Journal of Health Economics 9 (1990) 273–288. North-Holland

THE DYNAMIC RELATIONSHIP BETWEEN LOW BIRTHWEIGHT AND INDUCED ABORTION IN NEW YORK CITY

An Aggregate Time-Series Analysis

Theodore JOYCE*

Baruch College and the National Bureau of Economic Research, New York, NY 10003, USA

Michael GROSSMAN*

City University of New York and the National Bureau of Economic Research, New York, NY 10003, USA

Received December 1989, final version received June 1990

We use a vector autoregression to examine the dynamic relationship between the race-specific percentage of pregnancies terminated by induced abortion and the race-specific percentage of low-birthweight births in New York City. With monthly data beginning in 1972, we find that induced abortion explains low birthweight for blacks, but not for whites. There is no evidence of feedback from low birthweight to induced abortion. The findings suggest that unanticipated decreases in the percentage of pregnancies terminated by induced abortion would worsen birth outcomes among blacks in New York City.

1. Introduction

There is speculation that the U.S. Supreme Court will overturn its 1973 decision in *Roe versus Wade*. A likely outcome is that some states will increase restrictions on the availability of legalized abortion. A possible consequence is that aggregate indicators of infant health may worsen or display a slowdown in their rate of improvement in states or regions where access to legal abortion is limited. Evidence for such a result comes from a number of cross-sectional studies in which area-specific abortion rates or measures of abortion availability are inversely related to the rate of low birthweight births and neonatal mortality [Grossman and Jacobowitz (1981), Corman and Grossman (1985), Joyce (1987)]. The underlying mechanism is that the abortion rate measures the extent to which unwanted births have been averted. Areas in which the shadow price of abortion is relatively low

*The research was supported by Grant Number 88-5567 from the Henry J. Kaiser Family Foundation to the National Bureau of Economic Research (NBER). We wish to thank two anonymous referees for their comments. The research is part of the NBER's program in Health Economics. Any opinions expressed are those of the authors and not those of the Kaiser Family Foundation or the NBER.

should have higher abortion rates and lower rates of unintended childbearing. This should lower optimal family size and increase the resources devoted to each child.

The notion that induced abortion acts as a selection mechanism in which less healthy or less wanted fetuses are terminated has recently been examined by two studies that used micro vital statistics from New York City [Joyce and Grossman (1990), Grossman and Joyce (1990)]. In both studies the authors argued that abortion greatly reduced the uncertainty associated with fertility control. Thus, estimates of the health technology or the demand for prenatal inputs may be biased if women who give birth represent a non-random draw from the population of pregnant women. Their results confirmed the presence of selection bias. In particular, a black woman who had a greater than expected probability of giving birth given she was pregnant had a smaller than expected delay in the initiation of prenatal care, and delivered a heavier than expected baby. Because the studies found no such effect for white women, the authors speculated that blacks may substitute abortion for contraception more frequently than whites.

The micro-level studies provide important additional evidence of a positive link between induced abortion and infant health. It should be emphasized, however, that for any pregnant woman abortion or birth are mutually exclusive choices. As a result, there is no research design which can directly test the hypothesized link at the individual level. With other interventions, such as early prenatal care, one can compare the birthweight of women with and without treatment. By contrast, Grossman and Joyce (1990) relied on the expected birthweight of women who chose to give birth as compared to the expected birthweight of a woman who aborted had they instead given birth in order to infer that induced abortion lead to improved birth outcomes. The upshot is that unlike other interventions that affect infant health, aggregate data remain an important means of testing whether increases in the use of induced abortion are positively related to improvements in aggregate birth outcomes.

The novelty of the present study is the use of aggregate time-series data to examine the relationship between induced abortion and infant health. In particular, we apply a vector autoregression (VAR) to monthly, race-specific data from New York City on the percentage of low-birthweight births, the percentage of births to women who began care in the first trimester and the percentage of pregnancies terminated by induced abortion. If there is a relationship between low birthweight and induced abortion, then one would expect changes in the percentage of pregnancies terminated by induced abortion to precede changes in the rate of low birthweight.

Time-series analysis with its emphasis on the dynamic relationship between induced abortion and infant health should be viewed as a complementary test of the cross-sectional relationship which emphasizes variations in long-run equilibrium at a single point in time. Because a direct test with individual data is not possible, a 'weight of evidence' approach should be used to address the question of whether induced abortion improves infant health. The present study adds to that body of evidence.

2. Empirical implementation

2.1. The data

The data on low birthweight, prenatal care and induced abortions are from micro-level vital statistics. Although birth certificate data are readily available, New York City is one of only 14 vital registration areas, the other 13 areas being states, that report micro-level data on induced terminations to the National Center for Health Statistics [National Center for Health Statistics (1988)].

Each year of data has been aggregated by month for blacks and whites separately from January 1972 through December 1988. January 1972 is the first year in which monthly data on induced terminations became available.[1] Only residents of New York City were included in the aggregation. The data for whites contain a large component which is Hispanic. Ethnicity was not identified on New York City birth certificates until 1978. In 1984, the first year in which data on Hispanics were published, 50% of all white women in New York City who gave birth were of Hispanic origin or descent; 85% of Hispanic mothers were white [New York City Department of Health (1985)].

Since the data are aggregated, we used the percentage of low-birthweight births as the indicator of infant health. Prenatal care is measured by the percentage of births to women who began care in the first trimester.[2] To capture the effect of induced abortion, we used the percentage of pregnancies terminated by induced abortion.[3] To insure that births and induced abortions originated from the same cohort of pregnancies, abortions were lagged by six months since approximately 90% of all induced abortions occur within the first trimester [NCHS (1989)].[4]

[1]The absolute number of births and abortions were large enough to give reliable monthly rates. New York City averaged approximately 100,000 live births and 10,000 induced abortions per year to black and white residents between 1973 and 1988. This is equivalent to approximately 5,000 births and 2,500 induced abortions per month for whites, and 3,000 births and 3,500 abortions per month among blacks.

[2]Before 1978, New York City birth certificates did not distinguish between women who received no care and women whose care was unknown. Thus, to create a consistent series the denominator excludes women who received no care.

[3]Throughout the paper pregnancies are defined as the sum of live births plus induced abortions since data on miscarriages are not well-reported.

[4]The percentage of pregnancies terminated by induced abortion, A_t, can be written as follows: $[a_{t-6}/(a_{t-6}+b_t)]\times 100$, where a_{t-6} is the number of induced abortions at time $t-6$, and b_t is the number of live births at time t. Births delivered in December (t) and abortions performed in June $(t-6)$ were assumed to have been conceived in March $(t-9)$. Thus, we test whether the rate of low birthweight in December (t) has been affected by the percentage of pregnancies that were conceived nine months earlier but have been terminated.

Table 1
Summary statistics.

	Low birthweight (%)		Early care (%)		Aborted (%)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Jan. 1972–Dec. 1979						
Blacks	12.00	0.90	30.96	4.50	51.45	2.71
Whites	6.39	0.43	51.86	2.93	33.33	2.85
Jan. 1980–Dec. 1988						
Blacks	11.32	0.96	41.68	3.76	52.69	2.45
Whites	6.06	0.43	57.87	2.23	40.79	4.27

Visual inspection of the data reveal a strong linear trend in prenatal care and less pronounced trends in birthweight and abortion. Summary statistics are presented in table 1. We also test whether any of the series contain a stochastic, as opposed to a deterministic trend. It is often referred to as a test of unit roots [Maddala (1988)]. Failure to account for stochastic trends can lead to spurious or misleading inferences [Stock and Watson (1988)]. Following the test suggested by Said and Dickey (1984), we rejected the null hypothesis that each of the series contains a unit root.[5] Thus, the variables are expressed in levels and the model is estimated with and without a trend term. To control for seasonality, each equation includes a set of 11 seasonal dummies.

2.2. *Methodology*

Although primarily a tool of macroeconomists, VARs represent a relatively unrestrictive means of highlighting important correlations in the data so as to empirically confirm or question various hypothesized relationships. Vector autoregressions are not a substitute for structural models; rather, they should be interpreted as dynamic reduced forms.[6] As such, each variable is regressed on its own lags as well as on the lags of all the other variables. Because the right-hand-side variables are the same in each equation, ordinary least squares is consistent and efficient [Zellner (1962)]. To facilitate the discussion, a simplified version of the model to be estimated can be specified as follows.

[5] We used the extension of the Dicky–Fuller test proposed by Said and Dickey (1984) because it allows for a more general ARMA process. With first differences as the dependent variable, the t-statistic on the first lag of the level of each variables was as follows: low birthweight, 19.9 blacks, 27.8 whites; prenatal care, 10.4 blacks, 11.3 whites; and abortion, 14.8 blacks, 5.3 whites.

[6] Since each equation includes only lagged values of the 'endogenous' variables, they are predetermined or exogenous in a temporal sense. Thus, the reference to VARs as dynamic reduced forms.

$$B_t = a_1 B_{t-1} + b_1 P_{t-1} + c_1 A_{t-1} + e_t, \tag{1}$$

$$P_t = a_2 B_{t-1} + b_2 P_{t-1} + c_2 A_{t-1} + u_t, \tag{2}$$

$$A_t = a_3 B_{t-1} + b_3 P_{t-1} + c_3 A_{t-1} + v_t. \tag{3}$$

Let B_t be the percentage of low-birthweight births at time t, let P_t be the percentage of births to women who began care in the first trimester, let A_t be the percentage of pregnancies terminated by induced abortion and let e_t, u_t, and v_t be the residuals. Note that the intercept and other deterministic components have been suppressed and the lag length has been set to one.

Eqs. (1)–(3) represent the autoregressive specification. With more lags the t-statistics become relatively uninformative due to multicollinearity. Thus, inferences as to temporal ordering are obtained by F-tests on the set of lags for each regressor.[7] For instance, rejection of the null hypothesis that $c_1 = 0$ (or $\sum c_{1i} = 0$ with more lags) indicates that variation in abortion precedes changes in low birthweight. If at the same time, we cannot reject the null hypothesis that $a_3 = 0$, then the data are consistent with the effect from abortion to low birthweight being unidirectional. On the other hand, rejection of the null hypothesis that $a_3 = 0$ indicates feedback from low birthweight to induced abortion. Because there is no strong rationale for why prior variations in the rate of low birthweight should explain changes in the percentage of pregnancies terminated by induced abortion, feedback from low birthweight to abortion would most likely be the result of an omitted variable.[8] This represents an important check on the specification of the system.

Another advantage of VARs is that they allow for a rich set of interactions across equations. Once the parameters in eqs. (1)–(3) have been estimated, the system can be expressed in its moving-average representation, and one can trace out the responses of each variable to a shock in one of the others.[9] These simulations or impulse response functions are a useful means of characterizing the dynamics of the system [Sims (1980)]. Often the impact

[7]Tests of whether lagged X explains Y in a regression that includes lagged Y are often referred to as tests of Granger causality [Granger (1969)]. However, following Leamer (1985), we avoid the use of the term, Granger causality, in order to minimize the confusion between Granger's definition of causality and causality as defined by philosophers of science.

[8]One possibility is to view lagged low birthweight as proxy for the aggregate health endowment. Thus, for example, an increase in the incidence of AIDS or substance abuse among pregnant women – a decrease in the aggregate health endowment – could increase the proportion of pregnancies terminated by induced abortion. However, surveys as to why women abort suggest that medical problems are not a major reason for selective abortions [Torres and Forrest (1988)].

[9]Examples of changes that would cause a shock in the percentage of pregnancies terminated by abortion include the closing of abortion clinics, a supply-side shock, or restrictions on the Medicaid financing of abortions, a demand-side shock.

of a variable in a dynamic model is expressed as the sum of its coefficients in the autoregressive specification. Impulse response functions offer a more general description of the system's dynamics because they allow interactions among variables and across equations. To illustrate, substitute recursively in the system above, and express each variable as a function of contemporary and lagged disturbances. Thus, a one-standard deviation shock in the percentage of pregnancies terminated by induced abortion, σ_{vt}, for example, causes low birthweight to change by $c_1\sigma_v$ in period $t+1$ and by $(c_1c_3+a_1c_1+b_1c_2)\sigma_v$ in period $t+2$. Changes in subsequent periods can be mapped out in a similar fashion. If the system is stable, the impulses will diminish to zero and the cumulative effect becomes the sum of the impulse responses. Because the impulses are a complicated combination of coefficients, Monte Carlo integration is used to approximate confidence bands around the impulses.[10]

A drawback to VARs is that the addition of each new variable consumes degrees of freedom exponentially. Consequently, the number of variables as well as the number of lags that can be employed are limited. In addition to the abortion measure, we include early prenatal care in order to control for medical inputs. The positive effect of early prenatal care on birthweight has been reported frequently [Institute of Medicine (1985)]. Moreover, the initiation of prenatal care is a decision made by the woman and not the physician, and thus may reflect other healthy behaviors that are not observed.

We lack data on risk factors for low birthweight such as smoking and substance abuse. Their absence will not bias the results if the inclusion of lagged low birthweight adequately controls for the omitted risk factors. Such an assumption is not unreasonable provided there are no strong contemporaneous relationships between the rate of low birthweight and the excluded risk factors. The relationship between low birthweight and smoking, for example, is likely to be lagged and dispersed over several months. An increase in aggregate smoking among pregnant women in month t, for instance, will effect the incidence of low birthweight over the ensuing nine months. Yet, contemporaneous effects are likely to be the weakest, since longer exposure is associated with greater risk. In short, the inclusion of lagged low birthweight should effectively control for the impact of risk factors given the plausible assumption of serial correlation among the omitted factors.

With respect to lag length, we use nine lags in the rate of low birthweight in order to approximate the length of a pregnancy. It seems reasonable to assume that changes that occurred prior to when a particular cohort of

[10]The Monte Carlo integration is based on the procedure described in the estimation package, RATS [Doan and Litterman (1987)].

women became pregnant would have little impact on birth outcomes 10, 11, or 12 months later. There is no rationale for justifying a particular lag length with respect to prenatal care. The numerator of the prenatal care variable – the number of births delivered in month t to women who began care in the first trimester – incorporates the time lag between when care is received and when the infant is born. Thus, we begin with a general lag length of nine months.

In the case of the abortion measure, the number of potentially relevant lags may be less than nine. For instance, 90% of all abortions occur in the first trimester, but over a third are performed within the first eight weeks [NCHS (1988)]. Thus, induced abortions at months $t-7$ and $t-8$ may have an important impact on the rate of low birthweight in month t. Similarly, induced abortions at month $t-11$ and earlier are unlikely to impact on the rate of low birthweight in month t. This suggests a lag length of between 3 to 5 months given that the abortion measure already incorporates a lag (see footnote 4).

We use the Chi-square test proposed by Sims (1980) to test whether models with less than nine lags are an appropriate restriction of a model with nine lags. Rejection of the null implies that specifications with less than nine lags are not appropriate. As a further check of the specification, we will test for first-, sixth-, and twelfth-order autocorrelation by means of a Lagrange Multiplier test [Godfrey (1978), Maddala (1988)].[11]

3. Results

The results from the autoregressive specification that includes a linear trend term are presented in tables 2 and 3. The trend had little impact on the autoregressive estimates. We display estimates from the specification with trend term because the impulse response functions from this model converged more quickly.

For blacks, we were unable to reject the null hypothesis at the 5% level that a system of 8, 7, 6 or 5 lags was a restriction on a system with 9 lags. For whites, a model with 5 or 6 lags was an inappropriate restriction. However, for comparative purpose we used a specification in which each variable is lagged 5 times. As will be shown below, the substantive conclusions for whites were unaffected by the truncated lag length.

Based on the F-tests on the set of lags for each variable, in the case of blacks, we can reject the null hypothesis that the percentage of pregnancies terminated by induced abortion has no power to explain the percentage of low-birthweight births. The F-statistic is 3.02 with a marginal significance

[11]The Durbin–Watson statistic is inappropriate when the regression includes lagged endogenous variables.

Table 2

Estimates of autoregressive equations – blacks.[a]

Independent variable	Lag	Dependent variable: % Low birthweight	% Early care	% Aborted
% Low birthweight	1	0.412***	−0.206	−0.038
		(5.37)	(−1.33)	(−0.22)
	2	−0.065	0.221	−0.069
		(−0.79)	(1.33)	(−0.36)
	3	0.188	0.074	−0.019
		(2.31)	(0.45)	(−0.10)
	4	0.005	−0.002	0.110
		(0.06)	(−0.10)	(0.58)
	5	0.191	−0.161	−0.197
		(2.52)	(−1.04)	(−1.12)
% Early care	1	0.065	0.366***	−0.031*
		(1.68)	(4.66)	(−0.35)
	2	−0.043	0.106	−0.051
		(−1.03)	(1.27)	(−0.53)
	3	−0.006	0.128	0.081
		(−0.15)	(1.54)	(0.85)
	4	0.083	0.068	−0.201
		(2.01)	(0.80)	(−2.08)
	5	−0.074	−0.0001	0.372
		(−1.90)	(−0.01)	(3.01)
% Aborted	1	−0.042**	0.009	0.369***
		(−1.29)	(0.14)	(4.95)
	2	−0.085	−0.045	0.052
		(−2.49)	(−0.65)	(0.66)
	3	−0.009	0.060	0.142
		(−0.25)	(0.86)	(1.78)
	4	0.002	−0.056	0.008
		(0.06)	(−0.79)	(0.10)
	5	0.016	−0.034	0.172
		(0.48)	(−0.50)	(2.23)
Constant		8.679	12.343	15.051
		(2.83)	(2.00)	(2.12)
Trend		−0.002	0.034	−0.009
		(−0.39)	(3.70)	(−0.88)
Adjusted R^2		0.52	0.95	0.62
LM-test, chi-square[b]				
H_0: $p_1=0$		0.76	2.39	0.17
H_0: $p_{12}=0$		0.03	3.48	0.61
H_0: $p_1,\ldots,p_6=0$		4.55	6.24	1.60
H_0: $p_1,\ldots,p_{12}=0$		17.40	13.80	7.16

[a] t-ratios are in parentheses. The coefficients on the seasonal dummies are not shown.
***p-value on set of lags is less than 0.01.
**p-value on set of lags is less than 0.05.
*p-value on set of lags is less than 0.10.

[b] In the LM test for serial correlation, the residuals from the equations shown above are regressed on the complete set of independent variables and a set of lagged residuals. The number (k) and order (p) of the lagged residuals depend on the degree of autocorrelation being tested. An F-statistic is computed for the coefficients on the lagged residuals. Multiplying F by k yields the chi-square statistic with k degrees of freedom. The critical value for chi-square at the 0.05 level for 1 degree of freedom is 3.841, for 6 degrees of freedom is 12.592, and for 12 degrees of freedom is 21.026.

Table 3

Estimates of autoregressive equations – whites.[a]

Independent variable	Lag	Dependent variable: % Low birthweight	% Early care	% Aborted
% Low birthweight	1	0.199**	−0.145*	−0.255
		(2.57)	(−0.62)	(−0.84)
	2	0.019	−0.186	0.405
		(0.24)	(−0.76)	(1.31)
	3	−0.007	−0.339	−0.122
		(−0.08)	(−1.39)	(−0.39)
	4	−0.132	−0.558	−0.125
		(−1.65)	(−2.29)	(−0.40)
	5	0.095	−0.115	−0.270
		(1.18)	(−0.46)	(−0.86)
% Early care	1	−0.040	0.429***	0.012
		(−1.55)	(5.51)	(0.12)
	2	0.018	0.183	0.121
		(0.66)	(2.17)	(1.13)
	3	0.021	0.083	−0.045
		(0.76)	(0.97)	(−0.41)
	4	−0.020	0.081	−0.065
		(−0.73)	(0.95)	(−0.59)
	5	−0.008	−0.077	0.124
		(−0.32)	(−0.99)	(1.25)
% Aborted	1	−0.001	0.018	0.478***
		(−0.05)	(0.30)	(6.11)
	2	−0.006	0.020	0.172
		(−0.26)	(0.30)	(2.00)
	3	−0.004	0.045	0.373
		(−0.22)	(0.70)	(4.57)
	4	0.004	−0.053	−0.132
		(0.18)	(−0.79)	(−1.52)
	5	−0.021	−0.027	0.055
		(−1.06)	(−0.44)	(0.71)
Constant		7.898	5.410	−3.900
		(4.44)	(4.20)	(−0.56)
Trend		0.001	0.011	−0.009
		(0.92)	(3.12)	(−2.02)
Adjusted R^2		0.30	0.90	0.91
LM-test, chi-square[b]				
H_0: $p_1 = 0$		0.64	0.16	0.90
H_0: $p_{12} = 0$		0.77	1.97	0.41
H_0: $p_1, \ldots, p_6 = 0$		6.56	1.02	16.78
H_0: $p_1, \ldots, p_{12} = 0$		14.63	6.24	33.40

[a] t-ratios are in parentheses. The coefficients on the seasonal dummies are not shown.

***p-value on set of lags is less than 0.01.

**p-value on set of lags is less than 0.05.

*p-value on set of lags is less than 0.10.

[b] In the LM test for serial correlation, the residuals from the equations shown above are regressed on the complete set of independent variables and a set of lagged residuals. The number (k) and order (p) of the lagged residuals depend on the degree of autocorrelation being tested. An F-statistic is computed for the coefficients on the lagged residuals. Multiplying F by k yields the chi-square statistic with k degrees of freedom. The critical value for chi-square at the 0.05 level for 1 degree of freedom is 3.841, for 6 degrees of freedom is 12.592, and for 12 degrees of freedom is 21.026.

level of 0.012. For whites, the F-statistics is 1.76 with a marginal significance level of 0.123.

There is no indication of feedback from low birthweight to abortion. That is, lagged values of low birthweight do not explain variations in the abortion measure. The absence of feedback lessens the likelihood that an omitted third variable is confounding the results. Nevertheless, the possibility remains, although we believe it to be small, that the unidirectional effect from induced abortion to low birthweight among blacks would not hold in a higher dimensional model.[12]

An examination of the individual coefficients in the low-birthweight equation for blacks reveals that the first three lags of the percent aborted variable are negative and the t-statistic of the coefficient on the second lag is greater than two in absolute value. Given the potential for multicollinearity, the standard errors are suspect, but the results are consistent with the expectation that first trimester abortions, or abortions performed in month $t-7$, $t-8$ and $t-9$ (the first three lags in the abortion measure), are the most relevant for the rate of low birthweight in month t.

The results from the tests of autocorrelation are shown at the bottom of tables 2 and 3. The test is based on a set of auxiliary regressions in which the residuals from the autoregressive specifications are regressed on the complete set of regressors and lags of the residuals. The number of lagged residuals determines the order of autocorrelation to be tested. A rejection of the null hypothesis that the coefficients on the set of residuals equals zero would indicate that the residuals are not white noise. In the low-birthweight specifications there is no indication of first-, sixth- or twelfth-order autocorrelation. Nor can we reject the null that the coefficient on the twelfth residual lag is different from zero, which is evidence that the seasonal dummies are effective. Overall autocorrelation was not a problem except in the white abortion equations.

The impulse response functions are presented in figs. 1 and 2. Fig. 1(a), (b) shows the changes in the rate of low-birthweight births (in percentage points) over thirty-six months given a shock of one standard deviation in the residual variance of the abortion measure [σ_{vt} from eq. (3)]. For blacks the shock represents a 1.55 percentage point decline in the percentage of pregnancies terminated by induced abortion and a 1.50 percentage point

[12]Lutkepohl (1982) has shown that a lack of feedback from y to x in a bivariate model does not preclude the possibility of feedback from y to x in higher dimensional model. He then shows with Canadian data that even though income does not explain money in the bivariate case, it does explain money in a three-equation system that includes interest rates. However, unlike in macroeconomics, many of the relationships in birth outcome models have well-defined temporal orderings. The decision to terminate a pregnancy, or the decision to initiate prenatal care, for example, clearly precedes delivery. Thus, feedback from low birthweight would most likely be an indication of an omitted variable. The absence of feedback in the abortion equations combined with the temporal ordering of events, increases our confidence in the results.

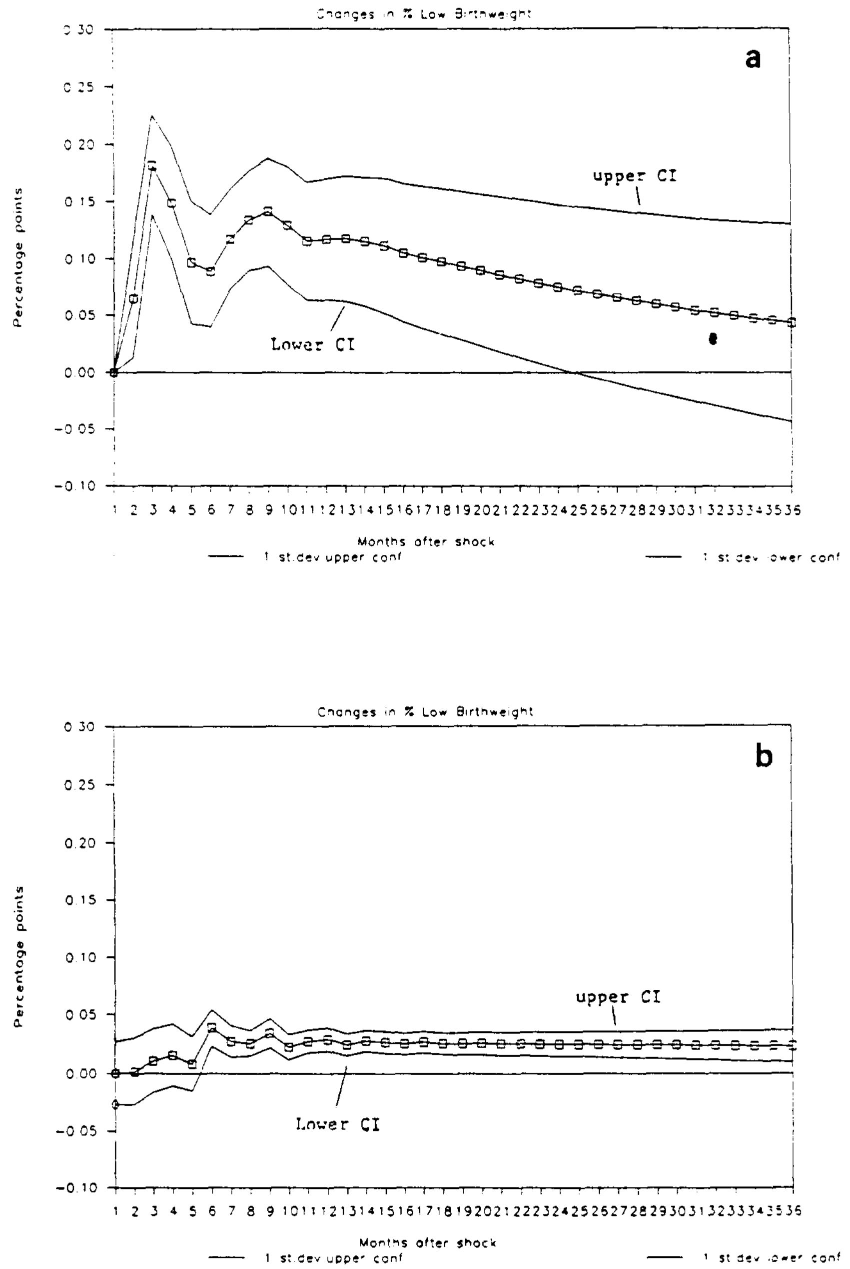
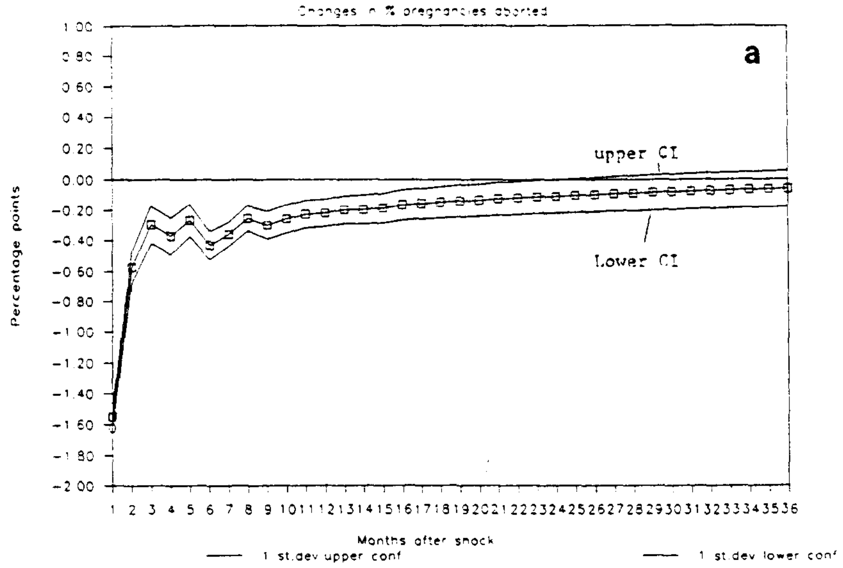


Fig. 1. (a) Impulse responses – blacks – for low birthweight. (b) Impulse responses – whites – for low birthweight.



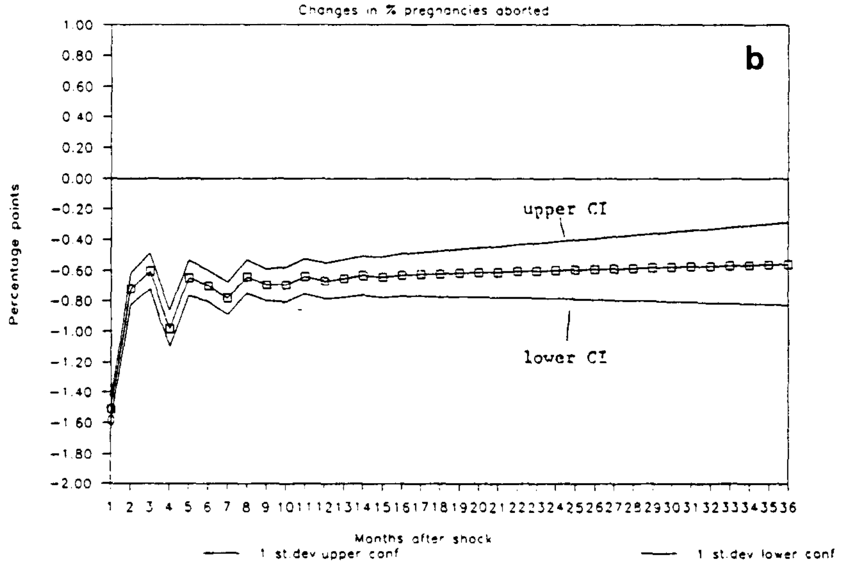


Fig. 2. (a) Impulse responses – blacks – for pregnancies aborted. (b) Impulse responses – whites – for pregnancies aborted.

decline in the case of whites.[13] The confidence bands are at plus and minus one standard deviation of each impulse.

As fig. 1(a) makes evident, an unanticipated decrease in the percentage of pregnancies terminated by induced abortion has a substantial effect on the rate of low birthweight among blacks. The increase in low birthweight is greatest in the third month following the shock after which the impulses slowly dampen towards zero. The results for whites [fig. 1(b)] are in marked contrast. There is no effect in the first four months and although the confidence bands are quite tight, the change relative to blacks is obviously small.

The cumulative increase in the rate of low birthweight among blacks over the 36 month forecast period is 3.16 percentage points relative to a mean of 11.6%; for whites, the cumulative increase is 0.83 percentage points relative to a mean of 6.2%. Although the change for blacks appears large relative to a 1.5 percentage point shock in the abortion measure, it should be noted that the change in the percentage of pregnancies that are terminated is not static. To illustrate, fig. 2 shows the changes in the percentage of pregnancies terminated by induced abortion after a one-standard deviation shock to itself [σ_{vt} from eq. (3)]. As fig. 2(a) makes clear, the change in the percentage of pregnancies that are aborted remains negative months after the initial shock. The cumulative decrease is 7.5 percentage points relative to a mean of 52.1%.[14]

The results from the impulse response functions suggest that legislative restrictions on the use of abortion would increase the incidence of low birthweight among blacks in New York City. Such an interpretation, however, should be made with caution. First, the impulse response functions present the short-run changes caused by a one-period shock whereas legislative changes are clearly more permanent. Similarly, the shock that generates the impulse responses is typical of the unanticipated changes in the abortion measure that occurred during the sample period. Legislative restrictions on abortion that would result from a reversal of *Roe versus Wade* may have much larger effects on the use of abortion. It is tempting, therefore, to argue that the impulse response functions represent lower bound estimates of the effect of severely restricted access to legal abortion. Yet, this is also questionable since we do not know the extent to which legislative changes

[13]To insure that the shocks to each variable are unique, Sims (1980) recommends that the residual variance–covariance matrix be transformed so that it becomes block recursive. However, we felt the transformation was unnecessary in this context since the temporal ordering of events minimized the likelihood of any important contemporaneous relationships. The data supported this expectation; the residual correlations across equations were inconsequential.

[14]The slow rate of convergence among the impulse response functions for whites is due in part to the short lag length which was an inappropriate restriction. However, none of the conclusions regarding the relationship between low birthweight and abortion were different in a specification that included 9 lags the results of which are available upon request.

can be considered 'unanticipated'. For some segments of population, such as teens or deeply disenfranchised women, the implications of a Supreme Court decision may not be well understood. Others, however, may change their contraceptive behavior quickly in response to a new legal environment regarding abortion. In sum, we believe the impulse reponse functions are instructive as to the qualitative changes in the rate of low birthweight among blacks due to a sudden drop in the use of abortion caused by restricted access. The magnitude of the change as well as its duration are more speculative.

4. Conclusion

We have examined the relationship between induced abortion and low birthweight with aggregate time-series data from New York City. For blacks, we found a unidirectional relationship between decreases in the percentage of pregnancies terminated by induced abortion and increases in the rate of low-birthweight births. There was no such relationship among whites. Simulations based on the model reveal that an unanticipated decrease in the percentage of pregnancies that are voluntarily terminated would have a substantial impact on the rate of low birthweight among blacks.

The estimates were obtained with a vector autoregression – a somewhat controversial technique used primarily by macroeconomists. A major disagreement is whether VARs are an appropriate means of testing causality or exogeneity [Leamer (1985)]. Nevertheless, even detractors of VARs acknowledge their usefulness as a means of summarizing data, testing temporal orderings, or uncovering relationships that ought to be confronted by researchers [Cooley and LeRoy (1985), Backus (1986)]. It is in this spirit that we felt a VAR represented a suitable methodology.

For several reasons, however, the relationship between low birthweight and induced abortion reported in this paper represents more than an interesting association. First, our findings are similar qualitatively with results obtained from a cross-sectional analysis of individual birth and abortion certificates from New York City [Grossman and Joyce (1990)]. Although models with individual data provide only indirect evidence, the concordance between two very different approaches is noteworthy. Second, unlike many of the relationships in macroeconomics, the temporal ordering between low birthweight and induced abortion is well defined. This provides an important check on the specification since feedback from low birthweight to abortion would strongly imply that a 'third' variable was missing. Third, the length of a pregnancy provides a meaningful restriction on the number of relevant lags, a constraint often lacking in macroeconomic applications.

The downside of VARs in the present context is that they offer little insight as to the underlying mechanism at work. Induced abortion may reduce unwanted childbearing, which may increase the aggregate level of care that is given to pregnancies that are not terminated [Joyce (1987), Joyce and Grossman (1990)]. Induced abortion may also reduce the proportion of unhealthy fetuses. Advances in prenatal diagnostic techniques have dramatically increased the information available to parents as to the well-being of the fetus. Although there is presently no evidence to suggest that fetal defects are a major reason for induced terminations, the spread of AIDS, the growth in substance abuse, and the wider diffusion of these diagnostic techniques may be increasing this type of selective abortion.

Finally, the results are important given the changes that may occur in the availability of legalized abortion. If legal restrictions on abortion increase the proportion of pregnancies that are carried to term, then our results suggest that the incidence of low birthweight among blacks in New York City will rise. However, even if *Roe versus Wade* were overturned, New York State is unlikely to inhibit access to legalized abortion. The question then becomes to what extent do these results generalize to other parts of the country. Any response is purely speculative because New York City is such a unique setting.[15] Nevertheless, in areas where induced abortion is a prevalent means of fertility control, restrictions on legalized abortion could have an adverse effect on birth outcomes.

[15]In 1984, for example, 46% of all pregnancies in New York City were terminated by induced abortion. Even among white, non-Hispanic women the figure was 45% [Joyce and Grossman (1990)]. The average of the 12 other states that reported induced abortion to the National Center of Health Statistics was 21% [NCHS (1987, 1986)].

References

Backus, D., 1986, The Canadian–U.S. exchange rate: Evidence from a vector autoregression, Review of Economic and Statistics 68, 628–637.

Cooley, T.F. and LeRoy, S.F., 1985, Atheoretical macroeconomics: A critique, Journal of Monetary Economics 16, 283–308.

Corman, H. and M. Grossman, 1985, Determinants of neonatal mortality rates in the U.S., Journal of Health Economics 4, 213–236.

Doan, T. and R. Litterman, 1987, Rats user's manual: Version 2.11 (VAR Econometrics, Evanston, IL).

Godfrey, L.G., 1978, Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables, Econometrica 46, 1303–1310.

Granger, C.W.J., 1969, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37, 424–438.

Grossman, M. and S. Jacobowitz, 1981, Variations in infant mortality rates among counties of the United States: The roles of public policies and programs, Demography 18, 695–713.

Grossman, M. and T. Joyce, 1990, Unobservables, pregnancy resolutions, and birthweight production functions in New York City, Journal of Political Economy 98, 983–1007.

Institute of Medicine, 1985, Preventing low birthweight (National Academy Press, Washington, DC).

Joyce, T., 1987, The impact of induced abortion on black and white birth outcomes in the U.S., Demography 24, 229–244.

Joyce, T. and M. Grossman, 1990, Pregnancy wantedness and the early initiation of prenatal care, Demography 27, 1–17.

Leamer, E.E., 1985, Vector autoregressions for causal inference?, in: K. Brunner and A. Meltzer, eds., Understanding monetary regimes (supplement to Journal of Monetary Economics 22) 255–304.

Lutkepohl, H., 1982, Non-causality due to omitted variables, Journal of Econometrics 19, 367–378.

Maddala, G.S., 1988, Introduction to econometrics (Macmillan, New York).

National Center for Health Statistics, 1986, Advanced report of final natality statistics, 1984. Monthly Vital Statistics Report 35, no. 4, Suppl. DHHS Pub. No. (PHS) 86–1120. (Public Health Service, Hyattsville, MD).

National Center for Health Statistics, E. Powell-Griner, 1987, Induced terminations of pregnancy: Reporting States, 1984. Monthly Vital Statistics Report 36, no. 5, Suppl. (2) DHHS Pub. No. (PHS) 87–1120. (Public Health Service, Hyattsville, MD).

National Center for Health Statistics, Kenneth D. Kochanek, 1989, Induced terminations of pregnancy: Reporting states, 1985 and 1986. Monthly Vital Statistics Report 37, no. 12, Suppl. DHHS Pub. No. (PHS) 89–1120. (Public Health Service, Hyattsville, MD).

New York City Department of Health, 1985, Summary of vital statistics, the City of New York (New York).

Said, S.E. and D.A. Dickey, 1984, Testing for unit roots in autoregressive moving-average models of unknown order, Biometrika 71, 599–607.

Sims, C., 1980, Macroeconomics and reality, Econometrica 48, 1–48.

Stock, J.H. and M. W. Watson, 1988, Variable trends in economic time series, Economic Perspectives 2, 147–174.

Torres, A. and J.D. Forrest, 1988, Why do women have abortions?, Family Planning Perspectives 20, 169–1176.

Zellner, A., 1962, An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, Journal of the American Statistical Association 57, 348–368.