Is Prenatal Care Really Ineffective? Or, is the 'Devil' in the Distribution?*

Karen Smith Conway
Department of Economics, University of New Hampshire

and

Partha Deb**
Department of Economics, Hunter College, CUNY

September 2004

Abstract

Prenatal care should improve infant health, yet research frequently finds only weak effects. If there are two kinds of pregnancies, ‘complicated’ and ‘normal’ ones, then combining these pregnancies may lead prenatal care to appear ineffective. Data from the NMIHS offers compelling evidence. The standard 2SLS approach yields obviously bimodal residuals and frequently insignificant prenatal care coefficients. In contrast, estimating birth weights with a finite mixture model yields estimates revealing that prenatal care has a substantial effect on ‘normal’ pregnancies. Our Monte Carlo experiment confirms that ignoring even a small proportion of ‘complicated’ pregnancies can lead prenatal care to appear unimportant.

JEL Codes: C51, D13, I12
Keywords: Infant health, birth weight, prenatal care, finite mixture models

---

* We wish to thank David Bradford, Amitabh Chandra, Ted Joyce, Steve Lehrer, Ellen Meara, Doug Staiger and Joe Newhouse, the editor, and an anonymous referee for suggestions that have improved the paper substantially. In addition, we thank participants at seminars at University of Alabama at Birmingham, Boston University, Dartmouth College, Hunter College, University of New Hampshire, and RAND and at the International Health Economics Association Conference in York, UK., for their comments.

** Corresponding Author: Department of Economics, Hunter College, City University of New York, 695 Park Avenue, New York, New York 10021, USA; Phone: (212)772-5435; Fax: (212)772-5398; Email: partha.deb@hunter.cuny.edu

It is widely believed that expanding prenatal care should improve infant health; indeed, increasing the prenatal care of low income women is part of the motivation behind the recent Medicaid expansions. Yet, economic research typically finds weak, if any, effects of prenatal care on infant health (e.g., Currie and Grogger, 2000; Kaestner, 1999; and Grossman and Joyce, 1990). Some of the reasons given include inadequate measures of prenatal care and difficulties in modeling its endogeneity. We offer another explanation – that there are essentially two kinds of pregnancy outcomes, 'complicated' and 'normal' ones. Our research investigates this possibility by extending the standard infant health production model to include a finite mixture of distributions of birth outcomes.

Combining 'complicated' and 'normal' pregnancies, as past research does, could lead prenatal care to appear ineffective for at least two reasons. First, complicated pregnancies typically entail a large number of prenatal care visits, but yield poorer outcomes. Most research considers these health complications in modeling prenatal care by including information on past pregnancies in both the prenatal care and birth weight equations. This remedy is most effective when the onset of prenatal care rather than the number of visits is used.

However, fully recognizing such complications requires different estimation of the birth weight equation. If there are two kinds of pregnancies, then the outcomes may not adhere to a single distribution, even after controlling for all observed factors. Furthermore, observed factors such as prenatal care and other maternal behaviors may have different effects on each type of pregnancy. This possibility has been strongly suggested in the medical literature (e.g., Paneth, 1995; and Alexander and Korenbrot, 1995). Low birth weight is typically due to retarded growth and/or preterm birth. Clinical evidence suggests that preterm births are quite difficult to prevent, and that even such typically important maternal behaviors as smoking and drinking are more closely linked to restricted fetal growth rather than shortened gestation (Goldenberg and Rouse, 1998). Prenatal care is likewise believed to do little to prevent preterm birth, but rather is most effective in preventing retarded growth in near full term infants (e.g., Alexander and

Korenbrot, 1995 and Shiono and Behrman, 1995). Therefore the lowest birth weights may be the most influential observations, yet may be least affected by prenatal care (and other factors).

By estimating the birth weight equation as a finite mixture model, we allow the observed factors to have different effects for each type of pregnancy and the unobserved factors to follow a different distribution. In this research, we focus on the effects of prenatal care on birth weight because that has been the factor of greatest interest. However, this empirical approach has implications that extend to other factors such as maternal smoking and drinking or policy changes, which may be likewise affected. More generally, we are estimating a model that is more in the spirit of the findings of the medical literature, regardless of the particular factor of interest.

Finite mixture models have wide appeal and applicability. McLachlan and Peel (2000) and Titterington, Smith and Makow (1985) provide excellent surveys of the statistical literature. Its growing popularity is reflected in an increase in the number of applications in labor economics (Heckman, Robb, and Walker, 1990; Gritz, 1993; Geweke and Keane, 1997), marketing (Wedel et al., 1993), development economics (Morduch and Stern, 1997), industrial organization (Wang, Cockburn, and Puterman, 1998), and other health economics contexts (Deb and Trivedi, 1997).

The empirical analysis uses data from the National Maternal and Infant Health Survey (NMIHS) for white and black women who experienced a live birth. The NMIHS is a widely used data set that contains information about the prenatal care received, including when it began, and the birth weight and gestation of the infant. Information on insurance coverage, health history, and many other demographic variables are also used. We impose standard identifying assumptions, such as including insurance coverage and income only in the prenatal care equation, to treat prenatal care as endogenous. More generally, we adhere to the methodology of past studies wherever possible, including using the typical instruments, so that we may isolate the impact of allowing for a finite mixture distribution. These instruments are not without their weaknesses, but this approach allows us to see if they are really to blame for prenatal care's weak effect on birth outcomes. In addition, we tackle the instruments issue, as well as abstract from other data issues altogether, by performing a Monte Carlo experiment using data generated under the assumption of a finite mixture distribution with *exogenous* prenatal care, which we then estimate using OLS. This exercise provides additional evidence about how ignoring the finite mixture distribution may affect the estimated effects of prenatal care in a situation that does not require finding appropriate instruments for prenatal care.

Our measure of prenatal care is the onset of prenatal care both because it is the most commonly used measure and so that we may avoid the problem that plagues visits -- i.e., that complicated pregnancies entail more visits. In this way, we can focus on the second issue -- that certain factors, such as prenatal care, may have different effects on the two types of pregnancies and that failing to account for this is misspecifying the model. Is there evidence of misspecification? Figure 1 presents the unweighted and weighted kernel densities of the 2SLS residuals from a birth weight regression estimated from our main samples, which can also be thought of as birth weights adjusted for demographic variables and prenatal care.[1] In light of the two modes, and the fact that multimodal distributions are the "textbook" representation of a finite mixture distribution, these figures present compelling evidence that the standard method is misspecifying the model and that a finite mixture is more appropriate. What impact does this misspecification have on the standard infant health model? And, what new insights are gained by using a finite mixture model? Answering these questions is the purpose of this paper.

2. Background

Previous economic research builds upon the infant health production model of Rosenzweig and Schultz (1982, 1983), which results in a birth outcome production function and input demand

[1]As we discuss in greater detail in Section 4, the NMIHS oversamples poor birth outcomes and blacks and so the sample weights provided by the NMIHS are frequently used in empirical analyses. We estimate our models both with and without the weights to verify that the oversampling is not driving our results. Bimodality is also apparent in the distribution of birth weights, regardless of whether the weights are used. We have presented the distribution of residuals because it eliminates modes due to observable characteristics.

functions, such as prenatal care utilization. A common specification is to estimate a simultaneous model that includes a prenatal care equation,

$$PNC_i = X_i\beta + u_i \quad \textbf{(1)}$$

and a birth weight equation,

$$BWT_i = PNC_i\gamma + Z_i\tau + \varepsilon_i \quad \textbf{(2)}$$

where *X* includes characteristics of the mother including age, education and her health endowment, and factors affecting her ability to obtain prenatal care such as income and insurance status.[2] The variables in *Z* are frequently, but not always, a subset of *X*, and the identifying restrictions tend to include variables that capture availability of prenatal care, such as insurance status, income and community-level variables.[3] As discussed in detail by Warner (1998), satisfactory restrictions are difficult to find and the structural birth weight equation may be weakly identified as a result. Indeed, such weak identification has been blamed in part for the lack of significant prenatal care effects (e.g., Currie and Grogger, 2000, who use policy changes in Medicaid and welfare as identifiers).

Further complicating the model are maternal behaviors such as smoking and drinking, which are strongly associated with poor birth outcomes, yet are likely a matter of maternal choice. Including these variables and treating them as endogenous further strains an already weakly identified birth weight equation. In addition, one of the objectives of prenatal care is to change these behaviors and so their inclusion may obscure the effects of prenatal care. Most studies therefore either omit these behaviors (e.g., Liu, 1998 and Currie and Grogger, 2000) or treat them as exogenous, sometimes using information on behaviors before the pregnancy, (e.g., Kaestner, 1999; Warner, 1998; and Grossman and Joyce, 1990). In order to focus on the effects

---

[2] Warner (1998) is the first to emphasize maternal anthropometric characteristics such as mother's birth weight and height, which he finds quite significant. We therefore include such variables also.

[3] For example, Grossman and Joyce (1990) use insurance status, the availability of WIC centers and prenatal care clinics in the area and the percentage in poverty to identify prenatal care. In the birth weight equation, however, they add the baby's gender, whether a private physician attended the birth and tobacco, alcohol and narcotics use, such that *Z* is not a subset of *X*.

of a finite mixture specification (and prenatal care), and to keep the model as uncomplicated as possible, we choose to omit these behaviors. However, using a finite mixture model may alter the estimated effects of these behaviors on birth weight and so is a worthwhile extension of this research.

All economic studies to date have ignored the consistent findings of the medical literature that one of the two main causes of low birth weight – preterm birth – is very difficult to prevent. The other cause, retarded growth, appears much more preventable, especially for infants at or near fullterm (greater than 37 weeks gestation). Shiono and Behrman (1995) in their excellent review of the problems of low birth weight note that while the medical profession has made huge strides in rescuing troubled infants, it has had little success in preventing troubled births, especially those due to preterm birth. They note that "the causes of most preterm births have not been identified," whereas low birth weight due to retarded fetal growth is more easily prevented by modifying maternal behaviors, which as discussed above is an objective of prenatal care. Indeed, Kogan et al. (1994) looks at the content of prenatal care and finds that it is the receipt of advice on healthy behaviors that appears to have the strongest effect. Shiono and Behrman (1995) go on to state "...even the best prenatal care alone cannot be expected to solve the dual problems of low birth weight and preterm birth." This appears to be borne out by several studies that have looked at attempts to increase access to prenatal care (e.g., Medicaid) and expand its content and have found no impact on preterm birth (Collaborative Group on Preterm Birth Prevention, 1993; Binstock and Wolde-Tsadik, 1995; and Ray, Mitchell and Piper, 1997).

Treating all pregnancies as the same, as the health economics literature has done, may therefore obscure the positive effects that prenatal care is having on a subgroup of pregnancies. Indeed, the medical literature strongly suggests that a finite mixture specification is more appropriate. Alexander and Korenbrot (1995) argue that researchers need to separate out the population of low birth weight births that are potentially modifiable. Paneth (1995, p.24) notes that "Several mathematically oriented investigators... have argued that the roughly normal birth weight curve in any population is really a mixture of two distributions, one of the normal population and the other of a pathological group of babies (referred to as the "residual" distribution) in whom small size is a reflection of some unhealthy maternal or fetal condition." Our intent here is to bring this wisdom to health economics research on infant health by estimating a finite mixture distribution model of birth weight. Our results allow us to see whether the pregnancy outcomes can be differentiated in such a way and, if so, whether prenatal care and other factors have differing effects on them.

3. Empirical Strategy

The basic model we estimate consists of equation (1) to predict onset of prenatal care and a birth weight equation derived from equation (2),

$$BWT_i = P\hat{N}C_i\gamma + Z_i\tau + \varepsilon_i \qquad \textbf{(3)}$$

where $P\hat{N}C$ is the predicted value of *PNC* obtained from equation (1). In choosing this specification, we attempt to capture the general spirit of previous studies rather than replicate a specific one, as each study differs slightly in the variables it includes. Likewise, we use typical data and instruments for prenatal care so that we may provide evidence of the impact of using a finite mixture model in a 'typical' infant health study – and address the issues of problematic instruments and specific data with our Monte Carlo experiment.

We choose the onset of prenatal care in weeks as the measure of prenatal care for several reasons. Foremost, it is the measure most used by past researchers, either directly (e.g., Grossman and Joyce 1990, Liu 1998, Warner 1995, 1998) or as part of a discrete measure or index of care (e.g., Joyce 1994 and Currie and Grogger 2000). Even those that focus on number of visits tend to condition on when prenatal care began (e.g., Joyce 1999 and Kaestner 1999).

Another advantage of onset of prenatal care is that, unlike visits, it is less likely plagued with also having a mixed distribution. 'Complicated' pregnancies are more likely to entail a greater number of visits than 'normal' ones. This suggests that visits should also be modeled with a finite mixture of distributions, which would greatly complicate our model. The onset of prenatal care is less likely to be affected by the type of pregnancy and its endogeneity can be more adequately remedied by including the woman's history of difficult pregnancies.[4] However, we later use number of visits to check the validity of our results by seeing if those women with a high probability of a complicated pregnancy did indeed have more visits. Finally, getting women early prenatal care has been a focal point in the push to improve infant health.

In the prenatal care equation, we include all of the variables in the system as is standard in a first stage regression within a simultaneous system. We therefore include the mother's anthropometric characteristics (her height and birth weight), her medical history (parity, number of prior fetal deaths), and her other characteristics (age, education, number of own children living with her, urban). The gender of the infant enters both equations for consistency, although it probably is not known at the time prenatal care is sought.

Ideal instruments for prenatal care are variables that capture the availability of prenatal care but that are not otherwise correlated with the errors, i.e., the instruments should be "relevant" and "exogenous". However, as discussed above and noted by Warner (1998) and others, finding such instruments for prenatal care is quite difficult. In order to focus on the merits of the finite mixture approach, we employ the typical identifying restrictions. Specifically, we use two state-level characteristics, a health care price index and population density, and several individual characteristics associated with the 'cost' of prenatal care and the mother's ability to pay for it – whether the mother cohabits with the father, her income and insurance status. These restrictions are not beyond criticism; for instance, one could argue that family income could directly affect birth weight. To address such criticism, we perform and report the results of two tests on the identifiers in the standard 2SLS model. First, we test their joint statistical significance in the prenatal care equation; this is the test for instrument relevance.

[4] For instance, a woman who has had difficult pregnancies in the past may be more likely to seek care early and may have more visits. This can be controlled for by including her medical history. Conversely, a woman with an otherwise normal history whose pregnancy becomes complicated is likely to have more visits, but is no more likely to have sought care early. Yet, the impact of having sought care early may differ by type of pregnancy.

Second, we perform a Hansen test on the over-identifying restrictions; this is the test for instrument exogeneity. Developing an over-identification test for our finite mixture model is beyond the scope of this paper, and so we must rely on the 2SLS results for reassurance. However, as a final inquiry into the impact that our instruments may have on the finite mixture results, we perform a Monte Carlo exercise that allows us to abstract from the endogeneity issue altogether.

A. A Finite Mixture of Distributions Birth weight Function

In the finite mixture model, the random variable of interest is assumed to be a draw from a population that is an additive mixture of distinct subpopulations or classes (*c*) in proportions p_c where $\sum_{c=1}^{C} p_c = 1, p_c > 0 \forall c = 1,2,\ldots,C$. For birth weight, *BWT*, the mixture density can be described in general by

$$g(BWT|\Theta) = p_1 g_1(BWT|\Theta_1) + p_2 g_2(BWT|\Theta_2) + \ldots + p_C g_C(BWT|\Theta_C) \qquad \textbf{(4)}$$

where the class densities $g_c(BWT|\Theta_c)$ are assumed to be normals, i.e.,

$$g_c(BWT|\Theta_c) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left(\tfrac{-1}{2\sigma_c^2}(BWT - X_c\beta_c)^2\right) \qquad \textbf{(5)}$$

The mixing probabilities, p_c, regression coefficients, γ_c and τ_c, and the standard deviation parameters, σ_c, vary across classes.[5]

The log likelihood function for the data is given by

$$l(BWT|\Theta) = \sum_{i=1}^{N} \log(g_i(BWT_i|\Theta)) \qquad \textbf{(6)}$$

[5] In principle, one could specify prenatal care, PNC, and BWT as a bivariate density (with BWT following a mixture density) and jointly estimate the parameters of both equations. Both for simplicity and to remain in the spirit of current research, we instead use a two-step approach that estimates prenatal care separately in a first stage.

The model is estimated by maximum likelihood using a quasi-Newton constrained maximization algorithm, the code for which is implemented in SAS/IML. Note that the mixing probabilities are jointly estimated with the class-specific regression coefficients and standard deviations.

B. Sampling weights

As we discuss in greater detail in the data section, the NMIHS is a complex, stratified random sample and includes oversampling of poor birth outcomes. In such designs, the use of sample weights can, in principle, substantially alter inference via estimates of standard errors of parameters and can change point estimates. In practice, point estimates tend not to change much although standard errors of estimates do change noticeably. To guard against the possibility that our inference is unduly affected by sampling weights, we estimate our models with and without them. In the weighted models, the 2SLS estimates are calculated using standard weighted least squares. For the finite mixture model, we maximize a weighted log likelihood function,

$$l(BWT|\Theta) = \sum_{i=1}^{N} w_i \log(g_i(BWT_i|\Theta) \quad \textbf{(7)}$$

where w_i denotes the weight associated with observation i. The sample weights are inversely proportional to the probability that the observation would be observed in the population, i.e., they are probability weights.

C. Inference

Asymptotic standard errors obtained using the ordinary least squares formula for the second stage in the 2SLS case, and formulae based on maximum likelihood or quasi maximum likelihood theory are incorrect because they do not account for the fact that a generated regressor, predicted prenatal care, is being used as a covariate in the birth weight regressions. In the case of 2SLS, a formula exists to adjusts the standard least squares estimates. In the case of finite mixture models, an analogous formula does not exist. Therefore, we use bootstrap methods to calculate standard errors for the parameter estimates in the finite mixture specification. For comparability, we also use bootstrap standard error estimates in the 2SLS specification.

Bootstrap estimates of standard errors are constructed in the following way. A random sample with replacement is drawn from the original sample, with the same sample size. The first stage regression is estimated using the random sample and predicted prenatal care is calculated. Least squares and maximum likelihood are used to calculate birth weight equations using predicted prenatal care as a regressor and the parameter estimates are stored. This process of drawing bootstrap samples and estimation is repeated 1000 times. The sample standard deviations of the 1000 sets of parameter estimates are the bootstrap standard errors of the point estimates of the parameters.

D. Robustness

In order to confirm that our results are generalizable, we check for robustness in our empirical analysis along three dimensions. First, as mentioned above, we estimate every variation of the 2SLS and finite mixture models both with and without using the sample weights provided with the NMIHS. Second, we guard against the possibility that our results might be due to a small number of outliers via the use of nonparametric bootstrapped standard errors for inference. Third, we estimate our models using our preferred sample that includes valid information for the birth weight of the mother and a larger sample which excludes this variable. In addition, as explained further below, in every variation of the model we stratify the sample by race. Finally, we confirm the plausibility of our findings by conducting a Monte Carlo experiment, described in Section 7, in which we can abstract from data and instrument specification issues altogether.

4. Data Description

Our primary data comes from the National Maternal and Infant Health Survey, 1988 (NMIHS), which is publicly accessible data published by the National Center for Health Statistics (NCHS). NMIHS contains information on women who were pregnant in 1988 and, following Kaestner (1999) and Warner (1998), we use only the data from the live birth sample. While Grossman and Joyce (1990) point out that this selection causes bias, using the standard Heckman correction is particularly problematic for us because it assumes an underlying normal distribution. We therefore follow Kaestner (1999) and Warner (1998) in omitting this correction. We also take comfort in the fact that Gray (2001), who uses the NMIHS and explores the impact of self-selection, finds that it has little impact. Finally, one could argue that by limiting our focus to live births we are, if anything, underestimating the importance of the 'complicated' class of the finite mixture inasmuch as fetal and infant deaths are likely disproportionately represented there.

Another issue is that the NMIHS is not a simple random sample. The stated purpose of the NMIHS is to study poor pregnancy outcomes and therefore such outcomes are oversampled. Within the live birth sample, it oversamples low and very low birth weight babies; black infants are oversampled as well. Because of this oversampling and because previous research finds important racial differences in birth weight (e.g., Warner 1995, 1998, Liu 1998 and Conway and Kennedy forthcoming), we stratify our sample into blacks and whites, and conduct our analysis using weighted and unweighted procedures.[6]

As most studies do, we eliminate multiple births because such babies tend to be born at shorter gestation and lower birth weights, and to also eliminate multiple birth observations that share one prenatal care observation. To construct the final samples, we begin with the 9146 live, singleton births. We then eliminate, in the following order, the small number of women who had no prenatal care (N=284)[7], those missing information for number of prenatal care visits (N=10), and those missing information for parity (N=79). We follow the typical practice of omitting teenagers (< 19 years old) and much older mothers (> 50 years), which eliminates 1168 observations. Following Warner (1998), gestations less than 20 weeks or greater than 45 weeks

[6] One might argue that this oversampling could be responsible for the 'bump' evident in the birth weight (residual) distributions presented in Figure 1. However, the fact that this 'bump' remains even after the sample weights are incorporated into the estimation procedure and the construction of the kernal density strongly suggests that this bimodality is not an artifact of the sampling process. Furthermore, as we discuss in the results section, the strong statistical evidence of a mixture of distributions remains even after taking the sample weights into account.

[7] Although this could be another potential source of self-selection, the small number of observations made controlling for it infeasible.

(N=48) and birth weights below 400 grams or above 6000 grams (N=35) were deleted. We also eliminate the 30 observations from Hawaii. After all of these exclusions, there are (9146-284-10-79-1168-48-35-30 =) 7492 observations, of which 3350 are nonHispanic blacks and 3245 are nonHispanic whites.

Maternal birth weight has been found to be an important predictor of the infant's birth weight (Warner 1998). However, a large number of observations have missing data for this variable. For this reason, we estimate the model with two samples for each race. A bigger sample (N=3350 and N=3245) is used when maternal birth weight is left out of the model, and a smaller sample (N=2312 and N=2905) is used when it is included. The results are very similar between the two samples and maternal birth weight is always statistically significant. For this reason, we report descriptive statistics and results for only this smaller sample.[8] As an additional check, we also re-estimate the 2SLS model for the smaller sample but excluding maternal birth weight as a variable and the estimates are again very similar. Thus, neither the sample nor the inclusion of maternal birth weight appear to have an impact on the results and so we feel comfortable emphasizing the results from the smaller sample that includes mother's birth weight.

Table 1 reports the names, definitions and sources of the variables used in the model. It also presents the weighted and unweighted sample means of the variables for nonHispanic white mothers and black mothers. Immediately evident is the difference between the races and the expected impact of using the weights. Black mothers consistently have infants that are on average 200-300 grams lighter than white infants, even after using the sample weights. Some of this may be due to socioeconomic circumstances as black mothers clearly appear more economically disadvantaged, regardless of whether the weights are used. Black mothers are younger, have slightly lower education, have lower family incomes and higher Medicaid participation. They are more likely to have never married, and are much less likely to be

[8]Results from the full sample are available upon request. In addition, the distributions of the 2SLS residuals – i.e., those reported in Figure 1 – are also very similar for this sample and are likewise available upon request.

cohabiting with the father yet are cohabiting with a greater number of children. They also receive prenatal care more than a week later than white mothers on average. In general, then, weighting the data increases birth weights, as expected, but otherwise has little impact on the means of the other variables and the racial disparities persist.

5. 2SLS vs. Finite Mixture Results

Table 2 reports the results from estimating the birth weight equation (measured in hundreds of grams) by standard 2SLS, both weighted and unweighted, for our main sample of white mothers and of black mothers.[9] Recall that prenatal care is predicted from another equation, in which cohabiting with the father, income, insurance status and state-level variables population density and health care price index, are identifying instruments.

Reported at the bottom of Table 2 are two tests of the quality of these instruments. The first is a test of the joint statistical significance of these identifiers in the prenatal care equation. Every model we estimate (including those unreported) strongly rejects that these identifiers are jointly equal to zero; in addition, most of the identifiers are individually statistically significant from zero as well. The second is a test of the over-identifying restrictions in the birth weight model, whether these variables are jointly uncorrelated with errors in the birth weight equation. Our instruments satisfy the usual requirements of a typical infant health production model, in that they contribute statistically significant explanatory power in the prenatal care equation, but are not correlated with the error in the birth weight equation. More generally, our prenatal care equation estimates suggest the usual pattern of disadvantaged women (e.g., lower income, Medicaid participants) receiving later prenatal care. The full results from the prenatal care equation are available upon request.

The estimates from the 2-class finite mixture models estimated for white mothers and black mothers are reported in Table 3. We restrict our attention to the 2-class mixture of normals

[9] The results from the bigger sample that includes observations with missing values of mother's birth weight are available upon request. All standard errors are bootstrapped estimates.

for several reasons. First, Figure 1 shows compelling evidence that there are two sub-populations in birth weights. In addition, our attempts to estimate three-class mixtures of normals lead to miniscule improvements in the maximized log likelihoods, improvements too small to justify their further analysis. Finally, there are the insights of Paneth (1995) and others in the medical literature who argue there is only a dichotomy between "normal" and "complicated" births, a view that is now supported by both our graphical view and our analysis of maximized log likelihoods. Our finite mixture model therefore yields two sets of coefficient estimates -- one for 'normal' pregnancies and one for 'complicated' pregnancies. It also yields an estimated probability, *p*, of having a 'normal' pregnancy."

The most striking feature of these results is the robustness of the finite mixture model across samples, races and weighting schemes, especially compared with the 2SLS results. Nowhere is this more evident than for the prenatal care coefficients. In the 2SLS models, the prenatal care coefficients are quite large (ranging from 68 to 75 grams improvement for getting care one week earlier) and statistically significant in the unweighted white results, but are completely washed out if one uses the sample weights. For blacks, the opposite pattern occurs -- prenatal care is only statistically significant if one uses the sample weights. While one might expect the results to be affected by the weighting scheme due to unobservables, it is not at all clear why there would be this strong difference by race.

In contrast, the message from the finite mixture models is clear -- getting prenatal care one week earlier significantly increases birth weights in 'normal' pregnancies by 30-60 grams per week. The gains appear slightly larger and more variable for blacks and also for the (unreported) models that use the bigger sample. For our reported samples, the estimates have a remarkably tight range of 30 to 35 grams per week across blacks and whites.[10] However, onset of prenatal care appears to have no significant effect on the birth weights from 'complicated' pregnancies.

---

[10] This compares to 12 to 37 grams *per month* for blacks and 4 to 23 grams for whites found by Grossman and Joyce (1990), 6 to 197 grams *per month* by Liu (1998), and -18 to 50 *per week* by Warner (1998). Our estimates therefore fall in the high end of the range and display much less variability. In our larger samples that exclude mother's birth weight, the estimated range is 42-60 grams a week.

There are other lessons from the finite mixture model as well. The majority of the characteristics found to affect birth weight in the 2SLS models (and other research) -- age, education, mother's birth weight and height, and number of children at home -- affect primarily 'normal' pregnancies. Number of previous infant/fetal deaths, however, only affects the birth weights of 'complicated' pregnancies. These results make sense and are consistent with the medical literature, as discussed earlier, that finds that little can be done to prevent premature births (e.g., Goldenberg and Rouse 1998), which account for a large proportion of low birth weight infants. Mother's age, birth weight and height are also occasionally significant, again being variables over which the physician has no control.

Finally, the estimated probability of having a 'normal' pregnancy is again remarkably robust. This parameter is estimated with great precision and ranges from 0.856 to 0.873 across samples, races and weighting schemes. Given that past research has found blacks to have a higher incidence of poor outcomes, it is noteworthy and somewhat surprising that the probabilities of a poor outcome (1-p) are very close although slightly higher for blacks.[11] But this stability across races is consistent with the view that racial differences in birth outcomes are largely due to economic and behavioral factors.

The estimated magnitudes also appear to be on target with the medical literature. Goldenberg and Rouse (1998) state that preterm birth occurs in 11 percent of all pregnancies. Given that many, but not all, low birth weight babies are also preterm one would expect the probability of a 'complicated' pregnancy to be close to but perhaps slightly larger than 11 percent. In addition, the predicted distribution of a ‘normal’ pregnancy yields a mean birth weight of about 3350 grams for whites and 3150 grams for blacks. These means fall comfortably within the median birth weight range of 3000-3499 grams given by the *National*

[11] For instance, in 1998 approximately 10% of nonHispanic white births were preterm (less than 37 weeks gestation) compared to approximately 17% of nonHispanic black births. Similarly, 6.6% of white births were low birth weight compared to 13.2% for blacks. (*National Vital Statistics Report, Vol. 48, No. 3, March 18, 2000.)*

Vital Statistics Report for 1998. In contrast, the fitted means for the 'complicated' pregnancies are consistent with the definition of very low birth weight babies (< 1500 grams).

6. What Does Our Model Predict with Respect to Birth Outcomes?

We can use our finite mixture parameter estimates to calculate the posterior probability of being a 'normal' pregnancy. Our model assumes that the prior (unconditional) probability of a 'normal' pregnancy is constant across observations (*p*). However, we can use Bayes Theorem to calculate the posterior probability of having a 'normal' pregnancy for each woman, conditional on the actual birth weight of her infant and her explanatory variables:

$$\Pr('normal'|BWT, X) = \frac{pg(BWT|\theta_1)}{pg(BWT|\theta_1) + (1-p)g(BWT|\theta_2)} \qquad \textbf{(8)}$$

With this calculated posterior probability, we can then classify pregnancies as 'normal' (Pr(*'normal'*) > 0.5) and 'complicated' (Pr(*'normal'*) < 0.5). By classifying the pregnancies in this way, we can examine how they are related to other characteristics and move closer to Alexander and Korenbrot's (1995) suggestion that researchers identify the low birth weight births that can be modified. We have calculated these posterior probabilities for the unweighted samples with mother's birth weights and associated parameter estimates. It is important to note that the predicted probabilities themselves are also bimodal such that very few observations have probabilites near 0.5; therefore, modifying our classification rule affects very few observations.

Our first exercise explores how well our model's classification is validated by the usual observed variables – gestational age and birth weight – used to classify troubled pregnancies. Using these observed variables, we classify pregnancies into four groups: 1) full-term, good outcome (normal birth weight and gestation > 37 weeks), 2) preterm but good outcome (normal birth weight and gestation < 37 weeks), 3) full-term, retarded growth infant (birth weight < 2500 grams and gestation > 37 weeks), and 4) pre-term, low birth weight (birth weight < 2500 grams and gestation < 37 weeks). For each of these groups, we report in Table 4 several characteristics, including the proportion of births classified as ‘normal’ by our model.

For both whites and blacks, it is evident that our predicted probabilites map very well against these observed variables. All of the normal birth weight births (the first two categories) are classified as ‘normal’ by our model, whereas only 30 percent of the preterm, low birth weight births are so classified. This adds credibility to our model. At the same time, these results suggest that one cannot isolate our predicted ‘normal’ (and therefore presumably ‘modifiable’) pregnancies by simply eliminating from samples of data all births that are low birth weight and preterm (category 4); some ‘complicated’ pregnancies will remain and some ‘normal’ pregnancies will be excluded. For example, the pre-term/low birth weight group – those we would think to be ‘complicated’ – consist of nearly one-third estimated “normal” births. This suggests that almost one-third of preterm/low birth weight births are modifiable and could therefore be affected by timely prenatal care. Likewise, between 10 and 20 percent of the full-term, low birth weight births (apparently ‘retarded growth’) are estimated to be ‘complicated’ pregnancies.

Other interesting differences emerge between these four groups of pregnancies. The ‘retarded growth’ births (category 3) did receive prenatal care somewhat later than the normal birth weight births, whereas the preterm/low birth weight births received it earlier. More generally, the mothers of preterm/low birth weight babies appear more similar to the mothers of normal birth weight births, in terms of education and onset of prenatal care, than do the mothers of the ‘retarded growth’ babies. The preterm/low birth weight births also received a much greater proportion of the recommended number of prenatal care visits, as expected.

The second exercise classifies pregnancies into three categories using the posterior probability and birth weight – 1) ‘normal’ pregnancies and normal birth weight, 2) ‘normal’ pregnancies and low birth weight, and 3) ‘complicated’ pregnancies (all of which are low birth weight). In this way, we can use our finite mixture estimates to identify the potentially modifiable, troubled pregnancies (group 2) as suggested by Alexander and Korenbrot (1995).

Key characteristics are reported for these three kinds of pregnancies for both whites and blacks in the right hand panel of Table 4.

Here again we see that preterm birth does not perfectly map against our estimate of a 'complicated' pregnancy – about half of the 'potentially modifiable,' low birth weight pregnancies are preterm. Likewise, between 5 and 10 percent of the 'complicated' pregnancies are not preterm. We also see that the 'potentially modifiable' group – group 2 – did indeed receive prenatal care later than the other two groups, whereas the 'complicated' births received it earlier. The second group is also less educated than the other two. They also received a higher proportion of recommended visits than the first group, which is consistent with the usual case of pregnancies with poor outcomes receiving more visits. However, the pregnancies we classify as 'complicated' (group 3) received a much greater proportion of the recommended visits, a further validation of our model.

7. Could a Finite Mixture Really be to Blame? A Monte Carlo Experiment

Our results thusfar suggest that 'normal' pregnancies are significantly affected by the onset of prenatal care, whereas 'complicated' ones are not. We have also asserted that including both types of pregnancies, as past studies do, may lead prenatal care to appear ineffective. However, our estimates suggest that the probability of a complicated pregnancy is only 10 to 15 percent. Is it really possible that such a small proportion of pregnancies are so influential as to make prenatal care appear ineffective? Are our findings unique to our specific data, our specific set of instruments?

To address these questions, we perform a Monte Carlo experiment in which we generate data under the assumption that our finite mixture specification (and resulting estimates) are 'correct'. We then investigate the effects of estimating a typical birth weight production model (2SLS) with this artificially generated data. In this way, we can see what results the 'usual' estimation method will yield if the data are indeed generated from a finite mixture distribution similar to the ones we estimated.

An additional complication to this experiment is the endogeneity of prenatal care. To abstract from its endogeneity and the problems introduced by using potentially poor instruments, we use predicted prenatal care to both generate the birth weights and to estimate the birth weight equation. Note that this eliminates the quality of the instruments as an issue in this exercise. For instance, if the instruments are really poor, using *actual* prenatal care to generate the birth weights and then *predicted* prenatal care in estimation should give poor results even when the data are generated under assumptions favorable to OLS – i.e., not mixture. By treating predicted prenatal care as the 'true' variable (so that it both generates the data and is used in estimation), we can focus on the importance of the finite mixture distribution assumption and make the OLS results the most likely to succeed. More importantly, it allows us to see whether a finite mixture distribution, by itself, can cause prenatal care to appear ineffective, even in a case where prenatal care is exogenous (or ideal instruments exist).

The specifics of our Monte Carlo experiment are most easily explained through a sequence of steps:

Step 1: Use the estimated coefficients from the finite mixture model as values for all nuisance parameters (all coefficients other than the prenatal care parameters and the mixing probability). We choose varying values for the prenatal care coefficients and the probability of a normal pregnancy, including p=1 which is a nonmixture case and the baseline case**.**

Step 2: Take the actual data for the explanatory variables (e.g., education, age) and the *predicted* prenatal care of prenatal care, and then randomly assign (according to probability p) each observation to one of the distributions.

Step 3: Generate the birth weight for each observation, which has been assigned to distribution j (=1,2), using 1) the finite mixture estimates and chosen values for the coefficients for that distribution (γ_j and τ_j), 2) the actual data for the explanatory variables ($P\hat{N}C$ and Z), and 3) a normally distributed random error (ε_j) drawn from distribution j with the finite mixture estimated variance (σ_j^2). In the case of the weighted sample, we use the coefficients and variance estimated from the sample-weighted, finite mixture model.

Step 4: With these generated birth weight data, estimate the birth weight equation with OLS. This estimation procedure assumes that all observations come from the same distribution. In the case of the weighted sample, we estimate the birth weight equation on the data generated with the weighted coefficients and use weighted OLS.

Step 5: Repeat steps 2-4 2000 times.

Step 6: We report the 5^{th}, 50^{th} and 95^{th} percentile values for the estimated prenatal care coefficient. This produces a range of estimated coefficients to compare to those commonly found in the literature. We also report the proportion of times that the estimates lead the hypothesis of no effect of prenatal care (i.e., the prenatal care coefficient is zero) to be rejected at 10% and 5% levels of significance (power of the test). This shows how likely one is to find a statistically significant effect of prenatal care.

Table 5 reports the different values of the prenatal care coefficients and mixing probability chosen, as well as the results of the experiment. Because the results are quite similar for both the black and white samples, we report only those for the white sample for brevity; those for blacks are available upon request. Focusing first on the baseline case in which there is not a finite mixture (p=1) validates our experiment. In these cases, the prenatal care coefficients are consistently negative and our tests have relatively high power. For instance, in the experiment based on the sample of white mothers when the prenatal care coefficient is set to -0.4, the range of estimates is between -0.17 and -0.64 and one has a 90% probability of rejecting the null hypothesis of no effect using a 10% significance level. And, as expected, the power of the test falls as either the magnitude of the prenatal care coefficient decreases to -0.3 or the significance level decreases to 5%. This baseline case with its generally high power also validates the explanatory power of our instrument set for prenatal care. If, instead, the identifiers had poor explanatory power in the prenatal care equation, then predicted prenatal care would be highly collinear with the included regressors and lead to low power.

Of primary interest is what happens to the estimated prenatal care coefficient and its statistical significance as we allow a finite mixture distribution. As the probability of a 'normal' pregnancy (*p*) decreases towards 0.85**,** the range of prenatal care coefficients widens to include *positive* values, and the power of the test decreases to where we are more likely to fail to reject the hypothesis of no effect of prenatal care than to reject it . For instance, consider the case in which prenatal care has the strongest effects (-0.4 for normal pregnancies and -0.1 for complicated ones) and look at the effects of going from a nonmixture case (p=1) to a finite mixture with probability of a normal birth of 0.85. On the one hand, the median values of the point estimates are what we would expect – roughly a weighted average of the two regimes' coefficients, which with a small 'complicated' regime does not lead to a dramatic reduction. On the other hand, the range of estimates dramatically widens and the probability of obtaining a statistically significant coefficient plummets. The power of the test is reduced from 0.90 to approximately 0.40 if the 10% level of significance is used. In the case of more modest prenatal care effects (-0.3 for normal and zero for complicated) and a 5% level of significance, the probability of finding a statistically significant effect of prenatal care is 21% or less.

This experiment suggests that ignoring the presence of a finite mixture distribution in the estimation procedure can make prenatal care appear statistically unimportant the majority of the time – even if only 10-15% of the pregnancies follow a different ('complicated') distribution. Furthermore, by allowing us to sidestep some of the model specification choices made in the empirical analysis (e.g., endogeneity and identifying restrictions, data selection, two-step estimation) and finding similar results, it reinforces the conclusions drawn from it. And, thus, even though our empirical analysis suggests that only a small proportion of pregnancies are 'complicated' and enjoy little benefit from prenatal care, our Monte Carlo experiment reveals that ignoring their presence may indeed be responsible for the commonly found weak effect of prenatal care.

8. What About Reduced Form Models?

Many researchers have rejected structural models of infant health in favor of reduced form models in order to explore the impact of the Medicaid expansions or other policy variables, such as welfare reform and cigarette taxes, on birth outcomes. Perhaps due to the difficulties of identifying the structural birth weight equation and therefore the indirect (through prenatal care) effects of policy changes, many estimate only a reduced form birth weight equation, in which the policy variables enter the birth weight equation directly and prenatal care does not. The total effect of a policy on birth weight can then be measured by the coefficients on the policy variables. For example, Currie and Grogger (2000) estimate both structural and reduced form birth weight equations, as well as prenatal care utilitization equations, for black and white women. They find that increasing Medicaid eligibility increases prenatal care utilization for both whites and blacks. These increases translate into modest improvements in the incidence of very low birth weight for whites, but has no real effect for blacks.

More generally, the estimated effects of the Medicaid expansions on birth outcomes are considered modest at best (e.g., Gruber 1997 and Kaestner 1999). If, however, the effects of prenatal care are obscured by a failure to acknowledge the two types of pregnancies, then Medicaid and other policies may appear to be ineffective as well. We therefore present, as an illustration only, results from a reduced form birth weight equation that includes Medicaid eligibility.[12] This exercise also reveals how a finite mixture model has implications that extend beyond prenatal care to other factors thought to affect birth weight; we choose Medicaid eligibility rules because they are more likely exogenous to the mother than maternal behaviors, which are almost certainly endogenous and thus plagued with the same issues as prenatal care.

Specifically, we estimate a reduced form birth weight equation that is designed to capture the effects of Medicaid policy. This is in the spirit of Currie and Gruber (1996) and Currie and Grogger (2000). Briefly, both prenatal care and insurance status are considered endogenous and

[12] We stress that this is only an exercise to see if our argument has relevance for the Medicaid expansion literature. The many complications of more thoroughly examining this issue, such as the possible endogeneity of Medicaid generosity and looking at changes over time, are beyond the scope of this paper.

are substituted out of the birth weight equation, leaving birth weight to be a function of all of the (other) variables in the prenatal care equation plus Medicaid *eligibility.* The coefficient on Medicaid eligibility therefore gives the net effect of expanding eligibility on birth outcomes. *If* expanding eligibility increases Medicaid participation (take-up rates are typically far less than 100%), *if* Medicaid participation improves access to prenatal care ("crowding out" of private insurance may act to negate this, e.g. Cutler and Gruber 1996), and *if* improved prenatal care improves birth outcomes, then the coefficient should be positive. All three conditions are required to find an effect of Medicaid on birth outcomes. Therefore, if the effects of prenatal care are obscured by a failure to recognize the finite mixture of distributions of birth weights, then these researchers may be incorrectly concluding that Medicaid has no effect on birth outcomes.

Table 6 reports the results of estimating a reduced form birth weight equation with 2SLS and a finite mixture. Our Medicaid eligibility measure is the income level below which the household's income would have to fall in order to be eligible for Medicaid. It therefore varies across state and across household size. We consider these results illustrative only as we are using only cross-sectional data, which may be confounded by unobserved state influences. (For example, a state with a high rate of uninsured people may enact generous eligibility guidelines.) Once again, the probability of a 'normal' birth, the predicted distributions of birth weights and the pattern of variables affecting 'normal' versus 'complicated' pregnancies are very similar to the previous models.[13]

What about the effects of Medicaid eligibility? In general, we find that Medicaid eligibility improves the birth outcomes for whites and that this effect is strengthened by using the finite mixture model. In addition, the finite mixture model suggests that greater access to

---

[13]One exception to this relative stability is the coefficient on whether the father cohabits with the mother in the white samples. It is negative (and against our expectations) for 'complicated' pregnancies and very significant. A closer look reveals that it is likely due to a combination of a small number of troubled pregnancies in this smaller sample and the high collinearity between it and the never married (Unmarry) variable, a variable whose coefficient is also quite large and significant. We can also take comfort in the fact that the asymptotic standard errors are about six times bigger for both of these variables' coefficients, which renders them statistically insignificant. No other coefficient's standard error is affected nearly as much as these two by bootstrapping.

Medicaid only has a significant effect on the outcomes of 'normal' pregnancies, just as prenatal care did in the structural model. In contrast, Medicaid eligibility is never statistically significant for blacks and is in fact negative. These results are similar to the reduced form results produced by Currie and Grogger (2000), in which eligibility cut-offs are found to have modest effects for whites and no effects for blacks. Our results suggest that the positive effects of expanded eligibility are further limited to white, 'normal' pregnancies.

9. Concluding Remarks

The basic premise of our research is that the standard approach to modeling birth outcomes — treating all births as coming from the same distribution — misspecifies the true model and ignores potentially important information. We argue that there are two types of pregnancies, 'complicated' and 'normal' ones, such that a finite mixture model is more appropriate. Our data from the NMIHS provides compelling evidence supporting our view. The standard approach yields residuals that are obviously bimodal. Estimating birth weights with a finite mixture model yields estimates that are much more robust across weighting schemes and races and that clearly show that most factors, including prenatal care, primarily affect 'normal' pregnancies. This is consistent with the medical literature that finds preterm birth, a main cause of low birth weight, to be very difficult to prevent, and with researchers such as Paneth (1995) that suggest that birth outcomes may follow two different distributions. The standard approach which combines these two types of pregnanies may therefore cause prenatal care to appear ineffective because the 'complicated' pregnancies are likely to be influential outliers that are relatively unaffected by such factors.

Indeed, that is what we find — that with a finite mixture model, prenatal care has a consistent, substantial effect (30-35 grams for each week sought earlier) on normal pregnancies. Note that our results are robust to choice of samples and outliers. As an ultimate robustness check, we conduct a Monte Carlo experiment in which we generate data under the assumption of a finite mixture distribution and exogenous prenatal care (such that valid instruments are no longer a concern), which we then estimate in the usual (i.e., OLS) way. Our Monte Carlo experiment confirms that ignoring even a small proportion of 'complicated' pregnancies (10-15%) can lead the onset of prenatal care to appear unimportant in the standard model.

Our main finding about prenatal care in a finite mixture model therefore contains both good news and bad news; the good news is that the onset of prenatal care *does* seem to improve birth weights, but the bad news is that it does *not* help the most troubled pregnancies. However, our analysis in section 6 suggests that nearly half of low birth weight births are the result of 'normal' pregnancies and therefore *can* be affected by timely prenatal care. Furthermore, the compelling evidence of a finite mixture distribution for birth outcomes has implications beyond the effects of the onset of prenatal care. As Kogan et al (1994), Witwer (1990) and others have stressed, the content of prenatal care is just as important, if not more, than its onset. Likewise, maternal lifestyle and mental state are other factors that may improve birth outcomes (e.g., Chomitz, Cheung and Lieberman 1995 and Conway and Kennedy forthcoming). Our framework may prove useful in identifying which of these factors *can* improve the outcomes of 'complicated' pregnancies. It also has implications for the common use of reduced form models in estimating the effects of a specific policy such as the Medicaid expansions or cigarette taxes, as we illustrate using our data and Medicaid eligibility. Another extension of this research could be to consider other birth outcomes, such as fetal/infant death, APGAR scores or excessive hospitalization, although the discrete nature of these alternative outcomes poses econometric challenges to the finite mixture model which are well beyond the scope of this paper. More generally, our results suggest that research investigating birth outcomes, including the growing research on the effects of the Medicaid expansions, should acknowledge these two types of pregnancies in its empirical approach.

References

Alexander, G. R. and C. C. Korenbrot, “The Role of Prenatal Care in Preventing Low Birth Weight,” *Future of Children,* 1995, 5(1), pp. 103-120.

Binstock, M. A. and G. Wolde-Tsadik, “Alternative Prenatal Care – Impact of Reduced Visit Frequency, Focused Visits and Continuity of Care,” *Journal of Reproductive Medicine,* 1995, 40(7), pp. 507-12.

Chomitz, V. R., L. W. Y. Cheung and E. Lieberman, “The Role of Lifestyle in Preventing Low Birth Weight,” *Future of Children,* 1995, 5(1), pp. 121-138.

Collaborative Group on Preterm Birth Prevention, “Multicenter Randomized, Controlled Trial of a Preterm Birth Prevention Program,” *American Journal of Obstetrics and Gynecology,* 1993, 169(2, part 1), pp. 352-66.

Conway, K. S. and L. D. Kennedy, “Maternal Depression and the Production of Infant Health,” *Southern Economic Journal,* forthcoming.

Currie, J. and J. Grogger, “Medicaid Expansions and Welfare Contractions: Offsetting Effects on Prenatal Care and Infant Health?” NBER Working Paper 7667, April 2000.

Currie, J. and J. Gruber, “Saving Babies: The Efficacy and Cost of Recent Changes in the Medicaid Eligibility of Pregnant Women,” *Journal of Political Economy,* 1996, 104(6), pp. 1263-96.

Cutler, D. and J. Gruber, “Does Public Insurance Crowd Out Private Insurance?” *The Quarterly Journal of Economics,* 1996, 111(2), pp. 391-430.

Deb, P. and P. K. Trivedi, “Demand for Medical Care by the Elderly in the United States: A Finite Mixture Approach”, *Journal of Applied Econometrics*, 12, 1997, pp. 313-336.

Geweke, J. and M. P. Keane, “Mixture of Normals Probit Models”, Federal Reserve Bank of Minneapolis, Staff Report 237, 1997.

Goldenberg, R. L. and D. J. Rouse, “Prevention of Premature Birth,” *The New England Journal of Medicine,* 339(5), 1998, pp. 313-20.

Gray, B., “Are Higher Medicaid Fees for Prenatal Services Associated with Better Birth Outcomes?” *Journal of Health Economics,* 20, 2001, pp. 571-590.

Gritz, M. , “The Impact of Training on the Frequency and Duration of Employment”, *Journal of Econometrics*, 57, 1993, pp. 21-51.

Grossman, M. and T. Joyce, “Unobservables, Pregnancy Resolution and Birth Weight Production Functions in New York City,” *Journal of Political Economy,* 98(5), October 1990, pp. 983-1007.

Heckman, J. J, R. Robb, and M. Walker “Testing the Mixtures of Exponentials Hypothesis and Estimating the Mixing Distribution by the Method of Moments”, *Journal of the American Statistical Association*, 85, 1990, pp. 582-589.

Heckman, J. J. and B. Singer “A Method of Minimizing the Distributional Impact in Econometric Models for Duration Data”, *Econometrica*, 52, 1984, pp. 271-320.

Joyce, T., “Self-selection, Prenatal Care, and Birth weight Among Blacks, Whites and Hispanics in New York City,” *Journal of Human Resources*, 29(3), 1994, pp. 762-94.

Joyce, T., “Impact of Augmented Prenatal Care on Birth Outcomes of Medicaid Recipients in New York City,” *Journal of Health Economics,* 18, 1999, pp. 31-67.

Kaestner, R., “Health Insurance, the Quantity and Quality of Prenatal Care, and Infant Health,” *Inquiry*, 36, Summer 1999, pp. 162-75.

Kogan, M. D., G. R. Alexander, M. Kotelchuck, and D. A. Nagey, “Relation of the Content of Prenatal Care to the Risk of Low Birth Weight – Maternal Reports of Health Behavior Advice and Initial Prenatal Care Procedures,” *Journal of the American Medical Association,* May 4, 1994, 271(17), pp. 1340-45.

Laird, N. “Nonparametric Maximum Likelihood Estimation of a Mixing Distribution”, *Journal of the American Statistical Association*, 73, 1978, pp. 805-811.

Liu, G. G., “Birth Outcomes and the Effectiveness of Prenatal Care,” *Health Services Research,* 32(6), February 1998, pp. 805-23.

McLachlan, G. J., and D. Peel, *Finite Mixture Models*, New York: John Wiley, 2000.

Morduch, J. and H.S. Stern "Using Mixture Models to Detect Sex Bias in Health Outcomes in Bangladesh", *Journal of Econometrics*, 77, 1997, pp. 259-276.

Paneth, N. S., "The Problem of Low Birth Weight," *Future of Children,* 1995, 5(1), pp. 19-34.

Ray, W. A., E. F. Mitchel, and J. M. Piper, "Effect of Medicaid Expansions on Preterm Birth,"*American Journal of Preventive Medicine,* 1997, 13(4), pp. 292-97.

Rosenzweig, M. R. and T. P. Schultz, "The Behavior of Mothers as Inputs to Child Health: The Determinants of Birth Weight, Gestation and Rate of Fetal Growth," in *Economic Aspects of Health,* edited by V.R. Fuchs, University of Chicago Press, Chicago, IL, 1982.

Rosenzweig, M. R. and T. P. Schultz, "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and their Effects on Birth Weight," *Journal of Political Economy,* 91(5), 1983, pp.723-46.

Shiono, P. H. and R. E. Behrman, "Low Birth Weight: Analysis and Recommendations," *Future of Children,* 1995, 5(1), pp. 4-18.

Titterington, D. M., A. F. M. Smith and U. E. Makow, *Statistical Analysis of Finite Mixture Distributions*, New York: John Wiley, 1985.

Wang, P., I. M. Cockburn, and M. L. Puterman "Analysis of Patent Data - A Mixed Poisson Regression Model Approach", *Journal of Business and Economic Statistics*, 16, 1998, pp. 27-36.

Warner, G. L. "Prenatal Care Demand and Birth weight Production of Black Mothers," *American Economic Review*, 85(2), May 1995, pp. 132-37.

Warner, G. L. "Birth weight Productivity of Prenatal Care," *Southern Economic Journal,* 65(1), 1998, pp. 42-63.

Wedel, M., W.S. Desarbo, J.R. Bult, and V. Ramaswamy "A Latent Class Poisson Regression Model For Heterogeneous Count Data", *Journal of Applied Econometrics*, 8, 1993, pp. 397-411.

Witwer, M. B., "Prenatal Care in the United States: Reports Call for Improvements in Quality and Accessibility," *Family Planning Perspectives,* 1990, 22(1), pp. 31-35.

Table 1: Variable Definitions and Sample Means

		White Mothers		Black Mothers	
Variable	Definition	unweighted	weighted	unweighted	weighted
Birth weight	Birth weight of child in 100 grams	30.715	34.449	28.589	31.699
Prenatal care	Number of weeks into pregnancy that prenatal care began	8.375	8.432	9.844	9.887
Mother's age	Woman's age at delivery	27.482	27.358	25.640	25.548
Mother's education	Mother's years of education	13.316	13.365	12.640	12.658
Mother's birth weight	Mother's own birth weight in 100 grams	31.782	32.119	30.321	30.587
Mother's height	Mother's own height in inches	64.847	64.932	64.676	64.738
Child is male	Whether infant is male	0.518	0.527	0.512	0.513
Parity	Whether woman experienced a previous pregnancy	0.671	0.671	0.716	0.715
No. of fetal deaths	Number of prior fetal deaths (abortions and miscarriages)	0.369	0.333	0.461	0.427
No. of kids cohabitating	Number of a woman's own children living in the household	0.844	0.871	1.121	1.129
Mother is unmarried	Whether woman was never married at the time of delivery	0.073	0.070	0.478	0.487
Urban	Whether woman lived in an urban county	0.753	0.746	0.792	0.789
Medical care price[a]	State level weighted average of 4 prices: hospital room, general medical & dental visits, bottle of aspirin	101.932	101.963	99.546	99.317
Population density[b]	State level population density per square mile	0.221	0.215	0.332	0.323
Father cohabits	Whether woman lived with the child's father	0.910	0.916	0.526	0.522
Family income	Annual family income in 1000 dollars	32.693	33.158	18.061	18.095
Private insurance	Whether the woman had private insurance	0.712	0.720	0.367	0.366
Medicaid	Whether the woman had Medicaid	0.110	0.106	0.435	0.437
Medicaid eligibility[c]	Medicaid dollar income eligibility threshold in 1988 (in thousands)	8.233	8.310	9.062	9.077
Prenatal visits	Number of visits divided by number recommended by the American College of Obstetrics and Gynecology and adjusted for gestation	1.122	1.002	1.069	0.967

Notes:
a. The medical care price is based on the Inter-City Cost of Living Index produced by the American Chamber of Commerce Researchers Association, and includes 1) average cost per day for semi-private room in hospital, 2) average charge for a GP office visit, 3) charge for teeth cleaning and inspection, no x-ray or flouride treatment for a dentist visit, and 4) 100-tablet bottle of Bayer brand aspirin. Data are from the 3rd or 4th quarter of 1988 for 256 cities, which we aggregate up to the state level.

b. Population density is from the U.S. Statistical Abstract.

c. Medicaid eligibility levels are from Table 1 of the National Governors' Association, MCH Update, State Coverage of Pregnant Women and Children - January 1990. They refer to to a period of time extending from April 1987-January 1989.

Table 2: 2SLS Estimates

	White Mothers		Black Mothers	
Coefficient	unweighted	weighted	unweighted	weighted
Prenatal care	**-0.678***	**-0.132**	**-0.315**	**-0.328+**
	(0.242)	**(0.118)**	**(0.306)**	**(0.189)**
Mother's age	-0.167*	-0.040	-0.053	-0.028
	(0.046)	(0.024)	(0.051)	(0.034)
Mother's education	0.392*	0.273*	0.326*	0.151+
	(0.096)	(0.049)	(0.130)	(0.080)
Mother's birth weight	0.251*	(0.172*	0.167*	0.130*
	(0.033)	(0.019)	(0.032)	(0.021)
Mother's height	0.326*	0.283*	0.242*	0.228*
	(0.065)	(0.037)	(0.065)	(0.038)
Child is male	1.516*	1.288*	0.747+	0.962*
	(0.351)	(0.189)	(0.393)	(0.248)
Parity	0.617	0.440+	0.064	-0.113
	(0.459)	(0.247)	(0.473)	(0.318)
No. of fetal deaths	-1.245*	-0.046	-1.137*	-0.269
	(0.290)	(0.145)	(0.305)	(0.172)
No. of kids	1.758*	0.681*	0.504*	0.285*
cohabitating	(0.284)	(0.145)	(0.184)	(0.130)
Mother is unmarried	1.116	0.126	-0.294	-0.436
	(0.940)	(0.494)	(0.506)	(0.302)
Urban	-0.143	-0.019	0.093	-0.000
	(0.445)	(0.235)	(0.514)	(0.350)
Instrument validity				
F-stat	11.05	8.40	5.88	5.41
	[0.000]	[0.000]	[0.000]	[0.000]
Hansen	5.865	6.378	6.186	3.572
	[0.271]	[0.320]	[0.271]	[0.289]

Notes:
Standard errors of coefficient estimates are in parentheses
* indicates that the coefficient is statistically significant at the 5 percent level.
+ indicates that the coefficient is statistically significant at the 10 percent level.
F-stat is the test for the joint significance of the instruments in the first-stage regression.
Hansen is Hansen's J overidentification test for all instruments.
p-values of the tests of validity of instruments are in square brackets.

Table 3: 2-class Finite Mixture Estimates

	White Mothers				Black Mothers			
	unweighted		weighted		unweighted		weighted	
Coefficient	'normal'	'complicated'	'normal'	'complicated'	'normal'	'complicated'	'normal'	'complicated'
Prenatal care	**-0.335*** **(0.140)**	**0.316** **(0.227)**	**-0.308*** **(0.133)**	**0.377+** **(0.221)**	**-0.302*** **(0.147)**	**-0.170** **(0.221)**	**-0.352*** **(0.153)**	**-0.210** **(0.229)**
Mother's age	-0.103* (0.029)	-0.123* (0.056)	-0.110* (0.030)	-0.114* (0.056)	-0.018 (0.031)	-0.003 (0.064)	-0.026 (0.033)	-0.009 (0.065)
Mother's education	0.370* (0.061)	0.237 (0.161)	0.386* (0.061)	0.234 (0.158)	0.183* (0.072)	-0.108 (0.135)	0.170* (0.076)	-0.119 (0.138)
Mother's birth weight	0.223* (0.024)	0.074+ (0.045)	0.222* (0.023)	0.076+ (0.045)	0.134* (0.021)	0.114* (0.033)	0.137* (0.022)	0.116* (0.032)
Mother's height	0.334* (0.030)	0.161* (0.063)	0.338* (0.029)	0.151* (0.062)	0.244* (0.033)	-0.034 (0.055)	0.250* (0.035)	-0.026 (0.058)
Child is male	1.493* (0.233)	0.440 (0.583)	1.510* (0.235)	0.367 (0.588)	1.166* (0.270)	0.235 (0.477)	1.162* (0.276)	0.232 (0.479)
Parity	0.872* (0.338)	0.555 (0.612)	0.870* (0.340)	0.575 (0.614)	0.021 (0.369)	-0.112 (0.492)	-0.049 (0.374)	-0.151 (0.495)
No. of fetal deaths	-0.183 (0.196)	-0.752* (0.343)	-0.143 (0.191)	-0.777* (0.347)	-0.179 (0.176)	-0.517* (0.202)	-0.150 (0.178)	-0.503* (0.200)
No. of kids cohabitating	0.953* (0.180)	0.555 (0.476)	0.971* (0.184)	0.483 (0.484)	0.266+ (0.149)	0.112 (0.236)	0.280+ (0.151)	0.124 (0.238)
Mother is unmarried	0.510 (0.617)	-1.036+ (0.566)	0.453 (0.599)	-1.249+ (0.653)	-0.505 (0.342)	0.146 (0.523)	-0.527 (0.333)	0.137 (0.518)
Urban	0.022 (0.296)	0.798 (0.582)	0.037 (0.298)	0.848 (0.571)	0.165 (0.358)	0.201 (0.487)	0.059 (0.379)	0.140 (0.490)
σ	5.502* (0.186)	4.103* (0.590)	5.504* (0.192)	4.087* (0.620)	5.360* (0.111)	3.542* (0.160)	5.361* (0.111)	3.533* (0.160)
***p*(normal)**	**0.866*** **(0.013)**		**0.867*** **(0.016)**		**0.856*** **(0.008)**		**0.856*** **(0.008)**	
Fitted means	33.584 (0.195)	11.995 (0.892)	33.582 (0.401)	11.973 (1.026)	31.584 (0.138)	10.794 (0.272)	31.582 (0.355)	10.788 (0.304)

Notes:
* indicates that the coefficient is statistically significant at the 5 percent level.
+ indicates that the coefficient is statistically significant at the 10 percent level.

Table 4: Sample Characteristics by Gestation and Birth weight

Variable	Full-term, normal birth weight Mean	Pre-term, normal birth weight Mean	Full-term, Low birth weight Mean	Pre-term, Low birth weight Mean	Normal birth weight / 'normal' Mean	Low birth weight / 'normal' Mean	Low birth weight / 'compli-cated' Mean
White Mothers							
Classified 'normal'	1.00	1.00	0.81	0.32			
Preterm					0.03	0.54	0.91
Prenatal care onset	8.37	8.82	9.70	7.90	8.38	8.70	8.08
Prenatal visits	0.98	1.26	1.00	1.74	0.99	1.25	1.80
Education	13.43	12.89	12.54	13.17	13.41	12.76	13.21
# of observations	2145	66	178	518	2211	313	383
Black Mothers							
Classified 'normal'	1.00	1.00	0.90	0.30			
Preterm					0.05	0.47	0.95
Prenatal care onset	9.83	9.70	10.57	9.66	9.82	10.41	9.47
Prenatal visits	0.91	1.36	0.90	1.62	0.94	1.08	1.73
Education	12.74	12.59	12.30	12.44	12.73	12.34	12.45
# of observations	1601	90	168	453	1691	287	334

Table 5: Monte Carlo Evaluation of Linear Regression Estimates

			Sample weights not used					Sample weights used				
			percentiles of OLS estimated coefficient			power of test		percentiles of OLS estimated coefficient			power of test	
γ_1	γ_2	p	5^{th}	50^{th}	95^{th}	10%	5%	5^{th}	50^{th}	95^{th}	10%	5%
-0.40	0.00	1.00	-0.62	-0.40	-0.19	0.90	0.83	-0.64	-0.40	-0.17	0.91	0.85
-0.40	0.00	0.90	-0.73	-0.36	-0.01	0.50	0.37	-0.76	-0.36	0.02	0.54	0.43
-0.40	0.00	0.85	-0.74	-0.33	0.05	0.37	0.26	-0.78	-0.34	0.09	0.43	0.32
-0.40	-0.10	1.00	-0.62	-0.40	-0.17	0.90	0.82	-0.64	-0.40	-0.16	0.90	0.85
-0.40	-0.10	0.90	-0.74	-0.37	-0.01	0.49	0.37	-0.79	-0.36	0.02	0.53	0.42
-0.40	-0.10	0.85	-0.77	-0.34	0.06	0.39	0.27	-0.81	-0.35	0.07	0.43	0.33
-0.30	0.00	1.00	-0.53	-0.30	-0.07	0.71	0.60	-0.54	-0.30	-0.05	0.74	0.65
-0.30	0.00	0.90	-0.64	-0.28	0.09	0.33	0.22	-0.67	-0.27	0.13	0.37	0.28
-0.30	0.00	0.85	-0.67	-0.25	0.15	0.25	0.16	-0.72	-0.25	0.20	0.29	0.21
-0.30	-0.10	1.00	-0.52	-0.29	-0.08	0.69	0.57	-0.55	-0.29	-0.06	0.72	0.63
-0.30	-0.10	0.90	-0.66	-0.28	0.11	0.33	0.23	-0.70	-0.28	0.14	0.39	0.29
-0.30	-0.10	0.85	-0.71	-0.27	0.16	0.27	0.18	-0.76	-0.27	0.19	0.32	0.23

Note:
Based on 2000 replications of artificially generated data and a birth weight equation estimated with OLS.
The nuisance parameters and covariates are based on the observed sample of White mothers.
The data generating process is a two-class finite mixture of normal densities. γ_1 and γ_2 denote the parameters on prenatal care in each class and p is the probability of being in class 1.

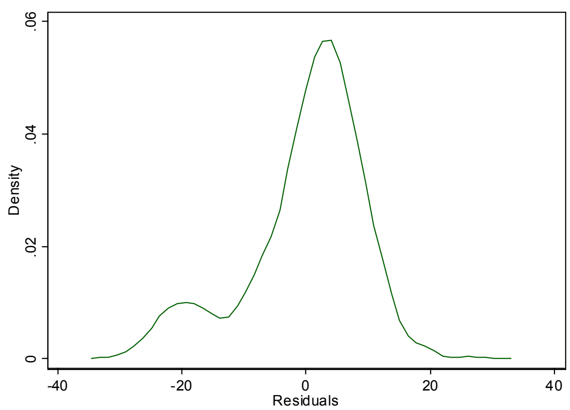
Table 6: Medicaid (Reduced form) Estimates

	OLS	OLS	Finite Mixture		Finite Mixture	
	White Mothers	Black Mothers	White Mothers		Black Mothers	
Coefficient			'normal'	'complicated'	'normal'	'complicated'
Mother's age	-0.029	-0.004	-0.093*	-0.140*	0.003	0.009
	(0.024)	(0.026)	(0.031)	(0.053)	(0.030)	(0.062)
Mother's education	0.284*	0.183*	0.390*	0.194	0.207*	-0.096
	(0.050)	(0.065)	(0.061)	(0.159)	(0.069)	(0.140)
Mother's birth weight	0.175*	0.123*	0.226*	0.065	0.129*	0.112*
	(0.018)	(0.020)	(0.023)	(0.046)	(0.021)	(0.032)
Mother's height	0.282*	0.229*	0.332*	0.036	0.250*	-0.021
	(0.036)	(0.034)	(0.022)	(0.048)	(0.023)	(0.039)
Child is male	1.268*	1.001*	1.450*	0.374	1.195*	0.190
	(0.189)	(0.231)	(0.260)	(0.444)	(0.272)	(0.431)
Parity	0.402	-0.062	0.807*	0.440	0.034	0.163
	(0.246)	(0.295)	(0.328)	(0.426)	(0.355)	(0.425)
No. of fetal deaths	-0.036	-0.257	-0.107	-0.837*	-0.144	-0.521*
	(0.143)	(0.168)	(0.189)	(0.324)	(0.169)	(0.199)
No. of children cohabiting	0.397*	0.289*	0.524*	0.635	0.310+	0.057
	(0.149)	(0.144)	(0.183)	(0.463)	(0.181)	(0.338)
Unmarried	-0.315	-0.427	-0.242	-1.701*	-0.482	-0.085
	(0.443)	(0.293)	(0.412)	(0.326)	(0.305)	(0.442)
Urban	0.064	0.207	0.156	0.547	0.203	0.505
	(0.227)	(0.315)	(0.305)	(0.423)	(0.361)	(0.406)
Medical care price	0.007	0.002	0.016+	-0.016	0.006	-0.033*
	(0.008)	(0.011)	(0.009)	(0.022)	(0.012)	(0.014)
Population density	-0.281	-0.194	-0.203	-0.018	-0.288+	-0.148
	(0.214)	(0.133)	(0.278)	(0.186)	(0.163)	(0.321)
Father cohabits	0.040	0.212	0.493	-2.075*	0.334	-0.216
	(0.449)	(0.284)	(0.482)	(0.409)	(0.300)	(0.399)
Family income	-0.001	0.014+	0.002	0.002	0.012	0.012
	(0.005)	(0.008)	(0.007)	(0.016)	(0.009)	(0.016)
Medicaid eligibility	**0.099***	**-0.027**	**0.136***	**0.083**	**-0.051**	**-0.008**
	(0.038)	**(0.040)**	**(0.044)**	**(0.095)**	**(0.042)**	**(0.091)**
σ			5.475	4.117*	5.361*	3.467*
			(0.178)	(0.563)	(0.112)	(0.166)
***p*(normal)**			**0.865***		**0.857***	
			(0.013)		**(0.008)**	
Fitted Means			33.600*	12.001*	31.575*	10.806*
			(0.398)	(0.905)	(0.356)	(0.313)

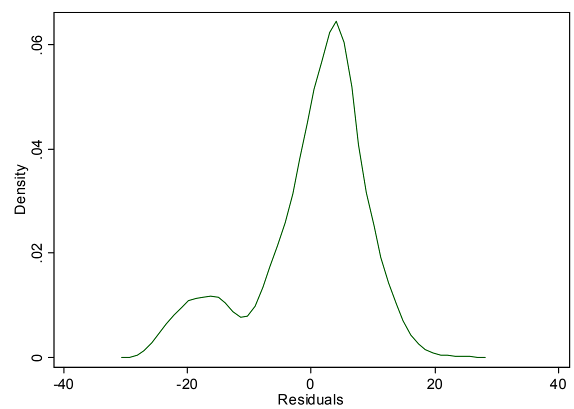
Notes:
Estimates from weighted models are reported.
* indicates that the coefficient is statistically significant at the 5 percent level.
+ indicates that the coefficient is statistically significant at the 10 percent level.

Figure 1: Unweighted and weighted kernel densities of 2SLS residuals from birth weight regressions

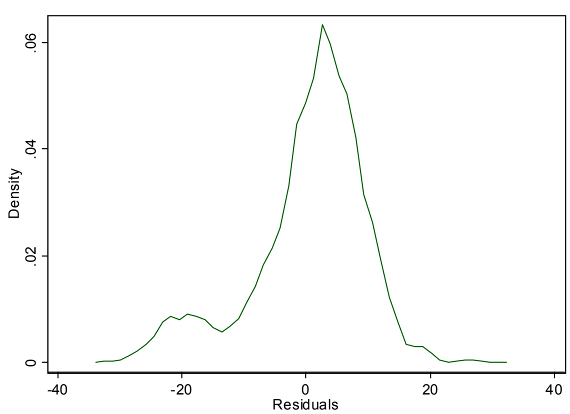
Unweighted kernel density of 2SLS residuals for Nonhispanic whites



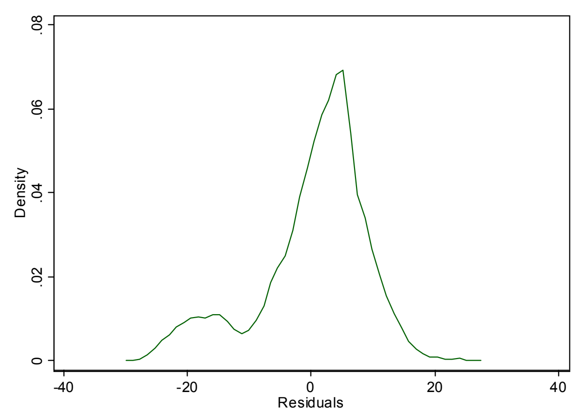
Unweighted kernel density of 2SLS residuals for Nonhispanic blacks



Weighted kernel density of 2SLS residuals for Nonhispanic whites



Weighted kernel density of 2SLS residuals for Nonhispanic blacks



Notes:
The data source is the National Maternal and Infant Health Survey (NMIHS) and the samples are stratified by race. Residuals are calculated from a birth weight equation estimated with 2SLS in which prenatal care is treated as endogenous.
Residuals are measured in 100's of grams.