

NBER WORKING PAPERS SERIES

UPDATED ESTIMATES OF THE IMPACT OF PRENATAL CARE
ON BIRTHWEIGHT OUTCOMES BY RACE

Richard G. Frank
Donna M. Strobino
David S. Salkever
Catherine A. Jackson

Working Paper No. 3624

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 1991

This research received financial support from grant #MCJ-240545 from the Bureau of Maternal and Child Health. We are grateful to Michael Grossman for helpful discussion of the issues in this paper. This paper is part of NBER's research program in Health Economics. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

UPDATED ESTIMATES OF THE IMPACT OF PRENATAL CARE
ON BIRTHWEIGHT OUTCOMES BY RACE

ABSTRACT

This paper estimates a quasi-structural birthweight production function using data on counties for the years 1975-1984. The analysis focuses on the effects of first trimester initiation of prenatal care, controlling for use of abortion services, cigarette smoking, birth order and income. Fixed effects model is used to control for unmeasured differences in health endowments across counties. The results indicate that early first trimester initiation of prenatal care leads to a reduction in low birthweight for both blacks and whites. Differences in use of prenatal care by race explain only a small part of the black-white differences in the fraction of low birthweight births.

Richard G. Frank
School of Hygiene and
Public Health Services R&D
The Johns Hopkins University
624 North Broadway
Baltimore, MD 21205

Donna M. Strobino
School of Hygiene and
Public Health
The Johns Hopkins
University
624 North Broadway
Baltimore, MD 21205

David S. Salkever
School of Hygiene and
Public Health - Room 429
624 North Broadway
Baltimore, MD 21205

Catherine A. Jackson
University of Maryland
College Park, MD 20742

I. Introduction

Infant mortality rates have declined rapidly in the United States during the past two decades. However, the rate of infant mortality for blacks remains roughly double that for whites.¹ The risk of infant mortality increases significantly with decreases in birthweight. The rate of low birthweight (LBW) (less than 2500 grams) births has fallen more slowly than has the infant mortality rate.² For this reason, policy makers and researchers have recently focused attention on reducing rates of low birthweight births as the primary strategy for obtaining further reductions in the infant mortality rate. Moreover, reducing differentials in (LBW) rates by race has been viewed as an important step in reducing the infant mortality differential, since blacks have rates of (LBW) births that are over twice the rate of whites.²

The 1985 Institute of Medicine Report on Preventing Low Birthweight identified problems of access to prenatal care as being of central importance for developing public policies to improve the birthweight distribution in the United States.³ In 1984, approximately 80% of white mothers initiated prenatal care in the first trimester of pregnancy compared to roughly 61% of black women.⁴ This differential in use of prenatal care by race has been posited to be a major contributor to the difference in birth outcomes by race.⁵ This paper examines the potential for closing the gap in rates of LBW births by equalizing use of prenatal care for black and white women.

The analysis presented below obtains estimates of the impact of early initiation of prenatal care on the rate of LBW births for white and black women aged 20 to 34 years. We use a panel of U.S counties as the unit of observation and estimate quasi-structural birthweight production functions. We

find that increasing the early initiation of prenatal care among black women to the levels experienced by whites makes only a small contribution to closing the gap in rates of LBW births between the races.

The paper is organized into 5 sections. Section II outlines our approach to modelling the birthweight production function. Section III describes the data used in the analysis. Section IV presents the estimation results. Concluding remarks are presented in the final section of the paper.

II. Approach to Model Specification

A number of previous studies by economists have developed empirical models of the birthweight production function. The work of Corman, Grossman and Joyce⁶, Joyce,⁷ and Rosenzweig and Schultz,⁸ is particularly relevant to our analysis. All three studies used cross sectional data. Corman, Grossman and Joyce and Joyce used aggregate data on counties in an approach similar to our own, while Rosenzweig and Schultz use data on individual women. The models used in these studies begin with a birthweight production function that takes the general form:

$$(1) b = b(v, a, s, f, g, x, e)$$

Equation (1) states that the rate of LBW births (b) depends on use of prenatal care (v), use of abortion services (a), smoking (s), other endogenous inputs (such as nutritional intake) (f), the rate of prematurity (g), exogenous risk factors (x) and mothers' biological endowment (e) which is presumably at least partially observable to the mother but not the researcher.

Rosenzweig and Schultz point out that with data on individuals the

direct estimation of the production function (1) will lead to biased coefficient estimates. The reason for the bias is that a pregnant woman may have information regarding her health endowment (e.g. genetic make-up) that may influence both her choice of inputs (e.g., prenatal care to monitor a possible problem) and her birth outcome. The implication of this potential bias is that ordinary least squares may not produce consistent parameter estimates for this type of model.

This issue is somewhat more complicated when aggregate data are used. If the woman's health endowment is randomly distributed in the population, then aggregating large numbers of individual birth records by county of residence should eliminate sample variance in mean values of e across counties unless women with similar values of e cluster geographically. However, one may not have sufficient numbers of births in most counties to assure that the variance in the mean of e is greatly attenuated. This potential problem has led to use of structural equation models to estimate birthweight production functions for both individual and aggregate cross-sectional data. Joyce, and Corman, Joyce and Grossman both tested for the exogeneity of the abortion, smoking and prenatal care inputs in the birth outcome production function. They rejected exogeneity in all cases.

The use of aggregate panel data allows for a different approach to specification and estimation of the birthweight production function. The primary source of bias in estimation of the birthweight production function using aggregate county data is likely to stem from geographic clustering of women with similar health endowments. To the extent that this is the case, and that average health endowment within a county does not vary over time, using a fixed effects estimator with aggregate panel data on birth outcomes would

permit one to obtain unbiased estimates of the birthweight production function directly.⁹ Therefore, the approach we adopt is to estimate the birthweight production function using separate intercepts for each county in the analysis. This approach will remove the impact of any systematic and temporarily stable geographic clustering by women with similar values of e . (This also precludes use of covariates that are temporarily stable).

We used a modified version of equation (1) in our empirical work. Not all relevant inputs into the birthweight production function are measured in our data set. For example, the quality of nutritional intake is not measured. Income enters the demand functions for such inputs. We include the per capita income of a county in order to proxy for the ability to buy certain unmeasured inputs. Since, income (I) would not ordinarily be included in a structural production function, we follow Corman, Joyce and Grossman and refer to the following modification of equation (1) as a quasi-structural production function:

$$(2) b = b(v, a, s, I, x, e)^{10}$$

Equation (2) is estimated using variables measured at the county level for the years 1975-1984. As mentioned above, an additional threat to our estimates may be the small number of births in some counties. This problem is dealt with in two ways. First, data are aggregated for two-year intervals. Second, we perform Wu tests for exogeneity of the potentially endogenous inputs. A detailed description of the data follows.

III. Data Used in the Analysis

The analysis is a pooled, time series, cross-sectional study for the years 1975 to 1984 with the county as the unit of observation. We analyze data on all counties in the United States with populations of 10,000 or more whites or 5,000 or more blacks (based on the 1980 census). The result is a panel of 2,137 study counties for whites and 660 counties in the case of blacks. Separate regressions are estimated for black and white birthweight outcomes for 20-34 year old women. The reasons for estimating race specific regressions is that race is thought to interact with a variety of other explanatory variables. Focusing on the 20-34 year old age group allows one to minimize the number of counties with very few births, since the vast majority of births occur to women in this age group.

The race specific regressions are estimated for samples of counties with at least 5000 blacks or 10,000 whites in 1980. We combine data from birth records for two years for each county; the two year pairs are 1975-76, 1977-78, 1979-80, 1981-82, and 1983-84. Counties which reported more than 30% of observations missing from items on individual birth records were eliminated from the analysis file. Arizona was eliminated from the analysis file since it did not have a Medicaid program during the study period.¹¹

The source of data for the measurement of birth outcomes as well as several of the explanatory variables discussed below is the national detailed natality files for the years 1975-1984 produced by the National Center for Health Statistics (NCHS). These files contain information reported on standard birth certificates, which are completed for all live births at the time of birth. During the 10 year study period, some states reported data for 100% of births while others reported a 50% random sample of births; moreover

the number of states reporting 100% rose from 1975 to 1984.

The race and age specific low birthweight rate (LBW) is used as the dependent variable in the analysis. LBW is defined by the percentage of infants weighing less than 2500 grams at birth in a county.

The age of the mother and the race of the infant are also obtained from information on the birth records. NCHS has investigated the reliability of the reporting of maternal age by making comparisons of these data with census data and has generally found a high degree of consistency between the data sets.¹² The infant's race is determined by the races of the parents. When both parents are of the same race the child is assigned that race. If neither parent is white the child is assigned the father's race. If the race of only one parent is known the child is assigned that race. A 1981 study showed that 99.4% of white birth records and 98.6% of nonwhite birth records were complete.¹³

The variable used to define the use of prenatal care in our analysis measures initiation of care in the first trimester of pregnancy, was also used by Corman, Joyce and Grossman. This measure has been widely used as one indicator of access to prenatal care.³ It has the advantage of being unrelated to the length of pregnancy for live born infants. Corman, Joyce and Grossman also noted that this measure does not reflect frequent use of prenatal care by women who develop complications during their pregnancy or who begin pregnancy with underlying medical problems.¹⁴

The use of abortion services is measured as a two year average of the predicted rate of abortions among residents of a county per fertile woman (15-44 years). The source of the data on the volume of abortions are surveys conducted by the Alan Guttmacher Institute (AGI). The AGI collects data on

the volume of abortions and the providers of abortions by county of occurrence. The reporting of data by occurrence is problematic because our analysis uses observations on birth outcomes by county of residence. We therefore used data on county of occurrence to create synthetic estimates of abortion rates by county of residence.¹⁵

The third input in the production function model is the level of smoking by women in a county. Unfortunately, data are only available on the volume of cigarette sales by state and year for the study period.¹⁶ These data have several important shortcomings. First they are not race or age specific and are collected at the state rather than the county level. Second, cigarette sales data can be misleading indicators of consumption because of border crossing from high-tax to low-tax states. Nevertheless, we believe the dangers of omitted variables bias are greater than those stemming from measurement error. The income measure was the county per capita income estimates from the U.S. Commerce Department Local Area Personal Income statistics.

There is a well established relationship between birth order and the probability of LBW births^{17,18,19} We measure this variable as the percentage of first births. These data are derived from the NCHS natality files. As noted above, any counties where more than 30% of birth records did not report information on this variable in any year were excluded from the analysis.

The dependent variable is the percentage of LBW births. This is equivalent to use of a linear probability model with micro-data; however, with grouped data, the use of percentages is less problematic than the linear probability model with micro data.²⁰ Maddala proposes the use of weighted least squares in order to obtain appropriate variance estimates. The weight

for the LBW equations is given by:

$$(3) W = N^{-.5} (p(1-p))^{-.5}$$

where N is the number of births in a county and p is the percentage that are LBW births. This weighing scheme serves to reduce the impact of random fluctuations by weighing counties with more births more heavily in the regressions. Variable descriptions and descriptive statistics are presented on Table 1 for the white and black counties.

IV. Results

Table 2 presents the weighted least squares, fixed effects production function estimates. Note that we included individual time dummy variables for each of the two year pairs to control for time trends in the data (T1-T4). We performed Wu tests to test for exogeneity of inputs in the quasi-structural production function.²¹ The Wu test for the black women clearly failed to reject exogeneity of the inputs. For the white women the Wu test failed to reject exogeneity at the 0.05 level but rejected it at the 0.10 level. When two stage least squares models were estimated the first stage estimates had rather weak explanatory power ($R^2=0.09-0.18$) which led to instability in the coefficient estimates of the production function. This finding is consistent with the Monte Carlo findings of Maeshiro who showed that two stage least squares estimators produced inferior estimates to OLS when the first stage models had low explanatory power.²² For this reason we focus our attention on the single stage estimates.

The first column of Table 2 presents estimates of the LBW production function for 20 to 34 year old white women. The prenatal care initiation

indicator (1st Trimester) and the per capita income measure were the only variables with coefficient estimates that were significantly different from zero at conventional levels. The second column reports the coefficient estimates for black women. In this model only the prenatal care initiation variable had a coefficient estimate that was significant at conventional levels.

The coefficients for the prenatal care indicators suggest that the marginal product of prenatal care is larger for blacks than whites. The difference is not statistically significant but amounts to roughly 1.5 standard deviations. The outcome elasticities of prenatal care initiation (evaluated at the means) are -0.11 and -0.10 for whites and blacks respectively. Previous studies, such as Joyce's analysis of 1974-1976, have reported substantially larger marginal products and outcome elasticities for whites when compared to blacks.⁷ Our data include the years from 1975 to 1984. During this period, there was an expansion of publicly supported Maternal and Child Health Projects in the United States, which were often targeted at minority "high risk" populations. Our results may reflect improvements in the types of services available to minority women. It is also possible that either 1) the differing estimation strategies or 2) our larger sample of counties led to the differences in coefficient estimates.

Table 3 reports projections of rates of LBW births using the coefficients from Table 2. The first row of Table 3 presents race specific rates of LBW births in 1984. The third column shows that the ratio of LBW rates for blacks relative to whites was 2.28 in 1984. The second row presents the actual race specific rates of first trimester initiation of prenatal care in 1984. Note, that the ratio of black to white early initiation of prenatal

care was 0.81. The third row of the Table projects the rates of LBW births that would occur if the black rate of first trimester initiation of prenatal care was set equal to that of whites (at 82.5%). The LBW rate for blacks falls from 12.08% to 11.79% by raising the early initiation rate from 66.5% to 82.5%. This reduction is expected to lead to a drop in the ratio of black to white LBW rates from 2.28 to 2.22. The fourth row of Table 3 presents projection of the LBW rates under the assumption that 100% of black women would initiate prenatal care in the first trimester of pregnancy. The projection suggests that the LBW rate for blacks would fall to 11.46%, with a corresponding drop in the black-white LBW rate ratio to 2.16.

V. Discussion and Conclusions

The reported estimates of the marginal product of prenatal care are considerably below those obtained from cross sectional studies, such as that of Joyce.⁷ The reasons for the differences may lie either with the sample of counties used, the approach to estimation or the time periods studied. In spite of these clear differences, our analysis leads us to conclusion that are also implied by Joyce's study and the Institute of Medicine study on low birthweight.⁸ The expansion of early initiation of prenatal care will make only a small contribution to reducing risk of low birthweight births. However, as a strategy for reducing differences in birthweight outcomes by race, expanding early initiation of prenatal care by blacks would have a minor impact on that differential.

The literature on prenatal care indicates that 1) assuring adequate prenatal care is more important than early initiation per se, 2) the content of prenatal care may have an impact on birth outcomes and 3) the packaging of

prenatal care with other services (such as nutritional services and maternal education) tends to have the largest impact on the birthweight distribution. Further exploration of these areas is important in the context of econometric studies of birth outcomes. An equally important aspect of this problem is the economics of supplying "efficient" packages of prenatal services to vulnerable populations of women.

1. Kleinman J.C. and Kessel S.S., "Racial Differences in Low Birth Weight: Trends and Factors" New England Journal of Medicine, 317(12):749-753 1987.
2. U.S. Congress, Office of Technology Assessment, Healthy Children: Investing in the Future, OTA-H-345 Washington D.C. USGPO 1988.
3. Institute of Medicine, Preventing Low Birthweight Washington D.C.: National Academy Press 1985.
4. National Center for Health Statistics, Advance Report on Natality Statistics, DHHS Pub No. 86-1120 1986.
5. Secretary's Task Force on Black and Minority Health, Infant Mortality and Low Birthweight, Washington D.C.: U.S. Department of Health and Human Services 1986.
6. Corman H., Joyce T., and Grossman M., "Birth Outcome Production Functions in the U.S." Journal of Human Resources 22(3) 1987.
7. Joyce T.J., "The Impact of Induced Abortion on Black and White Birth Outcomes in the U.S." Demography 24(2) 1987.
8. Rosenzweig M.R. and Schultz T.P., "The Behavior of Mothers as Inputs to Child Health: The Determinants of Birthweight, Gestation, and Rate of Fetal Growth" in The Economic Aspects of Health V. Fuchs' (ed) Chicago: University of Chicago Press 1982.
9. Chamberlain G., "Heterogeneity, Omitted Variable Bias and Duration Dependence" in Longitudinal Analysis of Labor Market Data J.J. Heckman and B. Singer (eds) Cambridge U.K.: Cambridge University Press 1986.
10. It is worth noting that in this specification we do not include the previously included variable g, representing prematurity. We exclude this variable for two reasons: 1) previous models have made rather arbitrary identifying exclusion restrictions such as excluding smoking from the prematurity production function. We are reluctant to make such an assumption. 2) On a more practical level much of the data on gestational age of a birth is missing from the birth records.
11. In testing for exogeneity of inputs in the production function we used Medicaid program characteristics, availability of publicly provided maternal and child health services and physician availability as instruments. Most of these variables had a significant impact in the first stage regressions. Thus Arizona was eliminated because we lost a number of our identifying restrictions.
12. National Center for Health Statistics, Advance Data: Final Natality Statistics 1983, Monthly Vital Statistics Report 34 Hyattsville, Maryland 1985.

13. National Center for Health Statistics, Advance Data: Final Natality Statistics 1979, Monthly Vital Statistics Report, 30 Hyattsville, Maryland 1981.
14. Standards for prenatal care include not only the early initiation of care but also scheduling of prenatal visits, the frequency of visits, of course, increases with the duration of pregnancy.
15. The specific algorithm involved the following steps: We aggregated abortion data to the regional level using the Health Systems Agency (HSA) regions as a measure of a distinct medical region. The number of abortions at the HSA level we refer to as A_j . We then estimated a regression of the form $A_j = A(\text{POP}, \text{PROV}, I, \text{MDPOP})$ where POP is the female population of the j th HSA between the ages of 15 and 44, PROV is the number of abortion providers in the HSA, I is the per capita income, and MDPOP is the physician to population ratio. Based on the regression results we used the estimated coefficients to predict the number of abortions in all counties that were in HSAs with positive numbers of abortions. This predicted value is given the symbol A_{ij} denoting the number of abortions in the i th county within the j th HSA. We then created the ratio $R_{ij} = A_{ij} / \sum A_{ij}$. R is therefore the share of the predicted total abortions within the j th HSA accounted for by the i th county. In order to obtain the final estimate of the number of abortions in a given county we added a constraint that the sum of abortions across counties within an HSA has to add up to the actual total number of abortions. Thus the estimated volume of abortions for the i th county is $R_{ij} A_j$.

We estimated the predicted levels of abortion for each sample county for all the even numbered years. In order to determine the validity of this approach to estimation we obtained data from vital statistics in Tennessee which records all abortions in the state. We used the vital statistics data on the volume of abortions by county for 1982. We compared the vital statistics data with our predicted estimates of the volume of abortions via calculation of a rank correlation statistic. The estimated correlation coefficient for the predicted and vital statistics volume of abortions was 0.875 which was significantly different from zero at the 0.01 level. The results suggest that we were able to develop rather accurate predictions of the volume of abortions by county of residence. It should be pointed out that this variable is neither age nor race specific. The final step involved dividing the predict volume of abortions in a county by the female population age 15 to 44 years.

16. These data are obtained from cigarette sales data from the Tobacco Institute which were made available to us by Michael Grossman of the National Bureau of Economic Research.

17. Chase H.C. Time Trends in Low Birth Weights in the United States. In The Epidemiology of Prematurity, edited by D.M. Reed and F.J. Stanley. Baltimore: Urban and Schwarzenberg, 1977, 17.37.
18. Selvin S. and Garfinkel J. Relationship Between Parental Age and Birth Order With Percentage of Low Birth Weight Infants. Human Biology, 44:501-510 (1972).
19. Strobino D.M. Trends in the Low Birth Weight Infants and Changes in Baltimore's Childbearing Population, 1972-1977. Public Health Reports, 97:273-282, (1982).
20. Maddala G.S. Limited Dependent and Qualitative Variable in Econometrics. Cambridge: Cambridge University Press, 1983.
21. Wu D., "Alternative Tests of Independence Between Stochastic Regressors and Disturbances" Econometrica 41:733-750 1973.
22. Maeshiro, A., "One More Look at the Choice Between OLS and TSLS: TSLS Can be Made as Inferior as One Wishes" American Statistical Association Proceedings of the Business and Economic Statistics Section Washington D.C.: ASA 1979 pp.302-306.

TABLE 1
Descriptive Statistics
Variable Description

<u>Variable</u>	<u>Variable Description</u>	Black mean (SD)	White mean (SD)
Smoking	Cigarette Sales/1,000 Population of State	137,339.19 (25,612.08)	135,046.16 (27,969.36)
Abortion	Abortion/Women of Childbearing Age (15-44)	0.0198 (0.0271)	0.0179 (0.0258)
1st Trimester	% Initiating Prenatal Care in 1st Trimester	64.11 (10.49)	81.15 (7.78)
First Birth	% of Births that are 1st Births	25.09 (5.73)	35.84 (5.06)
Income	Per Capita County Income	7663.08 (2,706.04)	7883.32 (2501.30)

TABLE 2
 Regression Results
 Quasi-Structural
 Production Functions¹

	White ²	Black
Smoking	6.30e ⁻⁷ (0.34)	4.78e ⁻⁶ (0.86)
Abortion	1.892 (1.09)	3.109 (0.59)
1st Trimester	-0.007 (2.01)	-0.0180 (2.00)
First Birth	-0.001 (0.21)	-0.0184 (1.08)
Income	-0.00005 (2.81)	-0.00009 (1.45)
T1	-0.179 (4.22)	0.125 (0.81)
T2	-0.207 (3.09)	0.333 (1.32)
T3	-0.155 (1.76)	0.469 (1.36)
T4	0.042 (0.39)	0.709 (1.65)
R ²	0.56	0.41

¹Fixed effects models t's are adjusted for degrees of freedom reduction.

²t's in parentheses.

TABLE 3
 Racial Differences in LBW and Prenatal Care
 White 1984

	Black Mean	White Mean	Ratio
LBW Rate in 1984	12.08%	5.30%	2.28
1st Trimester Initiation 1984	66.50%	82.50%	0.81
LBW Rate With Equalized 1st Trimester Initiation at 82.5%	11.79%	5.30%	2.22
LBW With 100% of Blacks Initiating in 1st Trimester	11.48%	5.30%	2.16