Expert Behavior Change in Response to Best Practice Changes: Evidence from Obstetrics*

Emily Oster, Brown University and NBER

December, 2018

Abstract

Guidelines for best practices in medicine change frequently. In this paper, I analyze to what extent such changes alter practice patterns and whether adoption of best practices differentially affects some patient groups. I consider these questions in the context of changes in practice guidelines in obstetrics. I find, first, that practice patterns do respond to new information, most notably to new research. Such response are slow, however, and do not seem to be sped up by changes in official organizational guidelines. I suggest that media coverage is an important part of catalyzing behavior change, even in this setting where the changes come from experts. Changes in practice are faster for more advantaged groups of patients. Related to this, I find evidence that changes occur more quickly for infants who are otherwise healthier, even conditional on observable features. I connect this last point to challenges in learning about the impacts of these practices. The results overall suggest practice changes would be faster if they were more widely promoted, and this might especially benefit infants who are otherwise disadvantaged.

1 Introduction

To what extent do doctors respond to changes in treatment guidelines by changing their practice patterns? Do some change their behavior more quickly than others, implying that some patients benefit first from improved treatment?

Practice guidelines in medicine change frequently (see, for example, changes in treatment thresholds for high blood pressure (Kristensen et al., 2014; Jackson et al, 2008; Gu et al, 2012) or changes

*James Okun, Jason Goettisheim and Cathy Yue Bai provided outstanding and crucial research assistance. I am also grateful to Valeria Zurla and Lauren To.
Changes in guidelines have some parallels to changes in technology, a topic in which economists have long been interested, beginning with the classic Griliches (1957) work on hybrid corn, and with many examples in the area of health (e.g. Skinner and Staiger, 2005; Freedman, Lin and Simon, 2015; Gold et al, 2014; Chandra, Cutler and Song, 2011). A key difference between practice changes and technology changes is the potential speed. In the case of new technology, adoption often requires learning about an unfamiliar procedure or practice, and possibly investing in new capital equipment. In contrast, many of the best practice changes simply change the recommended balance of procedures, or the threshold for some prescription or treatment. As a result, the only significant barriers to adoption are learning the information and reacting to it. In principle, changes could be immediate.

In this paper I address the question of how medical practitioners respond to different types of information in changing treatment patterns, focusing on obstetrics. Accepted best practices around childbirth have evolved considerably even in the past two decades, and there may be considerable welfare gains to improved practices, given that US outcomes for mothers and babies lag the rest of the world (National Research Council, 2013; Chen, Oster and Williams, 2016).

I analyze the patterns of behavior around four medical decisions which are common during childbirth: (1) whether to attempt a vaginal birth after cesarean section (VBAC); (2) whether to deliver a breech baby by planned cesarean section; (3) use of vacuum extraction for vaginal birth; and; (4) whether to discourage elective labor induction at 37 or 38 weeks of gestation.

The paper has two primary goals. The first is to analyze how these practices have changed over time, and connect changes to new information or other events. The second is to explore heterogeneity in the speed of response across groups of patients. Both analyses make use of the Natality Detail Files, which contain detailed microdata on all births in the United States over time.

Focusing on the first question, I find that there are significant changes in all four behaviors over time, and they do seem to reflect responses to changes in best practices. These responses are, however, slow. During this period there are two major research findings - on VBAC and breech delivery - which do change practice, but over a period of years, not months. I find limited evidence that changes in official guidelines increase this speed of change.

Notably, there is one example of a sharp and sustained decline in behavior - a large drop in the use of vacuum extraction over a one month period in 1999. This appears to be a result not, however,
of important changes in knowledge but instead a reaction to an episode of the show 20/20.

The slow change has welfare consequences. If new practices achieve (as we expect) better outcomes, then faster change would benefit more patients. The slow change also opens the related question of whether some patients systematically benefit from new information first.

The technology adoption literature has been quite interested in variation across space in adoption (e.g. Skinner and Staiger, 2005), so I also begin with this question. In practice in this case I see very little evidence of variation across space. Although the levels of the behaviors differ across space, the trends are the same.

Turning to demographics, however, there is evidence that changes in behavior are experienced differentially by different patients. More specifically, it appears that the changes in behavior initially benefit more positively selected patients. The changes occur more slowly for groups who have, on average, worse birth outcomes - in particular, African-American mothers and those with less education. To give one example of this, prior to the change in VBAC recommendations in 1996 those with less than a high school education were slightly less likely to have a VBAC than those with more education. By 2005 they were 3 percentage points more likely.

I use regression evidence to show that the selection gradient with respect to race and education moves systematically along with the level of the behavior. Further, I show similar conclusions when I focus on predicted mortality - predicted based on demographics - rather than race or education as the selection variable. In summary, the benefits of new information seem to accrue to mothers and infants in groups who are initially more advantaged.

I extend this point to show that selective behavior change favors better-off infants even conditional on demographics. Specifically, I explore whether the relationship between infant outcomes and the treatment varies with recommendations; I find that it does. To continue the VBAC example: after the change in recommendation, those infants born by VBAC have a lower APGAR score, relative to the difference before, even conditional on their observables. This suggests that there is also negative selection of behavior change on unobservables. This is true for all four behaviors, and when I consider infant mortality as the outcome rather than APGAR score.

Aggregating these facts paints a consistent overall picture. Despite the fact that these changes could be made quickly and universally, I see patterns similar to technology adoption: slow diffusion of change, with the earlier changes benefiting the better off patients. This suggests that it might be possible to do more to encourage faster changes, and doing so would benefit more vulnerable patients. Although it seems surprising since the actors here are experts, the evidence suggests that
media attention may be a way to prompt more significant and faster changes.

As a final note, the last piece of evidence on the changes in relationship with infant outcomes over time relates to the general issue of dynamic selection which I explore elsewhere. Oster (2018) focuses on consumer health recommendations and shows that changes in these recommendations can generate dynamic selection, with otherwise healthier consumers adopting newly recommended behaviors first, changing the apparent relationship between the behavior and health. The results here show a similar dynamic occurs even with expert decision-makers. Similar to the central point in that paper, these dynamic changes have the potential to make it more difficult to learn about relationships in the data.

This paper relates closely to the literature on health technology adoption. I focus in particular on the role of different types of information in driving changes in behavior. A large literature looks at other drivers, including (among others) peer effects, hospital structure and insurance policy (Burke, Fournier and Prasad, 2007; Bradley et al, 2006; Gold et al, 2014; Bokhari et al, 2008; Freedman, Lin and Simon, 2015; Sacarny, 2014; Domino, 2012; Agha and Molitor, 2018; Snieliuskas, 2011; Yu, 2015).

The analysis also relates to a smaller literature, primarily in medicine, on how doctors respond to changes in practice guidelines. Notable here are a number of papers on response to changes in hypertension and statin prescribing guidelines (Kristensen et al., 2014; Jackson et al, 2008; Gu et al, 2012; Ma et al, 2006; Pauff et al, 2015). These papers tend to find doctors do respond to changes in prescribing guidelines, although they do not focus on heterogeneity across patient types.

Finally, I relate to a large literature on understanding use of interventions in obstetrics, perhaps most notably papers on differential use of cesarean sections across patient types (i.e. Johnson and Rehavi, forthcoming).

The rest of the paper is organized as follows. Section 2 describes the background on the changes in practice guidelines, and the data I use. Section 3 shows results on changes in practice patterns over time. Section 4 shows the evidence on heterogeneity and Section 5 concludes.

2 Background and Data

This section begins by describing the context for each of the procedures I consider, and briefly reviewing the changes in recommendations. In the second subsection I describe the data used in the analysis.
2.1 Background on Contexts Considered

There are four practices considered here: (1) vaginal birth after cesarean section; (2) delivery method for breech births; (3) use of vacuum extraction for vaginal birth and; (4) elective labor induction at 37 or 38 weeks of gestation (“early term”).

**Vaginal Birth after Cesarean (VBAC)**

Vaginal birth after cesarean (VBAC) refers to the case in which a woman who has previously delivered a child by cesarean section delivers vaginally a subsequent birth. In 1988, the American College of Obstetricians and Gynecologists (ACOG) recommended VBAC as an option for prior cesarean patients. In 1996, however, a large study (McMahon et al, 1996) suggested outcomes were better with planned cesarean section in the case of a prior cesarean. Following this, in 1998 an ACOG bulletin cautioned that VBAC should only be attempted at hospitals readily equipped to provide emergency care in the case of uterine rupture. ACOG bulletins after 1998 continued to focus on the risks of VBAC and the necessity of immediately available emergency cesarean sections.

**Breech Delivery**

Most infants are head down in the uterus at the time of labor and delivery, a position referred to as “vertex presentation.” Breech presentation, in contrast, refers to a case where the infant is in another position, typically either feet down or butt down. Some breech babies can be delivered vaginally, but it is a more complex birth procedure. In 2000, results of a large trial (TERM-BREECH) suggested that outcomes for breech infants were better with a planned cesarean section than a trial of vaginal birth (Hannah et al, 2000). In 2001, ACOG followed this with a bulletin which recommends cesarean sections for breech births. They note that most physicians are not experienced in vaginal breech births and that vaginal birth is generally not appropriate for this position of the fetus.

**Vacuum Extraction**

Vacuum extraction is a method to assist in the delivery of a baby using a vacuum device which attaches to the child’s head. It is used in the second stage of labor if the infant appears to be stuck. Vacuum extraction has been variously recommended and rejected over time. In 1994, ACOG released a bulletin that detailed the qualifications for vacuum extraction and encouraged it
under appropriate circumstances. However, in 1998, the FDA released a report on the dangers of vacuum extraction, citing reports of deaths, which was followed by significant negative press. Back in the other direction, subsequent ACOG opinions and bulletins reaffirmed the safety of vacuum extraction and noted that complications are rare. The widely used alternative to vacuum extraction is cesarean section.

Redefinition of “Full Term Pregnancy”

On average, a pregnancy lasts 40 weeks, and a full-term pregnancy has generally been considered to be any delivery that falls between 37 and 42 weeks. A baby born before 37 weeks is considered premature. However, in 2013, ACOG redefined full-term to be between 39 and 40 weeks, and divided full term pregnancies into four categories: early-term (37 to 38 weeks), full-term (39 to 40 weeks), late-term (41 to 42 weeks), and post-term (more than 42 weeks). This had implications for induced labor, as part of the change was to discourage elective induction in the “early term” period.

2.2 Data

The primary data source in this paper is the US Natality Detail Files, from the National Center for Health Statistics (NCHS). The analysis uses both linked birth-death files and, where those are not available, birth files only. For most treatments I focus on the period from 1995 through 2005, where we have consistently coded information on treatments and demographics and during which time most of the major changes occur. For the redefinition of term pregnancy I extend the data through 2015.

The NCHS files are a Census of births in the United States based on birth and death certificates filed in vital statistics offices in all States and the District of Columbia. The data provides information on infant death within a year of birth (through 2009), various demographic characteristics such as race, age and level of education, the health of both mother and child post-partum, and physician practices carried out during labor.

In most cases it is necessary to recode variables to capture the outcomes of interest. The VBAC variable takes the value 1 if the infant is delivered vaginally when the mother had delivered another infant by cesarean and 0 if the infant was delivered by other means and the mother previously had a cesarean delivery. The breech delivery variable is 1 if the infant is breech presenting and is delivered by cesarean section and 0 if the infant is breech presenting and delivered vaginally. For vacuum extraction, the indicator variable is 1 if a vacuum delivery is performed and 0 otherwise.
Lastly, the early term birth variable takes the value 1 if the infant is delivered between 37 and 38 weeks, inclusive, and 0 if the birth occurs after 38 weeks. Births before 37 weeks are excluded from this analysis. Table 1 presents summary statistics on treatments, demographics and outcomes.

3 Evidence on Changes Over Time

This section reports the initial results on changes in practice patterns over time.

Figure 1 shows the evolution of each procedure over time. There are four sub-figures, each corresponding to one of the outcomes and showing the monthly means of behavior. Vertical lines in the figures indicate major research studies, ACOG bulletins or opinions and other events. Table 2 lists the changes in recommendations indicated by the vertical lines.

Below, I discuss these figures in turn.

Vaginal Birth after Cesarean

The trends over time for vaginal birth after cesarean (VBAC) are shown in Figure 1a.

Broadly, this outcome is trending down over time. In the first year of the data - 1995 - the series shows little trend and the mean is around 27%. To interpret, this implies that of women who had a prior cesarean section, 27% of them went on to deliver a second pregnancy vaginally. By 2005, however, only about 7% of women with a prior cesarean section are delivering vaginally.

The change appears to be precipitated by McMahon et al (1996), which indicated that major maternal complications were almost twice as common for trial of labor after cesarean than for a repeat cesarean.

The overall changes in this behavior are large. It is notable, however, that they are slow to diffuse, despite the fact that practitioners were clearly familiar with the procedures required at the time the change was made. Notable also is that the two official statements from the American College of Obstetricians and Gynecologists (ACOG) - in 1998 and 1999 - seem to have had little effect, changing neither the level nor the trend. Both statements echoed the 1996 paper and urged repeat cesarean sections rather than VBAC in most cases. To the extent that the goal of these guidelines was to increase the speed with which changes were made, the data does not show strong evidence of that.
Breech Delivery Method

Figure 1b shows the share of breech births delivered by cesarean section.

This series is initially flat, or trending slightly down, and then begins trending up over time in the 2000s. There is one sharp spike downwards, but overall the trend is towards universal use of cesarean section for breech deliveries. The changes over time seem, again, to be related to the timing of research findings. In 2000, researchers published the first results from the TERM-BREECH trial, a large randomized trial of delivery methods for breech births (Hannah et al, 2000). The authors concluded that short-term outcomes for breech infants were better after a planned cesarean section than after a planned vaginal birth.

Similar to the case of VBAC, these findings do seem to prompt an upward trend in the use of cesarean delivery although, again, the change is not sudden. Later changes in ACOG guidelines, published in 2001, reinforced the preference for cesarean section for breech delivery. Again, similar to VBAC, these do not seem to cause a change in level or trend.

In 2003, an abstract (Hannah et al, 2003) was released with follow-up data from the TERM-BREECH trial. This later follow-up was somewhat more reassuring about the outcomes for the planned vaginal birth group, showing no differences in outcomes after two years. This seems to cause a short term revision in the use of cesarean section for breech deliveries, although this is very short lived.

Vacuum Extraction

Figure 1c shows the share of births in which vacuum extraction is used, over time.

This figure has a sharp decline in the middle - over a one month period in 1999 there is a one percentage point drop in the use of vacuum extraction in births, from a base of about 6%. This continues, and is continued by, a downward trend.

The downward trend in use of this procedure seems to begin with a 1998 FDA report suggesting there were significant dangers to using vacuum extraction, including the possibility of infant injury or death. This alert was sent to all practitioners. The report begins a period of decline for this procedure, which had until then been trending upwards. The report does not, however, cause an immediate spike down. Indeed, the spike down does not coincide with any specific official events.

The timing is, however, consistent with the drop resulting from a sensationalist episode of the TV news program “20/20” which focused on the risks of vacuum extraction for babies, and featured
graphic images of injured or deceased infants. The news program was motivated by some of the earlier findings from the FDA, but the timing is clearly later. This program aired at the very end of January, 1999. While it is difficult to prove the two events are linked, we do not see any other events which line up in terms of timing.

It is notable that there are a number of information shocks which supported the use and safety of vacuum extraction. These include an ACOG bulletin in 1998 (the second dotted line) and a significant review study release in 1999 (Towner et al, 1999). Both argued for the safety of vacuum extraction. Towner et al (1999), for example, showed similar rates of infant injury for use of vacuum, forceps or cesarean section.

These releases do not seem to have any impact on stemming the reduction in use of vacuum. The downward trend continues through these releases.

**Early Term Birth**

Finally, I consider Figure 1d, which shows changes in the rate of early term delivery over time. This graph is limited to the universe of infants who were born at or after 37 weeks of pregnancy, and it shows the share of these births which were at 37 or 38 weeks. This share is increasing until the mid to late 2000s and then declines. This slow decline is hard to trace to any particular event; it seems, rather, to reflect a growing consensus over this period that early term births result in worse outcomes than later term births (Reddy, Ko and Willinger, 2006; Engle and Kominiarek, 2008; Fleischman, Oinuma and Clark, 2010).

In 2013, ACOG issued concrete guidelines around these issues - specifically, classifying 37 or 38 weeks as “early term” rather than “full term” as an approach to making clear that deliveries should wait for 39 weeks. The timing of this release is shown in the figure. Again, we see no evidence of an acceleration of change. If anything, the decline in early term births which had been continuous up to that point is arrested, and the trend flattens, after the change in guidelines.

**Discussion**

Visually, Figure 1 makes clear that these practices do change over time and the discussion above shows that, at least to some extent, these changes appear to be a response to changes in best practice.

Notably and unsurprisingly, we find that major research studies - in particular, those on VBAC and breech delivery - do play a significant role in catalyzing change. However, this information
does not seem to result in immediate changes. It takes years for the decline in the VBAC rate to taper off. If the appropriate level of this treatment is (say) 7% – where it stabilizes – there is a full decade in which a faster decline would have been feasible.

One natural approach to speeding up adoption is to change official guidelines to reinforce study findings. However, the evidence here suggests that is not an effective approach.

It is finally worth noting that changes which recommend more conservative behavior seem to have greater take-up. This is notable in the case of vacuum extraction, where there are large responses to both the FDA and, seemingly, the media, when they indicate that this procedure may be dangerous. There is, however, no corresponding recovery in response to either official guidelines or a major research study promoting safety.

This suggests, overall, that it may be more challenging to change some behaviors than others, and there is clearly space for trying to better publicize new findings about best practices. The fact that the behaviors change slowly leaves open the question of who benefits first, which I turn to now.

4 Heterogeneity in Behavior Change

I focus now on heterogeneity, beginning with geography and demographics. There is significant focus in the health economics literature on variation in practice patterns across space and demographic groups; this practice variation has implications for varying health outcomes across these groups. Here, I ask whether the changes in response to changes in best practice seem to differ across groups.

4.1 Geography and Demographics

A large literature has focused on variation across areas in the US in health outcomes in general (Skinner, 2011) or in changes in health technology (e.g. Skinner and Staiger, 2005). In contrast to this literature - which generally finds significant variation across space in health - I do not see much variation in behavior change across regions of the US. This is visualized in Figure 2, which replicates Figure 1 but includes separate lines for each region of the country. Note that due to limited information on region (namely, it is not available after 2004) the early birth graph covers a shorter time period here.

There are differences in levels of the behaviors - for example, vacuum extraction is used more commonly in the West than elsewhere - but the changes over time are remarkably consistent. There
is neither a coming together nor a fanning out over time across space. The one notable feature is that the short-lived drop in the use of C-section for breech deliveries in 2003 seems to be entirely due to the South. It is not clear why this is.

In contrast, I do see evidence that changes in behavior occur faster for some patient demographic groups. To visualize this, consider Figure 3. This, again, replicates Figure 1 but adds a second series in each figure. This second series shows the relationship between a dummy for low education and the procedure over time - effectively, it shows the difference in the level of the procedure between those with less than a high school degree and those with a high school degree or more.\(^1\)

If we consider the VBAC example (Figure 3a) we see the series clearly moving in opposite directions. As the procedure becomes less frequent, it becomes relatively more common among those with less than a high school degree. Stated conversely, the change over time - the adoption of VBAC - is faster among those with more education. We see similar patterns elsewhere. Appendix Figure A1 shows the same graphs with a dummy for the mother reporting African-American race rather than education. Finally, Appendix Figure A2 uses the predicted death rate - predicted based on demographics - as the demographic measure. This provides a way to summarize the relative speed of changes for better or worse off groups \textit{ex ante}.

In many of these figures and sub-figures we see the same patterns – that is, changes which suggest the change in practices occur earlier for more advantaged demographic groups. To get a sense of whether these patterns are systematic, the first three columns of Table 3 show regression evidence on the relationship. Specifically, in each case I regress the gradient (the coefficient from a regression of the treatment on the demographic measure) on the level of treatment. A negative coefficient here implies that as the level of the treatment goes up, the gradient goes down. In the case of education, for example, the outcome in this regression is the excess procedure rate for those with less education relative to more and the independent variable is the level of procedure. A negative coefficient indicates that as the procedure becomes more common, it becomes relatively less common among those with less education.

Column (1) of Table 3 shows these regression results for education, Column (2) for an indicator for African-American and Column (3) for the predicted mortality measure. In all three cases, the gradient is negative. Since each of these gradients are defined such that they take as the baseline the less advantaged group, this implies that more advantaged women systematically see faster changes.

\(^1\)These figures are computed at the quarterly, rather than the monthly level, to limit noise.
4.2 Outcomes

Changes accrue first to those who are already more advantaged, at least along observable dimensions. To explore the evolving relationship with infant outcomes, Figure 4 echoes Figure 3, but here the coefficient relationship is between treatment and whether the infant has a low APGAR score. To create these coefficients, I regress the low APGAR indicator on procedure measures, and include in the regression a full set of demographic controls.

Figure 4 looks very similar to Figure 3. As procedures become (for example) less common they become relatively more associated with low APGAR scores. Appendix Figure A3 shows the same graph for neonatal mortality (deaths in the first month of life) and Columns (4) and (5) of Table 3 shows the corresponding regressions. As in the observables, the regressions show significant negative coefficients.

This suggests that as the procedures become less common they are more commonly experienced by infants who have poor outcomes. This can be seen as an extension of the observable selection seen above, although since these gradient regressions are conditional on observables, the selection here must be on unobserved characteristics. It should be noted that there is another possible interpretation - namely, that there are heterogeneous treatment effects in these cases, and when the behavior becomes less common the infants who are still treated are those with the most negative treatment effects. This seems possible, but unlikely, as it would suggest that doctors are targeting these treatments to the mothers and infants who benefit least. More plausible is the interpretation that infants who are otherwise disadvantaged also get these practice improvements later.

5 Discussion and Conclusion

This paper opens by asking whether doctors respond to changes in best practices, and whether these changes exacerbate or ameliorate health inequality. The results above suggest that some types of information do prompt response, although even in these cases where in principle behavior change could be very fast, actually change is generally slow. The slow pace of change appears to favor positively selected patients, both based on observable demographics and on examining the outcomes associated with the procedures over time.

This analysis provides added motivation for improving compliance with changing best practices. Universal adoption of improved procedures would disproportionately benefit \textit{ex ante} worse off mothers and infants. Given the poor record of the US overall in maternal and infant health
(National Research Council, 2013) and the particular inequality in outcomes (Chen, Oster and Williams, 2016) these issue are of policy concern.

These results provide less guidance in terms of how to speed up adoption. Official guidelines seem to play little or no role, at least over and above the research studies they are based on. Relying on the example of vacuum extraction, it would seem that alarmist publicity might increase adoption speeds, but this is unlikely to be a widely acceptable policy response. It is also worth saying that the approach here is unable to separate slow behavior change resulting from doctors not learning about new evidence from slow behavior change due to resistance to change even in the face of evidence. Understanding this distinction better empirically could allow development of more effective approaches.

As a final point, the dynamics in Section 4.2 are notable because they highlight (as in Oster, 2018) that these changes in recommendations may increase the challenge of further research on the impacts of these treatments. For research which relies on observational data, there may be biases introduced by endogenous behavioral response. In the case of VBAC, for example, using these data to estimate the risks in 2005 would yield a different estimate than in 1995, with the later data suggesting a more negative outlook on VBAC. Given that much of our scientific knowledge relies on observational data, these dynamics are worth considering when evaluating research results.
References


Arroll, Bruce and Tim Kenealy, “Antibiotics for the common cold and acute purulent rhinitis,” *Cochrane Database of Systematic Reviews*, 2005, (3).


Figure 1: Changes in Procedures Over Time

Panel A: VBAC Changes and Events
Panel B: Breech Section Changes and Events
Panel C: Vacuum Changes and Events
Panel D: Early Birth Changes and Events

Notes: These figures show changes in obstetric procedures over time with each dot representing a month. VBAC is defined only for women with a previous cesarean section. Early term births are births at 37 or 38 weeks, as a share of all births 37 weeks or later. Dotted vertical lines represent the release of research studies related to the procedures. Solid vertical lines represent ACOG committee opinions and practice bulletins. Details of these events appear in Table 2.
Figure 2: Changes by Geographic Area

Panel A: VBAC by Region

Panel B: Breech C-Section by Region

Panel C: Vacuum by Region

Panel D: Early Birth by Region

Notes: These figures show the level of behavior over time by region. The four areas correspond to standard census regions: North East, South, West and Mid-West. Dotted vertical lines represent the release of research studies related to the procedures. Solid vertical lines represent ACOG committee opinions and practice bulletins. Details of these events appear in Table 2.
Figure 3: Variation in Procedure Use by Education Over Time

Panel A: Low Education and VBAC

Panel B: Low Educ and Breech C-Section

Panel C: Low Educ. and Vacuum

Panel D: Low Educ and Early Birth

Notes: These figures show the level of behavior and its relationship with education over time. The levels replicate Figure 1. To create the education series I estimate the difference in treatment levels between those with less education (defined as less than high school) and those with more education than that. To limit noise I do these calculations the quarterly level. Vertical lines indicate events as in Figure 1, with details in Table 2.
Figure 4: Variation in Procedure Effects on Outcomes: APGAR

Panel A: Effect of VBAC on low APGAR

Regression of Low APGAR on VBAC + Controls

Panel B: Effect of Breech on low APGAR

Regression of Low APGAR on Breech x C-Section + Controls

Panel C: Effect of Vacuum on low APGAR

Regression of Low APGAR on Vacuum + Controls

Panel D: Effect of Early Birth on low APGAR

Regression of Low APGAR on Early Birth + Controls

Notes: These figures show the level of behavior and its relationship with APGAR score over time. The levels replicate Figure 1. To create the APGAR series I regress, at the quarterly level, an indicator for low APGAR score (<7) on the treatment and demographic controls. The graphs show evolution of this coefficient over time. Vertical lines indicate events as in Figure 1, with details in Table 2.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Sample Size</th>
<th>Time Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBAC</td>
<td>0.19</td>
<td>0.39</td>
<td>4,768,286</td>
<td>1995-2005</td>
</tr>
<tr>
<td>Vacuum Extraction</td>
<td>0.05</td>
<td>0.22</td>
<td>43,771,430</td>
<td>1995-2005</td>
</tr>
<tr>
<td>Breech Delivered by Cesarean</td>
<td>0.86</td>
<td>0.34</td>
<td>1,694,848</td>
<td>1995-2005</td>
</tr>
<tr>
<td>Early Term</td>
<td>0.29</td>
<td>0.45</td>
<td>74,898,956</td>
<td>1995-2015</td>
</tr>
<tr>
<td>White</td>
<td>0.57</td>
<td>0.49</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Black</td>
<td>0.15</td>
<td>0.35</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.22</td>
<td>0.41</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Other Race</td>
<td>0.06</td>
<td>0.24</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Unknown Education</td>
<td>0.01</td>
<td>0.12</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Less than High School</td>
<td>0.10</td>
<td>0.29</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>High School</td>
<td>0.46</td>
<td>0.50</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>College</td>
<td>0.43</td>
<td>0.50</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Age &lt; 20</td>
<td>0.11</td>
<td>0.31</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>19 &lt; Age &lt; 30</td>
<td>0.52</td>
<td>0.50</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>29 &lt; Age &lt; 40</td>
<td>0.34</td>
<td>0.47</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>Age &gt; 39</td>
<td>0.03</td>
<td>0.16</td>
<td>65,072,301</td>
<td>1995-2015</td>
</tr>
<tr>
<td>North East</td>
<td>0.17</td>
<td>0.38</td>
<td>39,929,013</td>
<td>1995-2004</td>
</tr>
<tr>
<td>South</td>
<td>0.36</td>
<td>0.48</td>
<td>39,929,013</td>
<td>1995-2004</td>
</tr>
<tr>
<td>Midwest</td>
<td>0.22</td>
<td>0.41</td>
<td>39,929,013</td>
<td>1995-2004</td>
</tr>
<tr>
<td>West</td>
<td>0.24</td>
<td>0.43</td>
<td>39,929,013</td>
<td>1995-2004</td>
</tr>
</tbody>
</table>

Notes: This table shows summary statistics on the data used from the NHCS. All variables are binary. VBAC share is the share of women with at least one prior cesarean who have a vaginal birth. Breech delivery by cesarean is limited to breech births. Early term delivery is defined as delivery at 37 or 38 weeks, among all births at or after 37 weeks.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Event 1</th>
<th>Event 2</th>
<th>Event 3</th>
<th>Event 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>VBAC</td>
<td>1996, Negative: A research study from McMahon is the first to show that major complications, including uterine rupture, were more likely for patients who attempted VBAC than those who had a repeat cesarean. This study influences the more restrictive ACOG guidelines later.</td>
<td>1998, Negative: ACOG publishes Practice Bulletin 2 citing the risk of uterine rupture. There is subsequent media coverage on the risks of VBAC.</td>
<td>1999, Negative: ACOG publishes Practice Bulletin 5, which suggests that physicians should be “immediately” available to provide emergency care to women attempting VBAC. This was seen as very restrictive, and many commented that there is no such thing as an immediate cesarean.</td>
<td></td>
</tr>
<tr>
<td>Breech Delivery</td>
<td>2000, Guideline: A research study, Hannah, et. al. (2000), conducts a trial of term breech births. It concludes that planned cesarean sections are better than vaginal births for term breech babies.</td>
<td>2001, Guideline: ACOG publishes a committee opinion encouraging c-sections for breech births.</td>
<td>2003, Guideline: A follow-up research study from Hannah, et. al (2003) finds that there were no substantial differences between breech presenting babies delivered by planned c-section and planned vaginal birth at term 2 years post partum.</td>
<td></td>
</tr>
<tr>
<td>Vacuum Extraction</td>
<td>1998/1999, Negative: The FDA releases a report about the dangers of vacuum extraction. Also media coverage of a NJEM article in which doctors protest a goal of lowering c-section rates.</td>
<td>1998, Positive: ACOG publishes a committee opinion in response to the FDA report to reassure physicians that complications from vacuum extraction are rare.</td>
<td>1999, Positive: Towner, et. al (1999) publish a study, widely cited in the media, demonstrating that the rates of intracranial hemorrhage were similar for vacuum extraction, forceps delivery, and cesarean delivery, refuting prior claims about the relative safety of vacuum extraction.</td>
<td>2000, Positive: ACOG published Practice Bulletin 17, which remarks that there is no difference in mortality between births achieved by forceps and vacuum.</td>
</tr>
<tr>
<td>Redefinition of Full Term Pregnancy</td>
<td>2013, Guideline: ACOG changes the definition of full-term pregnancy, creating four categories.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table shows the information events identified for each outcome.
Table 3: Selection of Behaviors

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Behavior</td>
<td>-0.073***</td>
<td>-0.047***</td>
<td>-0.00017**</td>
<td>-0.025***</td>
<td>-0.034***</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.008)</td>
<td>(0.00008)</td>
<td>(0.005)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Treatment FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Year FE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Number of Obs.</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>600</td>
<td>552</td>
</tr>
</tbody>
</table>

Notes: This table shows statistical evidence on the co-movements between levels of treatment and the gradient with respect to demographics (education and race), the predicted death rate based on demographics and two outcomes (low APGAR score and infant mortality). Each cell represents a different regression of gradients on the level of treatment. An observation is a treatment-month. All regressions include fixed effects for each treatment and year. *indicates significance at the 10% level, **indicates significance at the 5% level, *** indicates significance at the 1% level.
Appendix A: Appendix Figures and Tables

Figure A1: Variation in Procedure Use by Race Over Time

Panel A: Black Race and VBAC
Panel B: Black Race and Breech C-Section
Panel C: Black Race and Vacuum
Panel D: Black Race and Early Birth

Notes: These figures show the level of behavior and its relationship with neonatal mortality (deaths in the first month of life) over time. The levels replicate Figure 1. To create the mortality series I regress, at the quarterly level, an indicator for infant mortality on the treatment and demographic controls. The graphs show evolution of this coefficient over time. Vertical lines indicate events as in Figure 1, with details in Table 2.
Figure A2: Variation in Procedure Use by Predicted Death Over Time

Panel A: Predicted Death and VBAC
Panel B: Predicted Death and Breech C-Section
Panel C: Predicted Death and Vacuum
Panel D: Predicted Death and Early Birth

Notes: These figures show the level of behavior and its relationship with neonatal mortality (deaths in the first month of life) over time. The levels replicate Figure 1. To create the mortality series I regress, at the quarterly level, an indicator for infant mortality on the treatment and demographic controls. The graphs show evolution of this coefficient over time. Vertical lines indicate events as in Figure 1, with details in Table 2.
Figure A3: Variation in Procedure Effects on Outcomes: Mortality

Panel A: Effect of VBAC on Mortality

Panel B: Effect of Breech on Mortality

Panel C: Effect of Vacuum on Mortality

Panel D: Effect of Early Birth on Mortality

Notes: These figures show the level of behavior and its relationship with neonatal mortality (deaths in the first month of life) over time. The levels replicate Figure 1. To create the mortality series I regress, at the quarterly level, an indicator for infant mortality on the treatment and demographic controls. The graphs show evolution of this coefficient over time. Vertical lines indicate events as in Figure 1, with details in Table 2.