testing for Weak Instruments in Linear IV Regression," in Stock, J.H. and D.W.K. Andrews, eds., *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*, Cambridge University Press, Cambridge, MA, pp. 80–108.
- Thorpe, K. E. and Florence, C. S., (1998), "Health Insurance among Children: The Role of Expanded Medicaid Coverage," *Inquiry* 35, 369-379.
- Van de Ven, W. P. M. M. and Van Praag, B. M. S., (1981), "The Demand for Deductibles in Private Health Insurance," *Journal of Econometrics* 17, 229-252.
- Yazici, E., Kaestner, R., (2000), "Medicaid Expansions and the Crowding Out of Private Health Insurance Among Children," *Inquiry* 37, 23– 32.
- Young, N., (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
Education of Highest Earner	12.68	12.69	12.84	12.74	12.92	12.95	12.95
FRACELIG	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>		Actual (3)	<u>Private Insurance</u>		Actual (6)	<u>Crowd-out</u>	
	<i>ATRE</i> (1)	<i>ATRESEL</i> (2)		<i>PRE</i> (4)	<i>PRESEL</i> (5)		No Selection (7)	Selection (8)
All	0.47*** (0.01)	0.51*** (0.02)	0.51	0.44*** (0.01)	0.34*** (0.03)	0.34	-0.02* (0.01)	-0.05 (0.03)
<i>Race</i>								
White	0.40*** (0.01)	0.44*** (0.02)	0.44	0.50*** (0.01)	0.39*** (0.03)	0.38	-0.01 (0.01)	-0.05 (0.04)
Non-White	0.64*** (0.01)	0.68*** (0.02)	0.67	0.30** (0.01)	0.23*** (0.02)	0.22	-0.02* (0.01)	-0.04 (0.03)
<i>Education</i>								
High School Drop-out	0.60*** (0.01)	0.62*** (0.02)	0.63	0.26*** (0.01)	0.19*** (0.02)	0.18	-0.005 (0.01)	-0.02 (0.03)
High School Graduate	0.44*** (0.01)	0.48*** (0.02)	0.49	0.49*** (0.01)	0.37*** (0.03)	0.36	-0.03*** (0.01)	-0.06 (0.04)
Some College	0.40*** (0.01)	0.44*** (0.02)	0.44	0.57*** (0.01)	0.44*** (0.03)	0.46	-0.04*** (0.01)	-0.07 (0.04)
College Graduate	0.22*** (0.01)	0.27*** (0.01)	0.16	0.79*** (0.01)	0.64*** (0.04)	0.69	-0.02 (0.02)	-0.06 (0.05)
<i>Family Structure</i>								
Female Head	0.69*** (0.01)	0.72*** (0.02)	0.71	0.27*** (0.01)	0.21*** (0.02)	0.21	-0.04*** (0.01)	-0.05* (0.03)
Male Head	0.28*** (0.02)	0.32*** (0.03)	0.30	0.49*** (0.02)	0.36*** (0.04)	0.37	-0.03 (0.02)	-0.06 (0.05)
Two parents	0.26*** (0.01)	0.30*** (0.02)	0.31	0.61*** (0.01)	0.47*** (0.04)	0.46	-0.004 (0.01)	-0.04 (0.04)
<i>Number of Earners</i>								
No earner	0.77*** (0.005)	0.79*** (0.02)	0.78	0.16 *** (0.04)	0.12*** (0.01)	0.13	-0.03* (0.01)	-0.03 (0.02)
One earner	0.28*** (0.01)	0.32*** (0.03)	0.33	0.62*** (0.01)	0.47*** (0.04)	0.46	-0.03* (0.01)	-0.06 (0.05)
Two earners	0.12*** (0.01)	0.18*** (0.02)	0.17	0.81*** (0.01)	0.61*** (0.05)	0.61	0.002 (0.01)	-0.04 (0.05)
More than two earners	0.08*** (0.01)	0.12*** (0.02)	0.14	0.85*** (0.02)	0.69*** (0.05)	0.63	0.03 (0.02)	0.01 (0.06)

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 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. Columns (1) and (2) are based on maximizing equation (5), and columns (4) and (5) are based on maximizing equation (8). *** 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	0.21*** (0.01)	0.27*** (0.01)	0.71*** (0.01)	0.49*** (0.05)	-0.02 (0.01)	-0.07 (0.06)
<i>Race</i>						
White	0.18*** (0.01)	0.24*** (0.01)	0.73*** (0.01)	0.52*** (0.05)	-0.02 (0.01)	-0.07 (0.06)
Non-White	0.32*** (0.01)	0.38*** (0.01)	0.62*** (0.01)	0.39*** (0.05)	-0.01 (0.02)	-0.06 (0.06)
<i>Education</i>						
High School Drop-out	0.28*** (0.01)	0.33*** (0.01)	0.55*** (0.01)	0.33*** (0.04)	0.02 (0.02)	-0.01 (0.05)
High School Graduate	0.21*** (0.01)	0.26*** (0.01)	0.72*** (0.01)	0.51*** (0.05)	-0.03* (0.02)	-0.07 (0.06)
Some College	0.21*** (0.01)	0.27*** (0.01)	0.76*** (0.01)	0.54*** (0.05)	-0.03** (0.01)	-0.09 (0.06)
College Graduate	0.11*** (0.01)	0.17*** (0.01)	0.90*** (0.01)	0.70*** (0.05)	-0.02 (0.01)	-0.09 (0.06)
<i>Family Structure</i>						
Female Head	0.39*** (0.01)	0.46*** (0.01)	0.58*** (0.01)	0.37*** (0.04)	-0.03* (0.02)	-0.09 (0.05)
Male Head	0.16*** (0.02)	0.21*** (0.02)	0.67*** (0.02)	0.42*** (0.05)	-0.02 (0.03)	-0.07 (0.07)
Two parents	0.15*** (0.01)	0.20*** (0.01)	0.76*** (0.01)	0.55*** (0.05)	-0.01 (0.01)	-0.06 (0.06)
<i>Number of Earners</i>						
No earner	0.63*** (0.01)	0.66*** (0.01)	0.23*** (0.01)	0.16*** (0.02)	-0.00 (0.02)	-0.02 (0.03)
One earner	0.22*** (0.01)	0.28*** (0.01)	0.69*** (0.01)	0.47*** (0.05)	-0.03* (0.01)	-0.08 (0.06)
Two earners	0.11*** (0.01)	0.17*** (0.01)	0.84*** (0.01)	0.60*** (0.05)	-0.01 (0.01)	-0.06 (0.06)
More than two earners	0.06*** (0.01)	0.11*** (0.02)	0.93*** (0.02)	0.72*** (0.06)	0.01 (0.02)	-0.03 (0.07)

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

Appendix A: Formulae for Calculating Medicaid Take-up, Private Insurance Coverage and Crowd-out among the Currently Eligible

Here we present the appropriate expressions for the currently eligible for take-up, private insurance coverage, and crowd-out. The predicted average take-up rate among all currently eligible individuals (the ATET) when ignoring selection is given by

$$ATRE = \frac{1}{N_e} \sum_{elig=1} [1 - \Phi_1(-X_i \hat{\mu})], \quad (A1)$$

where N_e denotes the number of currently eligible children. Private insurance coverage rates among the currently eligible in the absence of selection are given by

$$PRE = \frac{1}{N_e} \sum_{elig=1} \Pr(priv_elig_i = 1) = \frac{1}{N_e} \sum_{elig=1} \Pr(X_i \gamma_e + u_{ei} > 0) = \frac{1}{N_e} \sum_{elig=1} [1 - \Phi_1(-X_i \hat{\gamma}_e)]. \quad (A2)$$

Private insurance coverage rates among the currently eligible if their eligibility was taken away are given by (in the absence of selection)

$$PRNE = \frac{1}{N_e} \sum_{elig=1} \Pr(priv_nelig_i = 1) = \frac{1}{N_e} \sum_{elig=1} \Pr(X_i \hat{\gamma}_{ne} + u_{nei} > 0) = \frac{1}{N_e} \sum_{elig=1} [1 - \Phi_1(-X_i \hat{\gamma}_{ne})]. \quad (A3)$$

Thus we can measure average crowd-out among the currently eligible when we ignore selection as $PRE - PRNE$. Note that we obtain $\hat{\gamma}_e$ and $\hat{\gamma}_{ne}$ from maximizing (8) and (9) respectively; since we use separate random samples the estimates are independent and thus we do not need to calculate their covariance in calculating a standard error for this crowd-out estimate.

When we allow for selection, the predicted average take-up rate among all currently eligible children is given by

$$\begin{aligned}
ATRESEL &= \frac{1}{N_e} \sum_{elig=1} \Pr(pub_i = 1 | elig_i = 1) = \frac{1}{N_e} \sum_{elig=1} \frac{\Pr(pub_i = 1, elig_i = 1)}{\Pr(elig_i = 1)} \\
&= \frac{1}{N_e} \sum_{elig=1} \frac{\Pr(X_i \hat{\mu} + \varepsilon_i > 0, Z_i \hat{\delta} + e_i > 0)}{\Pr(Z_i \hat{\delta} + e_i > 0)} = \frac{1}{N_e} \sum_{elig=1} \frac{\Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon, e})}{\Phi_1(Z_i \hat{\delta})}.
\end{aligned} \tag{A4}$$

The private insurance coverage rate among the currently eligible is given by

$$\begin{aligned}
PRESEL &= \frac{1}{N_e} \sum_{elig=1} \Pr(priv_elig_i = 1 | elig_i = 1) = \frac{1}{N_e} \sum_{elig=1} \frac{\Pr(priv_elig_i = 1, elig_i = 1)}{\Pr(elig_i = 1)} \\
&= \frac{1}{N_e} \sum_{i \in New} \frac{\Pr(X_i \hat{\gamma}_e + u_e > 0, Z_i \hat{\delta} + e_i > 0)}{\Pr(Z_i \hat{\delta} + e_i > 0)} = \frac{1}{N_e} \sum_{elig=1} \frac{\Phi_2(X_i \hat{\gamma}_e, Z_i \hat{\delta}, \hat{\rho}_{13})}{\Phi_1(Z_i \hat{\delta})}.
\end{aligned} \tag{A5}$$

If they were ineligible for Medicaid, the currently eligible children would have private coverage given by

$$\begin{aligned}
PRNESEL &= \frac{1}{N_e} \sum_{elig=1} \Pr(priv_nelig_i = 1 | elig_i = 1) = \frac{1}{N_e} \sum_{elig=1} \frac{\Pr(priv_nelig_i = 1, elig_i = 1)}{\Pr(elig_i = 1)} \\
&= \frac{1}{N_e} \sum_{elig=1} \frac{\Pr(X_i \hat{\gamma}_{ne} + u_{ne} > 0, Z_i \hat{\delta} + e_i > 0)}{\Pr(Z_i \hat{\delta} + e_i > 0)} = \frac{1}{N_e} \sum_{elig=1} \frac{\Phi_2(X_i \hat{\gamma}_{ne}, Z_i \hat{\delta}, \hat{\rho}_{23})}{\Phi_1(Z_i \hat{\delta})}.
\end{aligned} \tag{A6}$$

We measure crowd-out accounting for selection among the currently eligible by calculating $PRESEL - PRNESEL$.

Appendix B: Variance Formulae for the Estimated Medicaid Take-up Rates among All of the Currently Eligible

In this section we provide the expressions for the variance of estimated Medicaid take-up rates among all of the currently eligible children when we do not, and do, account for selection. The variance terms for all other treatment effects are calculated in an analogous fashion. Using the Delta method, the estimated variance of $ATRE$ (equation A1) is

$$\text{var}(ATRE) = \left[\frac{1}{N_e} \sum_{elig=1} \frac{\partial \Phi_1[X_i \hat{\mu}]}{\partial \hat{\mu}} \right] [\text{var}(\hat{\mu})] \left[\frac{1}{N_e} \sum_{elig=1} \frac{\partial \Phi_1[X_i \hat{\mu}]}{\partial \hat{\mu}} \right].$$

The derivative in the equation above can be written explicitly as

$$\frac{1}{N_e} \sum_{elig=1} \frac{\partial \Phi[X_i \hat{\mu}]}{\partial \hat{\mu}} = \frac{1}{N_e} \sum_{elig=1} X_i \phi(X_i \hat{\mu}).$$

The variance of the take-up rate accounting for selection (*ATRESEL*) is

$$\text{var}(ATRESEL) = D'VD,$$

$$\text{where } D = \begin{bmatrix} \frac{\partial ATRESEL}{\partial \hat{\mu}} \\ \frac{\partial ATRESEL}{\partial \hat{\delta}} \\ \frac{\partial ATRESEL}{\partial \hat{\rho}_{\varepsilon,e}} \end{bmatrix} \text{ and } V = \begin{bmatrix} \text{var}(\hat{\mu}) & \text{cov}(\hat{\mu}, \hat{\delta}) & \text{cov}(\hat{\mu}, \hat{\rho}_{\varepsilon,e}) \\ \text{cov}(\hat{\mu}, \hat{\delta}) & \text{var}(\hat{\delta}) & \text{cov}(\hat{\delta}, \hat{\rho}_{\varepsilon,e}) \\ \text{cov}(\hat{\mu}, \hat{\rho}_{\varepsilon,e}) & \text{cov}(\hat{\delta}, \hat{\rho}_{\varepsilon,e}) & \text{var}(\hat{\rho}_{\varepsilon,e}) \end{bmatrix}$$

Furthermore, the derivatives can explicitly be written as

$$\frac{\partial ATRESEL}{\partial \hat{\mu}} = \frac{1}{N_e} \sum_{elig=1} \frac{\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho})}{\partial \hat{\mu}}}{\Phi_1(Z_i \hat{\delta})}$$

$$\frac{\partial ATRESEL}{\partial \hat{\delta}} = \frac{1}{N_e} \sum_{elig=1} \frac{\left(\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\partial \hat{\delta}} \right) \Phi_1(Z_i \hat{\delta}) - \left(\frac{\partial \Phi_1(Z_i \hat{\delta})}{\partial \hat{\delta}} \right) \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\{\Phi_1(Z_i \hat{\delta})\}^2}$$

$$\text{and } \frac{\partial ATRESEL}{\partial \hat{\rho}_{\varepsilon,e}} = \frac{1}{N_e} \sum_{elig=1} \frac{\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\partial \hat{\rho}_{\varepsilon,e}}}{\Phi_1(Z_i \hat{\delta})}, \quad \text{where}$$

$$\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\partial \hat{\mu}} = X_i \phi(X_i \hat{\mu}) \Phi[(Z_i \hat{\delta} - \hat{\rho}_{\varepsilon,e} X_i \hat{\mu}) / \sqrt{1 - \hat{\rho}_{\varepsilon,e}^2}],$$

$$\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\partial \hat{\delta}} = Z_i \phi(Z_i \hat{\delta}) \Phi[(X_i \hat{\mu} - \hat{\rho}_{\varepsilon,e} Z_i \hat{\delta}) / \sqrt{1 - \hat{\rho}_{\varepsilon,e}^2}]$$

$$\frac{\partial \Phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e})}{\partial \hat{\rho}_{\varepsilon,e}} = \phi_2(X_i \hat{\mu}, Z_i \hat{\delta}, \hat{\rho}_{\varepsilon,e}) \text{ and } \frac{\partial \Phi_1(Z_i \hat{\delta})}{\partial \hat{\delta}} = Z_i \phi(Z_i \hat{\delta}).$$

Appendix Table C1: Estimated Coefficients for Medicaid Eligibility, Medicaid Participation and Private Insurance Coverage from the Switching Probit Model

	Medicaid Eligibility (1)	Medicaid Participation (2)	Private Coverage When Eligible (3)	Private Coverage When Not Eligible (4)
Size of Household	0.27*** (0.01)	0.15*** (0.01)	-0.13*** (0.01)	-0.12*** (0.01)
White	-0.29*** (0.02)	-0.34*** (0.03)	0.29*** (0.03)	0.29*** (0.02)
Male	-0.01 (0.01)	0.004 (0.01)	0.01 (0.01)	0.01 (0.01)
Two parents	-0.84*** (0.03)	-0.85*** (0.03)	0.57*** (0.03)	0.42*** (0.03)
Male head only	-0.24*** (0.04)	-0.77*** (0.06)	0.21*** (0.06)	0.17*** (0.04)
No earners	2.44*** (0.06)	1.86*** (0.10)	-1.89*** (0.11)	-1.79*** (0.05)
One earner	0.98*** (0.05)	0.73*** (0.10)	-0.71*** (0.11)	-0.60*** (0.03)
Two earners	0.10** (0.05)	0.35*** (0.09)	-0.21* (0.11)	-0.14*** (0.03)
Highest earner's age	-0.03*** (0.001)	-0.02*** (0.002)	0.02*** (0.001)	0.01*** (0.001)
Highest earner's education	-0.15*** (0.003)	-0.07*** (0.01)	0.14*** (0.01)	0.16*** (0.003)
FRACELIG	0.48*** (0.01)	-	-	-

Notes: All index functions include demographic main effects, year, age, and state dummies. Standard errors have been corrected for repeated observations across the same children. The estimates in columns (1) and (2) are obtained maximizing the likelihood function in equation (5), while the estimates in columns (3) and (4) are obtained by maximizing (8) and (9), respectively. Medicaid eligibility estimates associated with column (3) and Medicaid eligibility estimates associated with column (4) not shown but are very similar to those reported in column (1).

*** significantly different from zero at the 1% level; ** significantly different from zero at the 5% level; * significantly different from zero at the 10% level.