Comparison of Alternative Propensity Score Approaches for Causal Inference from Observational Data*

Sarah R. Brauner
University of Michigan

* Direct correspondence to Sarah R. Brauner at the Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, Michigan 48106-1248. E-mail: sbrauner@umich.edu. This research was supported by a grant from the National Institute of Child Health and Human Development (R01-HD32912) and the Fogarty International Center, International Training and Research Program in Population and Health. I would like to thank Ben Hansen, William Axinn, and Dirgha Ghimire for their assistance with this project. Any errors are the responsibility of the author.

Comparison of Alternative Propensity Score Approaches for Causal Inference from Observational Data

Abstract

Demographers are growing increasingly concerned with the problems associated with making causal inferences from observational data. Statisticians and social scientists have both developed numerous approaches for circumventing the problem of the counterfactual including measurement adjustments and alternative analytic strategies. In this paper we apply linear-model measurement adjustments and matching techniques, along with propensity score analysis, to studying the relationship between health services and family formation behavior in the developing country Nepal. We find that estimates of the effect of the availability of various maternal and child health services on the use of permanent contraceptives are highly sensitive to measurement and model choices. None of our measures were found to have significant effects across all of the models we present. However, availability of child vaccinations and oral rehydration therapy were found to have positive and significant effects on permanent contraceptive use in our most restrictive model.

Demographers are growing increasingly concerned with the problems associated with making causal inferences from observational data (Moffitt 2001, 2005; Smith 2003). Demographic theory is typically about causal relationships. However, researchers are just as often limited by the data in their ability to test causality. Policy makers and program designers are concerned when this issue refers to questions of program effectiveness and the like. In practice, developing countries have to make difficult investment decisions. Ideally, they would like to conduct randomized, experimental trials to determine which programs should receive investment funds. However, this is often not a realistic option. Ethical concerns serve as a major deterrent against undertaking randomized experiments regarding health and family programs or related behaviors. Additionally, because experiments are difficult to implement successfully, the data obtained from them is often open to the same critiques as those from observational studies (Rosenbaum 2002). Consequently, it is imperative to develop tools for improving our ability to make causal inferences from observational data.

This paper investigates several measurement and model based approaches to enhancing our causal arguments in social demography. Social scientists and statisticians alike have developed many methods that attempt to compensate for this fundamental characteristic of observational data—that one cannot make causal inferences based on it. These methods range from making assumptions regarding the counterfactual to complex statistical procedures. This paper attempts to bring together several different approaches and apply them to an important substantive area. Specifically, I consider multiple empirical models of the relationship between health services and family formation behavior in a developing country. First, I explore measurement solutions to the problem of causal inference in this context. Second, I explore alternative analytic strategies to assess the stability of the effect estimates. By applying these multiple approaches to this single problem I am able to compare the different strategies and examine the extent to which they yield consistent conclusions.

This investigation focuses on a setting in rural Nepal that experienced a transition from virtually no use of birth control in 1945 to the widespread use of birth control to limit fertility by

1995. Over the same period as this change in fertility behavior, this setting experienced an equally dramatic increase in the availability of health services, accompanied by the proliferation of education, wage labor, markets, and the media (Axinn and Barber 2001; Axinn and Yabiku 2001; Barber and Axinn Forthcoming). We take advantage of a vastly diverse, rich dataset that allows incredible flexibility, detail, and depth of measurement, and investigate the relationship between health services, specifically the availability of child vaccinations, and the adoption of birth control to limit fertility. These data describe all health service providers, including hospitals, health clinics, family planning clinics, natural healers, and pharmacies that ever existed in our study area and document the many different services these facilities provided. Thus, this setting is ideal for investigating the impact of child health services on fertility behavior using multiple approaches.

Theoretical Background on Causal Inference

Because discussions of causal inference typically borrow the language of experiments, any such discussion should begin by clarifying the specific language that will be used. In a basic causal model we posit that X effects Y. In a traditional experiment, to determine whether X (here after referred to as the treatment) effects Y (the outcome), individuals are randomly assigned into one of two groups (Campbell and Stanley 1963). The “control” group does not experience the treatment whereas the “treatment” group does.

For the purposes of this paper, we would like to know if a woman is more likely to use contraceptives if she has access to maternal and child health services. Ideally, we would like to give a woman access to these services and then observe whether she uses contraceptives. We would also like to look at the same woman but *without* access to health services and again observe her contraceptive behavior. Unfortunately we can only observe individuals in one state. That is, we can only observe whether the woman had health services available at that moment or did not have health services available; we cannot observe her in both situations described above. This unobserved other state is often referred to as the counterfactual (Harding 2003; Winship and Morgan 1999).

An alternative conceptualization of the problem of the counterfactual is to consider it a missing data problem (Holland 1986). In order to make causal arguments researchers would like to have data regarding an individual's behavior in two, mutually exclusive, circumstances. In reality, it is possible to obtain behavioral data from only one of those circumstances. The missing data is the other circumstance.

One way researchers attempt to overcome this inability to observe people in two different states is by randomly assigning them to one state or the other. Under randomization, any difference in the outcome, contraceptive use in our example, is due to the treatment, having maternal and child health services available, and not due to some other factor. This occurs because in successfully implemented randomized experiments the control group and the treatment group should be identical in all observed and unobserved covariates. Consequently after treatment exposure, any observed differences in the outcome between those in the treatment group and those in the control group are assumed due to the treatment itself. That is, the estimated treatment effect can be considered an unbiased estimate of the actual treatment effect.

Without randomization we cannot draw the conclusion that the observed effects are a result of the treatment and not other factors. One problem with estimating treatment effects from observational data is that the units of observation who are exposed to the treatment, i.e. those who have health services nearby, may be systematically different from those who are not exposed to the treatment, those who do not have health services nearby (Holland 1986; Rosenbaum 2002; Rosenbaum and Rubin 1983, 1985; Winship and Morgan 1999). The factors we are particularly concerned with are those that influence both the treatment and the outcome. In our example of child vaccinations, these types of factors may be whether a given area has markets, the education of the individual or her parents, or the individual's previous experiences with health services. All of these factors may influence whether or not an individual has access to child vaccinations and whether the individual uses contraceptives. Areas with markets may be more likely to have health services and individuals living near markets may be more likely to use contraceptives.

Of additional concern is the possibility that the placement of health services, and the decisions to offer specific services at specific places, may be a result of non-random processes such as targeted program placement. Another option is that communities that have more of one kind of resource may be more likely to have more of another kind as well—communities with schools may be more likely to have health services. Part of this may be because these other institutions may actually cause the health services to be built and to provide specific services (Caldwell 1986; Whyte and Parish 1984). Caldwell argues that education and employment produce a better informed, more politically active, and more resourceful populace that is more effective at demanding and obtaining health care services (Caldwell 1986). If this is true, then access to education, wage labor, and changes in other aspects of community context may in fact lead to subsequent changes in health services. In that situation, or with targeted program placement,, the observed relationship between health services and fertility behavior may actually be spurious, reflecting instead the impact of changes in other non-health related services on fertility behavior.

It is worthwhile noting that experiments often do not in fact satisfy the condition that the treatment and the control groups are identical before treatment occurs. Problems with the implementation of the experiment may result in the control and treatment groups being different. In these situations, the estimated treatment effects are not unbiased estimators and the data here more closely resemble observational data than what we typically think of when we describe experimental data.

Because of the demands and challenges of running a well-implemented randomized experiment and the potential ethical conflicts associated with experiments in the social sciences, random assignment is often not a practical option. Consequently, researchers are often limited to using observational data. However, researchers still need to be able to draw conclusions from the data they have available. By presenting multiple approaches to one research question this paper provides important information about the techniques researchers can use to support their causal arguments.

Empirical Approaches

Statisticians and social scientists have developed numerous approaches for circumventing the problem of the counterfactual. While it is beyond the scope of this paper to review them all, we do describe some of those commonly used and a few that are gaining notoriety. The approaches considered here are ways to use observational data to aid our understanding of causal relationships. None of them claims to prove causality. They are designed to give the researcher increased confidence in his estimates.

These methods roughly fall into two groups: sensitive analyses and stability analyses. Sensitive analyses, quite simply, determine how stable your effect estimates are (Rosenbaum 2002). Common approaches involve weakening assumptions and empirically testing alternatives (Harding 2003; Rosenbaum 2002; Smith 1997). For example, a sensitivity analysis might involve estimating how large an effect would need to be attributed to an unobserved measure in order to explain the observed relationship between the treatment and the outcome. Stability analyses, on the other hand, examine how different analytical approaches affect the effect estimates (Rosenbaum 2002). This paper falls into the latter group.

The first type of stability analyses presented in this paper is linear model based adjustments. When designing observational studies social scientists attempt to deal with the challenge of causal inference by measuring those concepts or ideas that theory reveals may influence the outcome. Researchers then include in their models those measures they believe influenced the treatment assignment. The onus of observational studies is that the researcher can never know if there exists some covariate that effects the outcome but has not been measured. I present an array of multilevel models to determine how robust the findings are to different measurement specifications.

The second type of analyses in this paper draw on methods statisticians have developed for improving the estimates made of causal relationships. Specifically, I consider the non-parametric approach of matching. Matching is a technique used essentially to simulate an experiment after the data have been collected. Individuals are grouped into matched sets according to their value for the specific measures used to match on. By creating the artificial treated and control groups in this way the researcher limits the sample to only those respondents who overlap on the measures used in the match. This reduces the bias in the effect estimates because those respondents at the extremes are not included when estimating the effects. Of course, this does result in the loss of some information since not all observations are included in the estimation. However, the benefits gained from ensuring the proper overlap in the treated and control group outweigh the loss of information from excluding some respondents (Smith and Todd 2001; Winship and Morgan 1999). This approach is also preferred over linear model based adjustments because it involves fewer assumptions, and those assumptions that do exist are more transparent (Rosenbaum 2002). Estimating treatment effects with matched samples based on the propensity score does not require assuming any functional form of the relationship between the treatment and the outcome unlike traditional methods of controlling for observed covariates, such as regression analysis (Rosenbaum 2002; Rosenbaum and Rubin 1984; Hansen 2004).

For both the linear model based and non-parametric approaches I will also examine the potential benefits of using propensity scores. Propensity scores are estimates of the treatment assignment based on observed data (Moffitt 2005; Winship and Morgan 1999). In our example, propensity scores reflect the probability that an individual has access to child vaccinations. Estimating propensity scores allows the researcher to control for many more measured characteristic than typically allowed in these approaches. This is because estimated propensity scores essentially compress multiple characteristics into one. For linear model based approaches this is especially beneficial in situations where including multiple covariates is not possible due to multi-collinearity. Traditionally researchers would have to choose among the related measures, thereby potentially loosing some information. It is not necessary to limit the number of covariates used to estimate the propensity scores. Using the propensity score as a covariate to a traditional model removes the part of the observed relationship between the treatment and the outcome that is due to the assignment of the treatment (Winship and Morgan 1999, p. 677).

Propensity score matching is preferable to other techniques that control for observed covariates for at least two reasons. First, propensity score matching removes the bias due to all observed covariates by balancing the distributions of those covariates before estimating treatment effects (Joffe and Rosenbaum 1999; Rosenbaum 2002; Rosenbaum and Rubin 1983; Rubin 1997; Rubin and Thomas 1992). This methodological approach has been used in the statistical and biomedical literature for some time and has become increasingly used in sociological research (Berk and Newton 1985; Cornfield et al. 1959; Hansen 2004; Harding 2003; Morgan 2001; Rosenbaum 1986; Smith 1997). Second, similar to the benefit of using propensity scores in linear models, matching on propensity scores allows the researcher to create matches based on many more observed covariates than matching based on the covariates themselves (Cochran 1965; Rosenbaum 2002; Rosenbaum and Rubin 1983, 1984; Rubin 1997). This is because the possible combinations of characteristics increases exponentially when not using the propensity score, thereby making it unrealistic to attempt to match on more than a couple covariates (Cochran 1965).

Of course, propensity score matching can only remove the bias caused by *observed* covariates. Any bias due to unobserved factors still exists. However, previous tests of the sensitivity of treatment effects derived from propensity score matching show that in cases with a large sample of control observations to select from and a large number of observed covariates, it is unlikely that a realistic unobserved covariate exists that could account for the observed treatment effect (Harding 2003; Smith and Todd 2001). Consequently, treatment effects derived from propensity score based matched analysis can be considered unbiased estimates of the actual treatment effect (Rosenbaum and Rubin 1983; Smith and Todd 2001). Sensitivity analysis, which is not covered in this paper, is necessary to assess the potential role of unobserved factors. However, the data set used here and described below is exceptionally rich across a broad range of relevant areas thereby providing us with a unique opportunity to in forming our propensity scores.

Before we turn to a more detailed discussion of the methods and the results it is necessary to describe the specific setting, research question, and data we will use.

Setting and research question

To apply these multiple empirical strategies we use data from the Chitwan Valley Family Study (CVFS) conducted in rural Nepal. This study combines survey and ethnographic methods to obtain detailed measures of community context and individual life histories. The Chitwan Valley is an ideal setting for this research because the spread of health service providers, changes in the services offered, and the transition toward fertility limitation occurred within the lifetimes of the valley's current residents. Also, change in the community context in the Chitwan valley is contained both temporally and geographically making it a model setting for evaluating the various measures of context discussed above.[1] The valley was virtually uninhabited until the 1950s, when the government cleared the virgin rainforest and families began settling there. Subsequently, services such as health services began to appear in the valley. It was only after 1950 that modern health care services were made available to the public (Justice 1986). Previously only natural healers and homeopathic services were available. Nepal's public health policy was not formulated until 1960 (Justice 1986). There is still tremendous variation in access to health services.

Although strong evidence demonstrates that health services influence fertility behaviors, evidence to evaluate the relative impact of different types of health services and closely related features of community context has been elusive. Maternal and child health services have been promoted as an important influence on couples' fertility related behaviors because they decrease child mortality (Caldwell 1986; Frankenberg and Thomas 2001; Pebley 1984). With reduced child mortality, couples can achieve their desired family size by having fewer children. However, the proliferation of health services often accompanies the spread of numerous other important

[1] It is true that residents would have been able to leave the valley to obtain medical services prior to the building of the first health service provider in the valley. However, those service providers probably had very weak effects on behavior because they were so far away.

social, economic, and institutional changes that affect fertility behaviors (Axinn and Yabiku 2001; Caldwell 1986; Hernandez 1981). Of specific concern is that maternal and child health services often exist in conjunction with family planning services and other related key aspects of community context, yet we know little of their independent influence on fertility behavior. Also, family planning and other health services are often provided nearby maternal and child health services, if not at the same facilities. Consequently, any observed effects of various maternal and child health services on fertility regulation may in fact be due to the provision of other services.

Data

In 1996, the CVFS collected information from residents of a systematic sample of 171 neighborhoods in Western Chitwan Valley. Neighborhoods were defined as clusters of approximately 5 to 15 households. In this rural setting, a neighborhood defines a group of individuals who have face-to-face contact on a daily basis. The CVFS also collected detailed accounts of neighborhood resources and health services available since 1954—the first year a health service provider opened in Chitwan—and interviewed every resident between the ages of 15 and 59 in the 171 sampled neighborhoods, and their spouses. The overall response rate of 97 percent yielded 5,271 completed interviews. All interviews were conducted in the most common language in Nepal, Nepali (questions presented below are translated). Life History Calendar techniques were used to collect reliable information regarding residents' contraceptive behavior, marital and childbearing behavior, education, and labor force participation (Axinn, Pearce, and Ghimire 1999). We analyze data from 1,395 women in the CVFS who were between the ages of 25 and 54 at the time of the interview, married, and had at least one child. We restrict the sample in this way because in the Nepalese context, never married women, childless women and young women are extremely unlikely to use permanent contraceptive methods (Acharya and Bennet 1981; Axinn 1992a; Tuladhar 1989).

Information regarding changes in health services and other important aspects of community context was collected using the Neighborhood History Calendar technique (Axinn, Barber and Ghimire 1997). This technique combines archival, ethnographic, and structured interview methods. The Neighborhood History Calendar technique was also used to collect Health Service History Calendars on all 113 health service providers that have existed in the Chitwan Valley since 1954, regardless of whether that health service provider was currently active.

The vast amount of information included in the CVFS make it an ideal dataset to use to test these multiple approaches. Previous research has found that effect estimates from matching are unbiased estimates of actual treatment effects when the data used satisfy three conditions (Heckman et al. 1997, 1998a, b; Smith and Todd 2001). First, the data used should all come from the same source. Second, treated and control individuals should come from the same local labor markets.[2] Third, the dataset is rich enough to model the assignment process. The CVFS data do in fact satisfy all three of these conditions.

Analytic Strategy

Linear models. We first present our linear-model based adjustments. We estimate logistic regressions models of the relationship between each maternal and child health service and the likelihood of using permanent contraceptive methods.

We estimate these models using logistic regression of the form:

$$\ln\left(\frac{p}{1-p}\right) = a + \sum (\beta_k)(X_k), \qquad \text{(Equation 1)}$$

where p is the probability of *ever* using a permanent contraceptive method, $p/1\text{-}p$ is the odds of using a permanent contraceptive method, a is a constant term, β_k represents the effects parameters of the explanatory variables, and X_k represents the explanatory variables in the model.

We begin with a zero order model of the effect of the treatment, the specific maternal or child health service, on the outcome, permanent contraceptive use. We then incorporate a broad

[2] The focus on labor markets in these criteria is probably due to the substantive nature of these previous research efforts—job training programs. However, the same principal applies to the problem considered here. All the respondents need to be making their behavioral choices in similar situations. Because the CVFS collected information from respondents living in a localized area, there are not obvious dramatic differences that could account for any observed effects.

collection of covariates. The third model includes the propensity score, described below in detail, as a covariate along with the treatment variable.

Because these models are clustered with several individuals living in the same community who all have the same community characteristics, such as access to maternal and child health services, we estimate these models taking into account the multilevel nature of the data. Specifically, we use the *GLIMMIX* macro for SAS 8.2. The results presented in the tables below have all been calculated using *GLIMMIX* and therefore properly specify the multilevel nature of the data.

Non-parametric models—matching. Matching involves creating matched sets where respondents are assigned to either the treatment or control group after the data have already been collected. We conduct optimal matching in the analyses here. With optimal matching a treated respondent can be matched to multiple control respondents, visa versa, or one treated respondent can be matched to one control respondent. The matching problem is converted into a minimum cost flow problem. Essentially, with optimal matching the matches that minimize the overall "cost," or difference between matched sets, are selected.

Matching serves to create two groups that are "balanced." These two groups are parallel concepts to the treated and control groups in randomized experiments. With randomization the treated and control groups should yield unbiased effect estimates because any other factor that may have otherwise caused the observed effect is equally distributed in the two groups. This is how two balanced matched groups should also be characterized. Once the matches have been made, it is necessary to confirm that the two groups are in fact balanced. If they are not, then adjustments to the matching algorithm are necessary. If they are balanced then you can compare the outcome across the two groups. We follow these steps, checking balance and making the appropriate adjustments. However, for parsimony we do not present those results in this paper.

In this paper we first match exactly on the year the respondent's first birth occurred. Year of first birth is extremely important because both the access to health services and general use and acceptance of permanent contraceptive use have changed dramatically over our study period.

The second match we perform matches only on the propensity score which is described further below. The third match uses both techniques, matching exactly on year of first birth (specifically requiring matched sets to have had their first births within five years of each other) and matching on the propensity score.

Measures

Measure of Outcome

Fertility limitation—use of permanent contraceptives. Although a variety of contraceptives has been available in Nepal for the past 20 years, Nepalese, like other South Asian populations prefer permanent contraceptive methods. Furthermore, among Nepalese women, contraceptive methods that may be used for fertility spacing such as Norplant, Depo-Provera and the IUD are typically used to stop childbearing (Axinn 1992a). Data from the CVFS reveal that 98.8 percent of the married women aged 25 to 54 who had at least one child and who had ever used any of these contraceptives said that they wanted no more children. Consequently, we consider sterilization, Norplant, IUDs and Depo-Provera to be contraceptive methods used to stop childbearing.[3] Additionally, this specification is ideal because it has been used in previous research that examines the impact of community context on fertility limitation (Axinn and Barber 2001, Axinn and Yabiku 2001). We code a dichotomous variable equal to 1 if the respondent ever used contraceptives and zero if she did not. Fifty-eight percent of women used permanent contraceptives in our study period. Table 1 presents descriptive statistics for all individual and neighborhood level measures used in the traditional linear models. Table 2 presents the additional measures used in the non-parametric models and in creating the propensity scores.

(Table 1 here)

[3] To assess the robustness of our findings we explored alternative definitions of fertility limitation. Specifically, we found that when considering any contraceptive use, whether permanent, such as those mentioned above, or temporary, such as birth control pills or condoms, our estimates were virtually the same. Similarly, using sterilization only as the dependent variable did not substantially alter the estimates.

Measure of Treatment

Maternal and child health services. We constructed four measures of maternal and child health services. We consider four different types of services that potentially could have been offered by the health service provider: child vaccinations (BCG, measles, DPT3, OPV3, Tetanus Toxoid), oral-rehydration therapy for children, prenatal care (regular exams and provision of iron supplements) and pregnancy/delivery assistance services (safe delivery kits and midwives). Our measures of the treatment equal 1 if the health service provider closest to the respondent offered that service the year before the respondent gave birth to her first child. Seventeen percent of women had child vaccinations, 82 percent had oral rehydration therapy, 79 percent prenatal care, and 52 percent delivery assistance offered by their closest health service provider the year before they had their first birth.

Other Measures for Linear Models

As we mentioned, for our linear models we will take a measurement adjustment approach to assess the robustness of our conclusions. The discussion below describes the measures that we include in these models. These measures are chosen because they are theoretically predicted to influence both using permanent contraceptives and having access to child vaccinations. In order to estimate the full treatment effect it is crucial that we do not include in our models measures of anything that could be a result of the treatment. To ensure against this we avoid including measures of anything that occurred in time after the treatment. The following measures are introduced into the second set of models.

Previous research has shown that neighborhood characteristics are strongly associated with contraceptive use to limit childbearing (Axinn and Barber 2001, Axinn and Yabiku 2001). Therefore, we include measures of other elements of community context in our analysis. Specifically, we include measures of both recent and childhood community characteristics.

We use an index measure of local, recent community context. This measure uses information collected in Neighborhood History Calendars to indicate whether the individual's recent, local community has a school, employer, market, or bus transportation within a five-minute walk. The index counts the number of such services and organizations. If an individual has a school, employer, market, and bus within a five-minute walk the index equals four for that year. The index equals one for a respondent who has only a school within a five-minute walk. This index of the institutional context of local communities refers to the respondent's community the year before she had her first birth.

We also include an index of the number of non-family institutions—schools, health service providers, bus routes, employers, and markets—within a one-hour walk of the respondent's neighborhood during childhood. As with recent community context, this measure is the sum of separate dichotomous measures of the presence of these institutions. For example, we created a dichotomous measure equal to one if there was a school within a one hour walk of that respondent. We then summed all of the dichotomous variables. This measure is created from information in the individual survey interview.

Parents with education or work experience may choose to move to neighborhoods with greater access to schools, health service providers, and other institutions. Therefore, we control for several aspects of family background: father's and mother's education (ever went to school), father's employment (ever had non-family employment before respondent's age 12), mother's children ever born and parents ever used a contraceptive method.

Individual's non-family experiences may influence both contraceptive use and the individual's choice in community later in life (i.e. individual's who had used health services early in life may be more likely to seek out communities with health services nearby). We use information gathered on the Life History Calendars to create a measure of measure non-family work (wage employment, salaried employment, or owning a business outside the home), non-family school (standard education or adult education), living away from home (alone, in dormitories, with unrelated individuals, and other non-family possibilities), use of non-family health services, and seeing movies in a movie hall. These measures code whether the respondent had each experience before her first child was born. We code each measure as a dichotomous indicator equal to one if the individual has had the experience in question and sum these indicators to create an index counting the total number of non-family experiences.

We control for the respondent's ethnicity. Ethnicity in Nepal is complex, multifaceted, and interrelated with religion. A full description of the ethnic groups in this setting is beyond the scope of this article (see Acharya and Bennett 1981; Bista 1972; Fricke 1986; and Gurung 1980 for detailed descriptions). We use dichotomous variables to control for five classifications of ethnicity: high-caste Hindu, low-caste Hindu, Newar, hill Tibeto-Burmese, and terai Tibeto-Burmese. Each group has different propensities to use permanent contraceptives and different access to health services. High-caste Hindu is the reference group in our analyses.

We also include a control for migration: a dichotomous measure for whether the respondent had moved before the birth of her first child. We also include a time-varying measure of whether the respondent was living with her spouse. This measure is a proxy for coital frequency, and controls for the risk of pregnancy.

Because the prevalence of contraceptive methods has increased over time we control for birth cohort. We create dichotomous variables for three birth cohorts: 1962-71 (ages 25-34 at the survey), 1952-61 (ages 35-44 at the survey), and 1942-51 (ages 45-54 at the survey). The 1962-71 birth cohort is the reference group for the analysis.

A final control is the distance between a respondent's neighborhood and the nearest health service provider. This measure is the log of the distance, in meters, between the neighborhood and the provider.

Measure for matching

The treatment and outcome measures used in our matching approach are the same as those described above. The first match we perform uses the year of the respondent's first birth as the variable to match on. The values of this measure are the actual year. The matching algorithm minimizes the differences between these years for the respondents.

Propensity Score Estimation

The propensity scores we use are the estimated probability of having the treatment—i.e. of having child vaccinations offered by your closest health service provider the year before your first birth. These are obtained by running a simple logistic regression where the dependent variable is our treatment. We include many neighborhood and individual level predictors, some of which are also included in the our models of permanent contraceptive use.[4] The same covariates are included in each of the separate models of the specific maternal or child health service.

(Table 2 here)

For neighborhood level predictors we include characteristics of the community that may make it more likely to have maternal or child health services available. The two measures we include to capture this possible effect are a measure of recent community characteristics similar to that described above except that it refers to the period five years before the respondent's first birth, as opposed to the year before and a measure the distance to Narayangat, the nearest major metropolitan area.

We also include measures of childhood community context similar to those described above. However, instead of including an index measure we include separate dummy variables. In addition to dummies for schools, employers, markets, and buses, we also include dummies for the presence of development programs, police stations, and temples within an hours walk. Because of the nature of the propensity score we do not have the collinearity and variance concerns we would normally have in a regression, thereby allowing us to include these additional measures.

We include several individual level measures that influence either the choice of neighborhood or earlier behavior that may have led to the creation of health services. Parental education, employment, and contraceptive use history, ethnicity, and migration as described

[4] The measures we include in both are those we consider to be especially relevant for the relationship between child vaccinations and permanent contraceptive use.

above, are all included. Additional measures of parental characteristics we include are a dummy variables for whether either parent had ever worked for pay, had ever traveled outside of Nepal before the respondent was 12 years old, and whether either parent can read. Finally, we include three additional measures of individual experiences. We include a dummy variable for whether the respondent's parents chose her spouse without any input from the respondent and a dummy for whether the respondent had traveled to Kathmandu before her marriage. The last measure is of the year of the respondent's first birth. We include this as series of dummies based on whether the birth occurred, before 1975, between 1976 and 1980, between 1981 and 1985, 1985 and 1990, or 1991 and 1995. The reference category is first birth occurred before 1975.

We also include one measure about the health service itself—specifically how it was established. As discussed above, why the health service was built may be cause the researcher to observe a spurious effect of child vaccinations on permanent contraceptive use. The measure we include equals 1 if the health service was established solely by private efforts.[5] If the health service was established through government, non-governmental organization (NGO), or public project efforts we coded it as zero.

Results

Linear Models

In Table 3 we present the results for the logistic model estimates of the relationship between child vaccinations and the likelihood of permanent contraceptive use. We transformed the raw coefficients by exponentiating them; the coefficients we display are estimates of the multiplicative effects on the odds of contracepting. A coefficient of 1.00 represents no effect, a coefficient greater than 1.00 represents a positive effect on the odds, and a coefficient less than 1.00 represents a negative effect on the odds.

[5] This also includes one health service that is actually a combination of institutions, some of which were founded by private efforts and some by government efforts. Since the largest of the institutions in the complex was established through private efforts, we coded it as 1. We tested the model with coding this one health service as zero and the effects did not change.

Model 1, the zero order model, shows a positive, significant coefficient for child vaccinations. Women who had child vaccinations available at their closest health service provider the year before she had her first child were 36 percent more likely to use permanent contraceptives than women who did not have child vaccinations available. Interestingly, as shown by Model 2, including other covariates actually increases the size of the coefficient from a 36 percent increased likelihood to a 56 percent increase. In Model 3, we add the propensity score to the zero order model. For the case of child vaccinations we see that adding the propensity score makes the estimate no longer statistically significant. However, it may be unwise to make much of this finding since there is in fact very little difference between the estimated effects in Models 1 and 3.

(Table 3 here)

Tables 4, 5, and 6 present similar results as Table 1 except the treatment variables are the availability of oral rehydration therapy (ORT), prenatal care, and delivery assistance respectively. As these tables show, estimates of the effect of the different health services responded differently to the model specifications. For the zero order models, Model 1 in all three tables, only the availability delivery assistance was found to be statistically significant. However, when covariates were added to the model, Model 2, the availability of delivery assistance was no longer statistically significant, but the availability of ORT was—women who had ORT available at their closest health post, the year before their first birth, were 50 percent more likely to use permanent contraceptives than woman who did not have ORT available. In the final models where we included the propensity score as a covariate, Model 3, the availability of ORT remained statistically significant and that of prenatal care became significant. The inconsistent trend in findings across treatments and model specifications is cause for concern when using these linear models to make causal inferences.

(Tables 4, 5, and 6 here)

Non-parametric models

In Table 7 we present the results for our non-parametric models. We present the odds ratios for the estimate of the effect of the treatment on the outcome. We use the Mantel-Haenszel χ-square as our test statistic (Mantel and Haenszel 1959). It has a large-sample chi-squared null distribution with 1 degree of freedom. Model 1 for each panel shows the results when respondents were matched based on the year of their first birth only, Model 2 for matches on the propensity score only, and Model 3 for matches on the propensity score and with year of the first birth within 5 years of each other. Because this approach is meant to mimic that of a randomized experiment we use language typical of those discussions.

(Table 7 here)

Panel A shows the models with availability of child vaccinations as the treatment. In Model 1, the match on year of first birth only, we find that the treated group, those who had child vaccinations available, used permanent contraceptives 58 percent more than the control group. This treatment effect remains in our matches using the propensity score and the combined match.

Panel B shows the models for the availability of ORT as the treatment. We again find a statistically significant treatment effect across match specifications. Based on the test statistics we are much more confident in this estimate of a treatment effect than we are of the effect of child vaccination availability.

Panels C and D show the models for prenatal care and delivery services respectively. We do not find evidence of any effect of prenatal care or delivery services on permanent contraceptive use for any of our match specifications.

Conclusion

Although there have been, and continue to be, efforts to study demographic issues such as the effect of health services on fertility behavior through experiments, these efforts are not always successful nor can they realistically be expanded to cover every research question. As a result, researchers must find ways to enhance their ability to make causal claims from observational data. This paper presents two approaches for doing this. The first approach uses alternative measurement strategies to reinforce our confidence in our estimates. The second approach uses alternative analytic strategies to do the same.

Substantively, our findings consistently show that the availability of child vaccinations and ORT are related to increased permanent contraceptive use. Essentially across all our models, and importantly in our non-parametric models, we found positive and significant effects of these child health services on the use of permanent contraceptives. For child vaccinations we found a range of effect sizes that translate into a roughly 20 to 60 percent increase in the likelihood of permanent contraceptive use for women who had child vaccinations available at their closest health service provider the year before their first birth. For ORT availability we obtained estimates of between a 30 and 90 percent increase in the likelihood of permanent contraceptive use. These are important findings for those concerned with lowering child mortality, increasing contraceptive use, or lowering fertility in extremely poor countries like Nepal.

Methodological we find that as we expect, measurement and model choices can make dramatic differences in the estimated effects. However, having a multitude of methodological techniques available allowed us to make some fairly confident claims about relationships between specific health services' availability and permanent contraceptive use.

While this approach does provide use with dramatically increased confidence regarding our estimates of the effect of the availability of maternal and child health services on permanent contraceptive use, that does not necessarily mean this will be the case for all research questions. The specific question we used in this paper is a general one—do maternal or child health services have an effect. For other types of questions, such as those about relative effects of different programs, or about the actual change expected in the true population, these techniques may not be sufficient. The results presented here are probably not precise enough if we were concerned with whether providing child vaccinations or ORT would result in more contraceptive use. Similarly, we cannot say for sure how much of an increase in contraceptive use we should expect if we made child vaccinations or ORT more widespread.

Future efforts should focus on finding methods that will allow these types of questions to be satisfactorily addressed from observational data. Statisticians and demographers need to work together to develop these new methods and ensure they are applicable to real world situations. Up until now, most the work in this area has been by statisticians connected to the medical field. Specific attention to the problems in the social sciences is crucial. This is especially true for other demographic questions that do not involve programs or services. It will never be possible for us to randomly assign divorce to couples so in order to understand the effects of divorce on individuals we must develop these alternative methods.

One very clear next step in these methods is to incorporate time and life histories into matching and propensity score approaches. Hazard modeling is becoming increasingly common in demography because it more correctly specifies peoples lives. By incorporating time and exposure these models can capture more realistic relationships. Finding a way to combine the benefits of matching, propensity scores, and hazard modeling would be highly beneficial to demographic research.

Reference List

Acharya, Meena and Lynn Bennett. 1981. *Rural Women of Nepal: An Aggregate Analysis and Summary of Eight Village Studies*. Kathmandu: Tribhuvan University.

Allison, Paul D. 1982. "Discrete-Time Methods for the Analysis of Event Histories." Pp. 61-98 in *Sociological Methodology*, editor S. Leinhart.

Allison, Paul D. 1984. *Event History Analysis*. Newbury Park, CA: Sage Publishing.

Axinn, William G. 1992. "Family Organization and Fertility Limitation in Nepal." *Demography* 29(4):503-21.

Axinn, William G. and Jennifer S. Barber. 2001. "Mass Education and Fertility Transition." *American Sociological Review* 66(4):481-505.

Axinn, William G., Lisa D. Pearce, and Dirgha J. Ghimire. 1999. "Innovations in Life History Calendar Applications." *Social Science Research* 28(3):243-64.

Axinn, William G. and Scott T. Yabiku. 2001. "Social Change, the Social Organization of Families, and Fertility Limitation." *American Journal of Sociology* 106(5):1219-61.

Barber, Jennifer S. and William G. Axinn. Forthcoming. "New Ideas and Fertility Limitation: The Role of Mass Media." .

Barber, Jennifer S., S. A. Murphy, William G. Axinn, and J. Maples. 2000. "Discrete-Time Multilevel Hazard Analysis." *Sociological Methodology* 30:201-35.

Berk, Richard A. and Phyllis J. Newton. 1985. "Does Arrest Really Deter Wife Battery? An Effort to Replicate the Findings of the Minneapolis Spouse Abuse Experiment." *American Sociological Review* 50(2):253-362.

Bista, Dor B. 1972. *People of Nepal*. Kathmandu : Ratna Pustak Bhandar.

Caldwell, John C. 1986. "Routes to Low Mortality in Poor Countries." *Population and Development Review* 12(2):171-220.

Campbell, Donald T. and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Co.

Cochran, W. G. 1965. "The Planning of Observational Studies of Human Populations." *Journal of the Royal Statistical Society* 128:234-66.

Cornfield, Jerome, William Haenszel, E. C. Hammond, Abraham M. Lilienfeld, Michael B. Shimkin, and Ernst L. Wynder. 1959. "Smoking and Lung Cancer: Recent Evidence and a Discussion of Some Questions." *Journal of the National Cancer Institute* 22:173-203.

Frankenberg, Elizabeth and Duncan Thomas. 2001. "Women's Health and Pregnancy Outcomes:

Do Services Make a Difference?" *Demography* 38(2):253-365.

Fricke, Thomas E. 1986. *Himalayan Households: Tamang Demography and Domestic Processes*. Ann Arbor, MI: UMI Research Press.

Gurung, Harka B. 1980. *Vignettes of Nepal*. Kathmandu: Sajha Prakashan.

Hansen, Ben B. 2004. "Full Matching in an Observational Study of Coaching for the SAT." *Journal of the American Statistical Association* 99.

Harding, David J. 2003. "Counterfactual Models of Neighborhood Effects: The Effect of Neighborhood Poverty on Dropping Out and Teenage Pregnancy." *American Journal of Sociology* 109(3):676-719.

Hernandez, Donald J. 1981. "A Note on Measuring the Independent Impact of Family Planning Programs on Fertility Declines." *Demography* 18(4):627-34.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81:945-70.

Joffe, Marshall M. and Paul R. Rosenbaum. 1999. "Invited Commentary: Propensity Scores." *American Journal of Epidemiology* 150(4):327-33.

Justice, Judith. 1986. *Polices, Plans and People: Culture and Health Development in Nepal*. Berkeley: University of California Press.

Moffitt, Robert. 2003. " Causal Analysis in Population Research: An Economist's Perspective." *Population and Development Review* 29:448-58.

———. 2005. "Remarks on the Analysis of Causal Relationships In Population Research." *Demography* 42(1):91-108.

Morgan, Stephen L. 2001. "Counterfactuals, Causal Effect Heterogeneity, and the Catholic School Effect on Learning." *Sociology of Education* 74:341-74.

Pebley, Anne R. 1984. "Intervention Projects and the Study of Socioeconomic Determinants of Mortality." *Population and Development Review* 10(Supplement: Child Survival: Strategies for Research):281-305.

Petersen, Trond. 1986. "Estimating Fully Parametric Hazard Rate Models With Time-Dependent Covariates: Use of Maximum Likelihood." *Sociological Methods and Research* 14:219-46.

———. 1991. "The Statistical Analysis of Event Histories." *Sociological Methods and Research* 19(3):270-323.

Rosenbaum, Paul R. 1986. "Dropping Out of High-School in the United-States: An Observational Study." *Journal of Educational Statistics* 11(3):207-24.

———. 2002. *Observational Studies*. New York: Springer.

Rosenbaum, Paul R. and Donald B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41-55.

Rosenbaum, Raul R. and Donald R. Rubin. 1985. "Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score." *The American Statistician* 39(1):33-38.

Rubin, Donald B. 1997. " Estimating Causal Effects From Large Data Sets Using Propensity Scores." *Annals of Internal Medicine* 127(8):757-63.

Rubin, Donald B. and Neal Thomas. 1992. "Characterizing the Effect of Matching Using Linear Propensity Score Methods With Normal Distributions." *Biometrika* 79(4):797-809.

Smith, Herb L. 1997. "Matching With Multiple Controls to Estimate Treatment Effects in Observational Studies." *Sociological Methodology* 27:325-53.

———. 2003. "Some Thoughts on Causation As It Relates to Demography and Population Studies." *Population and Development Review* 29(459-69).

Smith, Jeffrey A. and Petra E. Todd. 2001. "Reconciling Conflicting Evidence on the Performance of Propensity-Score Matching Methods." *The American Economic Review* 91(2, Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association):112-18.

Tuladhar, Jayanti M. 1989. *The Persistence of High Fertility in Nepal*. New Dehli: Inter-India Publications.

Whyte, Martin K. and William L. Parish. 1984. *Urban Life in Contemporary China*. Chicago: University of Chicago Press.

Winship, Christopher and S. L. Morgan. 1999. "The Estimation of Causal Effects From Observational Data." *Annual Review of Sociology* 25:659-707.

Table 1. Descriptive Statistics for Measures Used in Linear Model Based Approach*

	Mean	SD	Min	Max
Outcome/Dependent variable				
Permanent contraceptive use………..…………………..……………..	0.58		0	1
Treatment/Independent variable				
Child vaccinations at closest provider in year before 1st birth……………	0.17		0	1
Oral rehydration therapy at closest provider in year before 1st birth…….	0.82		0	1
Prenatal care at closest provider in year before 1st birth ………………..	0.79		0	1
Delivery assistance at closest provider in year before 1st birth…………..	0.52		0	1
Controls				
Recent community characteristics………………………………………	1.04	1.16	0	4
Childhood community characteristics within one hour....…….……..…..	2.79	1.72	0	5
Father's education (ever went to school)………………..………………	0.21		0	1
Mother's education (ever went to school)……………….………………	0.02		0	1
Father's employment (ever had paid employment)………..………….....	0.38		0	1
Mother's children ever born ………………………………...………	6.25	2.91	1	19
Parental contraceptive use (parents ever use)…………….……….…….	0.14		0	1
Non-family experiences before first birth………..………..………….	1.64	1.21	0	5
High-caste Hindu……………………………………………...……….	0.47		0	1
Low-caste Hindu………………………………..…………………..	0.11		0	1
Newar………………………………………..…………………..	0.07		0	1
Hill Tibeto-Burmese…………………………....…………………..	0.18		0	1
Terai Tibeto-Burmese…………………………..…….……………..	0.17		0	1
Ever moved since first birth…………………...…………………….	0.73		0	1
Living with Spouse……………………………..……………………	0.90		0	1
Born 1962-71 (age 25-34 at the survey)……………….………………	0.41		0	1
Born 1952-61 (age 35-44 at the survey)……………….………………	0.34		0	1
Born 1942-51 (age 45-54 at the survey)……………….………………	0.25		0	1
Distance from current neighborhood to closest health service (log of meters)	9.34	3.13	3.07	14.94

*Summary statistics for N=1,395

Table 2. Descriptive Statistics for Measures Used in Non-parametric Approach*

	Mean	SD	Min	Max
Measures used in matching models				
Year of first birth……………………………………………………………..	1977	8.81	1953	1994
Additional measures used in propensity models				
Who established the service provider				
Private effort[a]…………………………………………………………….	0.90		0	1
Government…………………………………………………………	0.07		0	1
Non-governmental organization (NGO) or development project………	0.01		0	1
Combination-government and NGO or development project…………	0.02		0	1
Distance to Narayangat…………………………………………...…….	8.59	3.90	0.02	17.70
Community characteristics five years before 1st birth				
Bus stop within 5 minute walk…………………………………………	0.22		0	1
Bus stop with service to Narayangat within 5 minute walk……………..	0.19		0	1
Employer within 5 minute walk ……………………………………….	0.11		0	1
Health service provider within 5 minute walk………………………….	0.11		0	1
Mill within 5 minute walk……………………………………………..	0.12		0	1
Market within 5 minute walk………………………………………….	0.25		0	1
School within 5 minute walk ………………………………………….	0.31		0	1
Temple within 5 minute walk………………………………………….	0.19		0	1
Childhood community characterisitics				
Bus stop within 1 hr walk before age 12 …………………...………..	0.38		0	1
Employer within 1 hr walk before age 12 ……………………….………	0.45		0	1
Development program within 1 hr walk before age 12 …………….……….	0.24		0	1
Health service provider within 1 hr walk before age 12 ……..………….	0.45		0	1
Market within 1 hr walk before age 12 ………………………….………	0.72		0	1
Police station within 1 hr walk before age 12 …………………..……..	0.40		0	1
School within 1 hr walk before age 12 ……………………...………..	0.79		0	1
Temple within 1 hr walk before age 12 ………………………….………	0.69		0	1
Parental characteristics				
Either parent traveled outside of Nepal before respondent's age 12 ………	0.33		0	1
Either parent can read …………………………………………………..	0.52		0	1
Either parent worked outside the home…………………………………..	0.41		0	1
Individual experiences				
Parents chose spouse ……. ……..………………………………………	0.80		0	1
Traveled to Kathmandu before marriage ……………………………….	0.03		0	1
Years of school before first marriage …………………………………….	1.99		0	19
Year of first birth				
First birth before 1975 …………………...……..………….	0.43		0	1
First birth between 1976-1980…………..…..…...………….	0.18		0	1
First birth between 1981-1985 …………...…………………….	0.18		0	1
First birth between 1986-1990 …………...…………………….	0.16		0	1
First birth between 1991-1995 …………...…………...………..	0.04		0	1

*Summary statistics for N=1,395

[a]Includes one compound that was greated by both private and government efforts.

Table 3. Linear Based Measurement Approaches: Multilevel Estimates of the Relationship Between Whether Closest Health Service Provider At Time of Respondents 1st Birth Offered Child Vaccinations and Rate of Permanent Contraception

	1	2	3
Child vaccinations	1.36*	1.56**	1.38
	(1.86)	(2.34)	(1.6)
Propensity score			0.84
			(-0.45)
Controls			
Number of non-family organizations in current community		1.08	
		(1.13)	
Number of non-family organizations in childhood community		1.12**	
		(2.54)	
Family Background			
Father's education (ever went to school)		0.84	
		(-1.07)	
Father's employment (ever had paid employment)		1.5**	
		(2.94)	
Mother's education (ever went to school)		0.63	
		(-1.03)	
Mother's children ever born		0.98	
		(-0.75)	
Parental contraceptive use (parents ever use)		0.86	
		(-0.78)	
Non-family experiences before first birth		0.91	
		(-1.40)	
Ethnicity			
Low caste Hindu		0.61*	
		(-2.21)	
Newar		1.14	
		(0.47)	
Hill Tibeto-Burmese		0.57**	
		(-2.95)	
Terai Tibeto-Burmese		0.31***	
		(-5.73)	
Ever moved since first birth		1.10	
		(0.63)	
Living with Spouse		1.33	
		(1.40)	
Birth cohort[a]			
Born 1952-61 (age 35-44 at the survey)		1.71***	
		(3.20)	
Born 1942-51 (age 45-54 at the survey)		0.96	
		(-0.20)	
Distance to closest health service provider		1.01	
		(0.42)	
Neighborhood level variance component (sigma sq)	0.36	0.20	0.37
Residual variance component	0.94	0.97	0.94
Interclass correlation	0.27	0.17	0.28

Note: Estimates are presented as odds ratios, with z-statistics in parentheses.

[a]Born 1962-71 is the reference group

* P < .05, one tailed test; ** P < .01, one-tailed test; *** P < .001, one-tailed test

Table 4. Linear Based Measurement Approaches: Multilevel Estimates of the Relationship Between Whether Closest Health Service Provider At Time of Respondents 1st Birth Offered Oral Rehydration Therapy and Rate of Permanent Contraception

	1	2	3
Oral Rehydration Therapy	1.3 (1.6)	1.5* (2.24)	1.65** (2.63)
Propensity score			0.31 (-2.8)
Controls			
Number of non-family organizations in current community		1.04 (0.7)	
Number of non-family organizations in childhood community		1.12** (2.61)	
Family Background			
Father's education (ever went to school)		0.86 (-0.93)	
Father's employment (ever had paid employment)		1.4** (2.55)	
Mother's education (ever went to school)		0.67 (-0.92)	
Mother's children ever born		0.98 (-0.8)	
Parental contraceptive use (parents ever use)		0.92 (-0.47)	
Non-family experiences before first birth		0.91 (-1.48)	
Ethnicity			
Low caste Hindu		0.65 (-2.08)	
Newar		1.14 (0.5)	
Hill Tibeto-Burmese		0.63 (-2.5)	
Terai Tibeto-Burmese		0.3 (-6.33)	
Ever moved since first birth		1.02 (0.16)	
Living with Spouse		1.26 (1.17)	
Birth cohort[a]			
Born 1952-61 (age 35-44 at the survey)		1.59** (2.97)	
Born 1942-51 (age 45-54 at the survey)		0.84 (-0.83)	
Distance to closest health service provider		1.03 (1.05)	
Neighborhood level variance component (sigma sq)	0.37	0.17	0.40
Residual variance component	0.94	0.98	0.94
Interclass correlation	0.28	0.15	0.30

Note: Estimates are presented as odds ratios, with z-statistics in parentheses.

[a]Born 1962-71 is the reference group

* P < .05, one tailed test; ** P < .01, one-tailed test; *** P < .001, one-tailed test

Table 5. Linear Based Measurement Approaches: Multilevel Estimates of the Relationship Between Whether Closest Health Service Provider At Time of Respondents 1st Birth Offered Prenatal Care and Rate of Permanent Contraception

	1	2	3
Prenatal Care	1.2 (1.27)	1.06 (0.29)	1.4* (1.74)
Propensity score			0.5 (-2.16)
Controls			
Number of non-family organizations in current community		1.04 (0.64)	
Number of non-family organizations in childhood community		1.11** (2.52)	
Family Background			
Father's education (ever went to school)		0.86 (-0.98)	
Father's employment (ever had paid employment)		1.39** (2.52)	
Mother's education (ever went to school)		0.68 (-0.88)	
Mother's children ever born		0.98 (-0.83)	
Parental contraceptive use (parents ever use)		0.92 (-0.43)	
Non-family experiences before first birth		0.91 (-1.48)	
Ethnicity			
Low caste Hindu		0.67 (-1.87)	
Newar		1.16 (0.58)	
Hill Tibeto-Burmese		0.64 (-2.4)	
Terai Tibeto-Burmese		0.3 (-6.21)	
Ever moved since first birth		1.03 (0.2)	
Living with Spouse		1.26 (1.17)	
Birth cohort[a]			
Born 1952-61 (age 35-44 at the survey)		1.54** (2.78)	
Born 1942-51 (age 45-54 at the survey)		0.82 (-0.91)	
Distance to closest health service provider		1.01 (0.49)	
Neighborhood level variance component (sigma sq)	0.39	0.20	0.39
Residual variance component	0.94	0.97	0.94
Interclass correlation	0.29	0.17	0.29

Note: Estimates are presented as odds ratios, with z-statistics in parentheses.

[a]Born 1962-71 is the reference group

* P < .05, one tailed test; ** P < .01, one-tailed test; *** P < .001, one-tailed test

Table 6. Linear Based Measurement Approaches: Multilevel Estimates of the Relationship Between Whether Closest Health Service Provider At Time of Respondents 1st Birth Offered Delivery Assistance and Rate of Permanent Contraception

	1	2	3
Delivery Assistance	1.23*	1.22	1.21
	(1.7)	(1.43)	(1.32)
Propensity score			0.9
			(-0.35)
Controls			
Number of non-family organizations in current community		1.05	
		(0.8)	
Number of non-family organizations in childhood community		1.12**	
		(2.57)	
Family Background			
Father's education (ever went to school)		0.87	
		(-0.9)	
Father's employment (ever had paid employment)		1.4**	
		(2.57)	
Mother's education (ever went to school)		0.69	
		(-0.85)	
Mother's children ever born		0.98	
		(-0.8)	
Parental contraceptive use (parents ever use)		0.92	
		(-0.43)	
Non-family experiences before first birth		0.91	
		(-1.44)	
Ethnicity			
Low caste Hindu		0.65	
		(-2.01)	
Newar		1.16	
		(0.57)	
Hill Tibeto-Burmese		0.64	
		(-2.47)	
Terai Tibeto-Burmese		0.3	
		(-6.23)	
Ever moved since first birth		1.02	
		(0.15)	
Living with Spouse		1.27	
		(1.2)	
Birth cohort[a]			
Born 1952-61 (age 35-44 at the survey)		1.59**	
		(2.93)	
Born 1942-51 (age 45-54 at the survey)		0.87	
		(-0.68)	
Distance to closest health service provider		1.02	
		(0.68)	
Neighborhood level variance component (sigma sq)	0.38	0.19	0.39
Residual variance component	0.94	0.97	0.94
Interclass correlation	0.29	0.16	0.29

Note: Estimates are presented as odds ratios, with z-statistics in parentheses.

[a]Born 1962-71 is the reference group

* P < .05, one tailed test; ** P < .01, one-tailed test; *** P < .001, one-tailed test

Table 7. Results From Optimal Matches Using Propensity Scores-Relationship Between Whether Closest Health Service Provider At Time of Respondent's First Birth Offered Specific Service and Rate of Permanent Contraception

A. Child Vaccinations

	Year of 1st Birth	Propensity score only	Propensity score and year of 1st birth
	1	2	3
Odds ratio	1.58	1.40	1.55
Mantel-Haenszel X-squared	4.23	2.71	4.13
p-value	0.04	0.10	0.04

B. Oral Rehydration Therapy

	Year of 1st Birth	Propensity score only	Propensity score and year of 1st birth
	1	2	3
Odds ratio	1.44	1.91	1.56
Mantel-Haenszel X-squared	3.09	12.19	6.71
p-value	0.08	0.00	0.01

C. Prenatal Care

	Year of 1st Birth	Propensity score only	Propensity score and year of 1st birth
	1	2	3
Odds ratio	0.81	1.32	1.14
Mantel-Haenszel X-squared	0.87	1.56	0.29
p-value	0.35	0.21	0.59

D. Delivery Services

	Year of 1st Birth	Propensity score only	Propensity score and year of 1st birth
	1	2	3
Odds ratio	1.12	1.14	1.11
Mantel-Haenszel X-squared	0.59	0.73	0.52
p-value	0.44	0.39	0.47