Women's Health

Birth Outcomes and the Effectiveness of Prenatal Care

Gordon G. Liu

Objective. To investigate pregnant women's self-selection effects on the estimation of birthweight production function. A particular emphasis is placed on assessing the effectiveness of prenatal care as a major medical input in the birthweight production function.

Data Sources. Primary data compiled from birth and abortion certificates for the Commonwealth of Virginia in 1984. Several area-specific socioeconomic variables were also employed from the Area Resource File 1984; Supplemental Food Program for Women, Infants, and Children (WIC) Local Agency Directory; and the family planning clinics data by the Alan Guttmacher Institute (AGI).

Study Design. Two types of self-selection effects are defined: selection effect due to sample censoring from the resolution of pregnancies as live births or induced abortions; and selection effect due to the use of prenatal care as an endogenous variable. Race- and location-specific birthweight production functions are estimated using models with and without correction for self-selection effects.

Principal Findings. The self-selection effect in the resolution of pregnancies is race-specific, being significant for African American women. The effectiveness of prenatal care in birthweight production is underestimated substantially by the selection bias from the use of prenatal care, and overestimated by the selection bias from pregnancy resolutions. On average, the overall estimated effectiveness of prenatal care is over five times higher after controlling for the selection effects.

Conclusions. Self-selection effects could be a very serious problem in measuring the effectiveness of birthweight determinants in general. The overall effectiveness of prenatal care, in particular, tends to be significantly biased downward without controlling for selection effects. The significance and scale of the bias depends crucially on specific data and cohorts of the population investigated.

Key Words. Birth weight, effectiveness of prenatal care, pregnancy, self-selection effects

Prenatal care has long been endorsed as a major means to identify and, one would hope, to reduce the risks of pre-term, low-birthweight, and other adverse pregnancy conditions and birth outcomes. Counseling about diet, smoking cessation, drug avoidance, and diagnosis and timely treatment of complications are the major components of prenatal care (Alexander and

Korenbrot 1995). While there is a common understanding that prenatal care is beneficial in general, how it works and how cost effective it could be remains an unresolved critical issue facing the healthcare communities providing prenatal care.

For nearly four decades, the effectiveness of prenatal care in the prevention of low birth weight has been a subject of serious dispute (Gortmaker 1979; Harris 1982; Rosenzweig and Schultz 1983; Corman, Joyce, and Grossman 1987; Brown 1989; Mustard and Roos 1994; Piper, Mitchel, and Ray 1994). Women's self-selection behavior in seeking prenatal care can be a major source of this dispute because demand for prenatal care is likely to be correlated with or indicative of pregnant women's behavior or health status. As a result, it is very difficult to determine the extent to which the observed difference in birth weights can be attributed legitimately and rigorously to the effectiveness of prenatal care. For example, if women who receive more prenatal care are also healthier and have better health behavior (i.e., favorable selection), the effectiveness of prenatal care may be overestimated. Conversely, if women receive more prenatal care because of poorer health conditions during pregnancy (i.e., adverse selection), then the effectiveness of prenatal care may be underestimated. As a result, finding ways to correct for or to minimize such an estimation bias in assessing the effectiveness of prenatal care has been a real challenge to healthcare researchers and providers.

Ideally, an experimental study could largely avoid such an estimation bias. However, ethical consideration has been a major barrier to carrying out an experimental study that must make random assignments of some pregnant women to a control group receiving no prenatal care. Alternatively, researchers have had to focus their attention on studies using nonexperimental data. In a recent study, Grossman and Joyce (1990) provide an important addition to the literature dealing with the self-selection issues in assessing the effect of prenatal care on birthweight production (e.g., Harris 1982; Rosenzweig and Schultz 1983; Corman, Joyce, and Grossman 1987). In their study, Grossman and Joyce present the first infant health production functions that simultaneously control both for self-selection biases due to pregnancy resolution[1] and due to the endogenous use of prenatal care. Using data from the vital statistics for New York City 1984, Grossman and Joyce find the self-

---

Address correspondence and requests for reprints to Gordon G. Liu, Ph.D., Assistant Professor, Department of Pharmaceutical Economics and Policy, University of Southern California, 1540 E. Alcazar Street, Room 140G, Los Angeles, CA 90033. This article, submitted to *Health Services Research* on November 14, 1996, was revised and accepted for publication on May 12, 1997.

selection bias to be significant for African American women, but not for white women. Their study also suggests that the cost of contraception appears to be predominant for African American women.

Following Grossman and Joyce's approach, this study presents a further empirical investigation of the self-selection bias to the relationship of prenatal care and birthweight production functions using 1984 vital statistics of all births and induced abortions for the Commonwealth of Virginia. To demonstrate the self-selection bias in assessing the effectiveness of prenatal care, the study presents results both with and without correction for such a self-selection bias. Moreover, the study also examines how the infant birthweight production technology may differ by race and residential location in Virginia. Thus, four empirical equations are estimated respectively for urban whites, rural whites, urban African Americans, and rural African Americans.

A MODEL CONTROLLING FOR SELF-SELECTION EFFECTS

Estimation of birthweight production function using data for women giving birth may involve two types of self-selection: one from pregnancy resolutions and the other from the use of prenatal care. Put in the context of household utility function (Becker and Lewis 1973), pregnancy resolutions may be treated as outcomes of decision making by pregnant women according to their satisfaction associated with two options: giving birth or having an abortion. Suppose the two option-specific satisfaction can be valued by a utility function of the form: U_{bi} and U_{ai}, where the former represents the utility associated with giving birth, and the latter is the utility associated with having an abortion. Following the notion of random utility function (Manski 1977; Nakosteen and Zimmer 1980; and Ben-Akiva and Lerman 1989), rational individuals facing competing options are assumed to choose the one with the highest utility. Thus, the decision rule in the case of pregnancy resolutions is that women will choose to give birth if the birth utility is larger than the abortion utility: $U_{bi} \geq U_{ai}$; or otherwise to obtain an induced abortion (Liu 1995). That is, women will give birth only if the difference between the two utilities is greater than 0 (i.e., $\Delta U_i = U_{bi} - U_{ai} \geq 0$).

Having chosen to give birth, parents would then determine the way of taking care of their fetus in order to obtain the best birth outcome. More formally, a birthweight production function is specified to capture the relationship between birthweight outcome h_i and various inputs including prenatal care y_i and a vector of parental behavior and characteristics X_{hi}. The underlying choice of giving birth and birthweight production functions may be written as follows:

$$\begin{aligned} \Delta U_i &= B^u X_{u_i} + \varepsilon_{u_i} \geq 0 \\ h_i &= B^h X_{hi} + B_y^h y_i + \varepsilon_{hi} \end{aligned} \tag{1}$$

where X_{ui} is a vector of observable variables of the birth probability function; B^u, B^h, and B_y^h are the parameters to be estimated; ε_{ui} and ε_{hi} are the two stochastic terms in each equation. Vector X_{ui} may contain some different variables than vector X_{hi}.

The decision making in pregnancy resolutions is highly characterized by women's preference. As a result those giving birth are a self-selected (censored) sample from the population of all pregnant women. Women with better health status, for instance, may be more likely to select into the birth sample than those with poor health status. Put differently, the sample of women giving birth may represent a special cohort of pregnant women whose health behavior may substantially differ from that of women obtaining an abortion. As a result, it is not ensured that the estimated birthweight production function using the self-selected sample is an unbiased representative of the total population.[2]

Another major source of self-selection bias is from the endogeneity of prenatal care in birthweight production function. As noted elsewhere, parental demand for prenatal care could increase because of either healthier parental behavior or poorer health status. With everything else constant, birth outcomes would be contributed to positively by healthier behavior, and negatively by poorer health status. This type of estimation bias, due to the use of prenatal care being endogenous, has been well documented in previous birth outcomes studies (e.g., Rosenzweig and Schultz 1983).

Clearly, to obtain unbiased and consistent estimates of birthweight production function, one must simultaneously control for the two types of self-selection effects discussed above. In order to do so, we first specify a probability function of giving birth, assuming women in the birth sample have had positive birth utility differentials (i.e., $\Delta U_i = U_{bi} - U_{ai} \geq 0$), while women in the abortion sample have had negative birth utility differentials. In the second stage, the birthweight production function is controlled for the sample selection bias by inserting an estimated Mill's ratio ($\hat{\lambda}_{ui}$) as a regressor, obtained from the birth probability function (Heckman 1976; Grossman and Joyce 1990). Meanwhile, the endogeneity effect of the demand for prenatal care is also controlled for by instrument variables approach, $\hat{y}_i$. As a result, the birthweight production function correcting for the self-selection effects is modeled as follows:

$$h_i = B^h X_{hi} + B_y^h \hat{y}_i + \left(\frac{\sigma_{uh}}{\sigma_u}\right) \hat{\lambda}_{ui} + \nu_{hi} \tag{2}$$

where $\hat{\lambda}_{ui} = \frac{f(B^u X_{ui}/\sigma_{ui})}{F(B^u X_{ui}/\sigma_{ui})}$; σ_{uh} is the covariance of ε_{ui} and ε_{hi}; and disturbance ν_{hi} has zero mean for the sample of women giving birth: $E(\nu_{hi}|\varepsilon_{ui} \geq -B^u X_{ui}) = 0$. With normal assumptions on the distributions of the error terms, the estimators of B^h and B_y^h will be unbiased and consistent (Lee, Maddala, and Trost 1980; Newey, Powell, and Walker 1990). The coefficient on $\hat{\lambda}_{ui}$ is self-defined as a ratio of the covariance σ_{uh} to the standard deviation σ_u.

In the following empirical work, birth weight is specified as a function of the demand for prenatal care and other social and demographic factors including parity dead (defined as children born alive but now dead), parity alive (defined as children born alive and now alive), late termination (defined as total previous abortions after 20 weeks), child sex (dummy), mother's educational level (dummy), mother's age (dummy), and the estimated Mills Ratio $\hat{\lambda}_i$. The demand for prenatal care is treated as an endogenous variable and is measured by the number of months a pregnant woman delays seeking prenatal care.[3] Women who received no prenatal care are treated as having ten months delay. Both birth probability and prenatal care demand equations are treated as reduced forms.

The structural relationship between biological and behavioral inputs and birth outcomes serves as the primary basis for the specification of each model. In order to achieve model identification, certain exclusions are applied to each equation. In particular, we define early termination as previous abortions committed before 20 weeks, and late termination as previous abortions that occurred after 20 weeks. Since there were no direct data on induced or spontaneous abortions, respectively, it was assumed that early termination is more associated with induced abortions, and late termination is more associated with spontaneous abortions. While early termination appears more relevant to the choice of giving birth in that women with a larger number of previous induced abortions tend to use induced abortion again as one of their contraceptive methods (Tietz 1978; Cates 1984), the medical literature contains little evidence that previous early abortions lead to subsequent poor outcomes. However, there is a lot more evidence that late spontaneous abortions are associated with subsequent poor birth outcomes (e.g., Institute of Medicine 1985). This suggests that late miscarriages are more indicative of mothers' poor reproductive health. Thus, in our model specification early termination is treated as a determinant of the birth probability function, while late termination is included only in the birthweight function to capture the associated variations in the mother's health status.

As defined earlier, parity dead and parity alive tend to capture some variations in both health endowment, parental fertility, and demand behavior. However, parity variables are included only in the birthweight equation. This is because the dependent variable of the birth probability equation (giving birth or not) is mathematically a direct function of parity: a woman would give birth if her desired number of children is greater than or equal to parity alive plus one (Grossman and Joyce 1990).

DATA AND EMPIRICAL FINDINGS

Data for the empirical work were compiled from all birth and induced abortion certificates for the Commonwealth of Virginia in 1984. In 1984, there were 84,747 live births and 32,606 induced abortions. Since it is well known that teenage fertility behavior and reproductive capacity may be considerably different from those of adults,[4] this study is restricted to adults age 20 or above. Excluding non-Virginia residents and missing values, the final data set for the analysis includes 20,819 African Americans with 14,582 births and 6,237 abortions. Since the total population sample for whites was much larger than for African Americans, a 30 percent random subsample was drawn for 20,059 whites, being comparable with the total African American sample size. The white sample contains 16,091 births and 3,968 abortions.

In order to examine some area-wide effects on household fertility behavior and birth outcomes, the data set also incorporated several county-specific socioeconomic variables, including race-specific poverty from the 1984 Area Resource File; Supplemental Food Program for Women, Infants, and Children (WIC) from the WIC Local Agency Directory; the number of abortion providers in 1984; and the number of family planning clinics in 1983 as reported by the Alan Guttmacher Institute (AGI). The number of abortion providers and clinics as well as WIC centers are measured at county level per 10,000 reproductive-age women (ages 15–44).[5]

An interesting research and policy issue raised in the study is whether and how women may differ by race and residence with respect to the relationship between the considered explanatory variables as determinants and the birth outcomes. To test the hypothesis, a Chow test was conducted (Chow 1960). The null hypothesis of no difference in slopes coefficients across four groups was rejected at the 1 percent level. As a result, four race- and residence-specific equations are defined.

In order to assess the self-selection bias from different procedures, each cohort-specific birthweight production function is estimated respectively using ordinary least squares (OLS), two-stage least squares (2SLS), and two-stage least squares controlling for selection (2SLS–select) methods. The definitions of variables and descriptive statistics are given in Tables 1 and 2. Estimates of birth selection probability functions and birthweight production functions are presented in Tables 3, 4, 5, 6, and 7.

With respect to the selection effect in the resolution of pregnancies, we find this effect to be statistically significant for African American women as shown in Tables 6 and 7, similar to that of Grossman and Joyce (1990). However, the coefficients of $\hat{\lambda}_{ui}$ are negative for all cohorts, suggesting a predominant effect of abortion cost on women's choice of giving birth in Virginia. This finding offers a policy implication that the cost of abortion could be a significant barrier in preventing women from terminating unwanted

Table 1: Definition of Variables

Name of Variables	*Descriptions*
Birth weight	Infant birth weight in grams
Prenatal care delay	Number of months delayed before seeking prenatal care
Early termination	Number of previous abortions before the 20th week of gestation
Late termination	Number of previous abortions after the 20th week of gestation
Parity alive	Number of previous living births who are still alive
Parity dead	Number of previous living births who have died
WIC centers	WIC centers per 10,000 females age 15–44 by county
Abortion providers	Abortion providers per 10,000 females age 15–44 by county
Family planning clinics	Family planning clinics per 10,000 females age 15–44 by county
Below poverty rate (%)	Race-Specific percentage of population below the poverty line in 1979
Male	Dummy variable that equals 1 if the infant is male
Married	Dummy variable that equals 1 if the woman is married
Mother age 35–39	Dummy variable that equals 1 if the woman is 35–39 years of age
Mother age ≥ 40	Dummy variable that equals 1 if the woman is 40 years or older
Mother education 10–12	Dummy variable that equals 1 if the woman has 10–12 years of education
Mother college	Dummy variable that equals 1 if the woman has college education
λ	The inverse of Mill's ratio

Table 2: Means and Standard Deviations of Key Variables for Women Giving Birth

	Whites		*African Americans*	
Name of Variables	*Urban (s.d.)*	*Rural (s.d.)*	*Urban (s.d.)*	*Rural (s.d.)*
Birth weight	3451.15	3404.47	3163.87	3151.07
(grams)	(573.77)	(585.06)	(637.93)	(645.20)
Prenatal care delay	2.29	2.55	2.79	3.06
(months)	(1.23)	(1.32)	(1.52)	(1.60)
Male	0.52	0.52	0.51	0.51
	(0.50)	(0.50)	(0.40)	(0.50)
WIC centers	0.25	0.28	0.28	0.36
	(0.31)	(0.77)	(0.27)	(0.91)
Parity alive	0.82	0.89	1.12	1.27
	(0.97)	(1.02)	(1.19)	(1.29)
Parity dead	0.02	0.02	0.03	0.04
	(0.16)	(0.15)	(0.20)	(0.22)
Abortion providers	0.35	0.08	0.49	0.13
	(0.31)	(0.20)	(0.38)	(0.24)
Family planning clinics	0.75	1.99	0.92	2.12
	(0.65)	(1.93)	(0.77)	(2.10)
Below poverty rate (%)	6.11	11.53	21.72	30.06
	(2.68)	(3.24)	(5.15)	(5.57)
Late termination	0.02	0.02	0.03	0.03
	(0.14)	(0.18)	(0.20)	(0.22)
Early termination	0.39	0.29	0.48	0.33
	(0.73)	(0.64)	(0.80)	(0.68)
Mother education 10–12	0.42	0.58	0.61	0.65
	(0.49)	(0.49)	(0.49)	(0.48)
Mother college (> 12)	0.55	0.34	0.34	0.24
	(0.50)	(0.49)	(0.47)	(0.43)
Mother age 35–39	0.08	0.05	0.05	0.01
	(0.27)	(0.23)	(0.21)	(0.20)
Mother age ≥ 40	0.01	0.01	0.01	0.01
	(0.09)	(0.08)	(0.08)	(0.09)
Married	0.94	0.92	0.54	0.49
	(0.24)	(0.27)	(0.50)	(0.50)
Observations	10,878	5,213	9,059	5,523

fetuses. In contrast, Grossman and Joyce's result shows a positive selection effect, emphasizing the significance of contraceptive cost for New York City residents. The identified difference between New York City and Virginia seems to be quite consistent with some real observations.[6]

Concerning the endogeneity of prenatal care, a Wu-Hausman test (Wu 1973; Hausman 1978) was conducted. The null hypothesis of no endogeneity

Table 3: Probit Probability Function of Giving Birth

Variables	*Urban Whites*	*Urban African Americans*	*Rural Whites*	*Rural African Americans*
Intercept	2.206	2.032	2.231	2.127
	(21.396)	(23.729)	(15.922)	(16.682)
Abortion providers	−0.124	−0.039	−0.109	−0.109
	(−2.396)	(−1.090)	(−0.886)	(−1.541)
Family planning clinics	0.064	0.075	0.031	0.007
	(2.124)	(4.056)	(2.333)	(0.798)
Below poverty rate (%)	−0.009	−0.003	−0.006	0.002
(Race-specific)	(−1.164)	(−1.164)	(−0.731)	(0.669)
Early termination	−0.118	−0.172	−0.194	−0.279
	(−6.353)	(−12.695)	(−6.438)	(−12.949)
Mother education 10–12	−0.529	−0.633	−0.566	−0.680
	(−5.778)	(−9.999)	(−5.427)	(−9.119)
Mother college	−0.552	−1.040	−0.867	−1.113
	(−6.020)	(−16.026)	(−8.187)	(−14.151)
Mother age 35–39	−0.213	−0.220	−0.064	−0.265
	(−3.882)	(−4.027)	(−0.607)	(−3.132)
Mother age ≥ 40	−0.781	−0.689	−1.172	−0.658
	(−6.484)	(−5.636)	(−6.543)	(−3.896)
Married	2.310	1.079	1.940	0.902
	(70.300)	(40.849)	(38.938)	(23.910)
Likelihood ratio (χ^2)*	6321.3	2374.4	1917.0	1011.6
Prediction ratio	90.3%	73.3%	89.6%	77.5%
Observations	13,863	13,528	6,196	7,291

* The critical $\chi^2(9)$ at the 0.5 percent level is 23.6.

of prenatal care is rejected at the 5 percent significance level across all cohorts, suggested by OLS estimates. To control for the endogeneity of prenatal care, a set of exogenous instruments are employed to estimate prenatal care as a reduced form. In principle, as exogenous variables these instruments should not be determined primarily by the birth production function, but should be correlated significantly as a set with prenatal care. Based on the F-statistics of the identifying instruments in the first-stage estimation,[7] the instruments for prenatal care are considered to be highly significant as a set (Bound, Jaeger, and Regina 1995; Staiger and Stock 1994).

The estimated effectiveness of prenatal care varies substantially depending on whether and how the selection bias is corrected for, although a positive effect on birth weight of prenatal care is shown consistently for all estimates. Comparison of the coefficients of prenatal care across three estimation procedures provides strong evidence that the effectiveness of prenatal care may be

Table 4: Birthweight Production Functions, Urban Whites

Variables	*OLS*	*2SLS*	*2SLS-Select*
Intercept	3244.67	3604.03	3438.23
	(90.001)	(45.983)	(27.965)
Parity dead	−115.538	−127.199	−113.445
	(−3.381)	(−3.589)	(−3.295)
Parity alive	50.664	67.074	57.9815
	(8.631)	(9.801)	(6.830)
Late termination	−127.042	−116.175	−121.667
	(−3.261)	(−2.877)	(−3.087)
Male	128.213	128.692	129.191
	(11.789)	(11.431)	(11.785)
Mother education 10–12	98.978	36.357	65.926
	(2.959)	(0.992)	(1.665)
Mother college	163.316	60.805	105.754
	(4.882)	(1.527)	(2.264)
Mother age 35–39	11.955	−3.839	6.923
	(0.584)	(−0.179)	(0.318)
Mother age $\geq$ 40	55.523	62.616	75.030
	(0.939)	(1.023)	((1.249)
Prenatal care delay	−13.055	−139.292	−75.071
	(−2.903)	(−5.650)	(−1.685)
Inverse of Mill's ratio	–	–	−61.430
			(−1.713)
F-statistic	32.935	–	11.451
Wu-test*	29.2	–	–
R^2	0.027	–	0.010
Sample mean	3451.15	3451.15	3451.15
Observations	10878	10878	10878

* The critical $F(1,\infty)$ at the 5 percent level is 3.84.

underestimated substantially due to the endogeneity of prenatal care. To be specific, across all cohorts the 2SLS estimates suggest that per month prenatal care delay results in an average loss of 160 grams in birth weight, while the OLS method would otherwise report only a moderate loss by 14 grams per month delay. Furthermore, an upward bias is identified for the effectiveness of prenatal care due to self-selection in pregnancy resolutions. Compared to the average 2SLS estimate of −160 grams per month delay, the 2SLS–select estimate suggests about −76 grams per month delay. Overall, the total effect of prenatal care may be underestimated using the OLS method by as much as five times, as compared to the 2SLS–select approach. This observation appears to be more promising for rural residents than for urban residents.

Two major risk factors are identified in the determination of current birth outcomes. Parity dead shows a very large and statistically significant adverse

Table 5: Birthweight Production Functions, Rural Whites

Variables	*OLS*	*2SLS*	*2SLS-Select*
Intercept	3153.11	3559.40	3458.81
	(84.073)	(42.674)	(30.793)
Parity dead	−153.750	−175.780	−159.563
	(−2.818)	(−3.078)	(−2.855)
Parity alive	38.246	64.002	57.376
	(4.512)	(6.396)	(5.128)
Late termination	21.778	30.023	32.221
	(0.493)	(0.650)	(0.711)
Male	147.455	136.239	147.986
	(9.213)	(8.094)	(9.036)
Mother education 10–12	144.058	81.112	99.440
	(4.649)	(2.365)	(2.709)
Mother college	224.732	114.533	146.784
	(6.854)	(2.890)	(3.130)
Mother age 35–39	−11.445	−42.465	−34.302
	(−0.312)	(−1.099)	(−0.894)
Mother age $\geq$ 40	−88.025	−123.351	85.111
	(−0.865)	(−1.159)	(−0.780)
Prenatal care delay	−6.265	−142.552	−107.413
	(−1.006)	(−5.580)	(−2.784)
Inverse of Mill's ratio	–	–	−64.955
			(−1.192)
F-Test	17.837	–	–
Wu-test*	33.37	–	–
R^2	0.030	–	–
Sample mean	3404.47	3404.47	3404.47
Observations	5213	5231	5231

*The critical $F(1,\infty)$ at the 5 percent level is 3.84.

effect on birth weight for all groups. Following the 2SLS–select estimates, the cohort-specific marginal effect of parity dead ranges from −96 to −185 grams per additional birth. In addition, the effect of parity dead appears much stronger for rural residents than for urban residents, particularly among African Americans. Another major risk factor is late termination being a negative correlate of birth weight. The results are statistically significant for the urban population. According to the 2SLS–select model, the marginal effect per late termination is about −170 grams for urban African Americans and −122 grams for urban whites.

Parity alive is found to be a positive correlate of birth weight. Based on 2SLS–select estimates, the effect of parity alive ranges from 23 to 67 grams of gains in birth weight for each additional parity alive. This is true perhaps because parity alive may capture some variation in parental reproductive

Table 6: Birthweight Production Functions, Urban African Americans

Variables	*OLS*	*2SLS*	*2SLS-Select*
Intercept	2992.23	3523.25	3115.37
	(89.119)	(40.034)	(27.394)
Parity dead	−100.604	−96.063	−96.221
	(−2.970)	(−2.604)	(−2.850)
Parity alive	25.463	49.876	26.570
	(4.224)	(6.629)	(3.207)
Late termination	−166.472	−182.331	−170.283
	(−4.917)	(−4.936)	(−5.020)
Male	134.805	133.333	134.783
	(10.164)	(9.232)	(10.192)
Mother education 10–12	100.144	78.627	131.528
	(3.433)	(2.463)	(4.218)
Mother college	184.969	87.091	216.152
	(6.003)	(2.376)	(5.121)
Mother age 35–39	23.791	−13.891	29.702
	(0.743)	(−0.393)	(0.888)
Mother age $\geq$ 40	44.837	44.974	100.801
	(0.542)	(0.499)	(1.219)
Prenatal care delay	−14.773	−197.146	−35.783
	(−3.330)	(−7.062)	(−0.877)
Inverse of Mill's ratio	–	–	−207.348
	–	–	(−5.087)
F-statistic	24.097	–	26.091
Wu-test*	52.46	–	–
R^2	0.023	–	0.028
Sample mean	3163.88	3163.88	3163.88
Observations	9059	9059	9059

*The critical $F(1,\infty)$ at the 5 percent level is 3.84.

capacity and health status; it may also proxy the mother's experience with pregnancy and birth. Moreover, in ways similar to the pattern of prenatal care effect, the marginal effect of parity alive also tends to be underestimated by the OLS method or overestimated by the 2SLS method, as compared to the 2SLS–select method.

The results also provide additional insight for the role of education. While education has generally been considered an effective means to better birth outcomes, the way in which education works is far from clear. For instance, Rosenzweig and Schultz (1983) argue that parental education affects the choice of health inputs but has no direct effect on birth weight. In this study, infant birth weight shows a strong direct relationship with the mother's

Table 7: Birthweight Production Functions, Rural African Americans

Variables	*OLS*	*2SLS*	*2SLS-Select*
Intercept	3058.12	3478.67	3283.31
	(84.255)	(35.758)	(28.007)
Parity dead	−181.877	−214.743	−185.227
	(−4.513)	(−4.983)	(−4.541)
Parity alive	29.640	50.892	37.837
	(4.055)	(5.693)	(3.891)
Late termination	−36.968	−21.960	−31.730
	(−0.957)	(−0.537)	(−0.809)
Male	127.643	131.909	127.811
	(7.416)	(7.253)	(7.352)
Mother education 10–12	55.703	42.688	77.417
	(1.896)	(1.372)	(2.393)
Mother college	126.409	63.052	141.905
	(3.781)	(1.670)	(3.074)
Mother age 35–39	−40.797	−93.873	−53.350
	(−0.902)	(−1.914)	(−1.082)
Mother age $\geq$ 40	−1.646	−38.764	13.232
	(−0.017)	(−0.376)	(0.132)
Prenatal care delay	−21.994	−160.262	−85.572
	(−4.015)	(−5.348)	(−2.163)
Inverse of Mill's ratio	–	–	−163.711
	–	–	(−2.689)
F-statistic	13.187	–	–
Wu-test*	24.73	–	–
R^2	0.021	–	–
Sample mean	3151.07	3151.07	3151.07
Observations	5523	5523	5523

* The critical $F(1,\infty)$ at the 5 percent level is 3.84.

educational level. The mother's education is measured by three dummy variables in terms of schooling years: greater than 12 (college), 10–12 (high school), and less than 10. Moreover, not only is the educational effect found to be positive, but the marginal effect of parental education appears stronger at the college level than at the high school level. For example, over all four cohorts the 2SLS–select estimates show that an infant born to a mother with a college education and 10–12 years of schooling would gain an average of about 153 and 94 grams, respectively, more than an infant born to a mother with less than 10–12 years of schooling. This finding is robust regardless of the estimation method and the population cohort.

Another notable result is that maternal age has no significant impact on the infant birthweight outcome. Although women over age 35 are believed to be at a higher risk for pregnancy complications, their children are not significantly lighter than those of women ages 20–35 in this study. This finding is also robust across all cohorts and estimation procedures. A similar result for New York City residents is also reported by Grossman and Joyce (1990). Finally, as expected, due to genetic or biological reasons infant birth weight is shown to be significantly greater for male children than for female children. The gender difference in birth weight is highly significant for all groups, ranging from 128 to 148 grams on average.

DISCUSSION AND SUMMARY

This study provides strong evidence on whether and how self-selection effects may bias the estimates of birthweight production functions. The empirical results are obtained using the 1984 vital statistics of all induced abortions and living births in the Commonwealth of Virginia. Two types of self-selection effects are identified: the selection from the pool of pregnant women into the censored sample of women giving birth (sample selection effect) and the selection from women's endogenous demand for prenatal care (endogeneity effect). Birthweight production function is estimated by OLS, 2SLS, and 2SLS–select procedures across four cohorts of pregnant women: urban whites, rural whites, urban African Americans and rural African Americans.

Major findings of this study are these: First, the self-selection effect from pregnancy resolutions is negative, implying that the unobserved factors that increase the probability of choosing to give birth tend to reduce birth weight. This indicates that in the Commonwealth of Virginia the cost of obtaining an induced abortion may be predominant compared to other unobserved factors such as contraception cost and health status. Based on this finding, in Virginia more policy efforts may be directed to the means reducing the multidimensional costs of abortion (e.g., economic and psychic costs), in order to decrease the incidence of women having unwanted births, and thus to improve overall birthweight outcomes.

Second, with respect to the estimation bias, strong evidence shows that both sample selection effect and endogeneity effect could produce a significant estimation bias in assessing the effectiveness of prenatal care. In particular, the Virginia data suggest that the marginal effectiveness of prenatal care tends to be underestimated by the endogeneity bias of prenatal care, and overestimated by the sample selection bias. Furthermore, the sample selection bias is shown to be race-specific, suggesting that women's fertility attitudes, health behavior, or health status may be more heterogeneous among African Americans than among whites.

Third, previous number of late abortions and parity dead are two major risk factors associated with lower birth weight. As a policy implication, the pregnant women with a large number of previous parity dead or spontaneous abortions are a riskier cohort; therefore, they should be closely monitored and targeted by whatever preventive efforts to reduce low-birthweight rates. As a modifiable policy variable, women's higher education status remains a strong determinant of birth outcomes regardless of race and residence. Another positive correlate of birth outcomes is parity alive.

Fourth, in contrast to conventional wisdom, the study finds no evidence that women older than 35 would be a riskier cohort compared to younger cohorts in terms of birth weight, holding everything else constant. Primarily for genetic or biological reasons, as expected, gender plays a great role in the determination of birth weight in that male infants are found to be substantially heavier than female infants at birth.

Finally, certain limitations and remaining concerns of the study deserve attention. First, the study is based on a decade-old data set drawn from 1984 vital statistics. Over that time some conditions such as the cost and availability of abortion may have changed, and such changes may have influenced women's decisions about giving birth. The major components of prenatal care and their associations with birthweight outcomes, however, have probably not changed significantly since 1984. Since the focus of the study is to demonstrate how the birthweight production function and the effectiveness of prenatal care can be misestimated due to the selection bias, the age of the data is not so crucial for the analysis. Moreover, the 1984 data make the results comparable to those of Grossman and Joyce (1990).

Another limitation is the extent to which the birth choice and birth-weight production functions could be identified more suitably using alternative exclusions of the exogenous variables. As presented, the current model is identified by the exclusion of a less relevant exogenous variable from each equation, a strategy imposed by the data. The model could have been identified on firmer grounds had the data set contained unique determinants of each of these equations.

Two further research issues may be addressed in future research efforts when data become available: (1) whether and how significantly the population of reproductive-age women would self-select into the sample of pregnant women; and (2) whether significant self-selection effects would be produced by other choice variables such as residential location, maternal age at birth, educational attainment, and marital status. These variables may be particularly endogenous to a model analyzing adolescents' birthing behavior.

ACKNOWLEDGMENTS

I am indebted to Mike Grossman, Bill Hsiao, Ted Joyce, Dick Ernst, and Joel Hay for their advice and suggestions. I am also grateful to the editors and two anonymous referees for suggesting substantial improvements. I benefited from the contributions of Tetsuji Yamada, Tadashi Yamada, Yanling Zhang, and the seminar participants at Yale, Johns Hopkins, and USC. The views in this article are those of the author, and do not necessarily reflect those of the Journal or USC.

NOTES

1. Throughout the study, pregnancy resolution is defined as the resolution of pregnancies as live births or induced abortions. Miscarriage or spontaneous abortion is not treated as an outcome option of a pregnant woman's birth decision model.
2. More formally, one can find that the expectation of birthweight production conditional on a self-selected sample may differ from its unconditional counterpart based on the total population. That is:

$$E(h_i|\Delta U_i \geq 0) = E(B^h X_{hi} + B^h_y y_i) + E(\varepsilon_{hi}|\varepsilon_{ui} \geq -B^u X_{ui}) = B^h X_{hi} + B^h_y y_i + E(\varepsilon_{hi}|$$
$$\varepsilon_{ui} \geq -B^u X_{ui}) \neq E(h_i), \;\; if\, cov(\varepsilon_{hi}, \varepsilon_{ui}) \neq 0$$

   Following this expression, whether and how the birth sample–based estimates would differ from their population-based counterpart depend on whether and how the two unobserved error terms ε_{ui} and ε_{hi} are correlated. As suggested by Grossman and Joyce (1990), it is likely that these two error terms are correlated. This is because some common components of the error terms (e.g., the cost of contraception, cost of abortion, and health status) tend to influence both birth probability and birthweight production functions.
3. Note that treating prenatal care delay as the only endogenous variable in this study yields no indication that other determinants of the health production function are purely exogenous. For instance, variables such as mother's age at birth, education, smoking, and parity may also be endogenously interacted with women's behavioral selection of birth and prenatal care demanded (Rosenzweig and Schultz 1983). Testing and controlling for the endogeneity of these variables require more comprehensive data sets.
4. For instance, compared to adults, adolescents are more likely to involve simultaneously decision making in pregnancy resolutions, demand for health inputs, and other endogenous statuses such as education and marriage. The considerable

interactions or endogeneity of these variables for teenagers make it substantially inappropriate to share the same model by both adults and adolescents.

5. Because all of the independent cities in Virginia are treated as equivalent to counties in this study, the total number of geographic units (counties and independent cities) ends up at 142. In addition, since Virginia has quite divergent residential areas, each county and independent city is classified as either urban or rural, according to the classification of metropolitan counties from the *State and Metropolitan Area Data Book 1986* of the U.S. Department of Commerce, Bureau of the Census.
6. In August 1977, the Hyde Amendment took effect and prohibited federal funds for induced abortions except in cases where the continuation of the pregnancy threatened the pregnant woman's life. Although some states still voluntarily continued to pay for abortions for Medicaid-eligible women or are under court order to do so, the total number of publicly funded abortions fell substantially from about 240,900 in 1977 to about 191,911 in 1978. Furthermore, almost 90 percent of the total publicly funded abortions in 1978 were done in California, New York, Pennsylvania, and Michigan. In contrast, Virginia experienced the sharpest drop in Medicaid abortions, from 4,000 in 1977 to 10 in 1978. The relatively high cost of abortion in Virginia also can be seen from the low availability of abortion services relative to availability in other areas such as New York City. In Virginia in 1984, the average number of abortion providers per 10,000 women ages 15–44 was 0.28 for all pregnant whites and 0.38 for all pregnant African Americans. New York City, by contrast, experienced almost twice these ratios.
7. The cohort-specific F-statistics of the identifying instruments for prenatal care are 77 for urban whites, 65 for rural whites, 57 for urban African Americans, and 34 for rural African Americans. The critical value of $F(12, \infty)$ at the 1 percent level is 2.18.

REFERENCES

Alexander, G. R., and C. C. Korenbrot. 1995. "The Role of Prenatal Care in Preventing Low Birth Weight." *The Future of Children* 5 (1): 103–20.

Becker, G. S., and H. G. Lewis. 1973. "On the Interaction Between the Quantity and Quality of Children." *Journal of Political Economy* 81 (2): S279–S288.

Ben-Akiva, M., and S. Lerman. 1989. *Discrete Choice Analysis.* Cambridge, MA: The MIT Press.

Bound, J., D. A. Jaeger, and B. Regina. 1995. "Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variables Is Weak." *American Statistical Association Journal* 90 (430): 443–50.

Brown, S. 1989. "Drawing Women into Prenatal Care." *Family Planning Perspectives* 21 (2): 73–80.

Cates, W., Jr. 1984. "Induced Abortion." In *Perinatal Epidemiology*, edited by M. B. Bracken. New York: Oxford University Press.

Chow, G. C. 1960. "Tests of Equality Between Sets of Coefficients in Two Linear Regressions." *Econometrica* 28 (3): 591–605.

Corman, H., T. Joyce, and M. Grossman. 1987. "Birth Outcome Production Function in the United States." *Journal of Human Resources* 22 (3): 339–60.

English, P., B. Eskenazi, and R. Christianson. 1994. "Black-White Difference in Serum Cotinine Levels Among Pregnant Women and Subsequent Effects on Infant Birth Weight." *American Journal of Public Health* 84 (9): 1439–43.

Frank, R., C. Jackson, D. Salkever, and D. Strobino. 1992. "Updated Estimates of the Impact of Prenatal Care on Birth Weight Outcomes by Race." *Journal of Human Resources* 27 (4): 629–42.

Gortmaker, S. L. 1979. "The Effects of Prenatal Care upon the Health of the Newborn." *American Journal of Public Health* 69 (5): 653–60.

Griliches, Z., B. H. Hall, and J. A. Hausman. 1978. "Missing Data and Self-Selection in Large Panels." *Annal de l'INSEE* (30–31): 137–76.

Grossman, M., and T. Joyce. 1990. "Unobservables, Pregnancy Resolutions, and Birth Weight Production Functions in New York City." *Journal of Political Economy* 98 (5): 983–1007.

Harris, J. E. 1982. "Prenatal Medical Care and Infant Mortality." In *Economic Aspects of Health*, edited by V. R. Fuchs. Chicago: University of Chicago Press.

Hausman, J. 1978. "Specification Tests in Econometrics." *Econometrica* 46 (November): 1251–71.

Heckman, J. 1976. "Sample Selection Bias as a Specification Error." *Econometrica* 47 (January): 153–61.

Institute of Medicine. 1985. *Preventing Low Birth Weight*. Washington, DC: National Academy Press.

Kleiman, J. C., and S. S. Kessel. 1987. "Racial Differences in Low Birth Weight: Trends and Factors." *The New England Journal of Medicine* 317 (12): 749–53.

Lee, L. F., G. S. Maddala, and R. P. Trost. 1980. "Asymptotic Covariance Metrics of Two-Stage Probit and Two-Stage Tobit Methods for Simultaneous Equations Models with Selectivity." *Econometrica* 48 (March): 491–503.

Liu, G. G. 1995. "An Economic Analysis of Pregnancy Resolution in Virginia Specific as to Race and Residence." *Journal of Population Economics* 8 (3): 253–64.

Manski, C. 1977. "The Structure of Random Utility Models." *Theory and Decision* 8 (2): 229–54.

Mustard, C., and N. Roos. 1994. "The Relationship of Prenatal Care and Pregnancy Complications to Birth Weight in Winnipeg, Canada." *American Journal of Public Health* 84 (9): 1450–57.

Nakosteen, R., and M. Zimmer. 1980. "Migration and Income: The Question of Self-Selection." *Southern Economic Journal* 46 (3): 840–51.

National Center for Health Statistics. 1993a. *Advance Report of Final Mortality Statistics, Monthly Statistics Report*. Hyattsville, MD: U.S. Public Health Service (1950 through 1991).

———. 1993b. *Advance Report of Final Natality Statistics, Monthly Statistics Report*. Hyattsville, MD: U.S. Public Health Service (1950 through 1991).

Newey, W., J. Powell, and J. Walker. 1990. "Semiparametric Estimation of Selection Models: Some Empirical Results." *American Economic Review* 89 (2): 324–28.

Paneth, N. S. 1995. "The Problems of Low Birth Weight." *The Future of Children* 5 (1): 19–34.

Piper, J., E. Mitchel, and W. Ray. 1994. "Presumptive Eligibility for Pregnant Medicaid Enrollees: Its Effects on Prenatal Care and Perinatal Outcome." *American Journal of Public Health* 84 (9): 1626–30.

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

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

Staiger, D., and J. H. Stock. 1994. "Instrumental Variables Regression with Weak Instruments." Technical Working Paper 151, National Bureau of Economic Research, Cambridge, MA.

Tietz, C. 1978. "Repeat Abortion—Why More?" *Family Planning Perspectives* 10 (4): 286–88.

World Bank. 1993. *World Development Report 1993,* New York: Oxford University Press.

Wu, D. 1973. "Alternative Tests of Independence Between Stochastic Regressors and Disturbances." *Econometrica* 41 (3): 733–50.