WORKING PAPER

Can Institutional Deliveries Reduce Newborn Mortality?

Evidence from Rwanda

A.V. Chari and Edward N. Okeke

RAND Labor & Population

WR-1072
December 2014
This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

RAND CORPORATION

CAN INSTITUTIONAL DELIVERIES REDUCE NEWBORN MORTALITY? EVIDENCE FROM RWANDA

A.V. CHARI[1] AND EDWARD N. OKEKE*[2]

ABSTRACT. Current global health policies emphasize institutional deliveries as a pathway to achieving reductions in newborn mortality in developing countries. There is however remarkably little evidence regarding a causal relationship between institutional deliveries and newborn mortality. In this paper we take advantage of a shock to institutional deliveries provided by the randomized rollout of a government performance-based financing (PBF) program in Rwanda, to provide the first estimates of this causal effect. Using a combination of difference-in-differences and regression discontinuity approaches, we find that program-induced increases in the rate of institutional delivery have not been successful in reducing the rate of newborn mortality. The findings suggest that attempts to increase institutional deliveries without addressing supply-side constraints are unlikely to result in the large reductions in mortality that policy makers expect.

Key words: health production; infant mortality; institutional births

JEL Codes: C93, D01, D03, I12.

[1]University of Sussex.
[2]RAND Corporation. Email: eokeke@rand.org. Corresponding Author.
We thank participants at the 2013 Northeast Universities Development Consortium Conference, and participants at various other conferences and seminars for helpful comments. This work was supported by an Investment Award from the Bing Center for Health Economics.

1. INTRODUCTION

Can policies promoting institutional deliveries in developing countries reduce newborn deaths? This is an important question with far-reaching implications because nearly four million infants die every year within a month of being born.[1] Newborn deaths alone accounted for 40 percent of mortality among children under five in 2010 (Rajaratnam et al., 2010). It is well known that in developing countries, a large number of births take place outside of health facilities, usually at home and unattended by formally trained doctors or midwives, and it is widely believed that this contributes to high rates of mortality (Darmstadt et al., 2009). In sub-Saharan Africa and South Asia, the two regions that account for most newborn deaths, more than half of all births take place at home (Montagu et al., 2011).

Within the last decade, there has been growing emphasis on increasing the rate of delivery in hospitals and other health facilities.[2] In keeping with this policy focus, several developing countries have implemented programs that explicitly incentivize institutional deliveries. A notable example is the Janani Suraksha Yojana (JSY) program in India, which provides cash transfers to women who give birth in a health facility. The JSY has expenditures in excess of 200 million US dollars annually (Mazumdar et al., 2011). Other similar programs include the Safe Delivery Incentives Program (SDIP) in Nepal and the Demand-side Financing Maternal Health Voucher Scheme (MHVS) in Bangladesh (Powell-Jackson et al., 2009).

[1]Nearly all of these deaths take place in developing countries.

[2]Rates of institutional deliveries are now assiduously tracked by international agencies as a progress indicator. See for example the World Health Organization Reports.

Proponents of institutional delivery point to the fact that mortality risk is concentrated around the time of birth – a quarter of all neonatal deaths occur within the first 24 hours, three-quarters within the first week (WHO, 2006). Newborn deaths have been strongly linked to intra-partum (i.e. at time of delivery) factors, and delivery complications such as prolonged or obstructed labor are consistently associated with a higher risk of mortality (Bartlett et al., 1993; Chalumeau et al., 2000; Kusiako et al., 2000).[3] The main rationale for situating births in health facilities is therefore that it gives women access to skilled providers who are better able to recognize and manage delivery complications and perform essential interventions (including newborn resuscitation), and also minimizes delays in getting help (Filippi et al., 2006).

While in the abstract, access to high quality obstetric care at the time of delivery is no doubt beneficial, numerous studies (Chaudhury et al., 2006; Banerjee and Duflo, 2006; Das et al., 2008) show that the reality of health service delivery in many developing countries departs significantly from this ideal. Weak systems coupled with poor incentives lead to problems in the quality of service delivery that introduce a wedge between access and outcomes. High rates of absenteeism are common (Chaudhury and Hammer, 2004),[4] and even when present, health workers are often not qualified, not competent, or both

[3]Intra-partum factors are in fact thought to play a larger role in perinatal or neonatal deaths relative to pre-pregnancy or antenatal factors (Lawn et al., 2005). Weiner et al. (2003) for example, in a study in Kenya, estimated that up to half of all newborn deaths were attributable to labour complications.

[4]This has been linked to poor health outcomes (Goldstein et al., 2013).

(Banerjee et al., 2004; Das et al., 2008). This is compounded by low effort.[5,6] This literature suggests that large improvements in outcomes are unlikely without concomitant improvements on the supply-side.

Identifying the effect of institutional delivery on health outcomes has however proven challenging in the absence of a credible policy experiment. The primary concern is self-selection into treatment: If women who deliver in a hospital are (predominantly) adversely selected (in terms of maternal risk) then estimates of the effect of institutional deliveries on mortality are likely to be biased towards zero. Favorable selection will on the other hand result in estimates that are too large. In addition to mother-specific heterogeneity, unobserved birth-specific heterogeneity must also be accounted for. Bharadwaj and Nelson (2013) for example present evidence that when a woman is carrying a male fetus, households are more likely to invest in other child health inputs. While the gender of the baby is observable and can be controlled for, the point is that systematic (unobserved) preferences for certain births may lead to overestimates of the treatment effect.[7] Given these challenges, it is not surprising that the existing evidence (reviewed in Section 2) is mixed, with some studies suggesting a positive effect of institutional deliveries while other studies suggest that the returns may be zero.

[5] Leonard and Masatu (2006) find that the median patient consultation among their sample of health providers in Tanzania lasted 5 minutes and consisted of just two history taking questions.

[6] These problems are not unique to the health sector by any means. Duflo et al. (2012) for example find that increased access to schooling in developing countries has not resulted in improvements in learning, attributing this in part to educational quality including high rates of teacher absence.

[7] Household income shocks have also been shown to affect household fertility decisions, and to affect the demand for child health inputs conditional on deciding to have a baby. Most household surveys that include retrospective birth histories do not collect information about household income at the time of pregnancy/delivery; if (unmeasured) household income is correlated with the demand for a hospital birth and independently affects infant mortality, this will bias estimates away from zero. Bhalotra (2010) shows that income shocks affect mortality through a selection effect - higher risk women are more likely to defer births when households experience negative shocks – and through altering the demand for complementary child health inputs.

In this paper, we provide the first quasi-experimental estimates of the effect of an institutional birth on neonatal mortality. To solve the identification problems described earlier, we exploit an exogenous shock to institutional deliveries provided by the implementation of a government performance-based financing (PBF) program in Rwanda. What makes this program unique, and strengthens our identification, is that rollout was designed to be random (although, as we explain in Section 3, the actual implementation was slightly different than planned, a feature that we account for in our identification strategy). Twelve districts were assigned to Phase 1 of the program in 2006; the remaining districts joined the program in 2008. Health facilities in treated districts received payments for providing various maternal and child services, with the highest unit payments reserved for deliveries in the clinic, and for emergency transfers to hospital for obstetric care.

Using data on 8,383 births between 2000 and 2008 from the Rwandan Demographic Health Surveys, we first show that the PBF program significantly increased rates of institutional delivery in Rwanda: Since the introduction of the PBF, the rate of institutional delivery has increased almost three-fold from a baseline rate of about 25%. Remarkably, there appears to have been no corresponding decline in newborn mortality. We then adopt a more formal analysis, combining difference-in-differences and parametric regression discontinuity approaches and show that, consistent with the descriptive findings, the causal estimates are small and not significantly different from zero.

Our results question the logic behind policies promoting institutional delivery programs as a mechanism for reducing newborn deaths. We argue that policies promoting institutional deliveries without addressing other key supply-side constraints are unlikely to result in the large gains that are posited.

The remainder of the paper is organized as follows: in Section 2 we discuss the existing literature, in Section 3 we describe the PBF program, in Section 4 we discuss the data, in Section 5 we lay out our empirical strategy and results, and in Section 6 we conclude.

2. Existing Literature

There is an extensive economic literature on infant health production. Classic papers in this literature include Rosenzweig and Schultz (1983), Corman and Grossman (1985), and Grossman and Joyce (1990). Many of these early papers focus on the timing of prenatal care onset and its effect on infant health. More recent papers also examine the effect of the number and quality of prenatal care visits (Rous et al., 2004; Evans and Lien, 2005).[8] Given the rarity of neonatal mortality in developed countries, the modal outcome of interest has tended to be child birthweight. The developing country literature on infant health production is more limited. Similar to the developed country literature, prenatal care is the most frequent input studied. Guilkey et al. (1989) find a positive relationship between the number of prenatal care visits and birthweight among Filipino women. Todd (2007) studies the effect of month of prenatal care initiation and number of prenatal care visits on birthweight using Demographic and Health Survey data from four South American countries and also finds positive effects of prenatal care.[9] Habibov and Fan (2011) finds positive effects of prenatal care quantity and quality on birthweight using data from Azerbaijan.

[8]Noonan et al. (2007) and Lien and Evans (2005) also study the effect of maternal inputs such as smoking and illicit drug use.

[9]They also find slightly larger effects relative to US studies.

Panis and Lillard (1994) and Maitra (2004) are the only two papers we are aware of that attempt to estimate the effect of institutional deliveries on infant mortality.[10] Both find a statistically significant effect of institutional deliveries on child mortality. The similarity in results is not surprising given that both papers apply the same methodology, jointly estimating the demand and health production functions within a system of simultaneous equations.[11] As Maitra (2004) acknowledges, omitted variable bias is a problem if there are inputs that affect child health that are not included in the mortality equation. If for example institutional deliveries are correlated with use of postnatal inputs and these are not captured in the estimating equation, then the effect of a hospital delivery on child mortality would be overestimated. Importantly, while mother-specific heterogeneity is accounted for, unobserved birth-specific heterogeneity remains a problem.

A few recent attempts have been made to study the effects of programs promoting institutional births. These impact evaluations have not shown promising results. Mazumdar et al. (2011) exploit spatial and temporal variation in introduction of the Janani Suraksha Yojana program in India at the district level and find no effect of the program on either neonatal mortality (defined as deaths within 28 days), or early neonatal mortality (defined as deaths within the first 24 hours), despite a statistically significant increase in the rate of institutional deliveries. Another evaluation by Debnath (2013) that exploits detailed information about program eligibility at the individual level also finds no significant effect on neonatal mortality, although he appears to find a small reduction in early neonatal

[10]There is a large epidemiological literature on skilled birth attendance in developing countries that is relevant for this paper. Skilled birth attendance and institutional deliveries are often used interchangeably since one, in most cases, implies the other (Campbell and Graham, 2006). A recent meta-analysis of this literature finds inconsistent evidence of a negative effect on neonatal mortality while concluding that the overall quality of the evidence is low (Darmstadt et al., 2005).

[11]Child mortality is modeled using a proportional hazard model in which the child is at risk of dying from the time of birth until the time of the survey. Panis and Lillard (1994) allows for fetal deaths i.e. miscarriages and abortions.

mortality (defined as a death with the first seven days) that is statistically significant at the 10 percent level.

A general weakness of these studies is that programs such as the JSY are usually not implemented in a randomized fashion, but are instead typically targeted towards areas where the public health impacts are likely to be greatest. Specific to incentive programs is that the financial incentives provided to mothers may be invested in other health inputs that influence health outcomes, confounding the estimated impacts.[12] Financial incentives may also affect fertility decisions and if this effect is heterogeneous, it may change the composition of births, confounding estimation of the treatment effect. In addition, such programs (including the JSY) are usually bundled interventions that are aimed at increasing utilization of maternal health services in general (including prenatal care). Isolating the role of institutional deliveries from that of other simultaneously changing health care inputs is therefore difficult. Finally, to the extent that they are effective at improving health outcomes, such programs may also produce long-term effects on mortality rates through their impact on women's health stock, and may also affect subsequent fertility decisions. These long-term effects may be positive or negative, biasing estimated effects in an unknown direction.

As we explain in detail in Section 5, we are able to provide more transparent identification by relying on a shock to institutional births provided by the rollout of a government performance-based financing (PBF) program in Rwanda. The near-random rollout of the program allows for the construction of a credible identification strategy; in addition, the

[12]Note that this can affect investment in prenatal inputs even if payments are not made until after delivery because households can borrow against the future income. Households may also alter consumption patterns or other behaviors, in anticipation of the future payment that may also impact health outcomes. For example, since live births are now more valuable to the household, work and household chores may be redistributed in order to reduce the responsibilities of the expectant mother (Bhalotra, 2010).

supply-side nature of the intervention rules out the operation of income effects that are usually associated with demand-side interventions. The design also allows us to rule out possible confounding effects of changes in prenatal care use. Lastly, we are able to abstract from long-term effects by utilizing a regression-discontinuity based approach to examine how outcomes changed immediately after the program was introduced.

3. The PBF Program in Rwanda

Performance-based financing (PBF) began in Rwanda in 2001 when international NGOs started pilot schemes in two provinces – Cyangugu in 2001, and Butare in 2002.[13] Based on the success of these initiatives, the Ministry of Health with funding from the United States Agency for International Development (USAID) decided to scale up PBF and expand it to all health facilities in Rwanda (Soeters et al., 2005). Due to capacity and financial constraints, it was not feasible to implement the program in all districts, and so the program was implemented in two phases. To avoid criticisms of unfairness and "hidden agendas" (and to facilitate a rigorous evaluation), districts were randomly assigned to Phase 1 (rollout starting in 2006) or 2 (rollout starting in 2008). Areas with no existing performance based contracting operations were paired based on similar characteristics and a coin flip was used to assign districts to Phase 1 or 2.[14] District boundaries were redrawn in late 2005 as part of a government decentralization effort. This resulted in some control districts being combined with districts with existing PBF pilot programs. As Basinga et al. (2011) note, because PBF programs could not be removed from health facilities in which they had already been implemented, the government enrolled all health facilities

[13]A third scheme was begun in 2005 in Kigali-Ngali, Kabgayi, and Kigali Ville by Belgian Technical Cooperation (BTC), a development cooperation agency.

[14]The process of allocation is fully described in the original evaluation of the PBF (Basinga et al., 2011). This earlier evaluation studied the impact of the program on use of maternal and child health services.

in newly formed districts that had existing pilot programs schemes into the first phase of the rollout.[15] The final sample consists of twelve Phase 1 districts and seven Phase 2 districts.

Treated (i.e. Phase 1) and control (i.e. Phase 2) health facilities differed in that the former received explicit incentives for the provision of certain health services, whereas the latter received an unconditional payout equivalent to the average amount of payouts given to treated districts. This design was intended to allow for an evaluation of the effect of incentivizing service provision, while abstracting from any income effects associated with the resulting payments. The PBF program incentivized provision of fourteen different maternal and child health services ranging from prenatal care to child preventive visits and immunization see Table A.1). The highest per unit payments were for institutional deliveries and for emergency transfers to hospital for obstetric care ($4.59). Hospital referrals for at-risk pregnancies were assigned a value of $1.83. Bonus payments were paid out to health facilities quarterly according to the following payment formula:

$$Payment_{it} = Q_{it} * \sum_{j} P_j U_{ijt}$$

where i indexes the facility, j indexes the service, and t indexes time; P is the per unit payment for each incentivized service and U is the number of units provided; Q is a weighted quality index that measures overall facility quality. Q is a composite of various structural and process indicators and facilities can receive a score between 0 and 1 (Rusa et al., 2009).

[15]Based on personal email communication from Paulin Basinga, one of the lead investigators on the original evaluation, this affected three districts.

We expect that the PBF would have increased provision of health services due to an income effect and an incentive effect. The income effect captures the effect of providing greater budgetary resources to health facilities, which may then have been spent on increasing quality or staff compensation, both of which may have had the effect of increasing utilization (the latter by reducing staff absenteeism). By design, this effect should have operated to an equal extent in treated and control districts. The incentive effect, on the other hand, would only obtain in treated districts, as health providers would re-optimize, allocating effort towards incentivized services, with the greatest effort being allocated to the most highly incentivized services (Eggleston, 2005; Dumont et al., 2008). Given that institutional deliveries were so highly incentivized, one would expect providers to attempt to increase rates of facility delivery. Consistent with this, Basinga et al. (2011) estimated an 8.1 percentage point (23%) increase in the probability of an institutional delivery due to the PBF and found evidence that providers not only encouraged women to deliver in the facility during prenatal care, but also commissioned community health workers to reach out to pregnant women in the community.

4. Data

The Rwandan Demographic and Health Survey (RDHS) is a nationally representative household survey conducted approximately every five years. Households are selected in two stages: first, villages (also referred to as clusters or enumeration areas) are selected with probability proportional to the village size; a household listing is then conducted in each village, and households are systematically sampled from the household list. Interviewers collect detailed information about all pregnancies and birth in the selected households within the five years preceding the survey date. All women between the ages of 15 and 49 who are permanent residents of the household, or visitors present in the household on the night before the survey, are eligible to be interviewed. For all births within the preceding five years, women are asked where the birth took place, and additional details about each birth are collected including the sex of the child, the date of birth, birth order, whether it was a single or multiple birth, and survival status. For babies who died, information is collected about the age of death. For deaths within the first month, women are asked to provide the specific day of death (if known). This enables us to construct an indicator for neonatal mortality. For the most recent live birth, women are asked if they used prenatal care and how many prenatal care visits they attended. Some information about the quality of prenatal care is also collected including whether the respondent was weighed, whether their blood pressure was taken, and whether they were told about pregnancy danger signs. In addition to detailed birth information, the RHDS also collects information about respondent and household characteristics.

We merge the 2005 and 2010 waves of the RDHS. This otherwise straightforward exercise was complicated by the fact that the 2005 RDHS contains province IDs but not district IDs.[16] To match enumeration area clusters in 2005 to the appropriate districts in 2010, we used geo-positioning data. Recent DHS surveys collect GPS coordinates for each village or cluster. These GPS readings are generally accurate up to 15-20 meters (ICF Macro, 2011). We matched each cluster in 2005 to a district in 2010 by calculating pairwise distances between each cluster in 2005 and all clusters in 2010 and then assigning each cluster to the district with the lowest sum of squared distances to cluster i. To put this formally, cluster i in 2005 was assigned to the district, k in 2010 that minimized $\min[A_1, A_2, ..., A_K]$

[16]In 2005, Rwanda was divided into 12 provinces. Following the decentralization, the 12 provinces were aggregated into 5 provinces.

where $A_k = \sum_j (C_i - C_j^k)$ and $C_i - C_j^k$ is the spherical distance between cluster i in 2005 and cluster j in district k in the 2010 DHS.

4.1. **Baseline statistics.** The study sample consists of 8,383 births between 2000 and 2008 (we exclude births that occurred during the second phase of rollout of the PBF when it was expanded to include all remaining districts). The key independent variable is a dummy for institutional births.[17] The primary outcome variables are neonatal mortality, a dummy indicator for whether the baby died within the first month (between Day 0 and Day 30), and early neonatal mortality, a dummy variable equal to 1 if the baby died within the first 7 days. We examine early neonatal mortality separately because it is thought to more directly reflect the quality of care received by the mother during childbirth (Ngoc et al., 2006).

Table 1 compares summary birth characteristics for home and institutional deliveries prior to PBF rollout. Rates of newborn mortality are an order of magnitude higher than in developed countries: For instance, the rate of 30-day mortality in the United States is 4 out of 1000 births. It is notable that rates of early neonatal mortality are significantly higher for institutional deliveries than for home deliveries, probably reflecting a process of selection whereby riskier births take place in health facilities. The observed characteristics, however, do not unambiguously indicate that institutional births are inherently more risky: Women who deliver in health facilities are more likely to be younger, but also more likely to be from urban, wealthier households. Table 2 presents the coefficients

[17] In the RDHS, women can select from one of the following choices for birth location: delivered at home (either the respondent's home or some other home); delivered in a health facility (either government or private), or elsewhere.

from descriptive regressions of the mortality outcomes on the institutional birth indicator (restricting the sample once again to the pre-program period) and the observed birth characteristics. The positive correlation between institutional delivery and mortality remains virtually unchanged, even after controlling for the various birth characteristics. Of the birth characteristics, bearing a male child or bearing twins are both positively correlated with mortality, consistent with the generally accepted risk factors (McMillen, 1979; Vogel et al., 2013). Mother's age is also a significant risk indicator, with the associated risk being higher at the two ends of the age spectrum, which is again consistent with the previous literature (Reichman and Pagnini, 1997). However, none of the other observed birth characteristics appears strongly correlated with mortality. Overall, the descriptive evidence suggests that unobserved risk factors may be important in determining selection into institutional delivery, and that these unobserved factors are not generally captured by observed risk characteristics.

Table 3 tests for covariate balance between Phase 1 and Phase 2 districts. Even though rollout of the PBF was random by design, the redrawing of district boundaries may have 'tampered' with random assignment. Table 3 compares average birth characteristics in Phase 1 and Phase 2 districts. The characteristics appear to be balanced across Phase 1 and Phase 2 districts, suggesting that concerns about the effectiveness of randomization may be unfounded. Nevertheless, as we will explain in the next section, we rely on a difference-in-difference strategy that is robust to imperfect randomization.

5. Empirical Analysis

5.1. **Effect of PBF on institutional births and mortality: Descriptive evidence.** We begin with a graphical examination of trends in the rates of institutional delivery and newborn mortality, and how these were affected by the introduction of the PBF. Figure 1 shows the proportion of births that took place in hospitals/clinics in each year for Phase 1 (left panel) and Phase 2 (right panel) districts respectively, along with standard error bars. The rate of institutional deliveries was relatively constant leading up to 2006 (the year of introduction of the PBF), but experienced a sharp jump in 2006 in both Phase 1 and Phase 2 districts. Consistent with the design of the PBF rollout (in which Phase 2 districts only exhibit an income effect, whereas Phase 1 districts exhibit an additional incentive effect) there is a statistically significant difference in the rate of institutional deliveries between the two sets of districts in the post-2005 period that persists until 2008 when the Phase 2 districts were phased into the PBF (we establish these results more formally in the next subsection). By the end of the study period, the rate of institutional delivery had increased by nearly 50 percentage points, reaching 74% and 72% respectively in Phase 1 and Phase 2, starting from 2005 rates of 28% and 24%.

Figures 2 and 3 show 7-day and 30-day mortality rates, respectively, for Phase 1 (left panel) and Phase 2 (right panel) districts. Despite the sharp upward trend in the rate of institutional delivery following the introduction of the PBF, there is no evidence of any corresponding decline in rates of newborn mortality.

5.2. **Overview of the estimation strategy.** Figures 1-3 provide compelling, albeit descriptive, evidence that situating births in hospitals/clinics may not (by itself) reduce the rate of newborn mortality. We recognize, however, that the observed lack of association between institutional births and mortality does not constitute definitive evidence of a zero causal effect, because trends in unobservable factors may have counteracted any positive effects of increased institutional deliveries. We therefore propose a more careful econometric analysis that takes advantage of the phased rollout of the PBF to control for trends in unobservable factors using a difference-in-differences (DID) methodology, using Phase 2 districts as a control group.

The advantage of using a DID strategy is that it is robust to the possibly imperfect randomization of the PBF. The DID strategy using Phase 2 districts as a control group allows us to address a second concern: In addition to its effect on the rate of institutional deliveries, the PBF had a small but significant effect on the use of prenatal care services. To the extent the PBF differentially increased prenatal care utilization in Phase 1 districts, this would likely lead us to overstate the effect of institutional delivery on mortality. In practice, however, because the incentive payment for prenatal services was quite small, there does not appear to have been any differential increase in prenatal care utilization in Phase 1 districts relative to Phase 2 districts (i.e. the observed increase in prenatal care utilization appears to largely reflect an income effect in both groups of districts). We establish this result formally in Section 5.5.

Next, we examine the robustness of the results to a sharper identification strategy that makes use of the discontinuous jump in institutional deliveries at the time of program introduction. In this method, we construct a "local" difference-in-differences that captures the relative change in outcomes for Phase 1 districts relative to Phase 2 districts at the point of program introduction. By examining how outcomes changed in the immediate period following the introduction of the program, we are able to abstract from long-term effects of the PBF on health outcomes: Such long-term effects may arise, for example, because giving birth in a hospital may allow mothers to recover faster, and thus improve maternal health stock (and, potentially, impact fertility decisions), which may in turn have implications for mortality risk in the future.

Lastly, we estimate the causal effect of institutional birth on mortality, relying on the exogenous variation in the rate of institutional births due to the PBF program.

5.3. **Econometric specification.** Our basic regression model is:

$$y_{idt} = \alpha_0 + \alpha_1 Phase1 + \alpha_2 Post + \alpha_3 Phase1 * Post + f(t) + \mathbf{X}'_{it}\beta + \eta_d + e_{it}$$

y_{idt} is an outcome for child i born in month t in district d (where $t = 0$ at the time of introduction of the program), $Post$ is an indicator that takes the value 1 for $t \geq 0$, $Phase1$ is an indicator that takes the value 1 for districts that were in Phase 1, $f(t)$ is a flexible function of time, $\mathbf{X}_{it}$ is a vector of controls, η_d represents a district fixed effect, and e_{it} represents an unobserved error term which is allowed to be arbitrarily correlated within districts.

The vector of controls $\mathbf{X}$ consists of the following observable maternal and child characteristics: (i) sex of the child, (ii) mother's parity (number of previous live births), (iii) an indicator for a multiple birth, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

We consider two different specifications of $f(t)$ that represent two different empirical strategies. When $f(t)$ is simply a vector of month dummies, the coefficient α_3 has the standard difference-in-differences interpretation. In the second specification $f(t)$ consist of separate quartic polynomial trends for pre- and post-periods.[18] In this specification, α_2 captures the jump in the proportion of institutional births at the discontinuity in Phase 2 districts, while α_3 captures the differential jump in institutional births in Phase 1 districts at the discontinuity. Thus, α_3 may be interpreted as a "local" difference-in-differences, based on a flexible parametric regression discontinuity design, which captures the differential change in the outcome at the time of program introduction in Phase 1 districts relative to Phase 2 districts.

5.4. **Effect of PBF on the rate of institutional delivery.** Table 4 presents the regression estimates with the dependent variable being an indicator for whether a birth took place in a health facility. The estimates are of comparable magnitude across the two specifications. The standard DID estimates in Column 1 indicate a 9.3 percentage point increase in the rate of institutional births in Phase 1 births relative to Phase 2 districts. The local DID estimate of this program effect is nearly identical at 9.4 percentage points. These estimated program effects are similar to those obtained by Basinga et al. (2011) using clinic-level data, which is reassuring.

5.5. **Validity of the DID strategy.** The validity of the standard DID requires the parallel trends assumption to hold, i.e. outcomes in the treatment and control groups would have evolved in a similar way in the absence of the PBF. This requires that the unobservable determinants of the outcomes should not have changed differentially in the treatment group

[18]The use of a high-order polynomial allows for a better fit to the data, but in practice, we have verified that the estimated coefficients retain similar magnitude and precision in the presence of lower-order polynomials (these results are available on request).

relative to the control group, following the introduction of the PBF. Similarly, the validity of the local DID strategy requires that the unobservables should not have changed differentially for the treatment group, at the time of introduction of the PBF.

These assumptions cannot be directly tested, but we can instead examine whether there is any evidence that the observable birth characteristics changed differentially for the treatment group, i.e. whether the estimated differences-in-differences (using the two specifications described above) are different from zero when the dependent variables are observed birth characteristics. Table 5 presents the corresponding regression results. There is no evidence of any differential change in any of the observable birth characteristics, either at the time of program introduction or after. These results increase our confidence in the validity of the empirical strategy.

5.6. **Effect of PBF on prenatal care.** Before examining the effect of the PBF on newborn mortality, we note that it is important for the interpretation of the estimates that the PBF affected newborn mortality only through its effect on institutional deliveries. The main concern in this respect is that the PBF program also incentivized the provision of prenatal inputs that could have had an impact on birth outcomes. Specifically, PBF facilities received incentive payments of 9 cents for first prenatal visits, and 37 cents for each woman who completed four prenatal visits (Table A.1). Additionally, PBF clinics received incentive payments of 46 cents for each pregnant woman that received a tetanus shot, and a similar amount for each pregnant woman that received malaria prophylaxis. These incentive amounts were small relative to the incentive for delivering babies in the clinic ($4.59), but to the extent that they may have differentially increased take-up of prenatal care in

Phase 1 districts, this would lead us to overstate the beneficial effects of an institutional birth.[19]

We test for differential utilization of prenatal inputs in Phase 1 districts, using the same empirical specifications as before. We consider four outcome indicators: (1) An indicator for any prenatal visit, (2) An indicator for four or more prenatal visits, (3) An indicator for tetanus vaccination, and (4) An indicator for malaria prophylaxis. Table 6 presents the regression results.[20] Consistent with our expectations, the relatively low-powered incentives for providing prenatal services do not appear to have resulted in any differential utilization of prenatal care in the Phase 1 districts. This result is robust across prenatal care categories and across specifications, and provides reassurance that any differential effect of the PBF on newborn mortality in Phase 1 districts is attributable to the PBF's effect on the rate of institutional delivery alone.

5.7. **Effect of institutional deliveries on newborn mortality.** As before, we consider two alternative measures of mortality: neonatal mortality (newborn deaths within 30 days of being born), and early neonatal mortality, which includes only deaths within the first seven days. The more restrictive definition limits any potential bias arising from program-induced increases in post-natal inputs. Table 7 presents the results from the reduced form regressions. The difference-in-difference estimate is precisely estimated, and is not significantly different from zero, confirming that the increase in the rate of institutional births resulting from the PBF program had no impact on newborn mortality.

[19]There is now a critical mass of studies demonstrating beneficial effects of prenatal care on infant health (see for example Evans and Lien, 2005; Jewell, 2007).

[20]The DHS only collects information about prenatal visits for the most recent birth. This explains why the sample sizes in Columns 1-4 are significantly smaller.

Next, we estimate two-stage least squares regressions, in which $Phase1 * Post$ is used to instrument for institutional birth. Table 8 presents the results from the OLS regression of mortality on facility delivery, along with the 2SLS regression results. The F-statistic measuring the strength of the instrument is 9.165 in the case of the standard DID specification, and 12.587 in the case of the local DID specification, indicating that the instrument has reasonably good predictive power. In keeping with the reduced form results, the point estimates of the effect of an institutional birth suggests that there is little effect on neonatal mortality, although it should be noted that the standard errors are much larger than in the case of the OLS results, as one would expect. Consistent with our earlier results (Tables 1 and 2), we find again that the OLS results in both cases are positive and in the case of early neonatal mortality, the coefficient approaches statistical significance (p=0.14).

5.8. **Describing the compliers.** We observed earlier that, since the introduction of the PBF, rates of institutional delivery have increased threefold, to the point that nearly 3 out of 4 births occur in a health facility, but that this has not been accompanied by any discernible decline in rates of newborn mortality. For the purpose of the formal identification strategy, however, we have relied on a smaller variation (about 9 percentage points) in the rate of facility delivery due to the PBF, namely the difference-in-differences in Phase 1 relative to Phase 2. Under certain conditions (outlined in Appendix A) the effect of the PBF on mortality (as well as the instrumental variables estimate of the effect of institutional delivery on mortality) can be attributed to the causal effect of institutional delivery on the mortality of ”compliers”, i.e. those births that would have taken place at home, but for the exogenous shock provided by the PBF program. In the context of the formal analysis, therefore, it is useful to examine the characteristics of the compliers in order to determine whether this group is significantly different from the general population in terms of its risk characteristics (and hence its potential outcomes).

Although compliers are not individually identifiable, it is possible to compare their average characteristics with that of the population (e.g. Angrist and Pischke 2009). In the Appendix, we derive a formula which is similar in form to that derived by Almond and Doyle (2011), that allows us to measure the difference in the average value of any characteristic, say X, between the compliers and the general population. Table 8 presents the results. The compliers appears to be composed of younger women who are more likely to be bearing their first child. Compliers are also more likely to be urban and to come from wealthier households. These distinguishing characteristics are also reminiscent of the characteristics that are predictive of take-up of institutional delivery in the general population (as we saw in Table 1).

6. Discussion

Current global health policies emphasize institutional deliveries as a pathway to achieving MDG targets of reductions in mortality. Given this emphasis, a growing number of countries are implementing programs to incentivize women to give birth in a health facility. Conditional cash transfers (CCT) that reward women for giving birth in a hospital are becoming increasingly popular. Whether these programs pass a cost-benefit test however depends crucially on the causal effect of hospital births on health outcomes. In this paper, we have taken advantage of the Performance Based Financing (PBF) program in Rwanda to shed light on this question. Our findings indicate that despite the substantial effect of the PBF on the rate of institutional deliveries, there is little evidence of a causal effect on newborn mortality.

Our results diverge from the earlier work by Maitra (2004). As we have argued however, earlier estimates are likely to be biased away from zero. A notable strength of our approach is that we rely on an exogenous shock to institutional deliveries that is plausibly uncorrelated with unobserved maternal as well as birth-specific characteristics. Our identification strategy is therefore more transparent.[21] These results are however consistent with recent null findings from evaluations of institutional delivery programs such as the JSY (Mazumdar et al., 2011). We note as well that these results are also consistent with several epidemiological studies that find no effect of place of delivery (or skilled birth attendance) on neonatal mortality (Hatt et al., 2009; Jehan et al., 2009; Titaley et al., 2012). [22]

Given what we know about the causes of neonatal mortality, these findings are not implausible. Prematurity and complications related to birth asphyxia (failure to initiate and sustain breathing at birth)[23] are two leading causes of neonatal mortality, accounting for just over half of all neonatal deaths (Lawn et al., 2005). Resuscitation is obviously crucial for the non-breathing infant, but a recent assessment of health facilities in six African countries found that 80-90% of facilities lacked resuscitation equipment (Wall et al., 2009). Harvey et al. (2007) in a similar assessment of skilled birth attendants in five countries documented significant knowledge and skill gaps; of the approximately 1,500

[21] It is possible also that differences in results may, at least in part, be due to differences in the study sample. We use data from Africa (as against Asia) and from a much more recent time period, and the patterns of health care utilization and the quality of obstetric care are likely to be different.

[22] We do not have the right data to be able to explore impact heterogeneity by facility quality in this analysis, but we plan to do this in future work.

[23] The preferred term is now intra-partum related neonatal deaths.

providers assessed, only half were competent to perform resuscitation. It is well known that delivery complications such as obstructed labor play a major role in neonatal deaths, but the capability to manage these complications (instrumental deliveries or emergency c-sections) do not exist in many primary health facilities. Ronsmans et al. (2009) find that midwives are not skilled at managing complications even when women seek help early.[24] As Darmstadt et al. (2005) notes, prevention of birth-related deaths requires not only recognition of obstetric complications, but functioning referral and transport systems, and timely access to comprehensive care, many of which are lacking in developing countries. We conclude that attempting to increase institutional deliveries without addressing supply-side constraints is unlikely to result in the large reductions in mortality that policy makers expect.

References

Banerjee, A., Deaton, A., and Duflo, E. (2004). Health Care Delivery in Rural Rajasthan. *Economic and Political Weekly*, 39(9):944–49.

Banerjee, A. and Duflo, E. (2006). Addressing absence. *Journal of Economic Perspectives*, 20(1):117.

Bartlett, A., Paz de Bocaletti, M., and Bocaletti, M. (1993). Reducing perinatal mortality in developing countries: high risk or improved labour management? *Health Policy and Planning*, 8:360–368.

Basinga, P., Gertler, P. J., Binagwaho, A., Soucat, A. L., Sturdy, J., and Vermeersch, C. M. (2011). Effect on maternal and child health services in rwanda of payment

[24]Walraven and Weeks (1999) have in fact argued that the "skilled" provider in the local clinic may be no more skilled than the traditional community midwife.

to primary health-care providers for performance: an impact evaluation. *The Lancet*, 377(9775):1421–1428.

Bhalotra, S. (2010). Fatal fluctuations? cyclicality in infant mortality in india. *Journal of Development Economics*, 93(1):7–19.

Bharadwaj, P. and Nelson, L. K. (2013). Discrimination begins in the womb: Evidence of sex-selective prenatal investments. *Journal of Human Resources*, 48:71–113. Mimeo, University of California, San Diego.

Campbell, O. M. and Graham, W. J. (2006). Strategies for reducing maternal mortality: getting on with what works. *The Lancet*, 368(9543):1284–1299.

Chalumeau, M., Salanave, B., Bouvier-Colle, M.-H., de Bernis, L., Prua, A., and Bréart, G. (2000). Risk factors for perinatal mortality in west africa: a population-based study of 20 326 pregnancies. *Acta Pædiatrica*, 89(9):1115–1121.

Chaudhury, N. and Hammer, J. S. (2004). Ghost Doctors: Absenteeism in Rural Bangladeshi Health Facilities. *World Bank Econ Rev*, 18(3):423–441.

Chaudhury, N., Hammer, J. S., Kremer, M., Muralidharan, K., and Rogers, H. F. (2006). Missing in action: Teacher and health worker absence in developing countries. *Journal of Economic Perspectives*, 20(1):91–116.

Corman, H. and Grossman, M. (1985). Determinants of Neonatal Mortality Rates in the U.S.: A Reduced Form Model. *Journal of Health Economics*, 4:213–236.

Darmstadt, G. L., Bhutta, Z. A., Cousens, S., Adam, T., Walker, N., and de Bernis, L. (2005). Evidence-based, cost-effective interventions: how many newborn babies can we save? *The Lancet*, 365(9463):977 – 988.

Darmstadt, G. L., Lee, A. C., Cousens, S., Sibley, L., Bhutta, Z. A., Donnay, F., Osrin, D., Bang, A., Kumar, V., Wall, S. N., Baqui, A., and Lawn, J. E. (2009). 60 million non-facility births: Who can deliver in community settings to reduce intrapartum-related deaths? *International Journal of Gynecology & Obstetrics*, 107, Supplement(0):S89–S112.

Das, J., Hammer, J., and Leonard, K. (2008). The quality of medical advice in low-income countries. *Journal of Economic Perspectives*, 22(2):93–114.

Debnath, S. (2013). Improving Maternal Health with Incentives to Mothers vs. Health Workers: Evidence from India. Technical report, Dissertation, University of Virginia.

Duflo, E., Hanna, R., and Ryan, S. P. (2012). Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1241–78.

Dumont, E., Fortin, B., Jacquemet, N., and Shearer, B. (2008). Physicians' multitasking and incentives: Empirical evidence from a natural experiment. *Journal of Health Economics*, 27(6):1436–1450.

Eggleston, K. (2005). Multitasking and mixed systems for provider payment. *Journal of Health Economics*, 24(1):211–223.

Evans, W. N. and Lien, D. S. (2005). The benefits of prenatal care: evidence from the pat bus strike. *Journal of Econometrics*, 125(1-2):207–239.

Filippi, V., Ronsmans, C., Campbell, O. M., Graham, W. J., Mills, A., Borghi, J., Koblinsky, M., and Osrin, D. (2006). Maternal health in poor countries: the broader context and a call for action. *The Lancet*, 368(9546):1535–1541.

Goldstein, M., Graff Zivin, J., Habyarimana, J., Pop-Eleches, C., and Thirumurthy, H. (2013). The effect of health worker absence and health clinic protocol on health outcomes: the case of mother-to-child transmission of HIV in Kenya. *American Economic*

Journal: Applied Economics, 5:58–85.

Grossman, M. and Joyce, T. (1990). Unobservables, Pregnancy Resolutions, and Birth Weight Production Functions in New York City. *Journal of Political Economy*, 98(5):983–1007.

Guilkey, D. K., Popkin, B. M., Akin, J. S., and Wong, E. L. (1989). Prenatal care and pregnancy outcome in cebu, philippines. *Journal of Development Economics*, 30(2):241 – 272.

Habibov, N. N. and Fan, L. (2011). Does prenatal healthcare improve child birthweight outcomes in azerbaijan? results of the national demographic and health survey. *Economics & Human Biology*, 9(1):56–65.

Harvey, S. A., Blandón, Y. C., McCaw-Binns, A., Sandino, I., Urbina, L., Rodríguez, C., Gómez, I., Ayabaca, P., Djibrina, S., and the Nicaraguan maternal and neonatal health quality improvement group (2007). Are skilled birth attendants really skilled? a measurement method, some disturbing results and a potential way forward. *Bulletin of the World Health Organization*, 85(10):783–790.

Hatt, L., Stanton, C., Ronsmans, C., Makowiecka, K., and Adisasmita, A. (2009). Did professional attendance at home births improve early neonatal survival in Indonesia? *Health Policy and Planning*, 24(4):270–278.

ICF Macro (2011). Incorporating Geographic Information into Demographic and Health Surveys: A Field Guide to GPS Data Collection. Technical report, Calverton, Maryland, USA: ICF Macro.

Jehan, I., Harris, H., Salat, S., Zeb, A., Mobeen, N., Pasha, O., McClure, E. M., Moore, J., Wright, L. L., and Goldenberg, R. L. (2009). Neonatal mortality, risk factors and causes:

a prospective population-based cohort study in urban Pakistan. *Bulletin of the World Health Organization*, 87(2):130–138.

Jewell, R. T. (2007). Prenatal care and birthweight production: evidence from south america. *Applied Economics*, 39(4):415–426.

Kusiako, T., Ronsmans, C., and Van der Paal, L. (2000). Perinatal mortality attributable to complications of childbirth in Matlab, Bangladesh. *Bulletin of the World Health Organization*, 78(5):621–627.

Lawn, J. E., Cousens, S., and Zupan, J. (2005). 4 million neonatal deaths: When? where? why? *The Lancet*, 365(9462):891 – 900.

Leonard, K. L. and Masatu, M. C. (2006). Outpatient process quality evaluation and the hawthorne effect. *Social Science and Medicine*, 63(9):2330–40.

Lien, D. S. and Evans, W. N. (2005). Estimating the impact of large cigarette tax hikes: The case of maternal smoking and infant birth weight. *Journal of Human Resources*, 40(2):373–92.

Maitra, P. (2004). Parental bargaining, health inputs and child mortality in india. *Journal of Health Economics*, 23(2):259–291.

Mazumdar, S., Mills, A., and Powell-Jackson, T. (2011). Financial Incentives in Health: New Evidence from India's Janani Suraksha Yojana. SSRN Working Paper. Available at SSRN: http://ssrn.com/abstract=1935442.

McMillen, M. (1979). Differential mortality by sex in fetal and neonatal deaths. *Science*, 204(4388):89–91.

Montagu, D., Yamey, G., Visconti, A., Harding, A., and Yoong, J. (2011). Where do poor women in developing countries give birth? a multi-country analysis of demographic and health survey data. *PLoS ONE*, 6(2):e17155.

Ngoc, N., Merialdi, M., Abdel-Aleem, H., Carroli, G., Purwar, M., Zavaleta, N., Campódonico, L., Ali, M. M., Hofmeyr, J. G., Mathai, M., Lincetto, O., and Villar, J. (2006). Causes of stillbirths and early neonatal deaths: data from 7993 pregnancies in six developing countries. *Bulletin of the World Health Organization*, 84(9):699–705.

Noonan, K., Reichman, N. E., Corman, H., and Dave, D. (2007). Prenatal drug use and the production of infant health. *Health Economics*, 16(4):361–384.

Panis, C. W. A. and Lillard, L. A. (1994). Health inputs and child mortality: Malaysia. *Journal of Health Economics*, 13(4):455–489.

Powell-Jackson, T., Morrison, J., Tiwari, S., Neupane, B., and Costello, A. (2009). The experiences of districts in implementing a national incentive programme to promote safe delivery in nepal. *BMC Health Services Research*, 9(1).

Rajaratnam, J. K., Marcus, J. R., Flaxman, A. D., Wang, H., Levin-Rector, A., Dwyer, L., Costa, M., Lopez, A. D., and Murray, C. J. (2010). Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970-2010: a systematic analysis of progress towards millennium development goal 4. *The Lancet*, 375(9730):1988–2008.

Reichman, N. E. and Pagnini, D. L. (1997). Maternal age and birth outcomes: data from New Jersey. *Family planning perspectives*, 29(6):268–295.

Ronsmans, C., Scott, S., Qomariyah, S., Achadi, E., Braunholtz, D., Marshall, T., Pambudi, E., Witten, K., and Graham, W. (2009). Professional assistance during birth and maternal mortality in two indonesian districts. *Bulletin of the World Health Organization*, 87:416–423.

Rosenzweig, M. R. and Schultz, T. P. (1983). Estimating a household production function: Heterogeneity, the demand for health inputs, and their effects on birth weight. *The Journal of Political Economy*, 91(5):723–746.

Rous, J. J., Jewell, R. T., and Brown, R. W. (2004). The effect of prenatal care on birthweight: a full-information maximum likelihood approach. *Health Economics*, 13(3):251–264.

Rusa, L., Schneidman, M., Fritsche, G., and Musango, L. (2009). *Performance Incentives for Global Health: Potential and Pitfalls*, chapter Rwanda: Performance-Based Financing in the Public Sector, pages 189–214. Number 10. Brookings Institution Press, Washington DC.

Soeters, R., Musango, L., and Meesen, B. (2005). Comparison of two output based schemes in Butare and Cyangugu provinces with two control provinces in Rwanda. Technical report, Global Partnership on Output-Based Aid (GPOBA), the World Bank.

Titaley, C. R., Dibley, M. J., and Roberts, C. L. (2012). Type of delivery attendant, place of delivery and risk of early neonatal mortality: analyses of the 1994–2007 Indonesia Demographic and Health Surveys. *Health Policy and Planning*, 27(5):405–16.

Todd, P. E. (2007). Evaluating social programs with endogenous program placement and selection of the treated. volume 4 of *Handbook of Development Economics*, chapter 60, pages 3847 – 3894. Elsevier.

Vogel, J. P., Torloni, M. R., Seuc, A., Betrán, A. P., Widmer, M., Souza, J. P., and Merialdi, M. (2013). Maternal and perinatal outcomes of twin pregnancy in 23 low- and middle-income countries. *PLoS ONE*, 8(8):e70549.

Wall, S. N., Lee, A. C., Niermeyer, S., English, M., Keenan, W. J., Carlo, W., Bhutta, Z. A., Bang, A., Narayanan, I., Ariawan, I., and Lawn, J. E. (2009). Neonatal resuscitation in low-resource settings: What, who, and how to overcome challenges to scale up? *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*, 107:S47–S64.

Walraven, G. and Weeks, A. (1999). The role of (traditional) birth attendants with midwifery skills in the reduction of maternal mortality. *Tropical Medicine and International Health*, 4(8):527–529.

Weiner, R., Ronsmans, C., Dorman, E., Jilo, H., Muhoro, A., and Shulman, C. (2003). Labour complications remain the most important risk factors for perinatal mortality in rural Kenya. *Bulletin of the World Health Organization*, 81:561 – 566.

WHO (2006). *Working Together for Health - The World Health Report 2006*. World Health Organization.

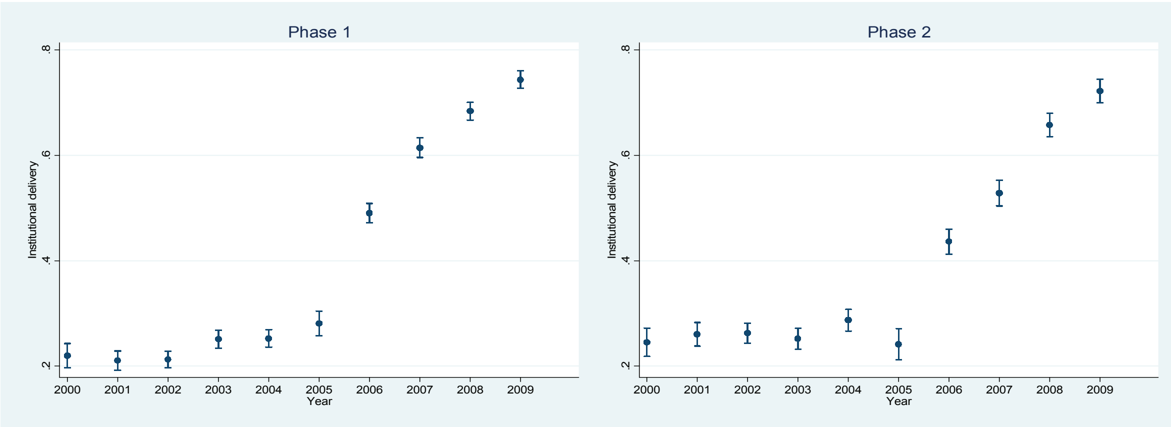

Figure 1: The graphs above show annual averages of the rate of institutional births, along with standard error bars. The left panel only includes births in Phase 1 districts; the right panel only includes births in Phase 2 districts.

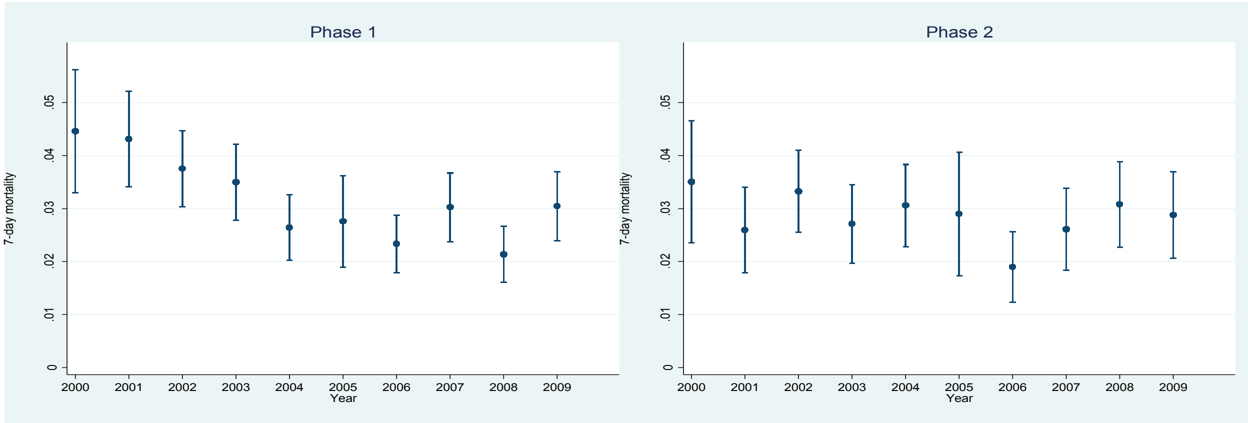

Figure 2: The graphs above show annual averages of the rate of 7-day mortality, along with standard error bars. The left panel only includes births in Phase 1 districts; the right panel only includes births in Phase 2 districts.

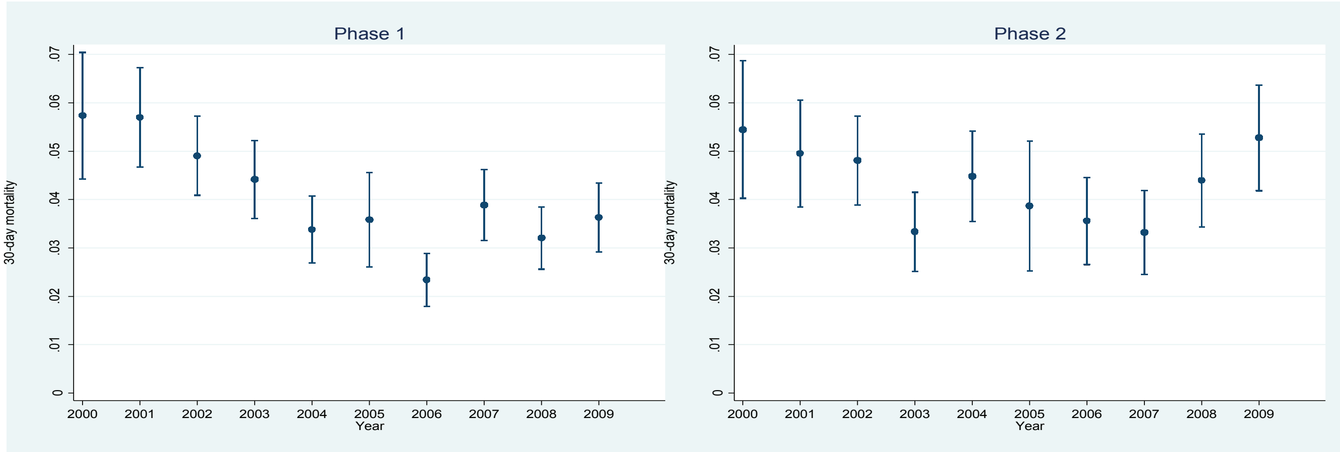

Figure 3: The graphs above show annual averages of the rate of 30-day mortality, along with standard error bars. The left panel only includes births in Phase 1 districts; the right panel only includes births in Phase 2 districts.

Table 1: Baseline characteristics by delivery location

	Home births	Facility births	*p-value*
30-day mortality	0.042	0.049	0.254
	(0.003)	(0.005)	
7-day mortality	0.029	0.040	0.027
	(0.003)	(0.005)	
Birth characteristics			
Male	0.495	0.527	0.034
	(0.007)	(0.013)	
Multiple births indicator	0.012	0.017	0.155
	(0.002)	(0.003)	
Mother's age at birth	29.054	27.313	0.000
	(0.103)	(0.167)	
Birth order	3.114	2.241	0.000
	(0.034)	(0.056)	
Prior facility use	0.059	0.258	0.000
	(0.004)	(0.013)	
Months since last birth	35.299	34.830	0.471
	(0.300)	(0.599)	
Household characteristics			
Urban household	0.094	0.206	0.000
	(0.004)	(0.010)	
Wealth index	2.630	3.363	0.000
	(0.019)	(0.036)	
Observations	4455	1561	

Notes: Standard errors in parentheses. p-values reported in the last column correspond to a two-tailed test of difference in means between facility births and home births.

Table 2: The relation between institutional (facility) births and mortality (baseline)

	(1) Neonatal mortality	(2) Early mortality
Facility birth	0.008	0.012**
	(0.007)	(0.006)
Male	0.011**	0.011**
	(0.005)	(0.005)
Multiple birth	0.175***	0.135***
	(0.046)	(0.042)
Mother's age at birth	-0.012***	-0.011***
	(0.004)	(0.003)
Mother's age at birth squared	0.000***	0.000***
	(0.000)	(0.000)
Birth order (parity)	0.000	0.001
	(0.002)	(0.002)
Prior facility deliveries	-0.001	-0.000
	(0.007)	(0.006)
Urban household	-0.007	-0.007
	(0.008)	(0.007)
Wealth index	-0.003	-0.002
	(0.002)	(0.002)
Constant	0.215***	0.190***
	(0.056)	(0.050)
Observations	6,016	6,016
R-squared	0.013	0.013

*** p<0.01, ** p<0.05, * p<0.1; Robust standard errors in parentheses. Neonatal mortality is defined as mortality within 30 days of birth; early mortality is defined as mortality within 7 days of birth. The regression samples are restricted to births that occurred before the PBF started.

Table 3: Baseline characteristics by treatment phase

	Phase 1	Phase II	*p-value*
Facility delivery	0.25	0.27	0.085
	(0.01	(0.01)	
Mortality			
30-day mortality	0.04	0.04	0.846
	(0.00)	(0.00)	
7-day mortality	0.03	0.03	0.417
	(0.00)	(0.00)	
Birth characteristics			
Male	0.50	0.50	0.971
	(0.01)	(0.01)	
Multiple births indicator	0.01	0.01	0.346
	(0.00)	(0.00)	
Mother's age at birth	28.6	28.6	0.980
	(0.12)	(0.13)	
Birth order	2.92	2.85	0.219
	(0.04)	(0.05)	
Prior facility delivery	0.11	0.12	0.190
	(0.01)	(0.01)	
Months since last birth	35.28	35.08	0.719
	(0.36)	(0.41)	
Household characteristics			
Urban	0.11	0.13	0.019
	(0.01)	(0.01)	
Wealth index	2.80	2.85	0.165
	(0.02)	(0.03)	
Observations	3495	2521	

Notes: Standard errors in parentheses. p-values reported in the last column correspond to a two-tailed test of difference in means between Phase 1 and Phase 2 births.

Table 4: The effect of PBF on the rate of institutional deliveries

Dependent variable: Institutional delivery indicator	(1)	(2)
Post		0.1192**
		(0.0424)
Phase1*Post	0.0933***	0.0936***
	(0.0308)	(0.0297)
Month fixed effects	Yes	No
Quartic trends	No	Yes
District fixed effects	Yes	Yes
Controls	Yes	Yes
Constant	-0.1858	-1.1757**
	(0.2371)	(0.4389)
Observations	8,383	8,383
R-squared	0.246	0.237

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses, clustered at district level. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 5: Testing for differential changes in birth characteristics

	Male		Multiple births		First-time mother		Birth order	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post		0.0033		0.0302		-0.0098		-0.4103
		(0.0903)		(0.0219)		(0.0111)		(0.3947)
Phase1*Post	-0.0103	-0.0109	-0.0085*	-0.0084*	-0.0264	0.0011	0.1436	0.1558
	(0.0186)	(0.0182)	(0.0045)	(0.0047)	(0.0223)	(0.0026)	(0.1174)	(0.1192)
Month fixed effects	Yes	No	Yes	No	Yes	No	Yes	No
Quartic trends	No	Yes	No	Yes	No	Yes	No	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	0.4404***	0.4777***	0.0141	0.0500*	0.2336***	-3.9459***	3.0873***	14.0367***
	(0.0508)	(0.1010)	(0.0082)	(0.0239)	(0.0524)	(0.0135)	(0.2993)	(0.3981)
Observations	8,383	8,383	8,383	8,383	8,383	8,383	6,836	6,836
R-squared	0.013	0.004	0.016	0.007	0.015	0.984	0.026	0.027

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses, clustered at district level. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 5 (contd): Testing for differential changes in birth characteristics

	Months since last birth		Mother's age at birth		Prior facility use		Urban household		Household wealth	
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Post		1.4662		0.4429		-0.0653		-0.0253		-0.1404
		(1.1466)		(1.1293)		(0.0636)		(0.0755)		(0.3539)
Phase1*Post	-0.8646	0.2110	0.2866	0.1214	-0.0018	-0.0072	0.0165	0.0154	0.0618	0.0525
	(1.0909)	(0.6634)	(0.2872)	(0.2956)	(0.0394)	(0.0386)	(0.0388)	(0.0393)	(0.2257)	(0.2252)
Month fixed effects	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Quartic trends	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Constant	37.6510***	-4.3757***	28.4419***	74.0375***	0.1172*	0.1982***	0.0186	0.1443*	2.5159***	2.2302***
	(2.3542)	(1.1473)	(0.7763)	(1.0617)	(0.0638)	(0.0552)	(0.0318)	(0.0796)	(0.2324)	(0.2803)
Observations	6,815	6,815	8,383	8,383	8,383	8,383	8,383	8,383	8,383	8,383
R-squared	0.025	0.562	0.020	0.286	0.043	0.064	0.136	0.130	0.059	0.052

*** $p<0.01$, ** $p<0.05$, * $p<0.1$; Standard errors in parentheses, clustered at district level. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 6: The effect of PBF on prenatal care utilization and quality

	Any prenatal care		At least 4 prenatal visits		Malaria		Tetanus	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post		0.0184		0.0524		0.1301*		0.1920***
		(0.0421)		(0.0938)		(0.0678)		(0.0473)
Phase1*Post	-0.0051	-0.0033	0.0339	0.0341	0.0362	0.0323	-0.0115	-0.0108
	(0.0141)	(0.0138)	(0.0456)	(0.0465)	(0.0303)	(0.0295)	(0.0225)	(0.0228)
Month fixed effects	Yes	No	Yes	No	Yes	No	Yes	No
Quartic trends	No	Yes	No	Yes	No	Yes	No	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Baseline rate	*0.95*		*0.12*					
Constant	0.9843***	0.8700***	0.2891***	-0.0740	0.6316***	0.5460***	0.8414***	-0.1225
	(0.0160)	(0.0624)	(0.0686)	(0.1087)	(0.0509)	(0.0561)	(0.0341)	(0.0711)
Observations	4,778	4,778	4,777	4,777	8,383	8,383	8,383	8,383
R-squared	0.059	0.039	0.154	0.136	0.329	0.330	0.081	0.101

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses, clustered at district level. In columns (1) and (2) "Any prenatal care" is a dummy variable that takes the value 1 for a birth where the mother received at least one pre-natal care visit. In columns (5) and (6), "Malaria" is a dummy variable that takes the value 1 for a birth where the mother received malaria prophylaxis during pregnancy. In columns (7) and (8), "Tetanus" is a dummy variable that takes the value 1 for a birth where the mother received a tetanus shot during pregnancy. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 7: Effect of PBF on infant mortality (reduced form)

	7-day mortality		30-day mortality	
	(1)	(2)	(3)	(4)
Post		0.0109		0.0147
		(0.0224)		(0.0202)
Phase1*Post	-0.0002	0.0004	-0.0031	-0.0026
	(0.0061)	(0.0065)	(0.0074)	(0.0078)
Month fixed effects	Yes	No	Yes	No
Quartic trends	No	Yes	No	Yes
District fixed effects	Yes	Yes	Yes	Yes
Baseline rate	*0.03*		*0.04*	
Constant	0.1455*	-0.7096**	0.1580	-0.8642**
	(0.0819)	(0.2731)	(0.0959)	(0.3051)
Observations	8,383	8,383	8,383	8,383
R-squared	0.039	0.028	0.041	0.029

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses, clustered at district level. Neonatal mortality is defined as mortality within 30 days of birth; early mortality is defined as mortality within 7 days of birth. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 8: The effect of facility delivery on mortality (OLS and IV)

	7-day mortality			30-day mortality		
	OLS	2SLS		OLS	2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Facility delivery	0.0054	-0.0332	-0.0157	0.0100	-0.0022	0.0113
	(0.0067)	(0.0754)	(0.0687)	(0.0059)	(0.0629)	(0.0599)
Month fixed effects	Yes	Yes	No	Yes	Yes	No
Quartic trends	No	No	Yes	No	No	Yes
District fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic (first-stage)		9.165	12.597		9.165	12.597
Observations	8,383	8,383	8,383	8,383	8,383	8,383

*** p<0.01, ** p<0.05, * p<0.1; Standard errors in parentheses, clustered at district level. Neonatal mortality is defined as mortality within 30 days of birth; early mortality is defined as mortality within 7 days of birth. All regressions include the following controls: (i) sex of the child, (ii) birth order (parity), (iii) an indicator for multiple births, (iv) mother's age at birth, (v) number of prior institutional births (mother-specific) that occurred before the program began, (vi) household religion (coded as dummies), (vii) household wealth as constructed in the DHS (representing an asset-based index that takes values 1-5), (viii) an indicator for rural households, and (ix) time interval since last birth (in months).

Table 9: Complier characteristics

	ΔE(X) (Compliers-Population)
Male	0.019
Multiple births	0.034
First-time mother	0.515
Birth order (parity)	-2.357
Months since last birth	-9.544
Mother's age at birth	-4.878
Prior facility births	0.893
Urban household	0.374
Wealth index	1.792

Notes: The figures show the deviation of average complier characteristics from population level averages, based on Equation A.4.

Appendix: Obtaining complier characteristics

In this appendix, we derive the relation between the average characteristics of the compliers and the characteristics of the general population. The reduced form estimates of program impact as well as the instrumental variables estimate of the effect of institutional birth on mortality may be attributed to the causal effect of institutional birth on the mortality of compliers, as long as the instrument is as good as randomly assigned and as long as the instrument has a monotonic effect, i.e. its only effect is to induce an additional set of women to deliver in health facilities.

Because our instrumental variable is $Phase1 * Post$, the exogeneity (i.e. random assignment) requirement can be assumed to hold conditional on the inclusion of month and district dummies. Thus, consider a particular birth characteristic, X. We write:

$$X_{idt} = \eta_d + \eta_t + e_{idt} \tag{A1}$$

where η_d and η_t represent district and time fixed effects, respectively, and the residual term e_{idt} is assumed to be distributed independently of the instrument. We will obtain e_{idt} as a residual from a regression of X on district and month dummies. In the following derivation we will refer to the residual e_{idt} as $\widetilde{X}$, and to the instrument as Z.

Let the binary treatment be denoted by D (facility delivery in the present context). Let Z represent the binary instrumental variable. We also define D_1 to represent treatment status when the instrument is "on" and D_0 to represent treatment status when the instrument is "off". Thus, an always-taker will have $D_0 = 1$ and $D_1 = 1$; a never-taker will have $D_0 = 0$ and $D_0 = 0$; and a complier will have $D_0 = 0$ and $D_1 = 1$.

We have assumed that the distribution of $\widetilde{X}$ is independent of Z. We can now write the following:

$$\begin{aligned} E(\widetilde{X}|D_1 &= 1) = E(\widetilde{X}|D_1 = 1, D_0 = 1).P(D_0 = 1|D_1 = 1) + E(\widetilde{X}|D_1 = 1, D_0 = 0).P(D_0 = 0|D_1 = 1) \\ &= E(\widetilde{X}|always - takers).\frac{\pi_A}{\pi_A + \pi_C} + E(\widetilde{X}|compliers).\frac{\pi_C}{\pi_A + \pi_C} \end{aligned} \tag{A.2}$$

Similarly:

$$\begin{aligned} E(\widetilde{X}|D_0 &= 1) = E(\widetilde{X}|D_0 = 1, D_1 = 1).P(D_1 = 1|D_0 = 1) + E(\widetilde{X}|D_0 = 1, D_1 = 0).P(D_1 = 0|D_0 = 1) \\ &= E(\widetilde{X}|always - takers) \end{aligned} \tag{A.3}$$

where we have assumed that the instrument satisfies monotonicity, so that $P(D_1 = 1|D_0 = 1) = 1$ and $P(D_1 = 0|D_0 = 1) = 0$. Combining Equations A.1 and A.2, we can write:

$$E(\widetilde{X}|compliers) - E(\widetilde{X}|D_0 = 1) = \frac{1 - \pi_N}{\pi_C}[E(\widetilde{X}|D_1 = 1) - E(\widetilde{X}|D_0 = 1)] \tag{A.4}$$

Equation A.4 above is also derived in slightly rearranged form by Almond and Doyle (2011), and can be used to calculate $E(\widetilde{X}|compliers)$. Lastly, we note that because $\widetilde{X}$ is a residual from regression A1 above, the population average is zero, i.e. $E(\widetilde{X}) = 0$, so that $E(\widetilde{X}|compliers)$ can be interpreted as the difference in the average value of $\widetilde{X}$ between the compliers and the general population.

Table A1: Schedule of incentive payments for services

Indicator	**Unit Payment (USD)**
Curative care visit	0.18
First prenatal care visit	0.09
Completion of 4 prenatal visits	0.37
First time family planning visit	1.83
Contraceptive resupply visit	0.18
Delivery in the facility	4.59
Child preventive care visits (0 - 59 months)	0.18
Tetanus vaccine received during prenatal care	0.46
Malaria prophylaxis received during prenatal care	0.46
At risk pregnancies referred to hospital for delivery	1.83
Emergency transfers to hospital for obstetric care	4.59
Completed child vaccinations	0.92
Malnourished children referred for treatment	1.83
Other emergency referrals	1.83