HSR Health Services Research

© Health Research and Educational Trust
DOI: 10.1111/1475-6773.12061
RESEARCH ARTICLE

# Midwifery Care at a Freestanding Birth Center: A Safe and Effective Alternative to Conventional Maternity Care

*Sarah Benatar, A. Bowen Garrett, Embry Howell, and Ashley Palmer*

**Objective.** To estimate the effect of a midwifery model of care delivered in a freestanding birth center on maternal and infant outcomes when compared with conventional care.

**Data Sources/Study Setting.** Birth certificate data for women who gave birth in Washington D.C. and D.C. residents who gave birth in other jurisdictions.

**Study Design.** Using propensity score modeling and instrumental variable analysis, we compare maternal and infant outcomes among women who receive prenatal care from birth center midwives and women who receive usual care. We match on observable characteristics available on the birth certificate, and we use distance to the birth center as an instrument.

**Data Collection/Extraction Methods.** Birth certificate data from 2005 to 2008.

**Principal Findings.** Women who receive birth center care are less likely to have a C-section, more likely to carry to term, and are more likely to deliver on a weekend, suggesting less intervention overall. While less consistent, findings also suggest improved infant outcomes.

**Conclusions.** For women without medical complications who are able to be served in either setting, our findings suggest that midwife-directed prenatal and labor care results in equal or improved maternal and infant outcomes.

**Key Words.** Midwives, prenatal care, cesarean section, low-income women, African American women

Organized around a commitment to nonintervention in the normal pregnancy and birth process, midwifery models offer personalized and holistic care and education, including prenatal care and attending to the physical, psychological, and social well-being of the mother (Rooks et al. 1989; American College of Nurse-Midwives 2004; Ickovics et al. 2007; Dominguez et al. 2008). In this article, we examine the impact of a midwifery model of care on birth

outcomes. We compare a set of maternal and infant outcomes reported on the birth certificate for women cared for by midwifes in a freestanding birth center with women who receive conventional care. Freestanding birth centers are typically staffed by nurse midwives and operate separately from hospitals—presenting women with an alternative setting in which to receive prenatal and labor care. Previous studies of midwife-led and birth center care in the United States have done little to address selection bias—whereby women served in these settings have an average lower risk of poor birth outcomes than the comparison group. In an attempt to address these risk selection problems, we use propensity score modeling techniques to control for observed risk factors, and instrumental variable analysis to address remaining concerns that unobserved differences could bias the estimated effects of birth center care.

## Background and Motivation

The conceptual framework that motivates our analysis draws on a behavioral model of service use and quality. Individual predisposing, enabling, and need (risk) factors (Anderson 1995), as well as system and environmental factors, drive health practices, use of services, and subsequent outcomes. Consequently, our research posits that high-quality, individualized services offered at the birth center lead to improved birth outcomes.

Existing literature suggests that these underlying factors can affect birth outcomes. Midwife-attended births have been associated with improved perinatal outcomes, including fewer interventions overall, fewer cesarean deliveries in particular, and improved patient satisfaction (Raisler and Kennedy 2005). It has been argued that these benefits can be explained by the fact that midwives generally see women with lower medical risk profiles. Declercq, Menacker, and MacDorman (2006), however, demonstrate through a birth certificate analysis that midwives caring for average risk women still see better than average results. Lenaway et al. (1998) find that midwife-directed care can be especially beneficial for an indigent population. There has been less research specifically on birth center care, or birth center care delivered to low-income or minority women (Davis et al. 2011), and a 2004 review of research

---

Address correspondence to Sarah Benatar, Ph.D., the Urban Institute, 2100 M St. NW, Washington, DC 20037; e-mail: sbenatar@urban.org. A. Bowen Garrett, Ph.D., is with McKinsey & Company, Washington, DC, and began this work while at the Urban Institute. Embry Howell, Ph.D., and Ashley Palmer, M.A., also are with the Urban Institute, Washington, DC.

on midwife care in birth center settings raises concerns about the quality of research on this topic and the need for well-designed studies (Walsh and Downe 2004).

Low-income and minority women are disproportionately at risk for experiencing more intervention at delivery and are more likely to experience poor birth outcomes overall (Kistka et al. 2007; Getahun et al. 2009; Blumenshine et al. 2010; Martin et al. 2010). Persistent racial and socioeconomic disparities in maternal and infant health have not been adequately explained by differences in demographic, medical, or behavioral risk factors (Goldenberg et al. 1996; Lu et al. 2003). In particular, African American women are more likely to deliver by C-section than other racial and ethnic groups in the United States, even when controlling for risk factors and insurance status (Aron et al. 2000; Getahun et al. 2009; Martin et al. 2010). Compared with white women, African American women are also nearly twice as likely to have a low-birth-weight infant, and twice as likely to initiate prenatal care late in pregnancy or not at all (Collins and David 2009; Martin et al. 2010).

C-sections are critical, at times, for the health of mother and child, but they are increasing at a rate that is not consistent with clinical determinations of necessity. World Health Organization guidelines published more than two decades ago suggest that C-section rates should not exceed 15 percent (WHO 1985). A recent review finds that C-section rates above 10 percent are in fact associated with an increase in maternal and neonatal mortality and morbidity (Althabe et al. 2006).

Since 1996, the C-section rate in the United States has risen more than 56 percent, climbing to 32.8 percent in 2010 (Hamilton, Martin, and Ventura 2011). Explanations for why C-sections are increasing include provider and patient preference, resistance to vaginal breech deliveries, greater tolerance for surgical intervention, an increase in primary C-sections, a disinclination toward vaginal birth after cesarean (VBAC), and malpractice concerns (Burns, Geller, and Wholey 1995; Yang et al. 2009; Kaimal and Kuppermann 2010). There remains little understanding of why C-section rates are higher among African American women, although overweight and obesity may be contributing factors (Getahun et al. 2009).

The increase in primary C-section rates is coupled with a precipitous decline in VBAC, down from a high of approximately 30 percent in 1996 to about 8 percent in 2007 (Curtin, Kozak, and Gregory 2000; Martin et al. 2010). This sharp decrease has been attributed to a combination of clinical recommendations, nonclinical factors, and legal concerns (Gregory, Fridman, and Korst 2010).

In addition to rising C-section rates, recent research points to increased infant morbidity among early term births, and generally cautions against unnecessary intervention prior to 39 weeks gestation (Wang et al. 2004; Engle and Kominiarek 2008; Petrini et al. 2009; Fleischman, Oinuma, and Clark 2010; Martin et al. 2010).

## METHODS

The purpose of this study is to better understand whether midwife-led birth center care can improve birth outcomes relative to traditional care. In light of the conceptual model cited earlier, we hypothesize that the birth center delivery model is an enabling factor that can provide improved access to culturally appropriate care and result in improved outcomes.

Producing convincing evidence of such effects, in the absence of a randomized control trial, requires statistical methods to isolate the causal effects of birth center care that are generally lacking in the existing literature, while controlling adequately for risk. We do this by comparing the maternal and infant outcomes of women cared for in Washington, D.C. by midwives from the Family Health and Birth Center (FHBC) with a very similar group of women who received usual care in the District.

The FHBC is the only freestanding birth center in D.C. Founded in 1998, the FHBC serves a primarily low-income, African American population and is housed with a child care center, a health care facility, and other supportive services. Approximately, 30 percent of women who receive prenatal care at the birth center choose to give birth at the FHBC, whereas the remaining 70 percent give birth at a local hospital with FHBC midwives. In some instances, giving birth at the hospital is medically indicated or necessary given other risk factors (e.g., homelessness), but often women choose to deliver at the hospital because of personal preference.

An earlier qualitative study characterizes the FHBC model of care, including a description of the services offered, the staff make-up, and typical care protocols (Palmer, Cook, and Courtot 2010). Based on case study interviews and focus groups with patients, the study authors observe that the FHBC provides comprehensive care to meet the varied needs of a socially high-risk group of women, and it works to address factors related to poor birth outcomes, including stress, inadequate social/emotional support, poor education, and poverty (Palmer, Cook, and Courtot 2010).

### *Outcomes*

We examine a set of maternal and infant outcomes, comparing women cared for at the FHBC and a similar group who receive usual care. The maternal outcomes we consider include (1) whether the mother had a vaginal birth or C-section; (2) whether forceps or vacuum extraction was used; (3) whether electronic fetal monitoring (EFM) was utilized during labor (EFM has been associated with increased C-section rates); (4) whether the mother had a vaginal birth after a previous cesarean (VBAC); and (5) whether the birth took place on the weekend. Finding the induction variable on the birth certificate to be unreliable, we construct a "weekend delivery" variable as a proxy to understand the extent to which there was interference with the birthing process. Prior research has found that most scheduled cesareans and inductions take place on weekdays because it is a more convenient time for medical staff (Burns, Geller, and Wholey 1995). Therefore, if delivery has an equal probability on each day in the absence of intervention (1 in 7), any deviation in the rate of weekend delivery from 2 in 7 suggests intervention.

The infant outcomes we examine include (1) preterm birth (gestational age <37 weeks); (2) whether the infant had a low APGAR score ($\leq 7$) at 5 minutes; (3) incidence of low birthweight (<2,500 g); and (4) average birthweight (at term and for all births).

### *Data*

We analyze birth certificate data between 2005 and 2008 for all women who gave birth in D.C. as well as District residents who gave birth in other jurisdictions. Birth certificates of women who received at least two prenatal care visits at the FHBC are flagged in the complete birth certificate file. FHBC staff provided mother's name, mother's date of birth, infant's date of birth, infant's sex, and birth site location for women who had at least two prenatal care visits at the FHBC to the D.C. Department of Health to facilitate this process.

Names were matched to birth certificates based on a combination of the first three letters of the mother's first and last name and her date of birth. Unmatched cases were manually matched. Data were linked only when there was an exact match, achieving a match rate of 92.8 percent.

### *Sample*

We limit our analysis sample to women with at least two prenatal visits, a singleton birth, and a gestational age $\geq 24$ weeks. These criteria match the

inclusion criteria of the FHBC, which does not deliver multiples or extremely preterm births.

In our study group, we retain all women who initiated care at the FHBC to avoid selecting out women who may have transferred care due to risk or other factors that are potentially correlated with the outcome measures. Therefore, the FHBC group includes women who delivered at a local hospital with FHBC midwives, women who delivered at the birth center with FHBC midwives, and women who initiated care at the FHBC but transferred for various reasons.

With an initial sample size of 890 FHBC observations and 61,071 usual care observations, we exclude 2,572 plural births, 762 births with fewer than two prenatal visits, and 233 births with gestational age <24 weeks. More than 14,200 cases are excluded as a result of propensity score modeling procedures (described below); specifically because one or more characteristics perfectly predict not being in the FHBC sample. We have a final analysis sample of 872 FHBC and 42,987 usual care births. Using all available usual care cases that meet our criteria, we obtain more statistically efficient estimates than if we were to force equal sample sizes.

### *Propensity Score Analysis*

We construct a comparison group with observable characteristics nearly identical to those in the birth center group using a propensity score reweighting approach.[1] Given available data, this approach controls as much as possible for pre-existing differences in women who receive prenatal care at the FHBC versus women who receive usual care. The FHBC specifically targets minority and disadvantaged women who have fairly low medical risks, and designs its program to provide supportive services to these women. Therefore, as outlined in our conceptual model, we control for a set of underlying predisposing demographic and health characteristics of the mother, as well as other observed measures of risk in estimating effects of midwife-directed care provided by the FHBC. These variables include mother's age, race, marital status, zip code, education, the presence of information about the father on the birth certificate, cigarettes smoked by the mother, parity, health risk factors exogenous to the mother's pregnancy (chronic hypertension, cardiac disease, lung disease, diabetes), pregnancy risk factors (previous birth of an exceptionally large or small baby, previous stillborn birth, previous preterm birth), and the month in which the mother initiated prenatal care. We match on all available predetermined characteristics, while avoiding matching on

characteristics that are endogenous to the pregnancy (e.g., total number of prenatal visits).

To compute the propensity score, we estimate a logistic regression model of the probability of having sought prenatal care at the FHBC as a function of the set of matching variables.[2] The propensity score for each observation is the predicted probability from the regression. There are several instances in which a control variable perfectly predicts an FHBC measure.[3] For example, if there are no FHBC cases from a particular zip code, the dummy variable for that zip code perfectly predicts not using the birth center. In these cases, women were assigned a propensity score weight of 0, resulting in their exclusion from the analyses.

To construct propensity score-adjusted weights, we divide the propensity score distribution into 20 groups (indexed by $j$) with approximately equal numbers of observations (i.e., 0–5th percentile, 5th–10th percentile, etc.). We then compute the share of FHBC cases within each of the 20 groups, which we denote as $P_{BC,\,j}$. FHBC cases all receive a weight ($W_i$) of 1. For all other cases, $W_i = P_{BC,\,j}/(1 - P_{BC,\,j})$. As a result, the weighted distribution of propensity scores for usual care cases is similar to the propensity score distribution for FHBC cases.

Our approach results in a (weighted) usual care comparison sample that is very similar to the FHBC sample on all observable characteristics (see Table 1). We regress each matching variable on the FHBC indicator using weighted logistic regression to confirm there are no statistically significant remaining differences. Here, and in other regressions reported below, we employ Huber-White heteroskedasticity-consistent standard errors. Because the sum of the weights in the comparison sample equals total $N$ of the birth center sample, the two samples receive equal weight in the regressions. As an additional check, we examine three pair-wise cross-tabulations on the three important predictors of FHBC use (being non-Hispanic, being less than 35 years old, and living in a specific zip code), and find that the cross-tabulations are nearly identical.

With FHBC and weighted usual care samples that are nearly identical on the dimensions accounted for by matching variables, the comparison of outcome measures between the two samples is straightforward. For binary outcomes, we estimate weighted logistic regression models of each outcome measure on the FHBC variable (and a constant) and report the odds ratios for the birth center variable. We obtain qualitatively similar results using weighted linear probability models (i.e., weighted least squares). For continuous outcome measures, we use weighted linear regression.

Table 1: Comparison of Family Health and Birth Center (FHBC) to Usual Care Samples Before and After Propensity Score Reweighting†

| *Sample Characteristics* | *FHBC Prenatal Care (%) N = 872* | *Usual Care (%) (Unweighted) N = 42,987* | *Usual Care (%) (Weighted) N = 42,987* |
|---|---|---|---|
| 19 and younger | 21.9 | 11.3** | 21.3 |
| Married | 22.6 | 43.6* | 23.2 |
| 35 and older | 7.7 | 19.1** | 7.9 |
| Smoker | 5.7 | 3.1** | 5.7 |
| First birth | 49.0 | 45.3* | 49.0 |
| Cardiac disease | 0.1 | 0.1 | 0.1 |
| Lung disease | 0.2 | 0.1 | 0.2 |
| Diabetes | 2.3 | 3.4 | 2.3 |
| Herpes | 1.1 | 1.0 | 1.1 |
| Chronic hypertension | 0.3 | 1.2* | 0.3 |
| White Non-Hispanic | 12.0 | 22.1** | 12.5 |
| Black Non-Hispanic | 85.0 | 53.8** | 84.8 |
| Hispanic | 1.5 | 21.1** | 1.2 |
| Other race | 1.5 | 3.1** | 1.5 |

*$p < .05$; **$p < .01$.
†Additional variables were included in the propensity score modeling, including parity, education level, zip code of residence, the presence of information about the father on the birth certificate, previous birth of an exceptionally large or small baby, previous stillborn birth, previous preterm birth, and the month in which the mother initiated prenatal care.

### *Instrumental Variable Analysis*

Although the propensity score reweighting approach is effective at controlling for observed characteristics, it does not control for unobserved characteristics that may affect outcomes. To address the concern that unobserved differences in risk could still bias the estimated effects of FHBC care using the propensity score approach, we also conduct an instrumental variable (IV) analysis.

An instrumental variable should (1) have a strong effect on FHBC use; and (2) only influence the outcome measures through its effect on FHBC use (after other covariates are held fixed). The instrument used here is the cube root of the distance to the FHBC. We base the distance measure on residential census tract where it is available on the birth certificate (43 percent of cases). Absent census tract distance, we use zip code distance where residential state matches the state listed for the mailing address (55 percent of cases). We set distance to missing for the remaining 2 percent where residential state and mailing address do not match. Taking the cube root provides a better fit

than linear distance (or the square or fourth root of distance). Distance has been used as an instrumental variable in several health services research studies.[4]

The first requirement for an instrument is clearly satisfied. In a simple linear probability model of being in the birth center as a function of the instrument and controls, the instrument is a strong predictor of FHBC use with an *F*-statistic of 321.5. The second requirement—that the instrument only affects the outcomes through use of the FHBC—cannot be fully tested, but we examine the assumption. The primary concern is that distance to the birth center is correlated with unmeasured risk factors that affect birth outcomes.

Although correlations between the instrument and unmeasured risk factors are untestable, we explore how the instrument correlates with observed risk factors. If distance is uncorrelated with observed risk factors, we can be more confident that it is not correlated with unobserved risk factors. After adjusting for age and race/ethnicity, we find no statistically significant relationship between the instrument and cardiac disease, diabetes, and hypertension. For smoking, lung disease, and herpes, however, we find statistically significant evidence that a lower prevalence of these risk factors is associated with greater distance from the FHBC. If such a relationship holds for unobserved risk factors as well (i.e., unobserved risk is higher near the FHBC), the IV estimates would be biased in the direction of unfavorable birth center effects. Thus, our IV results are likely conservative in terms of demonstrating the benefits of FHBC care.

For our application, given that most of our outcome measures are binary, we use bivariate probit models in which an outcome measure equation and a birth center status equation are estimated simultaneously with correlated error terms. The instrument is included in the birth center equation only, and a common set of other control variables are included in both equations. For continuous outcome measures, we use two-stage least squares (2SLS) with logit-predicted birth center status as the instrument, where the logit equation contains all control variables and distance to the FHBC.

In addition, the bivariate probit and 2SLS models are propensity score reweighted (using all control variables in the propensity score approach except the zip code variables, given that we are using distance as an instrument), so that the treatment and weighted comparison sample are balanced on covariates as a starting point. The N for the comparison sample in this analysis is greater than the N for the propensity score analysis (56,945 vs. 42,987) because we did not include zip code as a matching variable when propensity score reweighting for the IV analysis. In turn, some cases that had been

dropped for perfectly predicting FHBC status in the propensity score analysis were retained in the IV analysis.

## RESULTS

### *Propensity Score Analysis*

Using propensity score reweighting, we estimate the effect of FHBC care relative to usual care on maternal and infant outcomes for the FHBC group versus a usual care group with nearly identical observed characteristics, and we report the results in Table 2.

Receiving prenatal care at the FHBC is associated with fewer obstetrical interventions. Specifically, women in the FHBC group are significantly less likely to deliver via C-section compared with women who receive usual care (19.7 percent vs. 29.4 percent), and they are significantly less likely to deliver with the assistance of forceps or vacuum extraction than the usual care group (2.1 percent vs. 4.4 percent). In addition, women in the FHBC group are

Table 2: Effect of Birth Center Use on Maternal and Infant Outcomes (All Women, Propensity Score Analysis)

| | *FHBC (%) (N = 872)* | *Usual Care (%) (N = 42,987)* | *Difference (Percentage Points)* | *Odds Ratio*[¶] |
|---|---|---|---|---|
| Logistic regression | | | | |
| Cesarean section | 19.7 | 29.4 | −9.7 | 0.59** |
| Use of vacuum or forceps | 2.1 | 4.4 | −2.3 | 0.45** |
| Electronic fetal monitoring | 78.1 | 82.2 | −4.3 | 0.77** |
| Weekend delivery | 28.6 | 23.9 | 4.8 | 1.28** |
| Preterm birth ($\leq$ 36 weeks) | 7.9 | 11.0 | −3.1 | 0.70** |
| APGAR 5 score <7[†] | 3.4 | 3.7 | −0.4 | 0.92 |
| Low birthweight (<2,500 g) | 8.0 | 10.0 | −2.0 | 0.81 |
| Vaginal birth after cesarean section[‡] | 26.7 | 9.4 | 17.3 | 3.50** |
| Linear regression | | | | *Marginal Effect* |
| Average birthweight (g) | 3,245 | 3,166 | 79 | 79** |
| Average birthweight at term (g)[§] | 3,325 | 3,282 | 43 | 43** |

*$p < .05$; **$p < .01$.
[†]N for APGAR score at 5 minutes is smaller than the other measures due to missing values.
[‡]This analysis is restricted to women who had a previous C-section ($N = 4{,}250$ for the usual care group and $N = 45$ for the Family Health and Birth Center [FHBC] group).
[§]Ns for term births are 803 for the FHBC group and 38,773 for the usual care group.
[¶]Usual care is the reference group.

significantly less likely to receive EFM during delivery (78.1 percent vs. 82.2 percent). It is, however, hospital policy that all women who deliver at the hospital receive at least intermittent EFM while they labor. As a result, women who delivered at the FHBC drive any difference in receipt of EFM.

The FHBC group is significantly more likely to deliver on a weekend (28.6 percent vs. 23.9 percent), suggesting higher rates of intervention among the usual care group. Finally, among women with a previous C-section, we find a significantly higher proportion of VBACs among women who receive care at the FHBC versus usual care (26.7 percent vs. 9.4 percent).

There are significantly fewer preterm births in the FHBC group (7.9 percent vs. 11.0 percent), but no significant differences in APGAR score or the incidence of low birthweight between the two groups. Average birthweight among the two cohorts, however, is significantly different. At all gestational ages, we find that babies in the FHBC group weigh on average 79 g more than the comparison group. When we restrict the data to term births ($N = 803$ for the FHBC group and $N = 38{,}773$ for usual care), we find a smaller but still statistically significant difference in birthweight remains (43 g). Furthermore, we find significant impacts on the gestational age distribution among the FHBC group when compared with the usual care group. Specifically, women in the FHBC group are more likely to carry their babies to term than women in the usual care group, and less likely to deliver during the "early term" period (37–39 weeks) that is associated with increased morbidity (Fleischman, Oinuma, and Clark 2010). In addition, we find that fewer C-sections are performed between 37 and 39 weeks for the FHBC group when compared with usual care (not reported in table—details available on request).

### *Instrumental Variable Analysis*

The findings using instrumental variable approaches (bivariate probit for binary outcomes and 2SLS for birthweight) are similar in direction, size, and statistical significance to the propensity score results, although the effects are generally larger with the IV analysis (Table 3). For example, the marginal effect from the bivariate probit model shows that the birth center sample has a 10.5 percentage point lower incidence of C-section (similar to the 9.7 percentage point lower incidence reported in the propensity score analysis). For preterm delivery, the differences are 6.2 percentage points in the IV analysis compared to 3.1 percentage points in the propensity score analysis. In both cases, the differences are statistically significant. The FHBC sample has an

Table 3: Effect of Birth Center Use on Maternal and Infant Outcomes (All Women, Instrumental Variables Analysis)

| | *FHBC (%) (N = 872)* | *Usual Care (%) (N = 56,073)†* | *Difference (Percentage Points)* | *Marginal Effect* |
|---|---|---|---|---|
| Bivariate probit regression | | | | |
| Cesarean section | 19.7 | 29.6 | −10.0 | −10.5** |
| Use of vacuum or forceps | 2.1 | 4.6 | −2.5 | −3.3** |
| Electronic fetal monitoring | 78.2 | 83.3 | −5.1 | −10.2* |
| Weekend delivery | 28.6 | 23.4 | 5.2 | 11.7* |
| Preterm birth ($\leq$36 weeks) | 7.9 | 11.0 | −3.0 | −6.2** |
| APGAR 5 score <7 | 3.4 | 3.6 | −0.27 | 0.209** |
| Low birthweight (<2,500 g) | 8.4 | 10.2 | −1.8 | −1.6 |
| Vaginal birth after cesarean section‡ | 26.7 | 9.3 | 17.3 | 57.7** |
| 2SLS | | | | |
| Average birthweight (g) | 3,245 | 3,164 | 81 | 76 |
| Average birthweight at term (g)§ | 3,325 | 3,278 | 46 | 45 |

*$p < .05$; **$p < .01$.
†We did not use zipcode in the propensity score reweighting for this analysis as it is the basis for the instrument, and as a result the *N* is larger than in Table 2.
‡This analysis was restricted to only women who had a previous C-section ($N = 4{,}171$ for the usual care group and $N = 45$ for the Family Health and Birth Center [FHBC] group).
§$N = 48{,}769$ for average birthweight at term.

11.7 percentage point higher likelihood of delivering on the weekend according to the IV analysis, compared to a 4.8 percentage point difference in the propensity score analysis; again in both cases, the differences are statistically significant.

In the bivariate probit analysis, we observe an unexpectedly positive and significant result for APGAR score <7 at 5 minutes, and an unexpectedly dramatic difference in VBAC rates. We note, however, that the magnitude of the difference for the low APGAR score is very small and both low APGAR and VBAC are rare in the two samples, which could lead to high variability in the IV estimates. Findings for EFM and the use of instruments are qualitatively similar in the IV analysis and the propensity score analysis. Birthweight differences are not statistically significant in the IV analysis, but they are of the same magnitude as in the propensity score analysis, suggesting that they are not biased.

### *African American Subgroup Analysis*

As noted earlier, the FHBC model is designed to support the care of low-income African American women, a group that experiences poor birth

outcomes beyond what can be explained by prenatal care, insurance status, or other medical risk factors. As the original mission of the FHBC is to improve outcomes for this group, we conducted a subgroup analysis that restricts the sample to African American women in the treatment and comparison groups. The propensity scores are re-estimated using only data for African American women, and results of this subgroup analysis are presented in Table 4.[5] The number of observations falls more than proportionately (relative to Table 2) in the comparison group because African Americans make up a lower share of the unweighted comparison group. We recomputed propensity scores and weights for this subsample analysis, so that the sum of weights in the comparison groups equals the total *N* in the FHBC group; thus, the two samples receive equal weight in the regressions.

African American women who receive care from FHBC midwives are significantly less likely to have a C-section than African American women who receive usual care (20.9 percent vs. 29.7 percent). In addition, they are significantly less likely to have a delivery assisted by forceps or vacuum extraction (1.9 percent vs. 4.7 percent). There is no statistically significant difference in EFM for this subgroup. In contrast to the full sample findings, rates

Table 4: Effect of Birth Center Use on Maternal and Infant Outcomes (African American Women, Propensity Score Analysis)

| | *FHBC (%) (N = 744)* | *Usual Care (%) (N = 27,095)* | *Difference (Percentage Points)* | *Odds Ratios* |
|---|---|---|---|---|
| Logistic regression | | | | |
| Cesarean section | 20.9 | 29.7 | −8.9 | 0.62** |
| Use of vacuum or forceps | 1.9 | 4.7 | −2.8 | 0.39** |
| Electronic fetal monitoring | 82.5 | 84.7 | −2.2 | 0.85 |
| Weekend delivery | 27.1 | 24.4 | 2.7 | 1.15* |
| Preterm birth ($\leq$36 weeks) | 8.6 | 11.8 | −3.2 | 0.71** |
| APGAR 5 score <7[†] | 3.4 | 3.7 | −0.3 | 0.90 |
| Low birthweight (<2,500 g) | 9.8 | 11.1 | −1.4 | 0.872 |
| Vaginal birth after cesarean section[‡] | 20.0 | 7.9 | 12.3 | 2.93** |
| Linear regression | | | | *Marginal Effect* |
| Average birthweight (g) | 3,198 | 3,130 | 67 | 67** |
| Average birthweight at term (g)[§] | 3,325 | 3,282 | 43 | 43** |

*$p < .05$; **$p < .01$.
[†]Ns for APGAR score at 5 minutes is smaller than the other measures due to missing values.
[‡]This analysis is restricted to women who had a previous C-section ($N = 2{,}302$ for the usual care group and $N = 40$ for the Family Health and Birth Center [FHBC] group).
[§]Ns for term births are 679 for the FHBC group and 23,322 for the usual care group.

of weekend delivery are not significantly different when analyses are restricted to African American women.

African American women in the FHBC group are less likely to have preterm babies than those in the usual care group (8.6 percent vs. 11.8 percent). As with the full sample analysis, there are few significant differences in the other infant outcomes we measure (APGAR score and low birth weight). We do, however, find a significant difference in average birthweight overall (67 g) and when we restrict the analysis to term births (43 g).

## Discussion

By reducing C-sections, increasing average birthweight, and prolonging gestational age, care delivered at Washington D.C.'s only freestanding birth center results in improved or "as good" maternal and infant outcomes. Our findings provide evidence that after controlling for other predisposing and risk factors, birth center care can be an important enabling factor contributing to improved birth outcomes. We posited that highly individualized prenatal care delivered in a culturally relevant and comfortable environment would have a positive effect on a population that is at greater risk of experiencing poor or undesirable birth outcomes, and our findings suggest that this setting, when compared with traditional care, has indeed led to improved birth outcomes for higher need infants and mothers touched by the FHBC—the enabling factor.

Using propensity score modeling to fully control for available data, retaining women who transferred care in the study sample, and obtaining similar results using distance to birth center in an instrumental variable analysis, we go well beyond existing studies in addressing selection bias, on both measured and unmeasured characteristics.

Given rising health care costs, and a continued trend of increasing C-sections, these results suggest that alternative models of maternity care can be safe and effective in promoting noninterventionist approaches, improving maternal and infant outcomes, and perhaps addressing the seemingly intractable problem of low-birthweight and preterm babies in the United States. In particular, lower incidence of C-section, higher likelihood of VBAC, and increases in gestational age among the FHBC group are important cost-saving and health-promoting outcomes. Proliferation of this model among women with low medical risk pregnancies, including those with increased social risk factors, could contribute to improved maternal and infant outcomes. This shift

in system and environmental factors could also result in cost savings that would be realized as a result of avoiding unnecessary obstetrical procedures and improved the health of mothers and infants. Again, we note that while this study controls for observed risk factors using propensity score modeling and seeks to correct for potential unobserved risk factors using an instrumental variable method, there are limitations. Where random assignment to the birth center setting is not available or feasible, ideally, the propensity score method would be able to match on a more complete set of risk factors and maternal characteristics not available on the birth certificate, such as direct income measures. Also, the requirement of the instrumental variable—that it only affects the outcome through its effect on birth center cannot be tested— must be assumed. If unobserved risk factors are consistent with observed risk factors, this would suggest that those living near the FHBC, and more likely to use it, are likely to be at increased risk, suggesting that the beneficial findings for birth center could be understated. Although it seems unlikely, we cannot rule out the possibility that unobserved risk factors run counter to observed ones, which could imply our results overstate the beneficial effects.

Finally, this study examines the effects of care delivered at one birth center, and it would benefit from replication in other settings with midwife-directed care. Facility-level factors of the FHBC, including staffing patterns, policies, and organizational arrangements, may be unique to this setting and may not be representative of all midwife-directed practices. This model, however, holds promise as a supportive method of prenatal care for many women, including racial and economic minorities who are at disproportionately higher risk of experiencing poor birth outcomes. While further study is necessary, this analysis indicates that birth center care can be safe and beneficial, improving care for many women and their infants.

## Acknowledgments

*Joint Acknowledgment/Disclosure Statement*: We would like to acknowledge Ruth Lubic and Madeleine Beebe of the Family Health and Birth Center as well as Fern Johnson-Clarke of the D.C. Department of Health for their help acquiring and matching the birth certificate data used for this study. In addition, we are grateful to the members of the Advisory Committee and Urban Institute colleagues who commented on the design of the research. Finally, we thank the Kellogg Foundation, whose funding made this study possible.

*Disclosures*: None.

*Disclaimer*: The views expressed are those of the authors and should not be attributed to the Urban Institute, its trustees, its funders, or to McKinsey & Company.

## NOTES

1. We use a propensity score reweighting approach rather than matching because reweighting uses the data more efficiently and does not require arbitrary assumptions about how many untreated observations to match to each treated observation.
2. Results of the logistic regression used to construct the propensity scores are available upon request.
3. This is because the matching variables are expressed as dummy variables, and because only 3.2 percent of cases in the analysis file are birth center cases.
4. See McClellan, McNeil, and Newhouse (1994), McConnell et al. (2005), and Brooks et al. (2006). These papers use a differential distance measure from the patient to alternative providers, whereas we use a direct distance measure to the birth center. The direct measure is more appropriate here because alternative providers of prenatal care are numerous and not easily identified.
5. Given the similarity between propensity score reweighted and instrumental variable results for all women, we focus on the propensity score reweighted results in the African American subsample analysis.

## REFERENCES

Althabe, F., C. Sosa, J. M. Belizán, L. Gibbons, F. Jacquerioz, and E. Bergel. 2006. "Cesarean Section Rates and Maternal and Neonatal Mortality in Low-, Medium-, and High-Income Countries: An Ecological Study." *Birth* 33 (4): 270–7.

American College of Nurse Midwives. 2004. "Philosophy of `the American College of Nurse Midwives: Standard Setting Document" [accessed on March 28, 2013]. Available at http://www.midwife.org/index.asp?bid=59&RequestBinary=True&rec=49&cat=2

Anderson, R. M. 1995. "Revisiting the Behavioral Model and Access to Medical Care: Does It Matter?" *Journal of Health and Social Behavior* 36 (1): 1–10.

Aron, D. C., H. S. Gordon, D. L. DiGiuseppe, D. L. Harper, and G. E. Rosenthal. 2000. "Variations in Risk-Adjusted Cesarean Delivery Rates According to Race and Health Insurance." *Medical Care* 38 (4): 35–44.

Blumenshine, P., S. Egerter, C. J. Barclay, C. Cubbin, and P. A. Braveman. 2010. "Socioeconomic Disparities in Adverse Birth Outcomes: A Systematic Review." *American Journal of Preventive Medicine* 39 (9): 263–72. doi: 10.1016/j.amepre.2010.05.012

Brooks, J. M., C. P. Irwin, L. G. Hunsicker, M. J. Flanigan, E. A. Chrischilles, and J. F. Pendergast. 2006. "Effect of Dialysis Center Profit Status on Patient Survival: A Comparison of Risk-Adjustment and Instrumental Variable Approaches." *Health Services Research* 41 (6): 2267–89.

Burns, L. R., S. E. Geller, and D. R. Wholey. 1995. "The Effect of Physician Factors on the Cesarean Section Decision." *Medical Care* 33 (4): 365–82.

Collins, J. W., and R. J. David. 2009. "Racial Disparity in Low Birth Weight and Infant Mortality." *Clinics in Perinatology* 36 (1): 63–73. doi: 10.1016/j.clp.2008.09.004

Curtin, S. C., L. J. Kozak, and K. D. Gregory. 2000. "U.S. Cesarean and VBAC Rates Stalled in the Mid-1990s." *Birth: Issues in Perinatal Care* 27 (1): 54–7. doi: Article.

Davis, D., S. Baddock, D. Tchng, S. Pairman, M. Hunter, C. Benn, D. Wilson, L. Dixon, and P. Herison. 2011. "Planned Place of Birth in New Zealand: Does It Affect Mode of Birth and Intervention Rates among Low-Risk Women?" *Birth* 38 (2): 111–20.

Declercq, E., F. Menacker, and M. MacDorman, 2006. "Maternal Risk Profiles and the Primary Cesarean Rate in the United States, 1991–2002." *American Journal of Public Health*, 96 (5): 867–72. doi: 10.2105/AJPH.2004.052381

Dominguez, T. P., C. Dunkel-Schetter, L. M. Glynn, C. Hobel, and C. A. Sandman. 2008. "Racial Differences in Birth Outcomes: The Role of General, Pregnancy, and Racism Stress." *Health Psychology* 27 (2): 194–203. doi: 10.1037/0278-6133.27.2.194

Engle, W. A., and M. A. Kominiarek. 2008. "Late Preterm Infants, Early Term Infants, and Timing of Elective Deliveries." *Clinics in Perinatology* 35 (2): 325–41. doi: 10.1016/j.clp.2008.03.003

Fleischman, A. R., M. Oinuma, and S. L. Clark. 2010. "Rethinking the Definition of 'Term Pregnancy.'" *Obstetrics & Gynecology* 116 (7): 136–9. doi: 10.1097/AOG.0b013e3181e24f28

Getahun, D., D. Strickland, J. M. Lawrence, M. J. Fassett, C. Koebnick, and S. J. Jacobsen, 2009. Racial and Ethnic Disparities in the Trends in Primary Cesarean Delivery Based on Indications. *American Journal of Obstetrics and Gynecology* 201 (10): 422.e1–7. doi: 10.1016/j.ajog.2009.07.062

Goldenberg, R. L., J. F. Culhane, J. Dlams, and R. Romero. 1996. "Epidemiology and Causes of Preterm Birth." *Lancet* 371: 75–84.

Gregory, K. D., M. Fridman, and L. Korst. 2010. "Trends and Patterns of Vaginal Birth after Cesarean Availability in the United States." *Seminars in Perinatology* 34 (4): 237–43.

Hamilton, B. E., J. A. Martin, and S. J. Ventura. 2011. "Births: Preliminary Data for 2010." *National Vital Statistics Report* 60 (2).

Ickovics, J. R., T. S. Kershaw, C. Westdahl, U. Magriples, Z. Massey, H. Reynolds, and S. S. Rising, 2007. "Group Prenatal Care and Perinatal Outcomes." *Obstetrics and gynecology* 110 (Pt 1): 330–9. doi: 10.1097/01.AOG.0000275284.24298.23

Kaimal, A. J., and M. Kuppermann. 2010. "Understanding Risk, Patient and Provider Preferences, and Obstetrical Decision Making: Approach to Delivery after

Cesarean." *Seminars in Perinatology* 34 (10): 331–6. doi: 10.1053/j.semperi.2010.05.006

Kistka, Z. A., L. Palomar, K. A. Lee, S. E. Boslaugh, M. F. Wangler, F. S. Cole, M. R. DeBaun, and L. J. Muglia, 2007. "Racial Disparity in the Frequency of Recurrence of Preterm Birth." *American Journal of Obstetrics and Gynecology* 196 (2): 131.e1–6. doi: 10.1016/j.ajog.2006.06.093

Lenaway, D., T. D. Koepsell, T. Vaughan, G. van Belle, K. Shy, and F. Cruz-Uribe. 1998. "Evaluation of a Public-Private Certified Nurse-Midwife Maternity Program for Indigent Women." *American Journal of Public Health* 88 (4): 675–9. doi: 10.2105/AJPH.88.4.675

Lu, M. C., V. Tache, G. R. Aleander, M. Kotelchuck, and N. Halfon. 2003. "Preventing Low Birth Weight: Is Prenatal Care the Answer?" *The Journal of Maternal-Fetal and Neonatal Medicine* 13: 362–80.

Martin, J. A., B. E. Hamilton, P. D. Sutton, S. J. Ventura, T. J. Matthews, and M. J. Osterman. 2010. "Births: Final Data for 2008." *National Vital Statistics Reports* 59 (1).

McClellan, M., B. J. McNeil, and J. P. Newhouse. 1994. "Does More Intensive Treatment of Acute Myocardial Infarction in the Elderly Reduce Mortality?" *Journal of the American Medical Association* 272 (11): 859–66.

McConnell, K. J., C. D. Newgard, R. J. Mullins, M. Arthur, and J. R. Hedges. 2005. "Mortality Benefit of Transfer to Level I versus Level II Trauma Centers for Head-Injured Patients." *Health Services Research* 40 (2): 435–58.

Palmer, L., A. Cook, and B. Courtot. 2010. "Comparing Models of Maternity Care Serving Women at Risk of Poor Birth Outcomes in Washington, DC." *Alternative Therapies in Health and Medicine* 16 (5): 58–70.

Petrini, J., T. Dias, M. Mccormick, M. Massolo, N. Green, and G. Escobar, 2009. "Increased Risk of Adverse Neurological Development for Late Preterm Infants." *The Journal of Pediatrics* 154 (2): 169–176.e3. doi: 10.1016/j.jpeds.2008.08.020

Raisler, J., and H. Kennedy 2005. "Midwifery Care of Poor and Vulnerable Women, 1925–2003." *Journal of Midwifery & Women's Health* 50 (2): 113–21. doi: 10.1016/j.jmwh.2004.12.010

Rooks, J. P., N. L. Weatherby, E. K. Ernst, S. Stapleton, D. Rosen, and A. Rosenfield. 1989. "Outcomes of Care in Birth Centers." *New England Journal of Medicine* 321 (26): 1804–11.

Walsh, D., and Soo. M. Downe. 2004. "Outcomes of Free-Standing, Midwife-Led Birth Centers: A Structured Review." *Birth* 31 (3): 222–9.

Wang, M. L., D. J. Dorer, M. P. Fleming, and E. A. Catlin. 2004. "Clinical Outcomes of Near-Term Infants." *Pediatrics* 114 (2): 372–6. doi: 10.1542/peds.114.2.372

World Health Organization. 1985. Appropriate Technology for Birth. *Lancet* 1985; 2 (8452): 436–7.

Yang, Y. T., M. M. Mello, S. V. Subramanian, and D. M. Studdert. 2009. "Relationship between Malpractice Litigation Pressure and Rates of Cesarean Section and Vaginal Birth after Cesarean Section." *Medical Care* 47 (2): 234–42. doi: 10.1097/MLR.0b013e31818475de

## Supporting Information

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

Appendix SA1: Author Matrix.