# Adoption of Diagnostic Technology and Variation in Caesarean Section Rates: A Test of the Practice Style Hypothesis in Norway

*Jostein Grytten, Lars Monkerud, and Rune Sørensen*

**Objective.** To examine whether the introduction of advanced diagnostic technology in maternity care has led to less variation in type of delivery between hospitals in Norway.

**Data Sources.** The Medical Birth Registry of Norway provided detailed medical information for 1.7 million deliveries from 1967 to 2005. Information about diagnostic technology was collected directly from the maternity units.

**Study Design.** The data were analyzed using a two-level binary logistic model with Caesarean section as the outcome measure. Level one contained variables that characterized the health status of the mother and child. Hospitals are level two. A heterogeneous variance structure was specified for the hospital level, where the error variance was allowed to vary according to the following types of diagnostic technology: two-dimensional ultrasound, cardiotocography, ST waveform analysis, and fetal blood analyses.

**Principal Finding.** There was a marked variation in Caesarean section rates between hospitals up to 1973. After this the variation diminished markedly. This was due to the introduction of ultrasound and cardiotocography.

**Conclusion.** Diagnostic technology reduced clinical uncertainty about the diagnosis of risk factors of the mother and child during delivery, and variation in type of delivery between hospitals was reduced accordingly. The results support the practice style hypothesis.

**Key Words.** Medical practice variation, Caesarean section, diagnostic technology, uncertainty, practice style

One of the most widely documented phenomena within health care is the substantial variation in utilization of medical services by different populations (for comprehensive reviews see Paul-Shaheen, Clark, and Williams 1987; Wennberg 2010). According to Wennberg, Barnes, and Zubkoff

(1982), variation in medical care to a large extent exists because of clinical uncertainty about diagnosis and treatment. This uncertainty has resulted in lack of unambiguous diagnostic criteria and appropriate standards for the care that is given. Differences among physicians in their beliefs about how to interpret diagnostic signs and the benefits of alternative interventions lead to differences in clinical practice between physicians—they have developed their own practice style. The mix of physicians with different practice styles varies between areas. Therefore, level of utilization of health care per capita also varies between areas (Grytten and Sørensen 2003). The proposition that different practice styles lead to differences in health care utilization per capita has been termed the practice style hypothesis.

The practice style hypothesis is difficult to test empirically, mainly because practice style and uncertainty are not directly observable (Stano 1991). In the present study we performed a fairly direct test. We exploit the fact that the introduction of advanced diagnostic technology in maternity care during the last 30–40 years has markedly reduced uncertainty about which type of delivery to choose. Diagnostic technology helps to provide objective and clear criteria for when a Caesarean section is indicated. The clinical basis for making decisions in maternity care is improved. Our research question is whether this reduction in uncertainty about diagnosis of risk factors of the mother and the child during delivery has led to less variation in type of delivery. If this is the case, this provides support for the practice style hypothesis.

Below we first describe the background for the study—among other things the role of advanced diagnostic technology in maternity care. We tested our research question using a large and unique set of data, which contained detailed medical information about all deliveries in Norway during the period 1967–2005 (about 1.7 million). As far as we know, there are no international studies in which practice variation has been analyzed using individual data over such a long period of time. Another strength of our study is that the data contained a large number of variables that described the health status of the mother and child. This reduces the possibility that the effect of diagnostic

Address correspondence to Jostein Grytten, Ph.D., Professor, Section of Community Dentistry, University of Oslo and Department of Gynecology and Obstetrics, Akershus University Hospital, Section of Community Dentistry, Box 1052, Blindern, 0316 Oslo, Norway; e-mail: josteing@ odont.uio.no. Lars Monkerud Researcher, Ph.D., is with the Department of Economics, Norwegian Business School, Oslo, Norway and the Department of Gynecology and Obstetrics, Akershus University Hospital, Oslo, Norway. Rune Sørensen, Professor, Ph.D., is with the Department of Economics, Norwegian Business School, Oslo, Norway.

technology on variation in Caesarean section rates is biased due to the lack of sufficient control for risk factors for Caesarean section.

## BACKGROUND

### *Caesarean Sections and Diagnostic Technology*

The number of Caesarean sections has increased in many countries during the last few decades (Mayor 2002; Betrán et al. 2007; MacDorman, Menacker, and Declerq 2008). This has also been the case in Norway, where the Caesarean section rate increased from 2.2 percent in 1970 to 16.4 percent in 2005 (Backe, Heggestad, and Lie 2003; Norwegian Institute of Public Health 2008). Numerous studies have shown wide variation in Caesarean sections within countries and between countries (Renwick 1991; McKenzie and Stephenson 1993; Hemminki and Gissler 1994; Baicker, Buckles, and Chandra 2006; Betrán et al. 2007; Clark et al. 2007; Bragg et al. 2010). For example, in the United States Caesarean section rates vary fourfold between low and high use areas, and in England the rates vary threefold between National Health Trusts (Baicker, Buckles, and Chandra 2006; Bragg et al. 2010). Nearly all studies are descriptive and cross-sectional, and the practice style hypothesis has not been tested in any of the studies within maternity care.

The rapid development in diagnostic technology has improved fetal monitoring both before and during delivery. The obstetrician is then less dependent on clinical signs, judgment, and interpretation of information from the mother for assessing whether the delivery is progressing without complications. The following diagnostic tools are important (Norwegian Institute of Public Health 2007; Norwegian Society of Gynecology and Obstetrics 2008): two-dimensional ultrasound, cardiotocography, ST waveform analysis (STAN) and fetal blood analysis. Ultrasound is used to determine the number of fetuses, and to check fetal circulation and anatomy, both before and during the delivery. Cardiotocography, STAN, and fetal blood analysis are used to register fetal distress, a condition that can lead to lack of oxygen, and for which Caesarean section may be indicated. Ultrasound and cardiotocography are expected to have the greatest effect on reducing practice variation in Caesarean section rates between hospitals. The reasons are twofold.

First, ultrasound and cardiotocography were introduced early in the period (Table 1). If these two types of diagnostic technology have been effective in reducing practice variation, there would be little variation left when STAN and fetal blood analysis were introduced. The inter-hospital variation in

Table 1:    The Percentage of Deliveries According to Type of Diagnostic
Technology and Year

| Diagnostic Technology | 1970 | 1975 | 1980 | 1985 | 1990 | 1995 | 2000 | 2005 |
|---|---|---|---|---|---|---|---|---|
| Ultrasound | 0 | 19 | 78 | 97 | 99 | 100 | 100 | 100 |
| Cardiotocography | 0 | 36 | 86 | 96 | 100 | 100 | 100 | 100 |
| ST waveform analysis | 0 | 0 | 0 | 0 | 15 | 30 | 38 | 77 |
| Fetal blood analysis | 0 | 0 | 19 | 25 | 36 | 47 | 62 | 72 |
| Technology index (0–4) | 0 | 0.54 | 1.82 | 2.18 | 2.50 | 2.77 | 3.00 | 3.48 |

Caesarean section rates would then be so low that the introduction of STAN
and fetal blood analysis would not contribute much to an additional reduc-
tion.

Second, for ultrasound and cardiotocography, there are fairly clear med-
ical criteria for the conditions and test values that are regarded as normal. For
example, for cardiotocography, the sign that something is wrong is based on
deviations from the following criteria: "a baseline fetal heart rate frequency
between 110 and 150 beats per minute, presence of periodic accelerations, a
normal heart rate variability with a bandwidth between 5–25 beats per minute
and the absence of decelerations" (Rooth, Huch, and Huch 1987; Van Geijn
1998). There is agreement that in the case of deviation from these test criteria,
the child must be delivered without delay, preferably by Caesarean section.
For both gestation length and fetal growth, which are determined by ultra-
sound, there are fairly clear criteria for normal values, and indications for
Caesarean section (Morrison, Rennie, and Milton 1995).

Before cardiotocography and ultrasound were used, the possibilities to
assess the condition of the fetus were limited. Obstetricians were dependent
on clinical assessment and judgment, which we can expect led to large varia-
tion in choice of method of delivery. For example, the absence of fetal move-
ments presented a diagnostic problem for which there was no consensus
among obstetricians about choice of method of delivery (Van Geijn 1998).

For STAN and fetal blood analysis it can be difficult to ascertain whether
a test value indicates that something is wrong with the fetus (Westgate, Gari-
baldi, and Greene 1994; Dunton 1999). For example, for STAN, deviation
from normal progression is assessed, among other things, on the basis of
T/QRS gradient in the electrocardiogram of the fetus. This is a complicated
technology that requires advanced skills both to use and to interpret the results
(Luzietti et al. 1999; Amer-Wahlin et al. 2007). In Norway, authorized educa-
tion and certification, and continual follow-up and training are required for

people who use STAN (Eikeland et al. 2008). STAN and fetal blood analysis is most appropriate for high-risk patients, and they are nearly always used as supplementary investigations to ultrasound and/or cardiotocography (Amer-Wahlin et al. 2001; Swedish Council on Health Technology Assessment 2006). For these patients, it has often been decided well before the delivery that the method of delivery will be Caesarean section; that is, there is hardly any variation in type of delivery.

*Institutional Setting – Maternity Services in Norway*

Studying variation in Caesarean section rates in Norway not only has some advantages but also some disadvantages. There are two advantages. First, in Norway, all health services, maternity care included, are financed through taxes, and everyone has free health care at the point of delivery and equal access given equal need (Ministry of Health 2002). There is little competition between hospitals for women giving birth. The country is divided into hospital areas in which the capacity of maternity units is planned according to the expected number of births within the catchment area. Thus, variation in Caesarean section rates between hospitals cannot be explained by differences in fees, income, insurance coverage, and rationing of deliveries. Second, all hospitals, where nearly all deliveries take place, are publically owned and financed, with obstetricians who receive a fixed salary. Neither the hospitals nor the obstetricians have economic incentives for carrying out a Caesarean section rather than a normal delivery. Thus, variation in Caesarean section rates between hospitals is not due to differences in the way obstetricians are remunerated.

     The main disadvantage of studying variation in Caesarean section rates in Norway is that it is not possible to examine the effects of other factors that may be more important than diagnostic uncertainty in explaining variation in Caesarean section rates between hospitals. For example, several studies have found that variations in utilization of health care per capita are due to differences in accessibility of physician services, rather than differences in practice style (Escarce 1992; Carlsen and Grytten 1998). Furthermore, several researchers argue that variation in utilization is due to how physicians are paid (Gaynor and Pauly 1990; Hellinger 1996; Grytten and Sørensen 2001). For example, incentive-based payment systems lead to a higher supply of services than fixed salary payment systems. The former type of contracts may even lead to supplier-induced demand. The mix of physicians with different payment systems varies between areas, and the level of utilization of health care

per capita may vary accordingly. It therefore follows that our analyses can only test whether a reduction in diagnostic uncertainty results in a reduction in the part of the variation in Caesarean section rates that is not attributable to patient demand or provider incentives.


## MATERIALS AND METHODS

### The Sources of the Data

The analyses were carried out on data from the Medical Birth Registry of Norway (MBRN) for the period 1967–2005 (www.fhi.no; Irgens 2000). Data from MBRN were merged with two data registers in Statistics Norway. The first register contains information about immigrant background for all first-generation immigrants (Statistics Norway 2009). The second register, the Norwegian Standard Classification of Education (Statistics Norway 2000), contains information about the highest education for all Norwegians.

Information about use of diagnostic technology was collected using a questionnaire that was sent to all senior consultants in every maternity unit in all the hospitals in the country.[1] The survey was carried out by the Norwegian Medical Association's Research Institute. The response rate was high. A total of 44 of 46 senior consultants replied. During the 39 years covered by our study, some maternity units have been closed down,[2] so that it was not possible to send a questionnaire to them. Therefore, analyses with the diagnostic technology variables could only be done for maternity units that have existed for the whole period 1967–2005. We have data for approximately 1.7 million deliveries distributed among 44 maternity units. Throughout the period 1967–2005, there were about 2.2 million births in Norway. Previous analyses have shown that our sample is representative of the whole population of mothers who gave birth in Norway 1967–2005 (for further details see Grytten, Skau, and Sørensen 2011).


### Main Analyses–Model Specification

Since the outcome (Caesarean section) is binary for the individual mother (subscript $i$), we apply a binary logit model. Moreover, as we are interested in the variation in outcomes between hospitals (subscript $j$), we specify a multilevel model in which (mean) hospital outcome is entered as a random effect. We assume the usual normal distribution for the random effects. This allows us to assess directly the relative variation in Caesarean sections at the individ-

ual and hospital levels:

$$\text{Individual level (level 1)} L_{ij} = \log\left[\frac{P(\text{Caesarean section} = 1)}{1 - P(\text{Caesarean section} = 1)}\right]$$
$$= \alpha_{0j} + \sum_k \alpha_k \text{INDCHAR}_{ij}^{(k)} + \sum_t \alpha_t D_{ij}^{(t)} \quad (1)$$

$$\text{Hospital level (level 2)} \alpha_{0j} = \beta_{00} + u_{0j}^{\text{DT}}, \text{ where } u_{0j}^{\text{DT}} \sim N(0, \sigma_{\text{DT}}^2) \text{ for all DT}$$
$$(2)$$

Our primary interest is the variation in the hospital-level error term. We specify a heterogeneous variance structure, in which the error variances are allowed to vary according to the number of types of diagnostic technology that are used in the individual hospitals $j$ (Littell et al. 1996). The variable diagnostic technology (DT) has high (DT = H) or low (DT = L) values. In this way the estimated variances ($\sigma_{\text{DT}}^2$) of the technology specific error terms ($u_{0j}^{(\text{DT})}$) can be interpreted as the *uncertainty* in hospital $j$'s practice—*given* all the characteristics of the health status of the mother and child (INDCHAR) and the year of delivery ($D$). Our main hypothesis is thus, $\sigma_{\text{DT}=\text{H}}^2 < \sigma_{\text{DT}=\text{L}}^2$, that is, we expect the error variance between hospitals with much use of diagnostic technology (DT = H) to be smaller than that between hospitals with little use of diagnostic technology (DT = L).

We employed two variants of the DT variable in the analyses, one quantitative variant (DT1) that measures the number of types of diagnostic technology that the hospitals use at any time, and one qualitative variant (DT2) that takes in information on two specific combinations of diagnostic technology. For the first specification, we used an additive index, which was constructed from each of the four technology variables. This index has values from 0 (no diagnostic technology is used) to 4 (all four types of diagnostic technology are used). This index does not distinguish between types of diagnostic technology that are in use. That is taken account of in the latter specification (DT2), in which we examine the impact of ultrasound and cardiotocography separately. We limit the qualitative analyses to these two types of diagnostic technology, since they make the main contribution to the reduction in the error variance between hospitals.

A convenient way of assessing the importance of the diagnostic technologies is to calculate the intraclass correlation coefficient: $\text{ICC}_{\text{DT}=\text{m}} = \frac{\sigma_{\text{DT}=\text{m}}^2}{\sigma_{\text{DT}=\text{m}}^2 + 3.29}$ where the figure 3.29 is the level-1 error variance given by $\pi^2/3$. The $\text{ICC}_{\text{DT}=\text{m}}$ measures the ratio of the between-hospital variation to total

variation when the level of diagnostic technology used is m (i.e., *m* can take on all values of DT1 and DT2) (Goldstein, Browne, and Rasbash 2002).

Aron et al. (1998) have shown how important it is to include risk factors of the mother and the baby when Caesarean section rates are compared between hospitals. Accordingly, in our analyses we included several characteristics of the health status of the mother and baby (=INDCHAR in equation (1)) that are associated with Caesarean section (Aron et al. 1998; Gregory et al. 2002; Kolås et al. 2003; Norwegian Institute of Public Health 2008; Kringeland, Daltveit, and Møller 2009). Variable definitions and descriptive statistics are given in Appendix 1. Most of the medical conditions mentioned below are correlated with slow or no progress in labor or signs of fetal distress. A Caesarean section can then be indicated to prevent damage to the child.

We included two variables that reflect the mother's preferences for Caesarean section. Immigrants often come from countries where it is more common to have a Caesarean section than in Norway (Betrán et al. 2007). These mothers can therefore have a higher rate of Caesarean section than Norwegian mothers (Vangen et al. 2000). Women with high education more often have a Caesarean section on maternal request than women with low education (Hurst and Summey 1984; Gould, Davey, and Stafford 1989). To control for common trends across all hospitals, we control for year of delivery ($t$) by year fixed effects ($D$).

To assess the goodness of fit for models (1) and (2), the Generalized Chi-Square value normalized by the numbers of degrees of freedom was used (Littell et al. 1996; Fernández et al. 2010). A value close to 1 implies that the variability in the data is properly accounted for by models (1) and (2).

### Additional Analyses–Model Specifications

To test the robustness of the estimations based on Models (1) and (2), we performed three additional analyses in which we tested for causality of the technology variable.

First, we did a placebo test. This analysis was performed by setting the time for getting the new technology 5 years before the hospital actually got it. If the technology variable has an effect on practice variation for Caesarean sections 5 years before the hospital actually got the diagnostic technology, this indicates that the technology effect is not causal. The technology variable is then most likely correlated with an unobserved variable which has been omitted from equations (1) and (2).

Second, we did a test of reversed causality. In models (1) and (2) we assume that the direction of the effect goes from diagnostic technology to prac-

tice variation. The opposite direction is also a possibility. We estimated the following model at the hospital level:

$$\Delta DT_{j(t,t+1)} = \beta_0 + \beta_1 \Delta CV_{j(t-1,t)} + \sum_t \alpha_t D_{jt}^{(t)} + \varepsilon_{jt} \qquad (3)$$

where $\Delta DT$ was formulated in two ways. In the first formulation $\Delta DT$ equals one if the hospital gets one or more new types of diagnostic technology from year $t$ to year $t+1$, zero otherwise. This was estimated using binary logistic regression. In the second formulation we distinguished between all possible combinations of changes in types of diagnostic technology (i.e., in DT2). This was estimated using multinominal logistic regression. $\Delta CV$ was measured as the change in the coefficient of variation over monthly Caesarean sections rates from 1 year $(t-1)$ to the next $(t)$ for each hospital (Paul-Shaheen, Clark, and Williams 1987; Diehr et al. 1990). In addition, we included fixed year effects to account for common trends across all hospitals.

      Third, we did *within* hospital analyses to check the robustness of the *between* hospital analyses as given by equations (1) and (2). The *within* hospital analyses were done by examining the change in the coefficient of variation from just before to just after each hospital introduced new diagnostic technology. We estimated the following model at the hospital level:

$$CV_j^{(k)} = \gamma_0^{(k)} + \gamma_1^{(k)} NEWDT_j^{(k)} + \sum_j \phi_j^{(k)} HOSP_j^{(k,j)} + \varepsilon_j^{(k)} \qquad (4)$$

where $CV_j^{(k)}$ is hospital $j$'s coefficient of variation over monthly Caesarean sections rates for a maximum of 60 months before, and a maximum of 60 months after the introduction of new diagnostic technology. $NEWDT_j^{(k)}$ is a dummy variable, which is given the value of 1 when the hospital gets new diagnostic technology. We ran $k$ regressions for each of our two types of technology variable (DT1 and DT2). This gave $k = 4$ regressions for unit changes along DT1 (from DT1 = 0 to DT1 = 1, from DT1 = 1 to DT1 = 2, etc.), and $k = 5$ regressions for all possible combinations of DT2. Hospital fixed effects (HOSP) were included to account for inter-hospital heterogeneity.

## RESULTS

*Main Analyses*

The variation in Caesarean section between hospitals was large at the beginning of the period; that is, up to 1973 (Figure 1). This variation diminished markedly until the late 1970s. The fall in the inter-hospital variation in

Figure 1:    Inter-Hospital Variation in Caeserean Section (logit) According to Year: Estimates of Annual Random Effects and 90% Confidence Intervals ($n = 1,708,681$)
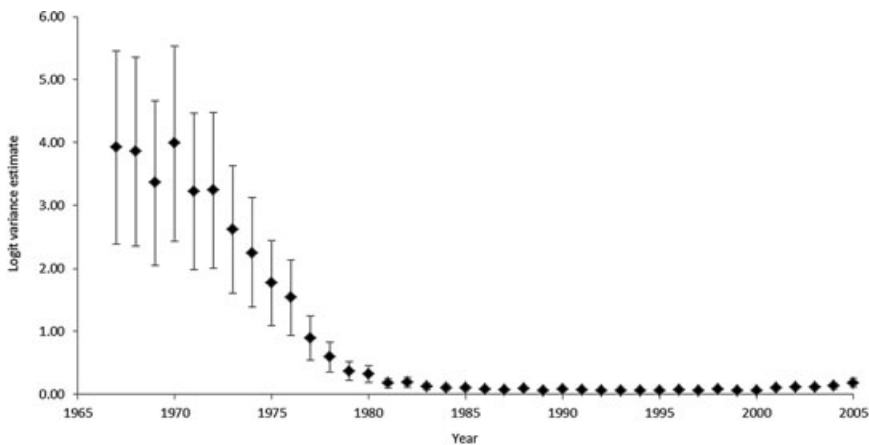


Caesarean section occurred at the time when there was a marked increase in use of ultrasound and cardiotocography. During this time period the use of ultrasound increased about fourfold, and the use of cardiotocography increased about threefold (Table 1). In 1985 nearly all the hospitals had ultrasound[3] and cardiotocography.

Appendix 1 column I presents logit coefficients for the individual-level variables (INDCHAR) for the main regression (i.e., with DT1 levels corresponding to 0–4 for the technology index). The goodness of fit for models (1) and (2) is good as indicated by a Generalized Chi-Square value of 0.98. Furthermore, all the coefficients are reasonable and statistically significant at conventional levels.

Appendix 1 column III presents the results from an additional regression where the effects that the number of types of diagnostic technology have on Caesarean section are estimated. The likelihood of a Caesarean section increases as the number of types of diagnostic technology increases. This is also shown in Appendix 2 top, where the hospitals are grouped according to the number of types of diagnostic technology they use.

The estimate for the error variance among hospitals that do not use any diagnostic technology is $\sigma^2_{dt1=0} = 0.61$ (Table 2). This value is substantially larger than the corresponding estimates for any other category of hospital grouped according to number of types of diagnostic technology used. More-

Table 2: Inter-Hospital Variation in Caesarean Section According to Use of Diagnostic Technology: Analyses at the Level of Each Delivery, Estimates of Hospital Random Effects with Confidence Intervals (CI), Intraclass Correlation Coefficients (ICC), and Hospital Percentiles for the Probability for Having a Caesarean Section

| Variable | Year (n) | Main Analyses | | | | Additional Analyses: Placebo Test* | |
|---|---|---|---|---|---|---|---|
| | I | Hospital Variance Estimate for the Log Odds of Having a Caesarean Section II | 95% CI for the Hospital Variance Estimate III | ICC IV | Lower (2.5%) and Upper (97.5%) Hospital Percentiles for the Probability of Having a Caesarean Section† V | Hospital Variance for the Log Odds of Having a Caesarean Section VI | 95% CI for the Hospital Variance Estimate VII |
| Number of types of diagnostic technology (DT1) | 1967–2005 (1,708,681) | | | | | | |
| 0 | | 0.61 | [0.32, 0.90] | 0.16 | (0.02, 0.34) | 0.15 | [0.04, 0.26] |
| 1 | | 0.18 | [0.08, 0.28] | 0.05 | (0.05, 0.20) | 0.15 | [0.03, 0.27] |
| 2 | | 0.04 | [0.02, 0.06] | 0.01 | (0.07, 0.14) | 0.03 | [0.01, 0.05] |
| 3 | | 0.03 | [0.01, 0.05] | 0.01 | (0.07, 0.13) | 0.02 | [0.01, 0.03] |
| 4 | | 0.06 | [0.01, 0.11] | 0.02 | (0.06, 0.15) | 0.02 | [0.00, 0.04] |
| Types of diagnostic technology (DT2) | 1967–1985 (704,127) | | | | | | |
| No technology | | 0.58 | [0.30, 0.86] | 0.14 | (0.01, 0.23) | 0.11 | [0.02, 0.20] |
| Ultrasound | | 0.22 | [0.00, 0.44] | 0.11 | (0.03, 0.14) | 0.11 | [0.01, 0.21] |
| Cardiotocography | | 0.16 | [0.05, 0.27] | 0.05 | (0.03, 0.13) | 0.11 | [−0.08, 0.30] |
| Ultrasound and cardiotocography | | 0.07 | [0.03, 0.11] | 0.02 | (0.04, 0.10) | 0.06 | [0.02, 0.10] |

*The analyses were done by setting the time for getting the new technology 5 years before the hospital actually got it.
†Calculations assume a mean probability equal to the overall rate for 1967–2005 for DT1 (=0.10) and the overall rate for 1967–1985 for DT2 (=0.06).

over, the 95% confidence interval reveals that it is also significantly larger at conventional levels of significance. The estimate for the error variance among hospitals that use one type of diagnostic technology is $\sigma^2_{dt1=1} = 0.18$. The 95% confidence interval shows that this variance is significantly larger than that for the hospitals that used two or three types of diagnostic technology. There is no statistically significant difference in the variance estimates between hospitals that used two, three, or four types of diagnostic technology.

The variances in Table 2 are measured on the logit scale. Another way of expressing the variation is to use probabilities. The mean Caesarean section rate in our material is 0.10. In Table 2 we present the 2.5 and 97.5 percentiles for the probabilities for having a Caesarean section around this mean value according to the number of types of diagnostic technology used. The results clearly show that variation in Caesarean section between hospitals is greatest when no or only one type of diagnostic technology is used. This is also confirmed by the intraclass correlation coefficients (Table 2 column IV).

Nearly all hospitals had ultrasound and cardiotocography by 1985. Therefore, the qualitative analyses (DT2) are limited to the years 1967–1985 (Table 2). The variance estimate for the error variance among hospitals that did not use any diagnostic technology was $\sigma^2_{dt2=0} = 0.58$. This value is substantially larger than the corresponding estimates for ultrasound and cardiotocography used alone or used in combination. The 95% confidence interval for the variance estimate of 0.58 overlaps with the corresponding interval for ultrasound, but not for cardiotocography alone, and for the two types of technology used in combination. The confidence intervals for ultrasound and cardiotocography used alone or used in combination all overlap. This is mainly because for most hospitals, both these types of diagnostic technology were introduced simultaneously (Table 1).

*Additional Analyses*

The result from the placebo test showed that there was no difference in the estimate for the error variance among hospitals according to number of types of diagnostic technology (DT1) or types of diagnostic technology (DT2) (Table 2 column VII). Furthermore, there was no sign of reversed causality: the change in the coefficient of variation over monthly Caesarean section rates from year $t-1$ to $t$ had no effect on the change in the number of types of diagnostic technology (DT1) or the change in types of diagnostic technology (DT2) from year $t$ to $t+1$ (Table 3).

Table 3:    Test of Reversed Causality: The Effect of a Change in Inter-Hospital Variation in Caesarean Section on the Change in Use of Diagnostic Technology.

| Independent Variable | Dependent Variables | | |
|---|---|---|---|
| | *Change in Number of Types of Diagnostic Technology from t to t + 1** | *Change in Types of Diagnostic Technology from t to t + 1†* | |
| Change in the coefficient of variation for Caesarean section from $t - 1$ to $t$ | | From none to cardiotocography | 0.001 (0.008) |
| | | From none to ultrasound | −0.025 (0.016) |
| | −0.005 (0.006) | From none to ultrasound and cardiotocography | −0.010 (0.011) |
| | | From cardiotocography to cardiotocography and ultrasound | −0.011 (0.012) |
| | | From ultrasound to cardiotocography and ultrasound | 0.001 (0.016) |

Regression coefficients with standard errors in brackets. Hospital-level analyses and year fixed effects are shown.
*Binary logistic regression. 1 = hospitals that get one or more new technologies from $t$ to $t + 1$; 0 = no change.
†Multinominal logistic regression. Reference category: no change in technology.
*$p \leq 0.01$, **$p \leq 0.001$.

The pre-post analyses at the hospital level as reported in Table 4 supported the analyses at the level of each delivery in Table 2. For example, an increase in the number of types of diagnostic technology from 0 to 1, and from 1 to 2 led to a statistically significant reduction in the coefficient of variation for Caesarean section rates. Conversely, an increase in the number of types of diagnostic technology from 2 to 3, and from 3 to 4 had no effect on the coefficient of variation for Caesarean section rates (Table 4 and Appendix 2, bottom). Changes in the different combinations of types of diagnostic technology all led to a reduction in the coefficient of variation for Caesarean section rates.

## DISCUSSION

According to the practice style hypothesis, uncertainty is the underlying cause for much of the variation in utilization of medical services by different populations. The results from both the main and the additional analyses support the

Table 4:   Pre-Post Test: Inter-Hospital Variation in Caesarean Section According to Use of Diagnostic Technology (Hospital-Level Analyses, Ordinary Least Squares Estimation, Hospital Fixed Effects Included)

| | | Number of Hosptials | Coefficient of Variation (Mean) | Regression Coefficient[†] (Standard Error) |
|---|---|---|---|---|
| Hospitals with a change in the number of types of diagnostic technology: | | | | |
| From | 0 | 25 | 78.1 | −15.6* |
| to | 1 | | 62.5 | (4.96) |
| From | 1 | 24 | 59.9 | −16.0* |
| to | 2 | | 43.8 | (4.12) |
| From | 2 | 24 | 33.3 | −0.005 |
| to | 3 | | 33.4 | (1.27) |
| From | 3 | 9 | 22.5 | −0.15 |
| to | 4 | | 22.3 | (0.91) |
| Hospitals with a change in types of diagnostic technology: | | | | |
| From | None | 17 | 82.1 | −16.7** |
| to | Cardiotocography | | 65.2 | (6.03) |
| From | None | 8 | 75.8 | −15.4 |
| to | Ultrasound | | 60.5 | (10.1) |
| From | None | 38 | 76.7 | −29.5* |
| to | Ultrasound and cardiotocography | | 46.9 | (3.69) |
| From | Cardiotocography | 17 | 59.7 | −13.6** |
| to | Cardiotocography and ultrasound | | 46.0 | (4.90) |
| From | Ultrasound | 8 | 55.9 | −16.9** |
| to | Ultrasound and cardiotocography | | 38.9 | (6.40) |

[†]Dependent variable: the coefficient of variation.
*$p \leq 0.01$,
**$p \leq 0.001$.

practice style hypothesis. The reason is that the introduction of diagnostic technology has reduced uncertainty about the diagnosis of risk factors of the mother and child during delivery, and variation in type of delivery between hospitals has been reduced accordingly. As expected, the reduction in practice variation was greatest for ultrasound and cardiotocography.

In our study, practice variation was measured at the individual level, whereas diagnostic technology was measured at the level of the hospital. A potential objection is that we should have connected the use of diagnostic technology to the individual obstetrician. This was not practically possible, since information about which maternity unit obstetricians work in is not available, either in MBRN, or from other sources in Norway, as far back as the 1970s and 1980s.[4] Also, even if it had been possible, it is not certain that

such a connection would have been an appropriate analytical approach. This is because obstetricians work in teams, and in principle diagnostic technology is available for all the obstetricians in a department, not just for some of them. This is supported by several studies, in which it has been found that practice variation within hospitals is small compared with variation between hospitals for physicians treating similar patients (Griffiths, Waters, and Acheson 1979; Westert, Nieboer, and Groenewegen 1993; De Jong et al. 2006). This is mainly the result of two factors. First, the leaders in a hospital department have a tendency to recruit doctors who have the practice norms that prevail in the department—there is a selection of doctors to the department. Second, if the medical-professional norms of one or several doctors in a department vary from the prevailing norms, these doctors will gradually conform to the prevailing norms. It is difficult to be a nonconformer, with the danger of being criticized or losing one's professional affiliation to one's colleagues (Westert, Nieboer, and Groenewegen 1993; De Jong, Groenewegen, and Westert 2003; Burke, Fournier, and Prasad 2010). Additional support is given by Westert et al. (2004), who examined practice variation for hospital discharge rates in the Netherlands for 12 diagnostic or surgical procedures for the period 1980–1997. For 11 of the 12 procedures, hospitals converged to a community standard over time.

The overall Caesarean section rate has increased, because several of the risk factors of the mother and the baby have increased over the last few decades (for example, the proportion of older women who give birth has increased, and babies are heavier) (Grytten et al. 2011). The research by Finsen, Storeheier, and Aasland (2008) and Lehmann et al. (2007) further suggests that obstetricians are important drivers behind the increase in the Caesarean section rate. This is because obstetricians and physicians have a higher rate of Caesarean section than the rest of the population. For example, Finsen, Storeheier, and Aasland (2008) observed that "Norwegian women do as obstetricians do—not as obstetricians say." Therefore, the authors argue that the Caesarean section rate in the general population is unlikely to fall as long as obstetricians have their own children delivered by Caesarean section. One way to deal with the increase in the Caesarean section rate is to use some sort of co-payment, as suggested by Fuglenes et al. (2010). Our findings show that the introduction of diagnostic technology in maternity care has made it easier to determine whether a Caesarean section is indicated on clinical grounds. This again implies that it has become easier to identify whether a mother requests a Caesarean section without any clear medical indications. In that case some sort of co-payment can be justified.

In conclusion, our results support the practice style hypothesis. In our study design we took advantage of the fact that hospitals introduced diagnostic technology in maternity units at different times. The study was carried out in a homogeneous population in which neither the mother nor the obstetrician had economic incentives that could influence their choice of method of delivery. Caution must be used in generalizing the findings to other countries where maternity care is organized differently. This applies particularly to countries where many births take place in private clinics, where obstetricians are remunerated on a fee-for-service basis, and where mothers have private insurance for deliveries. Here both the mother and the obstetrician may be tempted to take private economic interests into account when deciding on type of delivery. Therefore, in these countries disparities in type of delivery may persist even though clinical uncertainty in diagnosis has been reduced.

## ACKNOWLEDGMENTS

## NOTES

1. The senior consultant is the obstetrician who is in charge at each maternity unit.
2. This corresponds well with the number of hospitals that have been closed down, since most hospitals have a maternity unit. From 1970 to 2000 the number of maternity units was reduced from 150 to 57 (Nilsen, Daltveit, and Irgens 2001). The greatest reduction has been for units with fewer than 500 deliveries per year.
3. These figures are similar to those found in other studies from Norway. For example, in 1986, 36 of the maternity units in our set of data had ultrasound. In a national survey of the occurrence of ultrasound in 1986, 35 maternity units reported that they had ultrasound (Backe and Buhaug 1986). In 1974, three maternity units in our set of data reported that they had ultrasound. This figure is identical to the national figure for that year (Backe et al. 1987).

4. This also means that it was not possible to control for characteristics of the obstetrician at the level of the maternity unit. American studies have shown that whether obstetricians are board certified, and whether they have foreign nationality, is of some importance for their practice profile (Tussing and Wojtowycz 1993; Burns, Geller, and Wholey 1995). Their age and gender are of little importance. In Norway, there is one authorization to be a specialist obstetrician, and obstetricians have all fulfilled the same standard requirements. Therefore, the issue of whether obstetricians are board certified is not relevant in Norway. Also, there are very few obstetricians who have foreign nationality. Therefore, according to our assessment, lack of control for characteristics of the obstetrician does not lead to much bias in our results.

## REFERENCES

Amer-Wahlin, I., C. Hellsten, H. Norén, H. Hagberg, A. Herbst, I. Kjellmer, H. Lilja, C. Lindoff, M. Månsson, L. Mårtensson, P. Olofsson, A.-K. Sundström, and K. Maršál. 2001. "Cardiotocography Only Versus Cardiotocography Plus ST Analysis of Fetal Electrocardiogram for Intrapartum Fetal Monitoring: A Swedish Randomised Controlled Trial." *The Lancet* 358: 534–8.

Amer-Wahlin, I., S. Arulkumaran, H. Hagberg, K. Maršál, and G. H. A. Visser. 2007. "Fetal Electrocardiogram: ST Waveform Analysis in Intrapartum Surveillance." *BJOG An International Journal of Obstetrics and Gynaecology* 114: 1191–3.

Aron, D. C., D. L. Harper, L. B. Shepardson, and G. E. Rosenthal. 1998. "Impact of Risk-Adjusting Cesarean Delivery Rates When Reporting Hospital Performance." *Journal of the American Medical Association* 279: 1968–72.

Backe, B., and H. Buhaug. 1986. *Bruk av ultralyd i svangerskapet.* Rapport 8/86. Trondheim: Norsk Institutt for sykehusforskning.

Backe, B., T. Heggestad, and T. Lie. 2003. "Har keisersnittsepidemien nådd Norge?" *The Journal of the Norwegian Medical Association* 123: 1522–4.

Backe, B., G. Jacobsen, L. S. Bakketeig, and P. Bergsjø. 1987. "Bruk av ultralyd ved norske fødeinstitusjoner." *The Journal of the Norwegian Medical Association* 107: 471 –3.

Baicker, K., K. S. Buckles, and A. Chandra. 2006. "Geographic Variation in the Appropriate Use of Cesarean Delivery." *Health Affairs* 25: w355–67. doi: 10.1377/ hlthaff.25.w355.

Betrán, A. P., M. Merialdi, J. A. Lauer, W. Bing-Shun, J. Thomas, P. Van Look, and M. Wagner. 2007. "Rates of Caesarean Section: Analysis of Global Regional and National Estimates." *Paediatric and Perinatal Epidemiology* 21: 98–113.

Bragg, F., D. A. Cromwell, L. C. Edozien, I. Gurol-Urganci, T. A. Mahmood, A. Templeton, and J. H. van der Meulen. 2010. "Variation in Rates of Caesarean Section among English NHS Trusts after Accounting for Maternal and Clinical Risks: Cross Sectional Study." *British Medical Journal* 341: c5065. doi: 10.1136/ bmj.c5065.

Burke, M. A., G. M. Fournier, and K. Prasad. 2010. "Geographic Variations in a Model of Physician Treatment Choice with Social Interactions." *Journal of Economic Behavior and Organization* 73: 418–32.

Burns, L. R., S. E. Geller, and D. R. Wholey. 1995. "The Effect of Physician Factors on the Cesarean Section Decision." *Medical Care* 33: 365–82.

Carlsen, F., and J. Grytten. 1998. "More Physicians: Improved Availability or Induced Demand?" *Health Economics* 7: 495–508.

Clark, S. I., M. A. Belfort, G. D. V. Hankins, J. A. Meyers, and F. M. Houser. 2007. "Variation in the Rates of Operative Delivery in the United States." *American Journal of Obstetrics & Gynecology* 196: 526.e1–526e5. doi: 10.1016/j. ajog.2007.01.024.

Diehr, P., K. Cain, F. Connell, and E. Volinn. 1990. "What Is Too Much Variation? The Null Hypothesis in Small-Area Analysis." *Health Services Research* 24: 741–71.

Dunton, T. A. 1999. *An Introduction to Time Waveform Analysis* [accessed on January 5, 2012]. Available at http://www.unitechinc.com/pdf/IntroductiontoTimeWaveformAnalysis.pdf

Eikeland, T., B. Haugeberg, T. Henriksen, S. Hjelle, B. M. Yli, E. Blix, and P. Øyan. 2008. Veileder i fødselshjelp 2008, kapittel 22. Overvaking under fødsel (CTG, føtale blodprøvar, CTG + ST-analyse av foster – EKG) [accessed on January 5, 2012]. Available at http://www.legeforeningen.no./id/148399.0

Escarce, J. J. 1992. "Explaining the Association between Surgeon Supply and Utilization." *Inquiry* 29: 403–15.

Fernández, E. A., E. P. Souza Neto, P. Abry, R. Macchiavelli, M. Balzarini, B. Cuzin, C. Baude, J. Frutoso, and C. Gharib. 2010. "Assessing Erectile Neurogenic Dysfunction from Heart Rate Variability through a Generalized Linear Mixed Model framework." *Computer methods and Programs in Biomedicine* 99: 49–56.

Finsen, V., A. H. Storeheier, and O. G. Aasland. 2008. "Cesarean Section: Norwegian Women Do as Obstetricians Do – Not as Obstetricians Say." *Birth* 35: 117–20.

Fuglenes, D., P. Øian, D. Gyrd-Hansen, J. A. Olsen, and I. S. Kristiansen. 2010. "Norwegian Obstetricians' Opinions about Cesarean Section on Maternal Request: Should Women Pay Themselves?" *Acta Obstetricia et Gynecologica Scandinavica* 89: 1582–8.

Gaynor, M., and M. V. Pauly. 1990. "Compensation and Productive Efficiency in Partnerships: Evidence from Medical Group Practice." *The Journal of Political Economy* 98: 544–73.

Goldstein, H., W. Browne, and J. Rasbash. 2002. "Partitioning Variation in Multilevel Models." *Understanding Statistics* 1: 223–31.

Gould, J. B., B. Davey, and R. S. Stafford. 1989. "Socioeconomic Differences in Rates of Cesarean Section." *New England Journal of Medicine* 321: 233–9.

Gregory, K. D., L. M. Korst, J. A. Gornbein, and L. D. Platt. 2002. "Using Administrative Data to Identify Indications for Elective Primary Cesarean Delivery." *Health Services Research* 37: 1387–401.

Griffiths, M., W. E. Waters, and E. D. Acheson. 1979. "Variation in Hospital Stay after Inguinal Herniorrhaphy." *British Medical Journal* 1: 787–98.

Grytten, J., I. Skau, and R. Sørensen. 2011. "Do Expert Patients Get Better Treatment Than Others? Agency Discrimination and Statistical Discrimination in Obstetrics." *Journal of Health Economics* 30: 163–80.

Grytten, J., and R. Sørensen. 2001. "Type of Contract and Supplier-Induced Demand for Primary Physicians in Norway." *Journal of Health Economics* 20: 379–93.

———. 2003. "Practice Variation and Physician-Specific Effects." *Journal of Health Economics* 22: 403–18.

Grytten, J., L. Monkerud, T. P. Hagen, R. Sørensen, A. Eskild, and I. Skau. 2011. "The Impact of Hospital Revenue on the Increase in Caesarean Sections in Norway. A Panel Data Analysis of Hospitals 1976–2005." *BMC Health Services Research* 11: 267. [accessed on January 5, 2012]. Available at http://www.biomedcentral.com/1472-6963/11/267

Hellinger, F. J. 1996. "The Impact of Financial Incentives on Physician Behaviour in Managed Care Plans: A Review of the Evidence." *Medical Care Research and Review* 53: 294–314.

Hemminki, E., and M. Gissler. 1994. "Variation in Obstetic Care within and between Hospital Levels in Finland." *British Journal of Obstetrics and Gynaecology* 101: 851–7.

Hurst, M., and P. S. Summey. 1984. "Childbirth and Social Class: The Case of Cesarean Delivery." *Social Science & Medicine* 18: 621–31.

Irgens, L. M. 2000. "The Medical Birth Registry of Norway. Epidemiological Research and Surveillance throughout 30 Years." *Acta Obstetricia et Gynecologica Scandinavica* 79: 435–9.

De Jong, J. D., P. P. Groenewegen, and G. P. Westert. 2003. "Mutual Influences of General Practitioners in Partnerships." *Social Science and Medicine* 57: 1515–24.

De Jong, J. D., G. P. Westert, R. Lagoe, and P. P. Groenewegen. 2006. "Variation in Hospital Length of Stay: Do Physicians Adapt Their Length of Stay Decisions to What Is Usual in the Hospital Where They Work?" *Health Services Research* 41: 374–91.

Kolås, T., D. Hofoss, A. K. Daltveit, S. T. Nilsen, T. Henriksen, R. Häger, I. Ingemarsson, and P. Øian. 2003. "Indications for Cesarean Deliveries in Norway." *American Journal of Obstetrics and Gynecology* 188: 864–70.

Kringeland, T., A. K. Daltveit, and A. Møller. 2009. "What Characterizes Women in Norway Who Wish to Have a Caesarean Section?" *Scandinavian Journal of Public Health* 37: 364–71.

Lehmann, S., P. E. Børdahl, S. A. Rasmussen, and L. M. Irgens. 2007. "Norwegian Midwives and Doctors Have Increased Cesarean Section Rates." *Acta Obstetricia et Gynecologica* 86: 1087–9.

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger. 1996. *SAS System for Mixed Models*, pp. 162–4. Cary, NC: SAS Institute Inc.

Luzietti, R., R. Erkkola, U. Hasbargen, L. Å. Mattson, J. M. Thoulon, and K. G. Rosén. 1999. "European Community Multi-Center Trial "Fetal ECG Analysis during Labor": ST plus CTG Analysis." *Journal of Perinatal Medicine* 27: 431–40.

MacDorman, M. F., F. Menacker, and E. Declerq. 2008. "Cesarean Birth in the United States; Epidemiology, Trends, and Outcomes." *Clinics in Perinatology* 35: 293–307.

Mayor, S. 2002. "Caesarean Section Rate in England Reaches 22%." *British Medical Journal* 324: 1118.

McKenzie, L., and P. A. Stephenson. 1993. "Variation in Cesarean Section Rates Among Hospitals in Washington State." *American Journal of Public Health* 83: 1109–12.

Ministry of Health. 2002. *Behovsbasert finansiering av spesialisthelsetjenesten.* Oslo, Norway: Ministry of Health.

Morrison, J. J., J. M. Rennie, and P. J. Milton. 1995. "Neonatal Respiratory Morbidity and Mode of Delivery at Term: Influence of Timing of Elective Caesarean Section." *British Journal of Obstetrics and Gynaecology* 102: 101–6.

Nilsen, S. T., A. K. Daltveit, and L. M. Irgens. 2001. "Fødeinstitusjoner og fødsler i Norge i 1990-årene." *The Journal of the Norwegian Medical Association* 121: 3208–12.

Norwegian Institute of Public Health. 2007. "Hvorfor øker keisersnitt i Norge?" [accessed on January 5, 2012]. Available at http://www.fhi.no/artikler?id=52713

Norwegian Institute of Public Health. 2008. "Keisersnitt–faktaark." [accessed on January 5, 2012]. Available at http://www.fhi.no/artikler?id=52705

Norwegian Society of Gynecology and Obstetrics. 2008. "Veileder i fødselshjelp 2008." [accessed on January 5, 2012]. Available at http://www.legeforeningen.no/id/131068.0%20%5b10

Paul-Shaheen, P., J. D. Clark, and D. Williams. 1987. "Small Area Analysis: A Review and Analysis of the North American Literature." *Journal of Health Politics, Policy and Law* 12: 741–809.

Renwick, M. Y. 1991. "Caesarean Section Rates, Australia 1986: Variations at State and Small Area Level." *Australian and New Zealand Journal of Obstetrics and Gynaecology* 31: 299–304.

Rooth, G., A. Huch, and R. Huch. 1987. "FIGO News. Guidelines for the Use of Fetal Monitoring." *International Journal of Gynecology & Obstetrics* 25: 159–67.

Stano, M. 1991. "Further Issues in Small Area Variations Analysis." *Journal of Health Politics, Policy and Law* 16: 573–88.

Statistics Norway. 2000. Norwegian Standard Classification of Education. Revised 2000. Oslo-Kongsvinger: Statistics Norway.

Statistics Norway. 2009. "StatBank Norway, 02 Population, Table 05184 Immigrants by Sex and Country Background." [accessed on January 5, 2012]. Available at http://statbank.ssb.no//statistikkbanken/default_fr.asp?PLanguage=1

Swedish Council on Health Technology Assessment. 2006. *STAN– ST Waveform Analysis Combined with Cardiotocography for Fetal Monitoring During Childbirth.* Report 2006-04. [accessed on 5 January, 2012]. Available at http://www.sbu.se/en/Published/Alert/STAN-ST-Waveform-Analysis-Combined-With-Cardiotocography-for-Fetal-Monitoring-During-Childbirth/

Tussing, A. D., and M. A. Wojtowycz. 1993. "The Effect of Physician Characteristics on Clinical Behavior: Cesarean Section in New York State." *Social Science & Medicine* 37: 1251–60.

Van Geijn, H. P. 1998. "Cardiotocography." In *Textbook of Perinatal Medicine*, Vol. 2, edited by A. Kurjak, pp. 1424–8. London: Parthenon Publishing.

Vangen, S., C. Stoltenberg, A. Skrondal, P. Magnus, and B. Stray-Pedersen. 2000. "Cesarean Section among Immigrants in Norway." *Acta Obstetricia et Gynecologica Scandinavica* 79: 553–8.

Wennberg, J. E. 2010. *Tracking Medicine. A Researcher's Quest to Understand Health Care.* New York: Oxford University Press.

Wennberg, J. E., B. A. Barnes, and M. Zubkoff. 1982. "Professional Uncertainty and the Problem of Supplier-Induced Demand." *Social Science & Medicine* 16: 811–24.

Westert, G. P., A. P. Nieboer, and P. P. Groenewegen. 1993. "Variation in Duration of Hospital Stay between Hospitals and between Doctors within Hospitals." *Social Science & Medicine* 37: 833–9.

Westert, G. P., P. P. Groenewegen, H. C. Boshuizen, P. M. M. Spreeuwenberg, and M. P. M. Steultjens. 2004. "Medical Practice Variations in Hospital Care; Time Trends of a Spatial Phenomenon." *Health & Place* 10: 215–20.

Westgate, J., J. M. Garibaldi, and K. R. Greene. 1994. "Umbilical Cord Blood Gas Analysis at Delivery: A Time for Quality Data." *British Journal of Obstetrics and Gynaecology* 101: 1054–63.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Appendix S1: Effect of Independent Individual-Level Variables on Caesarean Section, 1967–2005: Analyses at the Level of Each Delivery, Control for Continuous Year and Hospital Random Effects, Logit Coefficients with Standard Errors in Brackets.

Appendix S2: The Proportion of Caesarean Sections (top) and Inter-hospital Variation in Caesarean Section (bottom) by Number of Types of Diagnostic Technology. Hospital Level Analyses.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.