International Journal of Technology Assessment in Health Care, 15:1 (1999), 123–135.

ESTIMATING THE EFFECT OF CESAREAN SECTION RATE ON HEALTH OUTCOME

Evidence from Swedish Hospital Data

Ingemar Eckerlund
Stockholm School of Economics and The National Board of Health and Welfare

Ulf-G. Gerdtham
Stockholm School of Economics

Abstract

This paper tests the null hypothesis of a zero effect of cesarean section rate on health outcome against the alternative of a positive effect. Using data from 59 hospitals in Sweden from 1988–92, we specify two separate linear regression models for health outcome, one with perinatal mortality, and the other with rate of asphyxia, as dependent variable. We estimate the models by single-year cross-section regressions and as pooled data systems. The null hypothesis cannot be rejected, i.e., we do not find any significant positive effect of cesarean section rate on health outcome. Thus, we conclude that an increase in cesarean section rate does not imply lower perinatal mortality or lower rate of asphyxia. This in turn indicates that the minimum cesarean section rate is optimal.

Keywords: Cesarean Section, Health Outcome, Regression Analysis, Economic Consequences

In Sweden, as in many other comparable countries, there has been a considerable and persisting variation in cesarean section rates among hospitals (obstetrical departments) (3;7;14). In a cross-sectional study of 59 hospitals in 1991, we identified some 20 determinants of cesarean section rate, demand-related as well as supply-related (5). A general model including all these regressors was specified. After reducing this model, we were able to explain about one-quarter of the variation. Since only a minor part of the variation could be explained by differences in demand- and supply-related variables, and since a cesarean section is more resource-demanding than a vaginal delivery, we concluded that the variation in cesarean section rates indicates inefficiency. What, then, is the optimal rate of cesarean section?

Financial support from the Swedish Council for Social Research and the Swedish Institute for Health Services Development is gratefully acknowledged.

To answer that question, we must estimate the effect of cesarean section rate on health outcome, i.e., estimate the health production function. While there have been numerous studies of the determinants of variation in cesarean section rates, there is little information on the consequences of differing cesarean rates on outcome variables. A critical issue is how to define appropriate outcome measures. The main reason to perform a cesarean section is to improve the health outcome for the newborn,[1] and the traditional measure used to assess outcome in obstetric care is perinatal mortality (15). Perinatal mortality rates have decreased in recent decades while there has been an upward trend in cesarean section rates. It seems logical to apply a cause-and-effect relationship between these phenomena. However, on closer investigation, there is not much to support such a relationship. Available data suggest that the cesarean section rate does not contribute much to explaining the variation in perinatal mortality (2).

International comparisons demonstrate that perinatal mortality has generally improved, regardless of whether the cesarean section rates have increased, decreased, or remained stable over time (16). Several studies have confirmed this finding and failed to show a strong correlation between cesarean section rate and perinatal mortality. Furthermore, it has been shown that it is possible to lower the cesarean section rate on a nationwide basis without increasing risks to newborn infants (15).

Joffe et al. (10) compared groups of obstetric units, categorized according to their cesarean section rates, to see what effect the rate has on fetal and maternal outcomes.[2] They found that for very low-birth-weight-infants, but not for other deliveries, mortality rates were lower in units with higher cesarean section rates.[3] Apgar scores showed no trend, but the onset of respiration after 1 minute was significantly more frequent in units with a cesarean section rate of less than 10%. Increased maternal postnatal blood transfusion was associated with higher cesarean section rates, but no trend was observed for the other maternal variables.

In summary, there is good reason to consider outcome measures other than perinatal mortality, such as a measure of morbidity. Furthermore, to assess the effectiveness of cesarean section with respect to outcome (mortality and morbidity), it is necessary to control for case-mix, or alternatively, to look at outcome for specific indications rather than overall outcome.

This paper tests the null hypothesis of a zero effect of cesarean section rate on health outcome against the alternative of a positive effect. We first briefly describe the data and the variables selected. A description of the estimation methods used are followed by the results. Finally, a discussion of our findings and some concluding remarks on the need for further research are presented.

DATA AND MEASUREMENT

Using data from the Swedish Medical Birth Registry from 1988–92 (14), we compiled a special database containing yearly data from the 59 hospitals—covering 97% of all newborns—in the abovementioned study (5) on the number of newborns, deliveries, perinatal deaths, cesarean sections, acute cesarean sections, mothers older than 35 years, multiple births (twins or more), cases with asphyxia (Apgar score < 7 at 5 minutes),[4] and cases with indications 1–11, respectively (placenta praevia = 1, ablatio = 2, disproportion = 3, pre-eclampsia = 4, dystocia = 5, fetal distress =

Table 1. Description of Dependent and Independent Variables, Average Value, and Coefficient of Variation, 1988–91

Variable	Description		Average value	CV
Dependent variables				
Permort	Percentage:	Perinatal deaths	0.59	0.44
Asphyx		With asphyxia	1.08	0.39
Independent variables				
C-section	Percentage:	Cesarean sections	10.74	0.20
Age		Mothers > 35 years	11.13	0.18
Twins[a]		Twin (or more) births	1.37	0.44
Ind1		Placenta previa	0.24	0.63
Ind2		Ablatio	0.46	0.61
Ind3		Disproportion	2.51	0.36
Ind4		Pre-eclampsia	1.76	0.51
Ind5		Dystocia	6.94	0.77
Ind6		Fetal distress	5.86	0.36
Ind7[a]		Multiple births	1.10	0.43
Ind8		Malpresentation	7.10	0.19
Ind9		Psychosocial ind	2.36	0.50
Ind10		Other indication	0.83	0.96

Abbreviations: CV = coefficient of variation; ind = indication.
[a] *Twins* and *Ind7* both measure multiple births (twins or more). However, while *twins* measure all cases, *Ind7* measures only those cases where it is the main indication.

6, multiple births = 7, malpresentation = 8, psychosocial indication = 9, prolonged delivery = 10, other indication = 11).

Based on this database, we defined the variables shown in Table 1. Besides the traditional outcome measure, perinatal mortality, defined as "percentage of stillbirths and newborns dead within first week," we also used a morbidity measure, rate of asphyxia, defined as "percentage of newborns with Apgar score less than 7 at five minutes." The independent variable in focus is the cesarean section rate. The variables age, twins, and indications 1–10 are intended to measure case-mix. Cesarean section rate is assumed to be negatively correlated with perinatal mortality and rate of asphyxia, while all case-mix variables are assumed to be positively correlated with the outcome measures.

Table 2 contains data on average values and coefficients of variation (CV) of perinatal mortality, rate of asphyxia, and cesarean section rates in Sweden from

Table 2. Perinatal Mortality, Rate of Asphyxia, and Cesarean Section Rate, 1988–92; Percent, Average Values, and Coefficient of Variation

	Perinatal mortality		Rate of asphyxia		Cesarean section rate	
Year	Average	CV	Average	CV	Average	CV
1988	0.60	0.45	1.29	0.32	10.89	0.18
1989	0.62	0.44	1.23	0.35	10.70	0.17
1990	0.61	0.41	1.03	0.36	10.73	0.18
1991	0.64	0.36	0.95	0.34	10.82	0.19
1992	0.50	0.50	0.92	0.46	10.73	0.25
1988–92	0.59	0.44	1.08	0.39	10.74	0.20

Abbreviation: CV = coefficient of variation.

1988–92. Obviously, the variation among hospitals has been large and persisting, especially regarding the outcome measures. This is confirmed by historical data. The average cesarean section rate increased rapidly during the 1970s, while the perinatal mortality and the rate of asphyxia decreased. Superficially, this seems to indicate that the increase in cesarean sections contributed to improved medical outcome. From the early 1980s on, however, the cesarean section rate stabilized and even showed a slight decrease, while the perinatal mortality and the rate of asphyxia continued to decrease. Notwithstanding, the variation among hospitals has remained large (14).

ESTIMATION METHODS

Obstetrical departments with higher cesarean section rates should, other things being equal, have better health outcomes in terms of lower perinatal mortality and/or lower rates of asphyxia. In other words, if we could find two obstetrical departments with exactly the same case-mix, we would expect the one with the higher cesarean section rate to have the best health outcome. Otherwise, if no such relationship can be established, the lower cesarean section rate would be optimal.

We first specify the following regression equations for health outcome, $HO_{i,t}$ (perinatal mortality or rate of asphyxia), at department i, in a specific time period, t (year):

$$HO_{i,t} = \beta_{1,t} + \beta_{2,t} * CSR_{i,t} + \beta_{3,t} * X_{3i,t} + ... + \beta_{14,t} * X_{14i,t} + \epsilon_{i,t} \qquad (i = 1,2,.....,59) \quad (t = 1988\text{–}92)$$

where $\beta_{1,t}$ are the intercepts, β_2,t...$\beta_{14,t}$ the coefficients, $CSR_{i,t}$ the cesarean section rates, and $X_{3i,t},...,X_{14i,t}$ a set of variables measuring case-mix, i.e., age, twins, and indications 1–10[5] (defined in Table 2). The stochastic disturbance terms, $\epsilon_{i,t}$, are assumed to be normally distributed with mean zero and variance σ_i^2, where $\sigma_i^2 = \sigma_*^2 N_i$, and N_i is the number of newborns at department i. By controlling for case-mix, we could, theoretically, eliminate the effect of a confounding factor that influences both the dependent and the independent variable. We test the null hypothesis, $\beta_{2,t} = 0$, against the alternative, $\beta_{2,t} < 0$. Note that a negative sign implies a positive effect on health outcome.

We estimate the models equation by equation by use of weighted least squares (WLS), and test for poolability in order to gain in degrees of freedom if pooling is appropriate and to assess the stability of the estimated regressions. The weighting variable is the reciprocal of the variance, $w_i = 1/(N_i)^2$.

We test for poolability by use of conventional F tests (Chow test) for non-normality of the residuals by Jarque-Bera's skewness-kurtosis test, and for homoscedasticity by the Breusch-Pagan test. Functional form misspecification is tested by the Ramsey RESET test (1;9;13). To test for robustness of the results, we reestimate the models using 5-year average values of all variables. We also estimate the models excluding all but two case-mix variables. These additional tests are discussed below.

RESULTS

The results of the regression analyses are summarized in Table 3 (perinatal mortality) and Table 4 (rate of asphyxia). Comparison between different years clearly demonstrates instability in the coefficients for cesarean section rate and for the case-mix variables. In both models, the coefficient for cesarean section changes sign as well as magnitude between periods. However, it does not differ significantly from zero except in one period (perinatal mortality, 1992). Thus, the findings do not indicate that higher cesarean section rates result in lower perinatal mortality or lower rates of asphyxia.[6]

The Jarque-Bera test (J-B), applied to test for normality in residuals, showed that in most cases the null hypothesis could not be rejected at the 10% level. For two models,[7] however, it was rejected at the 1% level. From the Breusch-Pagan test (B-P) it appeared that most of the models showed heteroscedastic residual variance.[8] The RESET test indicates functional form misspecification for three models.[9]

We also ran WLS regressions of reduced models with two control variables only, age and twins, which—in contrast to indications 1–10—are objectively measurable. This change influences the coefficients for cesarean section, with respect to sign as well as significance. There is a significant[10] positive effect of cesarean section rate on perinatal mortality in two single-year cross-section regressions (1988, 1990) and in the pooled regression (1988–92), and on rate of asphyxia in the single-year cross-section regression for 1988. Otherwise, the main difference was that the coefficient for twins became significantly different from zero in most periods. The hypothesis that all coefficients excluded in the reduced models are equal to zero (slope restrictions) was tested using a joint F test, and was rejected for all but three periods (perinatal mortality, 1988; rate of asphyxia, 1988 and 1991). Detailed estimation results are available from the authors.

An assessment of multicollinearity is presented in Table 5.[11] The first column reports the internal coefficients of determination for the regressor variables. The remaining columns contain coefficients of correlation between the regressor variables in the pooled regression equation. The highest variance inflation factor (VIF) for the regressors included is 3.846.[12] A rule of thumb is that a VIF greater than 10 indicates strong multicollinearity (12). Our conclusion is that the individual *t* ratios should not be too much affected by multicollinearity. However, the correlation coefficient between the variables twins and indication 7 is high, 0.82, and the internal R^2s for twins and indication 7 is 0.68 and 0.74, respectively. This made us re-estimate the models excluding either of these variables. However, this did not change the results very much. The coefficient for cesarean section rate changed from negative to positive in 1988 for perinatal mortality, and became significantly positive in 1992 for rate of asphyxia.

The Chow tests for equal slopes across equations show that the null hypotheses could be rejected at the 1% level for perinatal mortality, $F(52;45) = 3.23$, and at the 5% level for rate of asphyxia, $F(52;45) = 1.79$. Tests for equal slopes *and* equal intercepts across equations show that the null hypotheses could be rejected at the 1% level for both perinatal mortality, $F(56;45) = 3.53$, and rate of asphyxia, $F(56;45) = 3.61$, which indicates that pooling is not justified in this case. Moreover, pooling does not change the results.[13]

DISCUSSION AND CONCLUDING REMARKS

In the introduction, we stated that only a minor part of the variation in cesarean section rates among obstetrical departments in Sweden is explained by differences in demand- and supply-related variables. This indicates inefficiency. Some departments perform too many cesareans and some other departments too few. To find out which rate is optimal, we have to estimate the health production function, i.e., the effect of cesarean section rate on health outcome. The objective of this study was to estimate the effect of cesarean section rate on health outcome in terms of perinatal mortality and rate of asphyxia, respectively.

https://doi.org/10.1017/S0266462399152127 Published online by Cambridge University Press

Table 3. Parameter Estimates of Perinatal Mortality: Equation-by-Equation WLS and Pooled WLS; Coefficients (*t* values) for Regressors

Variable	1988	1989	1990	1991	1992	Pooled
Constant	0.777E-01	0.919E-01	−0.665[b]	0.546[c]	0.360	0.147
	(0.201)	(0.221)	(−2.378)	(1.469)	(1.135)	(1.080)
CSR	−0.409E-02	0.205E-01	0.847E-02	0.848E-02	−0.225E-01[c]	−0.224E-02
	(−0.165)	(0.722)	(0.453)	(0.584)	(−1.418)	(−0.238)
Age	0.700E-01[b]	0.188E-01	0.383E-01[b]	0.87E-02	0.126E-01	0.385E-01[a]
	(2.333)	(0.651)	(1.974)	(0.496)	(0.601)	(3.573)
Twins	−0.201	0.957E-01	−0.344[a]	0.101	0.713E-01	−0.113[a]
	(−0.846)	(0.876)	(−3.854)	(0.727)	(1.064)	(−2.950)
Ind1	0.277E-01	−0.335	0.166	0.288[b]	0.235	−0.178E-01
	(0.075)	(−0.748)	(0.953)	(1.960)	(1.219)	(−0.170)
Ind2	−0.272E-01	0.303[c]	0.29[b]	−0.217E-01	0.137	0.136[b]
	(−0.145)	(1.503)	(2.092)	(−0.231)	(1.166)	(2.193)
Ind3	0.479E-01	0.507E-02	0.492E-01[c]	−0.231E-01	−0.858E-01[a]	−0.299E-02
	(1.081)	(0.088)	(1.604)	(−0.686)	(−3.010)	(0.168)
Ind4	0.449E-02	0.163[a]	−0.669E-01[b]	−0.659E-01[b]	−0.457E-01	0.178E-01
	(0.098)	(3.007)	(−2.204)	(−1.917)	(−1.089)	(−0.912)
Ind5	−0.912E-02	−0.861E-02	−0.351E-02	0.182E-01[a]	0.105E-02	−0.601E-02[c]
	(−0.758)	(−0.692)	(−0.507)	(2.777)	(0.176)	(−1.391)
Ind6	−0.583E-01[b]	−0.358E-01[c]	0.227E-01[c]	0.410E-01[a]	−0.294E-01[c]	−0.209E-01[b]
	(−2.263)	(−1.621)	(1.431)	(2.518)	(−1.624)	(−2.252)
Ind7	0.703[a]	0.137	0.301[b]	0.216	0.146[c]	0.338a
	(2.536)	(0.830)	(2.388)	(1.209)	(1.489)	(6.336)
Ind8	−0.546E-01	−0.227E-01	0.620E-01[a]	−0.109[a]	0.412E-01	−0.837E-02
	(−1.378)	(−0.443)	(2.481)	(−2.479)	(1.094)	(−0.670)
Ind9	−0.692E-01	−0.185E-02	0.671E-01[b]	0.509E-03	−0.315E-01	0.408E-02
	(−0.901)	(−0.027)	(2.347)	(0.016)	(1.111)	(0.200)
Ind10	−0.743E-02	−0.581E-01	0.150E-01	0.124a	0.158E-01	0.406E-02
	(−0.111)	(−1.044)	(0.380)	(3.274)	(−0.506)	(0.200)

(*Continued*)

Table 3. (*Continued*)

Variable	1988	1989	1990	1991	1992	Pooled
Evaluation:						
df	45	45	45	45	45	281
R	0.615	0.476	0.753	0.614	0.562	0.281
Standard error	0.25	0.29	0.17	0.18	0.19	0.28
J-B χ^2 (2)	0.534	0.230	0.902	11.438[a]	0.308	3.301
B-P χ^2 (13)	14.07	34.47[a]	23.77[b]	39.70[a]	24.12[b]	49.16[a]
RESET F(3;42)	4.640[a]	0.726	0.914	2.404	1.053	9.256[a]

Abbreviations: WLS = weighted least squares; CSR = cesarean section rate; J-B = Jarque-Bera test; B-P = Breusch-Pagan test; ind = indication.
[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 10% level.
All tests are one-tailed.

Table 4. Parameter Estimates of Rate of Asphyxia: Equation-by-Equation WLS and Pooled WLS; Coefficients (*t* values) for Regressors

Variable	1988	1989	1990	1991	1992	Pooled
Constant	0.872[c]	1.833[a]	−0.659	0.846	1.039[b]	0.956[a]
	(1.581)	(2.866)	(−1.133)	(1.288)	(2.105)	(4.355)
CSR	0.307E-01	−0.191E-01	0.124E-01	0.497E-02	0.177E-01	0.851E-02
	(0.868)	(−0.438)	(0.318)	(0.194)	(0.714)	(0.558)
Age	0.242E-01	−0.101[b]	0.119E-01	−0.161E-01	−0.724E-01[b]	−0.379E-02
	(0.567)	(−2.274)	(0.296)	(−0.517)	(−2.226)	(−0.218)
Twins	−0.179	0.514E-01	0.159E-01	0.116	0.225[b]	−0.100[c]
	(−0.531)	(0.305)	(0.085)	(0.472)	(2.162)	(−1.611)
Ind1	−0.557E-01	0.152	0.485[c]	−0.929E-01	0.938[a]	0.261[c]
	(−0.106)	(0.220)	(1.336)	(−0.357)	(3.131)	(1.544)
Ind2	−0.270E-01	−0.861E-01	−0.463E-01	−0.179E-01	0.133	−0.124
	(−0.102)	(−0.277)	(−0.160)	(−0.108)	(0.724)	(−1.237)
Ind3	−0.348E-01	−0.129[c]	−0.175E-02	−0.139[b]	−0.164[a]	−0.236E-01
	(−0.551)	(−1.452)	(−0.027)	(−2.339)	(−3.694)	(−0.824)
Ind4	0.727E-01	−0.220[a]	−0.141E-01	−0.179E-01	−0.164[a]	−0.325E-01
	(1.119)	(−2.620)	(−0.224)	(−0.295)	(−2.508)	(−1.035)
Ind5	0.262E-01[c]	0.340E-01[b]	−0.226E-02	0.113E-01	0.222E-01[b]	0.915E-02[c]
	(1.526)	(1.774)	(−0.157)	(0.970)	(2.380)	(1.312)
Ind6	−0.693E-01[b]	0.286E-01	0.634E-01[b]	0.248E-01	−0.160E-02	0.158E-01
	(−1.889)	(0.842)	(1.923)	(0.863)	(−0.057)	(1.058)
Ind7	0.931[b]	0.319	−0.280	0.258	−0.204E-01	0.478[a]
	(2.358)	(1.250)	(−1.067)	(0.819)	(−0.133)	(5.542)
Ind8	−0.714E-01	0.163[b]	0.105[b]	0.143E-02	0.351E-01	−0.473E-01[a]
	(−1.265)	(2.069)	(2.011)	(0.018)	(0.599)	(−2.347)
Ind9	−0.668E-01	−0.129	0.220[a]	−0.249E-01	0.603E-01[c]	0.185E-01
	(−0.610)	(−1.237)	(3.707)	(−0.456)	(1.368)	(0.560)
Ind10	−0.763E-01	−0.613E-01	−0.321E-01	−0.451E-01	−0.795E-01[c]	−0.462E-01[c]
	(−0.798)	(−0.714)	(−0.390)	(−0.674)	(−1.631)	(−1.412)

(Continued)

https://doi.org/10.1017/S0266462399152127 Published online by Cambridge University Press

Table 4. *(Continued)*

Variable	1988	1989	1990	1991	1992	Pooled
Evaluation:						
df	45	45	45	45	45	281
R	0.559	0.458	0.496	0.451	0.675	0.231
Standard error	0.35	0.45	0.36	0.32	0.29	0.44
J-B χ^2 (2)	1.088	0.356	0.929	1.420	0.177	11.866[a]
B-P χ^2 (13)	38.18[a]	22.18[c]	14.92	15.38	30.63[a]	43.16[a]
RESET F(3;42)	3.890[c]	0.115	0.957	2.236	0.890	0.057[a]

For abbreviations, see Table 3.
[a] Significant at the 1% level.
[b] Significant at the 5% level.
[c] Significant at the 10% level.
All tests are one-tailed.

https://doi.org/10.1017/S0266462399152127 Published online by Cambridge University Press

Table 5. Assessment of Multicollinearity by Internal Coefficients of Determination and the Correlation Matrix of Regressors

		R^2	1	2	3	4	5	6	7	8	9	10	11	12	13
1	CSR	0.36	1												
2	Age	0.30	.35	1											
3	Twins	0.68	.22	.30	1										
4	Ind1	0.24	.32	.27	.30	1									
5	Ind2	0.26	.17	.15	.19	.27	1								
6	Ind3	0.24	.33	*.01*	−.19	*−.01*	*.05*	1							
7	Ind4	0.20	*.09*	*.10*	.30	.23	.30	*−.02*	1						
8	Ind5	0.14	*.00*	.22	.12	*.03*	*−.02*	−.12	.16	1					
9	Ind6	0.23	*.01*	.25	.22	.15	.34	*.06*	.18	*.05*	1				
10	Ind7	0.74	.28	.39	.82	.38	*.27*	−.17	.30	.14	.25	1			
11	Ind8	0.31	.19	.16	.43	.21	.33	*−.04*	.29	*.04*	.26	.49	1		
12	Ind9	0.18	.14	.24	.17	*.06*	*.02*	−.16	*.09*	.28	*.02*	.28	*.07*	1	
13	Ind10	0.05	*.05*	*−.10*	*−.03*	*.05*	*−.03*	*−.03*	*−.03*	*.01*	−.14	*−.07*	*.01*	*−.02*	1

Abbreviations: CSR = cesarean section rate; ind = indication.
Figures in italics indicate nonsignificance at the 5% level.

An obstetrical department may have a high cesarean section rate because of a high prevalence of complicated deliveries, which makes it necessary or desirable to perform relatively many cesarean sections. When controlling for case-mix, however, we should expect to find a negative correlation between outcome (perinatal mortality, rate of asphyxia) and cesarean section rate. As shown above, this is generally not the case. The null hypothesis of a zero effect could not be rejected: the findings did not indicate that a higher cesarean section rate implies better health outcome in terms of perinatal mortality or rate of asphyxia. Thus, our results are in accordance with those of Joffe et al. (10), who found the same lack of correlation using a different methodological approach.

A critical issue in our analysis is whether the outcome and case-mix measures are appropriate. Regarding the outcome measures, we chose perinatal mortality and rate of asphyxia, both of which seem to be valid and readily available from the Medical Birth Registry (14;15). To be sure, these measures are relatively crude. Still, they are the only available measures suitable for comparisons across departments. Ideally, we should use a broader range of outcome measures, preferably combined into a health outcome index. With regard to case-mix, we tried alternative measures (reduced models), but these did not change the results.

The instability of the models is an obvious problem. The coefficients for cesarean section rate as well as for the case-mix variables change sign and magnitude from year to year. The reason for this is not clear. One possible explanation could be the fact that we deal with low-incidence events. However, the fact that several coefficients are significantly different from zero indicates that the observed variation is larger than would be expected by chance alone. This was confirmed by chi-square tests, showing that the hypothesis of no variation in the dependent variables could be rejected.

Another conceivable explanation has to do with the statistical power of the test. The fact that the sample is relatively small implies a high probability of type II error, i.e., not rejecting the null hypothesis when it is in fact false. By expanding the sample size, we could in principle reduce the variance of an estimator. However, it is not quite clear whether this would influence the problem with changing signs of the regressor coefficients.

Alternatively, the instability may indicate that the models are misspecified due to omission of variables or nonlinearity. The RESET test indicates functional form misspecification for some of the models. To check for nonlinearity, we standardized the dependent variables in the reduced models with respect to the variables age and twins, and looked at scatter diagrams showing the correlation between cesarean section rate and the standardized outcome variables. These inspections did not indicate the existence of either nonlinearity or linearity.

Assuming that our conclusion about inefficiency due to variation in cesarean section rates is correct, a natural next step is to estimate the economic consequences of this inefficiency, i.e., to measure the welfare loss from undesired medical practice variation (4;8). Since we did not find any significant effect of cesarean section rate on health outcome, there is in principle no reason for performing cesarean section rates above the minimum level (adjusted for case-mix). However, with regard to the problems discussed above, there is good reason for caution in assessing the validity of our results.

If we assume that, in fact, there is a significant positive effect of cesarean section rate on health outcome, we are back to the problem of estimating the optimal rate as a basis for estimating the welfare loss from undesired variation. In our previous study, we tried to do so, using a model developed by Phelps and Parente (17). One of the problems that remains to be solved is to estimate the marginal value (demand) curve for cesarean sections. Since there is no real market for health care in Sweden, there is no self-evident method to estimate the demand. Possibly, we could use the contingent valuation approach, e.g., by asking patients who are about to go through a delivery about their willingness to pay for (valuation of) a cesarean section, given specified risks and benefits (11).

Finally, we should also look closer at available policy options, i.e., the opportunities for changing practice towards the best practice as defined by science and proven experience. It is a well-known fact that it is difficult to influence practice by information alone (6). However, we are convinced that improved knowledge and information about the economic consequences of different medical practices would be a valuable basis for influencing practice.

NOTES

[1] More than two-thirds of cesarean sections are performed exclusively for the newborns' sake (15).

[2] The outcome measures used were: perinatal mortality, Apgar scores at 1 and 5 minutes, onset of respiration after 1 minute, postnatal transfusion, postnatal infection, thromboembolism, low hemoglobin concentration at discharge, and puerperal psychosis.

[3] This is contradicted by Rydhström et al. (18), who found no such correlation in twin deliveries.

[4] Apgar score is a method of judging the condition of a newborn baby. The baby is given a maximum of two points on each of five criteria: color of the skin, heartbeat, breathing, muscle tone and reaction to stimuli.

[5] Indication 11, "other indication," is a residual, and is not included in the analysis.

[6] Since perinatal mortality as well as rates of asphyxia are generally very low in Sweden, they are extremely sensitive to random variation. Therefore, we also ran WLS regressions based on average values of all variables in the 5-year period. This did not change the results concerning the coefficient for cesarean section rate, but gave a higher R^2 (0.792 for perinatal mortality and 0.670 for rate of asphyxia).

[7] 1991 for perinatal mortality and the pooled model for rate of asphyxia.

[8] We also estimated t values with the covariance estimator, which allows for heteroscedasticity, but this did not change the results (13).

[9] 1988 and pooled model for perinatal mortality, and 1988 for rate of asphyxia.

[10] At the 5% level for perinatal mortality in 1990, otherwise at the 10% level.

[11] Table 5 is based on the pooled data, but the results are similar for every year.

[12] The VIF for a regressor X_k is calculated as $(1-R_k^2)^{-1}$, where R_k^2 is the R^2 from regressing the kth independent variable on all the other independent variables.

[13] We also estimated the five equations as a set of linear equations, using Zellner's seemingly unrelated regressions (SUR, or error-related regressions), i.e., one set of equations with perinatal mortality as dependent variable and the other with rate of asphyxia as dependent variable. The disturbances across the regressions were allowed to be freely correlated. Equation-by-equation OLS were used to obtain an estimate of the disturbance covariance matrix Ω and then generalized least squares (GLS) was used on the "stacked" set of equations (19). However, neither did the SUR estimations change the main conclusion of the WLS regressions.

REFERENCES

1. Breusch, T., & Pagan, A. A simple test for heteroscedasticity and random coefficient variation. *Econometrica,* 1979, 47, 1287–94.
2. Broadhead, T. J., & James, D. K. Worldwide utilization of cesarean section. In B. L. Flamm & E. J. Quilligan (eds.), *Cesarean section: Guidelines for appropriate utilization.* New York: Springer-Verlag, 1995.

3. Chalmers, I., Enkin, M., & Keirse, M. *Effective care in pregnancy and childbirth. vol. 2.* Oxford: Oxford University Press, 1991.
4. Clark, L., Mugford, M., & Paterson, C. How does the mode of delivery affect the cost of maternity care? *British Journal of Obstetrics and Gynaecology,* 1991, 98, 519–23.
5. Eckerlund, I., & Gerdtham, U.-G. *Econometric analysis of variation in cesarean section rates: A cross-sectional study of 59 obstetrical departments in Sweden. International Journal of Technology Assessment in Health Care*, 1998, 14, 774–87.
6. Evans, R. G. The dog in the night: Medical practice variations and health policy. In T. F. Andersen & G. Mooney (eds.), *The challenges of medical practice variations.* London: Macmillan, 1990.
7. Flamm, B. L., & Quilligan, E. J. (eds). *Cesarean section: Guidelines for appropriate utilization.* New York: Springer-Verlag, 1995.
8. Gardner, L. B. Economic considerations in cesarean section use. In B. L. Flamm & E. J. Quilligan (eds.), *Cesarean section: Guidelines for appropriate utilization.* New York: Springer-Verlag, 1995.
9. Jarque, C. M., & Bera, A. K. A Test for normality of observations and regression residuals. *International Statistical Review,* 1987, 55, 163–72.
10. Joffe, M., Chapple, J., Paterson, C., & Beard, R. W. What is the optimal cesarean section rate? An outcome based study of existing variation. *Journal of Epidemiology and Community Health,* 1994, 48, 406–11.
11. Johansson, P.-O. *Evaluating health risks: An economic approach.* Cambridge: Cambridge University Press, 1995.
12. Kennedy, P. *A guide to econometrics,* 3rd ed. Oxford: Blackwell, 1992.
13. Kmenta, J. *Elements of econometrics.* 2nd ed. New York: Macmillan, 1990.
14. *Medical birth registration.* Stockholm: National Board of Health and Welfare, 1988–92.
15. Nielsen, T. F., Otterblad Olausson, P., & Ingemarsson, I. The cesarean section rate in Sweden: The end of the rise. *Birth,* 1994, 21, 1.
16. O'Driscoll, K., & Foley, M. Correlation of decrease in perinatal mortality and increase in CS rates. *Obstetrics and Gynecology*, 1983, 61, 1–5.
17. Phelps, C. E., & Parente, S. T. Priority setting for medical technology and medical practice assessment. *Medical Care,* 1990, 28, 703–23.
18. Rydhström, H., Ingemarsson, I., & Ohrlander, S. Lack of correlation between a high CS rate and improved prognosis for low-birthweight twins (<2500 g). *British Journal of Obstretrics and Gynaecology*, 1992, 97, 229–33.
19. Zellner, A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association,* 1962, 57, 348–68.