Econometric Analysis of Prenatal Advice as a Preventive Measure for Fetal Alcohol Syndrome

by

Donald S. Kenkel

Tsui-Fang Lin

Department of Policy Analysis and Management
Cornell University
Ithaca, NY 14853-4401
dsk10@cornell.edu
tl61@cornell.edu

Shinichi Sakata
Department of Economics
The University of British Columbia
997-1873 East Mall
Vancouver, BC, Canada V6T 1Z1
ssakata@interchange.ubc.ca

and

Joseph V. Terza*
Center for Health Economic and Policy Studies
Medical University of South Carolina
Charleston, SC 29425
terza@musc.edu

(July, 2004)

This work was supported by a grant from the National Institute on Alcohol Abuse and Alcoholism (Grant # RO1-AA12487-01A1).

*Corresponding Author. Phone: 843-792-2202; Fax: 843-792-1358

Abstract

We conduct an econometric investigation of the impact of *prenatal care-giver advice* on alcohol consumption by pregnant women as a means of evaluating its potential effectiveness as a preventive measure for fetal alcohol syndrome [FAS] – a medical diagnosis based on a constellation of fetal abnormalities due to maternal alcohol consumption including alcohol-related birth defects [ARBD] and alcohol-related neurodevelopmental disorders [ARND]. The Institute of Medicine states that FAS "is arguably the most common known nongenetic cause of mental retardation, and that FAS, ARBD and ARND constitute a major public health concern." The econometric model and method that we implement accounts for three complicating features of the data. First, as is manifested in our estimation sample which was drawn from the 1988 National Maternal and Infant Health Survey (NMIHS), a large proportion of pregnant women do not drink at all. This may be indicative that a mother's choice to drink or not during pregnancy differs systematically from her decision regarding the level of alcohol consumption if she chooses to drink. To accommodate this possibility, the model incorporates the two-part approach of Duan et al. (1983). Second, in the NMIHS, respondents were only required to report their consumption up to a specified range of values (e.g., 1 to 2 drinks per week, 2 to 5 drinks per week, and so on). To deal directly with this aspect of the data, our model is cast in the grouped regression framework of Stewart (1983). Third, the binary physician advice variable is likely to be endogenous (see Kenkel and Terza 2001). For example, unobservable mental and emotional factors may exist that make an expectant mother more likely to *both* drink *and* receive advice about the adverse effects of alcohol on her baby's health. Conventional estimation methods that fail to account for such confounding unobservables will falsely attribute their effects to the advice variable. Our model explicitly accounts for the unobservable determinants of prenatal drinking and thereby avoids such endogeneity bias. Our results show that care-giver advice influences prenatal drinking, and that correcting for endogeneity is important.

1. Introduction

Although the abuse of alcohol has a variety of public health consequences, alcohol consumption by pregnant women is of special concern because of the health risks for the fetus. Fetal alcohol syndrome [FAS] is a medical diagnosis based on a constellation of abnormalities due to maternal alcohol consumption. FAS can be viewed as being at the extreme end of a spectrum of fetal alcohol effects [FAE], which also include alcohol-related birth defects [ARBD] and alcohol-related neurodevelopmental disorders [ARND]. The Institute of Medicine [IOM] (1996) states that FAS "is arguably the most common known nongenetic cause of mental retardation, and that FAS, ARBD and ARND constitute a major public health concern." A recent review concluded that the prevalence of FAS in the general population is between 0.5 and 3 per 1,000 births, while the prevalence of FAS and ARBD is at least 10 per 1,000 births (May and Gossage 2001). To prevent FAS and FAE, the ideal is for women to abstain completely from alcohol use while pregnant because no safe level of consumption has been identified. However, because the adverse consequences are in part influenced by the level of fetal alcohol exposure, prevention strategies to reduce alcohol consumption by pregnant women are steps in the right direction.

In this paper we conduct an econometric investigation of a type of brief intervention that has the potential to reduce alcohol consumption by pregnant women and thus help prevent FAS and FAE. We estimate the effect of prenatal care-giver advice on alcohol consumption by pregnant women. The effectiveness of advice regarding the adverse effects of alcohol use has been demonstrated in other contexts. To estimate the advice effect, and the model parameters relating to other control variables, we propose a two-stage maximum likelihood method that accounts for three complicating features of the data. First, a large proportion of pregnant women do not drink at all. This may be indicative that a mother's choice to drink or not during pregnancy differs systematically

from her decision regarding the level of alcohol consumption if she chooses to drink. To accommodate this possibility, the model incorporates the two-part approach of Duan et al. (1983). Second, in the data we use, respondents were only required to report their consumption up to a specified range of values (e.g., 1 to 2 drinks per week, 2 to 5 drinks per week, and so on). To deal directly with this aspect of the data, our model is cast in the grouped regression framework of Stewart (1983). Third, the binary physician advice variable is likely to be endogenous (see Kenkel and Terza 2001). For example, unobservable mental and emotional factors may exist that make an expectant mother more likely to *both* drink *and* receive advice about the adverse effects of alcohol on her baby's health. Conventional estimation methods that fail to account for such confounding unobservables will falsely attribute their effects to the advice variable. Our model explicitly accounts for the unobservable determinants of prenatal drinking and thereby avoids such endogeneity bias.

Our study examines the effectiveness of prenatal advice and, therefore, will provide some of the information policymakers need on the relative benefits and costs of this approach to FAS and FAE prevention. Section 2 places our study in the context of previous research on this issue. Section 3 develops the econometric model. The effect of physician advice is likely to be nonlinear (i.e. heterogeneous across the population). Therefore, in the econometric modeling of these effects we take account of *all* determinants of prenatal alcohol consumption, including *unobservable influences*. These issues are discussed in Section 3, along with details of the two-stage estimation method.

Section 4 describes the data. We analyze data from the 1988 National Maternal and Infant Health Survey (NMIHS) conducted by the National Center for Health Statistics. The public use

national files provide a sample of almost 19,000 pregnant women. The data contain measures of women's alcohol consumption, prenatal care, receipt of prenatal advice about drinking while pregnant, and relevant control variables. For the purpose of including it as an additional control variable in our alcohol demand framework, we merged data on state-level alcoholic beverage price with the NMIHS sample. The estimation results are discussed in section 5. The final section summarizes and concludes.

2. Literature Review

Our approach views alcohol consumption by pregnant women in the context of the standard economics model of consumer demand. The demand model has been extended to consider the implications of imperfect information about the health consequences from consumption of health-related goods. Economic studies that find that health information is an important determinant of consumer decisions about smoking, diet, drinking, and exercise include Viscusi (1990), Ippolito and Mathios (1990, 1995), Kenkel (1991), and Kenkel and Terza (2001). In the present paper we extend this line of research to examine the role of prenatal advice as a potentially important source of information for pregnant women. In doing so, we also contribute to non-economics based evaluations of the effectiveness of physician advice. In the following we discuss the literature on this topic.

The IOM (1996) and the U.S. Department of Health and Human Services [USDHHS] (1997) describe the advantages of multilevel prevention strategies that include education programs to increase general awareness of the hazards of drinking during pregnancy, screening for risky drinking, and interventions among high-risk pregnant women. In our study we will evaluate the effectiveness

of one form of brief intervention – prenatal advice from care givers about the potential dangers that drinking poses for the fetus. As the IOM (1996, p. 131) notes, justification for this type of intervention "comes from the belief that health messages provided by physicians, as sources of credible information, can affect behavioral change in alcohol use." Randomized clinical trials have demonstrated the efficacy of physician advice in the context of specific interventions in populations of male and female at-risk drinkers. Bien, Miller, and Tonigan's (1993, p. 319) review of the clinical trials literature "places brief counseling among the most strongly supported intervention modalities for alcohol problems..." Similar conclusions are drawn by the U.S. Preventive Services Task Force Report (1995, p. 572) and the Eighth Special Report to Congress on Alcohol and Health (USDHHS, 1993). More recently, in a multinational trial, the World Health Organization [WHO] Brief Intervention Study Group (1996) found that at-risk drinkers who received advice reported drinking approximately 17 percent less on average than those in the control group. Fleming *et al.* (1997) also found that in primary care settings brief physician advice for problem drinkers resulted in significantly larger reductions in alcohol use and binge drinking in the intervention group than the control group.

It is notable that in these two recent studies, pregnant women were excluded from participation in the trials. The IOM (1996, pp. 133-134) reviews studies by Rosett *et al.* (1981a, b) and Smith *et al.* (1987) that suggest that interventions during pregnancy can lead some women to stop drinking, but these studies were not controlled clinical trials of advice. The IOM (1996) further points out that pregnancy is known to be an opportune time for many health interventions, giving more reason to believe that prenatal advice can reduce drinking by pregnant women. However, the IOM also stresses that there are no clear data to predict who will respond to such interventions and

who will continue to drink. Chang et al. (1999) report the results of a clinical trial in which pregnant women were randomly assigned to either receive a comprehensive assessment of alcohol use only, or the comprehensive assessment with a brief intervention. Women in both groups on average decreased their drinking by about the same amount. The authors speculate that the failure to find an additional impact of the brief intervention might be due to the fact that the comprehensive assessment was itself a fairly intense stimulus to reduce drinking.

In clinical trials, the data are generated by randomizing the subjects into two groups – the *treatment group* comprised of those who are to receive physician advice, and the *controls* from whom the treatment is to be withheld. Subsequent drinking behavior is then observed and the desired treatment effect can be estimated using a simple difference-of-means approach. Controlled clinical trials are considered the gold standard for evaluating the efficacy of an intervention, i.e., evaluating whether the intervention can work under a set of tightly controlled conditions. There are, however, a number of important drawbacks to using such experimental data in the present context. As noted above, pregnant women are typically excluded from controlled clinical trials due to ethical and/or human subject concerns. In any case, demonstrated efficacy in a trial does not necessarily translate into effectiveness in everyday practice (Teutsch and Harris 1996, p. 5). Two aspects of the protocol for the WHO (1996) trial and other trials limit their general applicability. First, the samples include only at-risk drinkers; at baseline, the typical daily consumption in the WHO trial averaged over 4 drinks (over 2 ounces of absolute ethanol). Secondly, in clinical trials data the content of the physician advice is carefully defined, whereas in nonexperimental data physician advice is observed as it is given in everyday practice. Moreover, experimental datasets are typically costly to produce relative to comparably sized nonexperimental surveys. Finally, it is argued by Heckman (1992) and

Heckman and Smith (1995), that randomization itself may lead to sampling bias. For example, individuals' decisions to participate in the study may be influenced by the knowledge that they will be subsequently randomized. This is referred to as *randomization bias*. In the present context, risk averse individuals who are more likely to receive prenatal advice in everyday practice, may decline to participate in the study when faced with the possibility of not being assigned to the advice treatment group. We complement the results obtained in studies of clinical trials data and avoid some of its shortcomings by analyzing a nonexperimental dataset via appropriate econometric technique.

3. The Econometric Model

We now turn to the modeling of the effect of prenatal advice on prenatal alcohol consumption. We seek a specification of this effect that is not only mathematically sound but also lends itself to relatively easy parametric estimation.

3.1 The Estimation Objective

The objective here is to estimate the *policy effect* of prenatal advice on drinking during pregnancy for a specified subgroup of the population We use the term *policy effect* in this context to mean the amount by which prenatal drinking would differ between two counterfactual scenarios – one in which all individuals in the specified population subgroup are given prenatal advice about the potential adverse consequences for the baby; another in which no one in the subgroup is so advised. Formally, the policy effect that we seek to estimate is

$$PE_g = E_g[y_1] - E_g[y_0] \quad (1)$$

where g denotes the population subgroup of interest,

$$A^* = \begin{cases} 1 \text{ if advice is given to everyone in subpopulation g} \\ 0 \text{ if advice is withheld from everyone in subpopulation g} \end{cases}$$

y_{A^*} denotes the amount of prenatal drinking in the counterfactual A* scenario, and $E_g[\ ]$ represents expectation with respect to g.

To fix ideas, let us consider a simple version of the model. Suppose that y_{A^*} is binary – $y_{A^*} = 1$ if the individual consumes any alcoholic beverages during pregnancy, $y_{A^*} = 0$ otherwise. Morever, assume that the "specified subgroup" in which we have interest is the population itself. Therefore in this simplified scenario, we seek to estimate the change in the population probability of drinking that would result from a policy that ensures that all pregnant women are advised of the negative effects of drinking. Consider the following population of 10 individuals

Table 1

y_0	y_1
1	1
1	0
0	0
0	0
0	0
1	1
1	1
1	1
1	0
0	0

The first column in Table 1 contains the drinking outcomes for the counterfactual "no advice" (A* = 0) scenario. The second column is similarly defined for the "advice" (A* = 1) scenario. It is

clear that the value of the desired policy effect (1) in this case is

$$PE_g = E_g[y_1] - E_g[y_0] = .4 - .6 = -.2 \quad (2)$$

indicating that a fully implemented policy in which all pregnant women are given advice regarding the ill effects of prenatal drinking would result in a .2 decline in the probability of prenatal drinking. The main problem in estimating the policy effect is that the population outcomes on y as displayed in Table 1 are not fully observable via survey sampling. In survey sampling, data on y_1 is only observable for those individuals who happened to get advice from their prenatal care givers. Likewise, observations on y_0 are only available for those who failed to receive advice. Let A be the random variable indicating receipt of advice where A = 1 if advice was given, and A = 0 otherwise. In the context of Table 1, let us suppose that the first five persons listed in the table did not get advice, while the remaining five did receive advice. Table 2 makes explicit this sampling constraint.

Table 2

Person	y_0 (A* = 0)	y_1 (A* = 1)
a (A = 0)	c	e
1	1	1
2	1	1
3	1	1
4	1	0
5	0	0
b (A = 1)	d	f
6	1	1
7	1	0
8	0	0
9	0	0
10	0	0

Individuals in cell a are those that were observed not to receive advice, while those in cell b did receive advice. The shaded cells c and f contain the observable (factual) data on y – i.e. the values of y in cell f correspond to people who are observed to have gotten advice (A=1); the values of y in c correspond to people who were not given advice (A=0). Unshaded cells d and e contain the unobservable (counterfactual) data on y – i.e. the values of y in d are those that *would have been observed* for individuals observed to have gotten advice (A=1) *had they not gotten advice (A*=0)*; the values of y in e are those that *would have been observed if advice were given (A*=1)* to the pregnant women who did not actually receive advice (A=0). In summary, the shaded cells contain factual (observable) data; the unshaded cells contain counterfactual (unobservable) data.[1]

Through survey sampling, we are unable to get unrestricted access to the relevant data for the estimation of the true (but counterfactual) policy effect as given in (1) [expression (2) in the context of our example]. In our example, we would like to freely sample from the full population, including the unshaded blocks of Table 2 (d and e). Unfortunately, we are restricted to sampling from the population data in the shaded blocks (c and f). Because of this sampling restriction, conventional unbiased methods applied to such data will be biased for the true policy effect. The reason for this is simple. When we sample from blocks e and f we are actually sampling from the distributions of $(y \mid A = 0)$ and $(y \mid A = 1)$, respectively. So, for instance, the conventional difference-of-means (DOM) estimator will be unbiased for $E_g[y \mid A = 1] - E_g[y \mid A = 0]$ which may not be identical to (1). For the example in Table 2 we have

[1]It should be noted that the observable y data for a corresponding observed value of A may not necessarily be identical to the y data obtained after forcing that same value of A* on the entire population. We restrict the data in this way to facilitate exposition.

$$E_g[y \mid A = 1] - E_g[y \mid A = 0] = .2 - .8 = -.6. \quad (3)$$

Clearly, in this example, the conventional DOM estimator will be biased for (2).

As this discussion makes clear, if conventional unbiased estimators applied to the observable data are to produce unbiased estimates of PE as defined in (1), the following condition must be true

$$E_g[y_{A*}] = E_g[y \mid A=A*]. \quad (4)$$

Why would (4) fail to hold? It fails because there exist confounders – variables that affect y and are correlated with A. So what can we do if (4) fails? We must find a way to control for the confounders. Let us now examine how that is accomplished via parametric regression methods. Without loss of the generality of condition (4), we can now allow y to be any measure of prenatal drinking (binary, count, continuous). We begin by specifying a counterfactual regression model of the form

$$y_{A^*} = H(A^*, x_o^*, x_u^*, \varepsilon, \tau) \quad (5)$$

where

y_{A^*} = the amount of prenatal drinking in the counterfactual A* advice status scenario

A* = 1 if advice is counterfactually imposed; 0 otherwise

x_o* = other observable confounding influences on prenatal drinking

x_u* = other unobservable confounding influences on prenatal drinking

ε = is a vector random error terms

τ = a vector of parameters

and the "*" superscript denotes values that are counterfactually determined rather than sampled at random.

Some points of clarification are in order. First note that the confounders x_o and x_u are counterfactually imposed in one of four ways:

a) they may be fixed and known values (6)

b) they may be components of random vector whose distribution is specified and known (7)

c) they may be components of a random vector whose distribution follows that determined by a specified subgroup of the observable population (8)

d) they may be defined as some combination of a), b) and c). (9)

Second, the random error ε is independent of the confounders x_o* and x_u*. Finally note that what we will call the *regression function* is defined as

$$J(A^*, x_o^*, x_u^*, \tau) = \int_{\varepsilon} H(A^*, x_o^*, x_u^*, \varepsilon, \tau)\, f_\varepsilon(\varepsilon)\, d\varepsilon \tag{10}$$

where $f_\varepsilon(\varepsilon)$ denotes the pdf of ε. Usually J( ) has a form that is similar to that of H( ).

Using (10) it is clear that

$$E_g[y_{A^*}] = E_g\left[J(A^*, x_o^*, x_u^*, \tau)\right] \tag{11}$$

where g denotes the relevant group of interest defined by:

a) the fixed and known values of x_o* and x_u*, in which case (11) becomes

$$E_g[y_{A^*}] = J(A^*, x_o^*, x_u^*, \tau)$$

or b) a designated subset of the support of the known and specified distribution of x_o and x_u

or c) a specified subgroup of the observable population

or d) a combination of (a), (b) and (c).

These four options for defining g correspond to (6) through (9), respectively. Now if we can find a consistent estimator of τ from the observable (survey) data, then (1) can be consistently estimated using

$$\text{a) } \hat{PE} = J(A^*=1, x_o^*, x_u^*, \hat{\tau}) - J(A^*=0, x_o^*, x_u^*, \hat{\tau}) \tag{12}$$

$$\text{or b) } \int_g \left\{ J(A^*=1, x_o^*, x_u^*, \hat{\tau}) - J(A^*=0, x_o^*, x_u^*, \hat{\tau}) \right\} f_x(x_o^*, x_u^*)\, dx_o^*\, dx_u^* \tag{13}$$

$$\text{or c) } \sum_{i=1}^{n_g} \frac{1}{n_g} \left\{ J(A^*=1, x_{oi}, x_{ui}, \hat{\tau}) - J(A^*=0, x_{oi}, x_{ui}, \hat{\tau}) \right\} \tag{14}$$

or d) some combination of (a), (b) and (c) (15)

where $\hat{\tau}$ denotes a consistent estimate of τ and $f_x(x_o^*, x_u^*)$ denotes the specified and known joint pdf of x_o^* and x_u^*. The estimators in (12), (13), (14) and (15) correspond, respectively, to the four ways given for imposing the values of x_o^* and x_u^* in (6), (7), (8), and (9). Note that there are no "*" superscripts on x_{oi} and x_{ui} because these represent values sampled from population.

The approach that we will take to the specification and estimation of (1) conforms to (9) and (15). In order to deal with the nonobservability of x_u, we will assume that it is a scalar comprising

all of the unobservable confounders. Moreover, its distribution will be implied by one of the structural assumptions of the model. The subpopulation of interest (g) will then be defined in terms of the observable confounders. This hybrid approach leads to the following specifications for PE and $\widehat{PE}$, respectively

$$PE = E_g\left[\int_{x_u}\left\{J(A^*=1, x_o, x_u, \tau) - J(A^*=0, x_o, x_u, \tau)\right\} f_{x_u}(x_u)\, dx_u\right] \tag{16}$$

$$\widehat{PE} = \sum_{i=1}^{n_g} \frac{1}{n_g} \int_{x_u}\left\{J(A^*=1, x_{oi}, x_u, \hat{\tau}) - J(A^*=0, x_{oi}, x_u, \hat{\tau})\right\} f_{x_u}(x_u)\, dx_u \,. \tag{17}$$

where $f_{x_u}(x_u)$ denotes the pdf of x_u and $\hat{\tau}$ is a consistent estimate of τ. It will be possible to obtain a consistent estimate of τ if the following analog to condition (4) holds[2]

$$E[y \mid A, x_o, x_u] = J(A, x_o, x_u, \tau). \tag{18}$$

Condition (18) holds if the set of confounders comprising the elements of x_o and x_u is comprehensive. We now turn to the specification of the function J( ).

3.2 Specification of the Regression Function

The outcome variable (prenatal drinking) is nonnegative and a large percentage of the sample

---

[2]We call (18) an analog to (4) because, given (10) we can write

$$E_g[y_{A*}] = E_g[H(A^*, x_o^*, x_u^*, \epsilon, \tau)] = E_g[J(A^*, x_o^*, x_u^*, \tau)] = E_g\left[E[y \mid A=A*, x_o^*, x_u^*]\right]$$

report no prenatal alcohol consumption.[3] We therefore allow the decision to drink (or not) during pregnancy to be determined by a parametric process that is distinct from that representing the strictly positive level of prenatal alcohol consumption for those who have decided to drink during pregnancy. In specifying the regression function we turn to the two-part model of Duan et al. (1983). The following binary variable q_{A*} (1 if drink, 0 otherwise) represents the first part of the model (the decision to drink or not)

$$q_{A*} = I(A*\gamma_1 + x_o*\beta_1 + x_u*\theta_1 + \varepsilon_1 > 0) \tag{19}$$

and the second part of the model is represented by

$$y_{A*, q_{A*}=1} = \exp(A*\gamma_2 + x_o*\beta_2 + x_u*\theta_2 + \varepsilon_2) \tag{20}$$

where I(C) denotes the indicator function which is equal to 1 if condition C holds, and 0 otherwise; n(a, b) denotes the normal variate with mean a and variance b; $\varepsilon_1 \sim n(0, 1)$; $\varepsilon_2 \sim n(0, \sigma^2)$ for any value of $\varepsilon_1 > -(A*\gamma_1 + x_o*\beta_1 + x_u*\theta_1)$ [i.e for any value of ε_2 such that $q_{A*} = 1$; and $y_{A*, q_{A*}}$ exists only if $q_{A*} = 1$. Using (19) and (20), in the spirit of equation (5) we obtain

$$y_{A*} = H(A^*, x_o^*, x_u^*, \varepsilon, \tau) = I(A*\gamma_1 + x_o*\beta_1 + x_u*\theta_1 + \varepsilon_1 > 0)\exp(A*\gamma_2 + x_o*\beta_2 + x_u*\theta_2 + \varepsilon_2). \tag{21}$$

Now using (10), (21) yields the regression function as[4]

$$J(A^*, x_o^*, x_u^*, \tau) = \int_{\varepsilon} H(A^*, x_o^*, x_u^*, \varepsilon, \tau)\, f_\varepsilon(\varepsilon)\, d\varepsilon$$

[3]In our sample the percentage of non-drinkers is 84%.

[4]The derivation of (22) is given in the Appendix.

$$= \Phi(A^*\gamma_1 + x_o^*\beta_1 + x_u^*\theta_1)\exp(A^*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \sigma^2/2) \quad (22)$$

where $\varepsilon = [\varepsilon_1, \varepsilon_2]$ and $\Phi(\ )$ is the standard normal cdf. Using (22) we can formulate the following fully specified version of the generic policy effect given in (16)

$$PE = E_g\left[\int_{x_u} \left\{ \Phi(\gamma_1 + x_o\beta_1 + x_u\theta_1)\exp(\gamma_2 + x_o\beta_2 + x_u\theta_2 + \sigma^2/2) - \Phi(x_o\beta_1 + x_u\theta_1)\exp(x_o\beta_2 + x_u\theta_2 + \sigma^2/2) \right\} \phi(x_u)\, dx_u \right] \quad (23)$$

where $\phi(\ )$ denotes the standard normal pdf.

3.2 Parameter Estimation

We assume that the vector x_o and the scalar random variable x_u capture all of the relevant confounders so that the pdf of observable random variable q, conditional on observable A, x_o, and x_u is

$$f_q(q \mid A, x_o, x_u) = \begin{cases} 1 - \Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1) & \text{for } q = 0 \\ \Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1) & \text{for } q = 1 \end{cases} \quad (24)$$

Also

$$f_{y,\, q=1}(\ln y \mid A, x_o, x_u) = N(A\gamma_2 + x_o\beta_2 + x_u\theta_2, \sigma^2) \quad (25)$$

where N(a, b) denotes the pdf of a normal random variable with mean a and variance b. Therefore, given that x_o, and x_u are comprehensive, condition (18) holds in this case because it is easy to show that

$$\mathbf{E[y \mid A, x_o, x_u] = J(A, x_o, x_u, \tau)} = \Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\exp(A\gamma_2 + x_o\beta_2 + x_u\theta_2 + \sigma^2/2). \quad (26)$$

Aside from its nonzero coefficients (θ_1 and θ_2) in (26), the defining feature of x_u as a confounder is its correlation with A. We formalize the relationship between x_u and A as

$$A = I(z\alpha + x_u > 0) \quad (27)$$

where α is a vector of parameters conformable with z, and $(x_u \mid z)$ is standard normally distributed.

Given (24), (25), and (27) the model is fully parametrically specified. There is, however, an additional aspect of the survey data that further complicates the estimation. The data on prenatal alcohol consumption are not directly observable. Instead, respondents are only required to report their weekly alcohol consumption up to a specified range of values. For individuals who report at least some drinking during pregnancy, there are ten consumption categories defined by the following nine thresholds: .25, .5, .75, 1, 2, 5, 8, 13, 20.[5] For these ten categories we define the corresponding binary variables C_j (j = 1, ..., 10) such that $C_j = 1$ if the jth category is observed, and $C_j = 0$ otherwise. We therefore extend assumption (22) in the following way

$$C_j = \begin{cases} 1 & \text{if and only if } \ln(\mu_{j-1}) < (\ln y \mid A, x_o, x_u, q = 1) \le \ln(\mu_j) \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

$\ln(\mu_0) = -\infty$, $\ln(\mu_{10}) = +\infty$, $\mu_1 = .25$, $\mu_2 = .5$, $\mu_3 = .75$, $\mu_4 = 1$, $\mu_5 = 2$, $\mu_6 = 5$, $\mu_7 = 8$, $\mu_8 = 13$, and $\mu_9 = 20$. Given (24), (25), (27) and (28) the joint pdf of A, q, and C_1 through C_{10} conditional on

[5]The three categories with fractional limits were actually stated on the survey questionnaire as: "less than 1 drink per month;" "1 drink per month;" and "2-3 drinks per month;" respectively.

x_o and x_u is

$f(A, q_A, C_1, ..., C_{10} \mid x_o, z) =$

$$\left[\left(\int_{-z\alpha}^{\infty}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)^{A}\left(\int_{-\infty}^{-z\alpha}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)^{(1-A)}\right]^{(1-q)}$$

$$\times\left[\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)^{A}\right.$$

$$\left.\times\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)^{(1-A)}\right]^{q} \tag{29}$$

where $\xi = 1/\sigma$. The corresponding likelihood function is

$L(\alpha, \gamma_1, \gamma_2, \beta_1, \beta_2, \theta_1, \theta_2, \xi) =$

$$\prod_{i=1}^{n}\left\{\left[\left(\int_{-z\alpha}^{\infty}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)^{A}\left(\int_{-\infty}^{-z\alpha}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)^{(1-A)}\right]^{(1-q)}\right.$$

$$\times\left[\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)^{A}\right.$$

$$\left.\left.\times\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)^{(1-A)}\right]^{q}\right\}$$

(30)

where x_o, A, and q should all have "i" subscripts but they were suppressed in the interest of notational parsimony. The log version of (28) can be optimized to obtain the full information maximum likelihood estimates of the parameters.

A computationally simpler two-stage estimator can, however, be implemented. To see this, note that the joint pdf in (27) can be rewritten as

$$f(A, q, C_1, ..., C_{10} \mid x_o, z) = f(A, q \mid x_o, z)\, f_C(C_1, ..., C_{10} \mid A, q = 1, x_o, z)$$

where

$$f(A, q \mid x_o, z) =$$

$$\left[\left(\int_{-z\alpha}^{\infty}\left[1 - \Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\right]\phi(v)\,dv\right)^{A}\left(\int_{-\infty}^{-z\alpha}\left[1 - \Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\right]\phi(v)\,dv\right)^{(1-A)}\right]^{(1-q)}$$

$$\times\left[\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\,\phi(x_u)\,dx_u\right)^{A}\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\,\phi(x_u)\,dx_u\right)^{(1-A)}\right]^{q} \quad (31)$$

and

$$f_C(C_1, ..., C_{10} \mid A, q = 1, x_o, z) =$$

$$\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1 + x_o\beta_1 + x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j) - A\gamma_2 - x_o\beta_2 - x_u\theta_2\right)\right) - \Phi\left(\xi\left(\ln(\mu_{j-1}) - A\gamma_2 - x_o\beta_2 - x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)\,dx_u\right)^{A}$$

$$\times\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)^{(1-A)}$$

$$\div\left[\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\,\phi(x_u)\,dx_u\right)^{A}\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\,\phi(x_u)\,dx_u\right)^{(1-A)}\right]. \quad (32)$$

The key here is that neither (31) nor the divisor in (32) involves γ_2, β_2, θ_2, or ξ. This implies that the likelihood function (30) is separable so that the following computationally simpler two-stage estimator can be implemented.

First Stage: Estimate α, β_1, θ_1 by maximizing the following log-likelihood function:

$$L_1(\alpha,\gamma_1,\beta_1,\theta_1\mid x_o,z)=\sum_{i=1}^{n}\left\{A(1-q)\ln\left(\int_{-z\alpha}^{\infty}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)\right.$$

$$+(1-A)(1-q)\ln\left(\int_{-\infty}^{-z\alpha}\left[1-\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\right]\phi(x_u)\,dx_u\right)$$

$$+(A\,q)\ln\left(\int_{-z\alpha}^{\infty}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\,\phi(x_u)\,dx_u\right)$$

$$\left.+(1-A)\,q\ln\left(\int_{-\infty}^{-z\alpha}\Phi(A\gamma_1+x_o\beta_1+x_u\theta_1)\,\phi(x_u)\,dx_u\right)\right\} \quad (33)$$

where x_o, z, A, and q should all have "i" subscripts but they were suppressed in the interest of notational parsimony.

Second Stage: Estimate ξ, γ_2, β_2, and θ_2 by maximizing the following log-likelihood function

$$L_2(\xi, \gamma_2, \beta_2, \theta_2 \mid x_o, z) =$$

$$\sum_{i=1}^{n}\left\{A\ln\left(\int_{-z\hat{\alpha}}^{\infty}\Phi(A\hat{\gamma}_1+x_o\hat{\beta}_1+x_u\hat{\theta}_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)\right.$$

$$\left.+(1-A)\ln\left(\int_{-\infty}^{-z\hat{\alpha}}\Phi(A\hat{\gamma}_1+x_o\hat{\beta}_1+x_u\hat{\theta}_1)\prod_{j=1}^{10}\left[\Phi\left(\xi\left(\ln(\mu_j)-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)-\Phi\left(\xi\left(\ln(\mu_{j-1})-A\gamma_2-x_o\beta_2-x_u\theta_2\right)\right)\right]^{C_j}\phi(x_u)dx_u\right)\right\} \tag{34}$$

where $\hat{\alpha}$, $\hat{\gamma}_1$, $\hat{\beta}_1$, and $\hat{\theta}_1$ are the first stage estimates; and x_o, z, and A should all have "i" subscripts but they were suppressed in the interest of notational parsimony. Using the consistent estimates obtained from (33) and (34), we can compute the following fully specified version of the policy effect estimator which was generically stated in (17)

$$\hat{PE}=\sum_{i=1}^{n_g}\frac{1}{n_g}\int_{x_u}\left\{\Phi(\hat{\gamma}_1+x_{oi}\hat{\beta}_1+x_u\hat{\theta}_1)\exp(\hat{\gamma}_2+x_{oi}\hat{\beta}_2+x_u\hat{\theta}_2+\hat{\sigma}^2/2)-\Phi(x_{oi}\hat{\beta}_1+x_u\hat{\theta}_1)\exp(x_{oi}\hat{\beta}_2+x_u\hat{\theta}_2+\hat{\sigma}^2/2)\right\}\phi(x_u)dx_u\ . \tag{35}$$

4. Data and Variable Specifications

4.1 Data

The primary data set used in this study is the 1988 National Maternal and Infant Health Survey (NMIHS) conducted by the National Center for Health Statistics. This public use data set

provides us with a unique opportunity to study factors related to pregnancy outcome. Valuable information about prenatal care, maternal drinking, maternal smoking, drug use, exercise, pregnancy and delivery complications is detailed in the survey. In addition, information regarding health care providers, such as physicians and hospitals, was recently released in 1998.

The NMIHS was completed by 18,594 mothers who had a pregnancy in 1988. After restricting the sample to those who received at least some prenatal care, and after imposing other winnowing criteria, 15,978 observations remained. Table 1 gives the breakdown of the sample with regard to drinking during pregnancy and pregnancy disposition. The NMIHS over-samples pregnancies that resulted in fetal and infant deaths; our analysis combines the live birth, fetal, and infant death samples.[1] The estimation sample is merged with state-level data on alcohol price.

4.2 Variable Specifications

Prenatal Drinking The dependent variable of interest in this study is y – the number of drinks consumed per week during pregnancy. One drink refers to 12 ounces of beer, 4 ounces of wine or 1.5 ounces of liquor. As was mentioned earlier, values of y other than 0 are not observed in the sample. For those who drank during pregnancy, only the values of the binary variables, C_j [as defined in (28)] are observed. Table 2 displays the sample frequencies for all possible responses regarding prenatal alcohol consumption.

---

[1]The NMIHS also oversampled low- and very low-birthweight infants and black infants. The oversampling distorts the conditional distribution of the variables related to the weight of the infant given other variables. This is a potential problem for our analysis, in which prenatal alcohol consumption is the dependent variable. Fortunately, the sample design of the NMIHS can be viewed as a special case of the complex survey design considered in Sakata (2000), and in future work we will extend the analysis to account for this feature of the data.

Advice The advice variable (A) is given a value of 1 if the individual responds affirmatively to a questionnaire item asking if they were advised to "cut down/stop drinking alcohol" at any of their prenatal visits.

Other Variables Included in x_o The observable confounders included in the vector x_o are defined in Table 3. Descriptive statistics for these variables and A are displayed in Table 4.

5. Estimation Results

The relevant vectors of regressors (x_o, and z) are specified as

x_o = [1 P AGE HISPANIC BLACK ASIAN NATIVAM COLLEGE LT-HISCH SOMECOLL INCOME MARRIED WIDOWED DIVORCED SEPARATE WORKING FATHER GRANDPAR CHILDREN]

z = [1 P AGE HISPANIC BLACK ASIAN NATIVAM COLLEGE LT-HISCH SOMECOLL INCOME MARRIED WIDOWED DIVORCED SEPARATE WORKING FATHER GRANDPAR CHILDREN TRAVTIME TRIPS OTHERPAY].

For the purpose of comparison, we took g to be the entire observable population and estimated the average policy effect of advice using (35); first with A assumed to be exogenous, and again accounting for the possibility that A is endogenous. When A is exogenous $\theta_1 = \theta_2 = 0$ and the

likelihood function in (30) becomes fully separable with respect to the two subsets of parameters (γ_1 β_1) and (γ_2 β_2 ξ). Specifically (30) becomes

$L(\alpha, \gamma_1, \gamma_2, \beta_1, \beta_2, \theta_1, \theta_2, \xi) =$

$$\prod_{i=1}^{n} \left[1 - \Phi(A\gamma_1 + x_1\beta_1)\right]^{(1-q)} \Phi(A\gamma_1 + x_1\beta_1)^q \times \prod_{i=1}^{n} \left(\prod_{j=1}^{10} \left[\Phi\left(\xi\left(\ln(\mu_j) - A\gamma_2 - x_2\beta_2\right)\right) - \Phi\left(\xi\left(\ln(\mu_{j-1}) - A\gamma_2 - x_2\beta_2\right)\right) \right]^{C_j} \right)^q . \tag{36}$$

Therefore γ_1 and β_1 can be estimated by applying simple probit analysis with q as the dependent variable and A and x_o as the regressors; and γ_2, β_2, and ξ can be estimated using simple grouped regression (Stewart, 1983) with $C = [C_1, ..., C_{10}]$ as the outcome and A and x_o as the regressors. The simple grouped regression estimator is applied to the subsample corresponding to q = 1. Results for simple probit estimation of γ_1 and β_1 are given in Table 5. Results for simple grouped regression estimation of γ_2, β_2, and ξ are given in Table 6. When A is exogenous the policy effect estimator in (35) becomes

$$\hat{PE}_{exog} = \sum_{i=1}^{n} \frac{1}{n} \left\{ \Phi(\hat{\gamma}_1 + x_{oi}\hat{\beta}_1) \exp\left(\hat{\gamma}_2 + x_{oi}\hat{\beta}_2 + \left(\frac{1}{2\hat{\xi}} \right)^2 \right) - \Phi(x_{oi}\hat{\beta}_1) \exp\left(x_{oi}\hat{\beta}_2 + \left(\frac{1}{2\hat{\xi}} \right)^2 \right) \right\}. \tag{37}$$

Substituting the estimates from Tables 5 and 6, and taking g to be the entire observable population, we get

$\hat{PE}_{exog} = 0.112$.

This result is of course contrary to expectation. It says that giving a pregnant mother advice to cut down or stop drinking will cause here to increase her weekly alcohol consumption by more than one-tenth of a drink.

Under the assumption that A is endogenous we applied the two-stage estimator defined in (33) and (34) and obtained the results given in Tables 7, 8, and 9. Table 7 displays the first-stage estimates of γ_1, β_1, and θ_1 – the parameters of the first part of the two-part model (19). The estimated coefficient of A in the first part of the two-part model is negative ($\hat{\gamma}_1 = -0.292$) but is not significantly different from zero. The marginally significant estimate of θ_1 (p-value = .106), along with the change in sign of the estimated coefficient of A when compared to its counterpart in Table 5 (A assumed exogenous), can however be taken as evidence of endogeneity. Table 8 gives the second-stage estimates of γ_2, β_2, ξ and θ_2 – the parameters of the second part of the two-part model (20). In comparison with the corresponding result in Table 6 (A assumed exogenous) the estimated coefficient of A changes from positive and significant to negative and significant ($\hat{\gamma}_2 = -0.969$), and the estimated coefficient of x_u ($\hat{\theta}_2 = .740$) is significantly different from zero. Endogeneity is apparent in this part of the model.

The estimates of the elements of α for the advice equation (27) are given in Table 9. Recall that TRAVTIME, TRIPS, and OTHERPAY are included in z as instrumental variables – i.e., they are excluded from x_1 and x_2. TRAVTIME is significant at better than a 10% level (p-value = .062), and TRIPS is significant at any reasonable level.

Plugging the estimates from Tables 7, 8, and 9 into (35) yields the following estimate of

the treatment effect of care giver advice on prenatal drinking

$$\hat{TE} = -0.39. \tag{38}$$

This value is much more intuitive than the one obtained using the naive estimator in (37). It implies that advice leads the typical expectant mother to curtail her prenatal drinking by nearly two-fifths of a drink.

As a point of reference, we computed the average weekly number of drinks for a typical pregnant woman in the population given that she does not receive advice (A* = 0). We do this by imposing the condition that A* = 0 on all sample members and then computing the average projected drinking level that would obtain if no one received advice ($APD_{A^*=0}$) . Specifically we used

$$APD_{A^*=0} = \sum_{i=1}^{n_g} \frac{1}{n_g} \left[\int_{-\infty}^{\infty} \Phi(x_{oi}\hat{\beta}_1 + x_u\hat{\theta}_1) \exp(x_{oi}\hat{\beta}_2 + x_u\hat{\theta}_2 + \hat{\sigma}^2/2)\, \phi(x_u)\, dx_u \right] = .534. \tag{39}$$

The estimated average policy effect of -.39, therefore, represents a 73% decrease in prenatal alcohol consumption. Given that the estimated coefficient of A in the first part of the two-part model was not significantly different from zero. We computed the average treatment effect using (35) after imposing the restriction that $\gamma_1 = 0$. We thus obtained

$$\hat{PE} = -0.331 \tag{40}$$

which, in light of (39), represents a 63% decrease in prenatal drinking attributable to the receipt of advice. This estimated effect is large in percentage terms although small in absolute value

because the estimated levels of prenatal drinking for both a typical woman who did not receive advice and a typical woman who received advice are so low (about half a drink per week or less). It is interesting to note that in a clinical trial where pregnant women received either an alcohol assessment or an alcohol assessment followed by a brief intervention, alcohol use fell in both groups by between 0.3 and 0.4 drinks per drinking day (Chang et al. 1999). This result, and evidence from clinical trials for other populations, provides evidence on the plausibility of our estimate of the treatment effect of advice.

6. Discussion

The National Institute on Alcohol Abuse and Alcoholism (NIAAA, 1998) concludes that "Because no safe level of alcohol consumption during pregnancy has been identified .... NIAAA defines its mission to be the support of research that targets any drinking and especially abusive drinking by pregnant women." In line with this conclusion, because our results indicate that prenatal care giver advice has a negative and significant effect on drinking while pregnant, it appears that this type of brief intervention during pregnancy is an effective preventive measure for FAS. Moreover, the results provide some evidence that, in practice, care giver advice may have different effects on different aspects of prenatal drinking behavior. Care giver advice is estimated to have a statistically significant negative impact on the amount of drinking by pregnant women, conditional on any drinking. But the impact of advice on the probability of any drinking is weaker (the estimated coefficient is still negative but not statistically significantly different from zero).

Evidence that care giver advice is effective in reducing alcohol consumption among

pregnant women is necessary but not sufficient evidence that these interventions are socially desirable. One approach to quantify the benefits of care-giver advice focuses on the potential cost savings from FAS prevention. Studies of the economic cost of alcohol abuse provide estimates of the present discounted value of lifetime medical expenditures and foregone earnings due to a case of FAS, PDV_{FAS}.[2] The dollar value of the potential benefits of prenatal advice for the typical expectant mother can then be quantified as

$$\mathbf{PE} \times \frac{\partial \mathbf{Pr\{FAS\}}}{\partial \mathbf{y}} \times \mathbf{PDV_{FAS}} \tag{41}$$

where Pr{FAS} denotes the probability of occurrence of FAS for a typical pregnancy. The econometric results reported above provide an estimate of PE the policy effect of advice on alcohol consumption. The missing piece of expression (41), however, is the dose-response relationship showing the extent to which the probability of occurrence of FAS falls for a marginal change in alcohol consumption. Although medical research convincingly demonstrates that alcohol consumption is the causal agent of the constellation of problems described by FAS, less is known about the shape and magnitude of the dose-response relationship. The lack of information about the dose-response relationship between alcohol and FAS makes it impossible to complete a definitive cost-benefit analysis of prenatal advice about alcohol use.

---

[2]Calculated this way, the costs of FAS can be seen as a lower bound approximation of societal willingness to pay for a small reduction in the risk of an FAS case. To date, there is only very limited evidence on the value of child health. Estimates of the value of a statistical life for a child range from $0.75 million to $7.74 million (Carlin and Sandy 1991, Mount et al. 2000, Jenkins, Owens and Wiggens 2001). Most of these values are substantially larger than the discounted value of lifetime earnings, supporting the argument that PDV_{FAS} is likely to be a lower bound to the value of a statistical case of FAS.

Instead, we ask the following question: How strong would the dose-response relationship between alcohol use and FAS have to be to make prenatal advice socially beneficial? The cost of advice comprises the cost of the time spent by the care giver in dispensing advice plus the value of the patient's time. Kenkel and Terza (2001) estimate this cost to be about $7.00. Equating this value with (41) and plugging in both the estimated value of PE, which is given as -.39 in (38), and the estimated value of PDV_{FAS} ($163,000) taken from Harwood and Napolitano (1985) we have

$$-.392 \times \frac{\partial \Pr\{FAS\}}{\partial y} \times \$163{,}000 = \$7.00. \tag{42}$$

Solving (42) yields

$$\frac{\partial \Pr\{FAS\}}{\partial y} = .001.$$

This implies that if the effect on the probability of FAS of one more drink per week during pregnancy is greater than .001 then the net benefit of prenatal advice is positive.

Appendix

$$\int_{\varepsilon} H(A^*, x_o^*, x_u^*, \varepsilon, \tau)\, f_\varepsilon(\varepsilon)\, d\varepsilon = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(A^*, x_o^*, x_u^*, \varepsilon, \tau)\, f_{\varepsilon_1,\varepsilon_2}(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} I(A*\gamma_1 + x_o^*\beta_1 + x_u^*\theta_1 + \varepsilon_1 > 0)\exp(A*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \varepsilon_2)\, f_{\varepsilon_1,\varepsilon_2}(\varepsilon_1, \varepsilon_2)\, d\varepsilon_1\, d\varepsilon_2$$

$$= \int_{-m_1}^{\infty} f_{\varepsilon_1}(\varepsilon_1)\left[\int_{-\infty}^{\infty} \exp(A^*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \varepsilon_2)\, f(\varepsilon_2 \mid \varepsilon_1)\, d\varepsilon_2\right] d\varepsilon_1$$

But for the values of ε_1 in the relevant range of integration we know that $f(\varepsilon_2 \mid \varepsilon_1)$ is the normal density with mean 0 and variance σ^2. Therefore, the integral inside the square brackets is

$$\int_{-\infty}^{\infty} \exp(A^*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \varepsilon_2)\, f(\varepsilon_2 \mid \varepsilon_1)\, d\varepsilon_2 = \exp(A^*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \sigma^2/2)$$

and

$$J(A^*, x_o^*, x_u^*, \tau) = \int_{\varepsilon} H(A^*, x_o^*, x_u^*, \varepsilon, \tau)\, f_\varepsilon(\varepsilon)\, d\varepsilon$$

$$= \Phi(A^*\gamma_1 + x_o^*\beta_1 + x_u^*\theta_1)\exp(A^*\gamma_2 + x_o^*\beta_2 + x_u^*\theta_2 + \sigma^2/2)$$

References

Bien, Thomas, William Miller, and J. Scott Tonigan (1993). "Brief Interventions for Alcohol Problems: A Review," *Addiction*, 88: 315-336.

Bloss, Gregory (1994). "The Economic Cost of FAS." *Alcohol Health & Research World* 18 (1): 53-55.

Carlin, P. And R. Sandy (1991). "Estimating the Implicit Value of a Young Child's Life," *Southern Economic Journal* 58 (1): 186-202.

Chang, Grace, Louise Wilkins-Haug, Susan Berman, and Margaret Ann Goetc (1999). "Brief Intervention for Alcohol Use in Pregnancy: A Randomized Trial." *Addiction* 94 (10): 1499-1508.

Duan, Naihua., Willard Manning, et al. (1983). "A Comparison of Alternative Models for the Demand for Medical Care," *Journal of Business and Economic Statistics* 1: 115-126.

Fleming, Michael F., Kristen Lawton Barry, Linda Baier Manwell, K. Johnson, and R. London (1997). "Brief Physician Advice for Problem Alcohol Drinkers: A Randomized Controlled Trial in Community-Based Primary Care Practices." *JAMA* 277: 1039-1045.

Heckman, J. (1992): "Randomization and Social Program Evaluation," in Manski and Garfinkel eds. *Evaluating Welfare and Training Programs*, Cambridge: Harvard University Press.

_____ and Smith, J. (1995): "Assessing the Case for Social Experiments," *Journal of Economic Perspectives*, 9, 85-110.

Institute of Medicine (1996). *Fetal alcohol syndrome : diagnosis, epidemiology, prevention, and treatment.*/ Kathleen Stratton, Cynthia Howe, and Frederick Battaglia, editors; Committee to Study Fetal Alcohol Syndrome, Division of Biobehavioral Sciences and

Mental Disorders. Washington, D.C. : National Academy Press, 1996.

Ippolito, Pauline and Alan Mathios (1990). "Information, Advertising and Health: A Study of the Cereal Market," *Rand Journal of Economics*, 21(3), (Autumn): 459-480.

Ippolito, Pauline and Alan Mathios (1995). "Information and Advertising: The Case of Fat Consumption in the United States," *American Economic Review: Papers and Proceedings*, 85(2): 91-95.

Jenkins, Robin, Nicole Owens, and Lanelle Wiggins (2001). "Valuing Reduced Risks to Children: The Case of Bicycle Safety Helmets," *Contemporary Economic Policy*, 19: 397-408.

Kenkel, D.S. (1991). "Health Behavior, Health Knowledge, and Schooling," *Journal of Political Economy*, 83, (April): 287-305.

_____ and Joseph Terza (2001): "The Effect of Physician Advice on Alcohol Consumption: Count Regression with an Endogenous Treatment Effect," *Journal of Applied Econometrics*, 16, 165-184.

Harwood, H.J, and Napolitano, D.M. (1985): "Economic Implications of the Fetal Alcohol Syndrome," *Alcohol Health and Research World*, 10, 38-43.

May, Philip A. And J. Phillip Gossage (2001): "Estimating the Prevalence of Fetal Alcohol Syndrome: A Summary." *Alcohol Research and Health*, 25, 159-167.

Mount, Timothy, W Weng, and William Schulze (2000). "Automobile Safety and the Value of Statistical Life in the Family: Valuing Reduced Risk for Children, Adults, and the Elderly." Working Paper, Department of Applied Economics and Management, Cornell University.

National Institute on Alcohol Abuse and Alcoholism (1998). "Preventing Fetal Alcohol Syndrome." Request for Applications number AA-99-002, U.S. Public Health Service, National Institutes of Health.

Rosett, H.L. *et al.* (1981a) "Identifying and Treating Pregnant Patients at Risk from Alcohol," *Canadian Medical Journal* 125: 149-154.

Rosett, H.L. *et al.* (1981b). "Strategies for Prevention of Fetal Alcohol Effects," *Obstetrics and Gynecology* 57: 1-7.

Sakata, S. (2000): "Quasi-Maximum Likelihood Estimation with Complex Survey Data," Unpublished mimeo.

Smith, I.E., *et al.* (1987). "Identifying High-Risk Pregnant Drinkers: Biological and Behavioral Correlates of Continuous Heavy Drinking During Pregnancy," *Journal of Studies on Alcohol* 48: 304-309.

Stewart, M. (1983): "On Least Squares Estimation when the Dependent Variable is Grouped," *Review of Economic Studies*, 50, 141-149.

Teutsch, Steven and Jeffrey Harris (1996). "Introduction" in *Prevention Effectiveness: A Guide to Decision Analysis and Economic Evaluation*, Anne C. Haddix, et al. (eds.), New York: Oxford University Press.

U.S. Department of Health and Human Services (1993). *Eighth Special Report to the U.S. Congress on Alcohol and Health*.

U.S. Department of Health and Human Services (1997), *Ninth Special Report to the U.S. Congress on Alcohol and Health.* Public Health Service, National Institute on Alcohol Abuse and Alcoholism.

U.S. Preventive Services Task Forces [PSTF] (1995). *Guide to Clinical Preventive Services Second Edition*. Report of the U.S. Preventive Services Task Force, U.S. Department of Health and Human Services.

Viscusi, W. Kip. (1990). "Do Smokers Underestimate Risks?" *Journal of Political Economy*, 98(5), (December): 1253-1269.

World Health Organization Brief Intervention Study Group (1996): "A Cross-National Trial of Brief Interventions with Heavy Drinkers," *American Journal of Public Health*, 86(7): 948-955.

Table 1: Drinking During Pregnancy

Drinking frequency	Live Birth Sample	Infant Death Sample	Fetal Death Sample	Total
None	7,455	3,483	2434	13,372
Some Drinking	1,509	656	441	2,606
Total	8,964	4,139	2,875	15,978

Table 2: Drinking Categories and Frequencies

Drinks per week	Total
$y = 0$	13,372
$C_1 = I(0 < y < .25)$	1,211
$C_2 = I(.25 \leq y < .5)$	392
$C_3 = I(.5 \leq y \leq .75)$	348
$C_4 = I(.75 < y \leq 1)$	185
$C_5 = I(1 < y \leq 2)$	197
$C_6 = I(2 < y \leq 5)$	135
$C_7 = I(5 < y \leq 8)$	70
$C_8 = I(8 < y \leq 13)$	70
$C_9 = I(13 < y \leq 20)$	26
$C_{10} = I(20 < y)$	19

Table 3: Definitions of Variables

Variable	Definition
AGE	Actual age of respondent
P	Defined as a weighted average of the prices of the various types of alcoholic beverages. Prices of beer, wine and liquor are obtained from the Inter-City Cost of Living index published by the American Chamber of Commerce Research Association (ACCRA). Quarterly prices of a six pack of 12-ounce Budweiser or Schlitz, a 750 ml bottle of Seagrams Crown 7 and a 1.5 liter bottle of Paul Masson Chablis are reported. The price index is a weighted average of these three prices. The information regarding the weights is calculated from the 1988 National Health Interview Survey (NHIS) alcohol supplement. The weights are 49.74, 25.43 and 24.83 for beer, wine and liquor, respectively.
HISPANIC	Binary variable equals 1 if respondent is of Spanish/Hispanic origin
BLACK	Binary variable equals 1 if respondent is black
ASIAN	Binary variable equals 1 if respondent is Asian/pacific islander
NATIVAM	Binary variable equals 1 if respondent is Eskimo/Aleut or American Indians
COLLEGE	Binary variable equals 1 if respondent has more than 16 years of schooling
LT- HISCH	Binary variable equals 1 if respondent has 1-11 years of schooling
SOMECOLL	Binary variable equals 1 if respondent has 13-15 years of schooling
INCOME* (1000$)	The total income for people living in your household during the 12 months before your delivery
MARRIED	Binary variable equals 1 if respondent is married
WIDOWED	Binary variable equals 1 if respondent is widowed
DIVORCED	Binary variable equals 1 if respondent is divorced
SEPARATE	Binary variable equals 1 if respondent is separated
WORKING	Binary variable equals 1 if respondent works during her pregnancy
FATHER	Binary variable equals 1 if respondent reports that baby's father lives with her during most of her pregnancy
GRANDPAR	Number of baby's grandparents that lived with mother during most of her pregnancy
CHILDREN	Binary variable equals 1 if respondent reports that her own children live with her during most of her pregnancy
TRAVTIME	Travel time (in minutes) to the primary care provider's office
TRIPS	number of trips to primary care provider's office during pregnancy
OTHERPAY	Binary variable equals 1 if prenatal care was paid for by someone other than respondent or spouse/partner.

* The family income variable is categorical. Interval midpoints were used.

Table 4: Descriptive Statistics

Variable	Mean	Maximum	Minimum
A	0.612	1.000	0.000
P	0.476	0.619	0.411
AGE	26.184	50.000	12.000
HISPANIC	0.090	1.000	0.000
BLACK	0.449	1.000	0.000
ASIAN	0.020	1.000	0.000
NATIVAM	0.008	1.000	0.000
COLLEGE	0.130	1.000	0.000
LT-HISCH	0.235	1.000	0.000
SOMECOLL	0.233	1.000	0.000
INCOME	23.640	90.127	0.500
MARRIED	0.580	1.000	0.000
WIDOWED	0.004	1.000	0.000
DIVORCED	0.045	1.000	0.000
SEPARATE	0.048	1.000	0.000
WORKING	0.586	1.000	0.000
FATHER	0.685	1.000	0.000
GRANDPAR	0.353	4.000	0.000
CHILDREN	0.537	1.000	0.000
TRAVTIME	21.434	150.000	1.000
TRIPS	10.083	30.000	1.000
OTHERPAY	0.735	1.000	0.000

Table 5: Estimates for First-Part of the Two-Part Model Advice Assumed Exogenous Endogenous – Equation (21)

Variable	Estimate	t-Statistic
A ($\hat{\gamma}_1$)	0.279	10.691
x_o ($\hat{\beta}_1$)		
CONSTANT	-0.970	-5.747
P	-0.821	-2.601
AGE	0.017	6.284
HISPANIC	-0.492	-9.901
BLACK	-0.398	-12.709
ASIAN	-0.868	-7.270
NATIVAM	0.032	0.253
COLLEGE	0.178	4.305
LT-HISCH	0.074	2.135
SOMECOLL	0.091	2.796
INCOME	0.003	5.050
MARRIED	-0.337	-8.283
WIDOWED	-0.155	-0.777
DIVORCED	0.028	0.465
SEPARATE	-0.021	-0.353
WORKING	0.075	2.704
FATHER	-0.044	-1.174
GRANDPAR	-0.136	-5.763
CHILDREN	0.092	3.412

Table 6: Estimates for Second Part of Two-Part Model
Advice Assumed Exogenous – Equation (21)

Variable	Estimate	t-Statistic
A ($\hat{\gamma}_2$)	0.304	4.065
x_o ($\hat{\beta}_2$)		
CONSTANT	-2.010	-4.361
P	-0.343	-0.412
AGE	0.045	6.185
HISPANIC	-0.190	-1.276
BLACK	0.289	3.248
ASIAN	0.016	0.038
NATIVAM	0.175	0.577
COLLEGE	-0.321	-2.998
LT-HISCH	0.377	4.073
SOMECOLL	-0.324	-3.676
INCOME	0.000	0.281
MARRIED	-0.505	-4.620
WIDOWED	-0.872	-1.605
DIVORCED	-0.263	-1.777
SEPARATE	0.242	1.628
WORKING	-0.434	-5.832
FATHER	-0.051	-0.529
GRANDPAR	0.055	0.850
CHILDREN	0.092	1.285
$\hat{\xi}$	0.666	----

Table 7: Stage 1 Estimates for First-Part of Two-Part Model
Advice Assumed Endogenous – Equation (19)

Variable	Estimate	t-Statistic
A ($\hat{\gamma}_1$)	-0.292	-0.823
x_o ($\hat{\beta}_1$)		
CONSTANT	-0.589	-2.062
P	-0.911	-2.718
AGE	0.017	6.104
HISPANIC	-0.518	-9.148
BLACK	-0.431	-10.265
ASIAN	-0.971	-6.745
NATIVAM	0.041	0.308
COLLEGE	0.182	4.201
LT-HISCH	0.056	1.487
SOMECOLL	0.098	2.869
INCOME	0.004	4.798
MARRIED	-0.362	-7.707
WIDOWED	-0.221	-1.045
DIVORCED	0.015	0.240
SEPARATE	-0.014	-0.225
WORKING	0.093	2.980
FATHER	-0.048	-1.235
GRANDPAR	-0.150	-5.624
CHILDREN	0.077	2.681
$\hat{\theta}_1$	0.352	1.616

Table 8: Stage 2 Estimates for Second Part of Two-Part Model
Advice Assumed Endogenous – Equation (20)

Variable	Estimate	t-Statistic
A ($\hat{\gamma}_2$)	-0.969	-3.000
x_o ($\hat{\beta}_2$)		
CONSTANT	-1.242	-2.390
P	-0.559	-0.686
AGE	0.046	5.910
HISPANIC	-0.260	-1.741
BLACK	0.199	1.835
ASIAN	-0.213	-0.479
NATIVAM	0.172	0.550
COLLEGE	-0.299	-2.666
LT-HISCH	0.343	3.559
SOMECOLL	-0.302	-3.290
INCOME	0.002	1.055
MARRIED	-0.562	-5.022
WIDOWED	-1.024	-2.092
DIVORCED	-0.290	-1.787
SEPARATE	0.248	1.642
WORKING	-0.399	-5.058
FATHER	-0.067	-0.595
GRANDPAR	0.018	0.235
CHILDREN	0.074	1.031
$\hat{\xi}$	0.710	----
$\hat{\theta}_2$	0.740	4.030

Table 9: Stage 1 Estimates of the Advice Model – Equation (19)

Variable	Estimate	t-Statistic
z ($\hat{\alpha}$)		
CONSTANT	0.443	3.144
P	-0.321	-1.232
AGE	0.001	0.356
HISPANIC	-0.057	-1.505
BLACK	-0.105	-4.098
ASIAN	-0.378	-5.212
NATIVAM	0.054	0.476
COLLEGE	-0.007	-0.192
LT-HISCH	-0.082	-2.954
SOMECOLL	0.021	0.784
INCOME	0.003	5.104
MARRIED	-0.095	-2.885
WIDOWED	-0.295	-1.802
DIVORCED	-0.072	-1.369
SEPARATE	0.028	0.543
WORKING	0.070	3.106
FATHER	-0.019	-0.614
GRANDPAR	-0.050	-2.793
CHILDREN	-0.086	-3.849
TRAVTIME	-0.001	-1.870
TRIPS	0.013	6.460
OTHERPAY	-0.034	-1.392