Estimating the Causal Effect of Maternal Education on Infant Mortality with DHS Data for Iran

S. Farrokh Mostafavi
Institute for Management and Planning Studies, Tehran, Iran.
(farrokh_mostafavi@hotmail.com)

August 2009

Paper submitted to the XXVI IUSSP International Population Conference, Marrakech 2009

Estimating the Causal Effect of Maternal Education on Infant Mortality with DHS Data for Iran

Most demographic research indicates that there is a strong statistical association between maternal education and infant mortality (Bicego and Boerma 1991; Hobcraft, McDonald, and Rutstein 1984; Mensch, Lentzer, and Preston 1985). These findings have led some researchers (Caldwell 1979; 1994) to conclude that there is a causal relationship between mother's education and child health and mortality. Other researchers (Desai and Alva 1995) reject the existence of a strong causal relationship. They maintain that the nature of disease during the first year of life is such that mothers cannot do much in order to prevent death of their children. Using fixed effect models, Desai and Alva (1995) show that the association between maternal education and infant mortality declines considerably in some countries and disappears in others when they control for socioeconomic status and community characteristics of mothers with such models. The question is important because if there is a causal relationship, then education of mothers can enhance the health of children and reduce infant mortality. This is particularly important for poor households and countries.

Mother's education may influence child health and mortality through different pathways (see, for example, Rosenzweig and Schultz 1982; Caldwell and Caldwell 1993; Hobcraft 1993): (1) Education enhances the acquisition and use of health knowledge. (2) Education enhances the use of health services. (3) Mother's education increases family resources, either through their own work or that of their husband, which in turn affect the health of family members. (4) Education affects preferences for child health and family size.

Generally, mothers with higher education also have higher incomes or marry husbands who have higher incomes. They also tend to live in rich communities with better access to health services (Palloni 1981). Consequently, when socioeconomic and community characteristics of mothers are controlled, the associations between maternal education and child health and mortality are attenuated, sometimes to negligible levels

(Hobcraft 1993; Desai and Alva 1998; Kravdal 2004). However, as Hobcraft (1993) points out, these socioeconomic and community characteristics may be capturing the pathways through which maternal education influences child health and mortality (the use of health services is one such pathway). Research also indicates that mother's education is less important for infants than for children (Cleland and van Ginneken 1988).

Although the controversy on this subject centers on the existence of a causal relationship between maternal education and infant mortality, researchers have failed to use the extensive literature on causal analysis. In this paper, I outline the framework for causal analysis and use it in order to obtain estimates of the causal effect of maternal education on infant mortality with Demographic and Health Survey (DHS) data for Iran. First, I use ordinary logistic regression and interpret the estimated coefficient of maternal education within the context of the literature on causal analysis. I also use matching method to obtain a semiparametric estimate of the effect.

Estimating the causal effect of maternal education on infant mortality

Let Y represent the survival status of the child at the end of his or her first year of birth, with Y=1 if the child is still alive and Y=0 if he or she is dead. Let D represent the educational status of mother. I consider D as a dichotomous variable with the value 1 representing educated mothers and the value 0 representing uneducated mothers. Y_D is the child survival status for the mother with educational status D. Thus, Y_1 is child survival status for educated mothers and Y_0 is child survival status for uneducated mothers. The causal effect of maternal education on child survival (infant mortality) is

$\Delta = Y_1 - Y_0$.

The fundamental problem of causal inference (Holland 1986) is that it is impossible to *observe* the value of Y_1 and Y_0 on the same unit and, therefore, it is impossible to *observe* the effect of D = 1 (education) on all mothers. In contrast to experimental studies, we cannot change the value of D (mother's education) and observe the outcome Y_D (child survival status). In observational studies, D is determined by a process outside the control of observer. If Y_1 is observed, Y_0 is the counterfactual unobserved outcome; if Y_0 is observed, Y_1, is the counterfactual outcome. Hence the causal effect Δ for any child is unobservable.

The average effect of maternal education on child survival is

$E(\Delta) = E(Y_1 - Y_0) = E(Y_1) – E(Y_0)$,

where E(.) is the expected value of the random variable. Since Y_1 and Y_0 are partially observed (on mutually exclusive subsets of the population), $E(\Delta)$ cannot be calculated. However, under certain conditions, we may have the following equalities:

$E(Y_1) = E(Y_1 | D = 1)$ and $E(Y_0) = E(Y_0 | D = 0)$. (1)

That is, the mean infant mortality for all children when they all have educated mothers is equal to the mean infant mortality for the subgroup of children in the population who have educated mothers. Similarly, the mean infant mortality for all children when they all have uneducated mothers is equal to the mean infant mortality for the subgroup of children in the population who have uneducated mothers. If the equalities in (1) are satisfied, we may use the difference

$E(Y_1 | D = 1) - E(Y_0 | D = 0)$

to estimate $E(\Delta)$. $E(Y_1 | D = 1)$ and $E(Y_0 | D = 0)$ can be estimated by the mean of Y for the educated and uneducated mothers in the sample, respectively. The important question is under what conditions the equalities in (1) are satisfied. For observational studies, a general sufficient condition for the equalities (1) is ignorability, which holds if

$(Y_1, Y_0) \perp D$.

That is, if Y_1 and Y_0 are independent of D ($\perp$ indicates statistical independence). In other words, the potential outcomes are independent of the assignment of children to the two groups with educated and uneducated mothers. Ignorability may hold within certain subgroups of the population. That is,

$(Y_1, Y_0) \perp D | \mathbf{X}$

Where X is a certain variable defined on the population.

Heckman (1992) has argued that for many policy questions the parameter of interest is average treatment effect for the treated (TT). In deciding whether a policy is beneficial, we are interested in whether on average it is beneficial for those individuals who are assigned to the treatment (not for all individuals). In the case of child mortality, we are interested in whether education enhances child health for educated mothers. In this case, the policy-relevant effect is

$TT = E(\Delta \mid D = 1) = E(Y_1 - Y_0 \mid D = 1) = E(Y_1 \mid D = 1) – E(Y_0 \mid D = 1).$

Estimation of TT requires a weaker ignorability assumption, namely

$Y_0 \perp D.$

This assumption will be satisfied if child health outcomes for educated mothers, if they were uneducated were the same as child health outcomes for uneducated mothers in the population. Again, the relationship can be conditioned on a set of covariates **X**.

In regression models, the counterfactual model is:

$Y_D = \beta_0 + \Delta D + \beta \mathbf{X} + e_D$

Where Y_D is infant mortality rate for every values of mother's educational level D, factual or counterfactual, X is a matrix of control variables with the corresponding coefficients β, and e_D is an error with mean 0. Δ is the average causal effect of mother's education (D) on infant mortality (Y_D).

In practice, we use the standard regression equation

$Y = \beta_0 + \beta_1 D + \beta_2 \mathbf{X} + \varepsilon$

This equation only pertains to the one value of Y and D that is observed for each individual in the data.

A sufficient condition for the regression equation parameter β_1 to become equal to the average causal effect Δ is that ignorability holds; that is Y_D and D should be independent of each other.

One of the ways for estimating the causal effect is through matching. If ignorability holds within certain subgroups of population, say **Z**, then one can estimate the causal effects within these subgroups and calculate the causal effect for the population from these estimates. However, if the number of variables in **Z** is large, we will encounter a dimensionality problem. As a solution, Rosenbaum and Rubin (1983) propose matching on propensity score, which is defined as

$P(D = 1 \mid \mathbf{Z}) = P(\mathbf{Z})$.

That is, propensity score is the probability of selection into treatment group (being educated) conditional on the set of covariates. Rosenbaum and Rubin (1983) prove that if ignorability holds within subgroups **Z** of the population, then

$(Y_1, Y_0) \perp D \mid P(\mathbf{Z})$.

Matching then can be performed on P(**Z**).

The setting: The distribution of health services in Iran

According to the 1996 census, Iran had a population of about 60 million persons, with over 38 percent (about 23 million persons) living in rural areas. In 2000, the country was divided into 28 provinces called Ostans and 282 sub provinces or districts called Sharestans (Statistical Center of Iran 2004). The government provides health service to the population through its Primary Health Care (PHC) services and public hospitals. A strong private sector also provides health services to the population of urban areas.

District health network is the smallest independent unit of Iranian health system (see Shadpour 2001; Naghavi et al. 2004 for the material presented in this section). Each district provides health services in three levels. The first level includes Health Houses in rural areas and Health Stations in cities. The main duties of the first level include provision of health education, family health care, disease control services, and environmental health activities. The second level of district health network includes rural and urban Health Centers. Health Centers support and supervise the activities of Health Houses and Health Stations. They also accept referred cases from the latter and can refer cases in need of more advanced services to higher levels (the district hospital). Other activities of the Health Centers include provision of outpatient care, basic environmental sanitation (including water testing), training, basic laboratory tests, and monitoring of environmental health in schools and workplaces. The third level of district health network includes district Health Center and district hospital. The district Health Center supports and supervises the activities of rural and urban Health Centers, Health Houses, and Health Stations. The district hospital accepts referrals from Health Centers into its specialty clinics, emergency department, and inpatient services. In 2003, there were 16,380 Health Houses, 2300 rural Health Centers, 3530 Health Stations, 2180 urban Health Centers, and 315 district Health Centers in Iran.

In terms of access to health and medical services, Iranian villages are divided into three groups: the main, satellite and mobile villages. The main villages have Health Houses that provide health services to the population of these villages. The satellite villages are located near the main villages within a distance of less than one hour's walk. The population of the satellite villages obtains their required health services from the Health Houses of the main villages. Health workers of Health Houses keep health records of households in the satellite villages and visit them once every week and provide them with necessary services. Compared to the main villages, the satellite villages generally have less population, are poorer, and have less welfare and infrastructure facilities. In 2003, about 24 percent of the population covered by Health Houses (over 4.8 million persons) were living in the satellite villages. The mobile villages, due to their small size and difficulty of access, do not have any established Health Houses. They are also far away from the main villages so that their populations cannot have access to the Health Houses of these villages. Mobile groups of health workers from the nearest Health Centers visit these villages within a period of 15 days or one month and provide health services to their population. These services include vaccination, identification of infected people for local diseases such as malaria, maternal and child care services, and family planning. In 2003, about 14 percent of the rural population in Iran resided in the mobile villages.

The variation in socioeconomic status and access to health services across the country offers an opportunity to control the effects of these variables.

Data and Method

Data

Data used are from Iran Demographic and Health Survey (DHS) that were collected in 2000 by the Ministry of Health and Medical Education in collaboration with the

Statistical Center of Iran. The population under study included all the members of settled households in Iran. Unsettled households, which constituted about 0.35 percent of the population in 1996 (Statistical Center of Iran 2004), were excluded from the study. In order to select the sample, each of the 28 provinces in the country, except Tehran, was divided into two domains of study: urban and rural areas. The province of Tehran was divided into three domains of study: the city of Tehran, other urban areas, and rural areas. In this way, the whole country was divided into 57 domains of study (28 rural areas and 29 urban areas), and 2000 households were selected from each domain. Systematic cluster sampling was used to select the sample. From each domain of study 200 clusters were selected, with each cluster containing 10 households. The original plan was to select 114,000 households in the country. In practice, a few households were not accessible, and the total number of households selected was 113,957. For each household selected, information on children born alive for married women in the age group 10-49 was collected (Ministry of Health, Treatment and Medical Education 2000). In this study, the analysis is restricted to those children who were born one to ten years before the survey, that is, during the period 1990-99. The total number of these children is 100200.

Method

Multivariate analysis

I Use logistic regression in order to estimate the effect of maternal education on infant mortality. Certain characteristics of mothers could be correlated with both her education and infant mortality and need to be included in the models. These include mother's biological and psychological characteristics, preferences, socioeconomic status (SES), and access to health services. Generally educated mothers have higher SES due to their own or their husband's higher income. Educated women are more likely to live in communities with access to health services, mainly due to their higher SES, but also due to their political activities (Palloni 1981) or health seeking behavior. Some biological and psychological characteristics of mothers may influence both their acquisition of cognitive skills and the health of their children (either through genetic transmission or their nursing practices). All these variables are unobserved.

I try to control for mother's socioeconomic status and access to health services through the variables place of residence, region, and household characteristics. We do not have any proxies for biological and psychological characteristics of mother and her values and beliefs. The variables place of residence and region registers, to some extent, variations in values and beliefs across the country. Lack of such controls is one of the main limitations of this analysis.

The categories of places of residence (urban areas, main villages, satellite villages, and mobile villages) indicate differential access to health services. Urban areas have better access to health services and the mobile villages do not have access to any established health services in their community. The main and the satellite villages lie in between. I use place of residence in order to control for access to health services. Of course, as we noticed, mobile villages receive some health services through mobile health workers. The variable place of residence also indicates other community and socioeconomic differences. As we move through the categories urban areas, main, satellite, and mobile villages, the socioeconomic status of households decline.

The 28 provinces in Iran have different cultural and socioeconomic characteristics. Hence, a control for region (Provinces) can also be considered as a control for some of these characteristics. I use household characteristics in order to control for the socioeconomic status of households.

As the description of the sample indicates, sample elements do not have equal probabilities of selection. Hence, data are weighted. The weight for each child is *k/p,* where *p* is the probability of selection for the child and *k* is a constant (Kalton 1983). The probability of selection for any child in a domain of study is *n/N*, where *N* is the number of households in the domain of study and *n* is the selected number of households in that domain of study. The constant *k* is calculated in such a way that the sum of the weights for the total sample is equal to the sample size.

Variables

I code mother's education into educated and uneducated (reference category). The category educated includes all levels of formal schooling (primary or higher), and informal education, religious studies, and literacy crusade. Mother's characteristics that are controlled are age at birth of the child (in years) and her migration status, which is coded as migrated during the period five years before the survey, and not migrated during this period (reference category). Child's characteristics that are controlled are parity, sex, and the incidence of multiple births. Parity refers to the birth order of the child. Sex is coded as female and male (reference category). Multiple births are coded as the child belongs to a multiple birth, or not (reference category). Each province is represented by a dummy variable, with the category for the city of Tehran as the reference group.

The DHS survey does not have any information on income or the socioeconomic status of households. I use household characteristics in order to control for the socio-economic status of the mothers. The household characteristics used are access to sanitary toilet, refrigerator, telephone, television, motorcycle, and bicycle. Each of these variables is coded as have access, or do not have access (reference category). There were other household characteristics in the data that were excluded because they were either insignificant or had small effects. The inclusion of the excluded characteristics in the models does not change the conclusions of the study.

Matching

I use matching on propensity scores in order to estimate the effect of mother's education on her infant mortality. We need determinants of selection of mothers into the two educational groups. The main determinants of mother's education are the socioeconomic status of her parents, their education, and access to educational services. DHS data for Iran do not provide any information about these variables. I use place of residence and region (provinces) as proxies for both access to educational facilities and the socioeconomic status of mothers' parents. These proxies can be justified by the fact that socioeconomic status of households vary by their place and region of residence and access to educational services also varies by place of residence.

Results

Descriptive statistics

Table 1 presents descriptive statistics for the variables used in the study. Over 64 percent of children had mothers with some education, about 54 percent lived in households located in urban areas, over 40 percent lived in the main and satellite villages and only 5 percent lived in the mobile villages. Most children (88.5 percent) lived in households that had access to refrigerator, but relatively few lived in households with access to telephone and sanitary toilet (37.8 and 28.3 percents, respectively). About 3.7 percent of children died during the period of study.

Table 1 Percentage distribution of Iranian children born during 1990-99, by some maternal, child, and household characteristics

Variable	Percent
Place of residence	
Urban areas	54.4
Main villages	31.3
Satellite villages	9.3
Mobile villages	5.0
Mother's education	
No education	34.7
Incomplete primary	20.1
Completed primary	20.3
Incomplete high school	11.1
High school diploma or above	13.1

Age at birth	26.1*
Migrated	12.5
Parity	3.4*
Sex (Female)	48.7
Multiple birth	2.1
Access to sanitary toilet	28.3
Access to refrigerator	88.5
Access to telephone	37.8
Access to television	86.5
Access to motorcycle	17.1
Access to bicycle	24.9
Survival status (dead)	3.7

Data are weighted. For some variables, the given percentages do not add to 100. The remaining percentages are missing values.

* The numbers for age at birth and parity are means.

Source: Children's data from Iran Demographic and Health Survey 2000.

Table 2 presents household characteristics of children in urban areas, and the main, satellite, and mobile villages. As can be seen, there is a considerable difference in the socioeconomic characteristics of these regions, as indicated by their access to household amenities. For example, over 43 percent of children in urban areas lived in households with access to sanitary toilet. The corresponding figures for the main, satellite, and mobile villages are about 12.3, 6.2, and 3.1 percents, respectively. Over 95 percent of children in urban areas lived in households that had access to refrigerator, compared to 48 percent in the mobile villages. It is clear that the socioeconomic status of households declines as we move through the urban areas, the main, satellite, and mobile villages.

Table 2 Percentage distribution of Iranian children born during 1990-99, by household characteristics and place of residence

Variable	Urban areas	Main villages	Satellite villages	Mobile villages
Household characteristics (Access to)				
Sanitary toilet	43.5	12.3	6.2	3.1
Bath	86.0	49.5	33.1	21.5
Refrigerator	95.4	86.0	77.8	48.1
Television	94.6	83.5	69.4	48.3
Telephone	54.4	23.8	7.2	2.3
Car	25.9	16.7	13.9	10.9
Motorcycle	16.2	19.7	15.5	14.2
Bicycle	31.7	19.8	12.0	7.1

Data are weighted. For some variables, the given percentages do not add to 100. The remaining percentages are missing values.

Source: Children's data from Iran Demographic and Health Survey 2000.

Regression

Table 3 shows the results of logistic regression. In order to interpret the estimated coefficient of maternal education as a causal effect, we need to assume the ignorability condition. The strong version of ignorability requires Y_1 and Y_0 to be independent of selection of mothers into the educated and uneducated groups. This means that if uneducated mothers were educated, their child health outcomes would be similar to the outcomes for educated mothers in the population. Such an assumption, however, could be challenged. Consequently, it is better to invoke the weaker assumption of ignorabilty, which only requires that Y_0 be independent of selection of mothers into the two educational groups. This implies that child health outcomes for educated mothers if they were uneducated would be similar to child health outcomes for uneducated mothers in the population. Of course, this assumption may also be challenged. One may argue that educated mothers are cleverer and they would have had lower infant mortality had they been not educated. Given our definition of the education variable as a dummy, such an argument is hard to defend. There is almost no self-selection of mothers into the two educational groups. Generally, parents prefer to send their children to school, in particular primary schools. The decision to send children to school is made by parents and not children and the factors that affect parents' decisions are mainly economic. It is also possible that there is lack of school in the community. This case, however, is also related to the socioeconomic status of the family who are unable to move to communities with schools for their children. Hence, there is no self selection and smarter women cannot select themselves into the two educational groups. There is, however, one group of women who may decide to get educated in order to benefit the cognitive skills acquired. This is the group of women who were uneducated during their childhood and decided to get educated when they were adults. Hence, it is better to exclude these women from analysis in order to justify the ignorability assumption. If we accept the weak version of ignorability, we have to interpret the estimated coefficient as the effect of education on the infant mortality of the educated mothers.

Regression estimates are obtained for the following data sets: (1) the "educated" category of the variable maternal education includes all the educational categories; (2) the categories "literary crusade" (basically adult education) and "religious studies" (mainly traditional education) are excluded; (3) in addition to the previous two categories, "incomplete primary and informal education" (which may include adult education) is also excluded. The exclusions in the second and third data sets are employed in order to rule out self-selection of mothers into the "educated" category of the maternal education variable. In the original DHS data set incomplete primary and informal education are included in one category. Hence, in order to exclude informal education, the category incomplete primary should also be excluded. For the third data set, therefore, infant mortality of educated mothers is compared to the infant mortality of mothers who have at least completed primary education. For each of these data sets another data file is constructed that includes only the first births of mothers. These data sets are used in order to overcome the clustering problems due to the inclusion of all children of households in the sample.

For each data set five models are estimated. Model 1 only includes the maternal education variable. In model 2, mother's characteristics (age, age squared, and migration status) and child's characteristics (parity, sex, and multiple birth) were added. Model 3 adds place of residence (urban areas as the reference category, main, satellite and mobile villages), model 4 adds the variables for provinces, and model 5 adds the variables indicating socioeconomic status (access to sanitary toilet, telephone, television, refrigerator, motorcycle, and bicycle).

Table 3 presents the estimated logistic regression coefficients of maternal education for all the data sets and models. All these data sets are weighted. For comparison, estimates for unweighted data sets were also obtained (table 4). As can be seen (table 3), with the addition of variables representing community characteristics, socioeconomic status, and cultural differences in models 3 to 5, the absolute values of the coefficients of maternal education decline considerably. For the data sets with first births the amount of decline is larger. In fact, in the final model 5, the coefficients of maternal education become almost insignificant. Given that we are using proxies for community characteristics, socioeconomic status, and cultural differences, if we were able to measure these variables and include them in our models, we would have seen further reduction in the size of coefficients and more probably insignificant effects. These results are consistent with the fixed effect estimates of the coefficients of maternal education (Desai and Alva 1995).

Table 3 Logistic regression coefficients of maternal education on infant mortality for various models, Iranian children born during 1990-99; Data are weighted.

Sample	Model 1	Model 2	Model 3	Model 4	Model 5
All educational categories, all births	-0.791 (0.034)	-0.601 (0.039)	-0.494 (0.041)	-0.425 (0.042)	-0.355 (0.044)
First births	-0.768 (0.080)	-0.744 (0.081)	-0.484 (0.086)	-0.363 (0.090)	-0.219* (0.094)
Literary crusade and	-0.838	-0.637	-0.521	-0.446	-0.366

religious studies excluded, all births	(0.036)	(0.042)	(0.044)	(0.046)	(0.048)
First births	-0.816	-0.795	-0.509	-0.376	-0.209*
	(0.082)	(0.083)	(0.089)	(0.093)	(0.099)
Literary crusade, religious studies, informal and incomplete primary excluded, all births	-0.929 (0.039)	-0.709 (0.046)	-0.595 (0.050)	-0.526 (0.051)	-0.441 (0.054)
First births	-0.926	-0.916	-0.628	-0.512	-0.332**
	(0.085)	(0.086)	(0.095)	(0.100)	(0.108)

Standard errors appear in parentheses.

* p<0.05 **p<0.01 no star p<0.001

Source: Children's data from Iran Demographic and Health Survey 2000.

Table 4 Logistic regression coefficients of maternal education on infant mortality for various models, Iranian children born during 1990-99; Data are not weighted.

Sample	Model 1	Model 2	Model 3	Model 4	Model 5
All educational categories, all births	-0.647 (0.033)	-0.503 (0.037)	-0.412 (0.039)	-0.343 (0.040)	-0.275 (0.041)
First births	-0.721	-0.700	-0.513	-0.430	-0.333
	(0.075)	(0.076)	(0.080)	(0.084)	(0.088)
Literary crusade and religious studies excluded, all births	-0.682 (0.035)	-0.533 (0.040)	-0.427 (0.042)	-0.351 (0.044)	-0.273 (0.046)

First births	-0.756	-0.742	-0.532	-0.445	-0.336
	(0.076)	(0.077)	(0.083)	(0.087)	(0.092)
Literary crusade, religious studies, informal and incomplete primary excluded, all births	-0.787 (0.039)	-0.802 (0.042)	-0.663 (0.046)	-0.580 (0.048)	-0.492 (0.051)
First births	-0.869	-0.862	-0.662	-0.596	-0.473
	(0.080)	(0.082)	(0.089)	(0.095)	(0.102)

Standard errors appear in parentheses.

* p<0.05 **p<0.01 no star p<0.001

Source: Children's data from Iran Demographic and Health Survey 2000.

Maternal education influences child health and infant mortality through different pathways, including the use of health services, the use of health knowledge, and increase in income. Consequently, when we control community characteristics and socioeconomic status through the proxies for these variables (as in models 3 to 5), the estimated coefficients represent only the partial effects of maternal education on infant mortality. The effect due to the use of health services and income is, to some extent, excluded. This argument is also true for fixed effect estimates (as in Desai and Alva (1995); see also Kravdal (2004) for a similar view). Hence Desai and Alva's assertion that maternal education does not have a strong effect on infant mortality cannot be defended because their models do not estimate the whole effect of maternal education. The estimated effects in such models, as here, is only the partial effects that can be attributed mainly to the use of knowledge.

I have interpreted the estimated coefficients as the partial effect of maternal education on the infant mortality of the educated mothers. The assumption behind this interpretation is that, due to the nature of data sets used, the definition of the education variable and the controls in our models, the condition of weak ignorability holds. This assumption, however, may be challenged. It may be argued that due to some selection mechanism (such as selection due to better socioeconomic status), educated mothers, if they were uneducated, would have lower infant mortality than uneducated mothers in the population. If this is true (which may not), then their infant mortality would be lower than the infant mortality of uneducated mothers and the estimated coefficients underestimate the actual partial effect of education on infant mortality of educated mothers. Hence, the true partial TT effect should be at least as large as estimated here, given our control variables.

Matching

In the population, infant mortality rates for the educated and uneducated mothers are 26 and 56 deaths per thousand live births, respectively. This gives an average causal effect of -.030, implying education reduces infant mortality rate by about 30 deaths per thousand live births.

I used logistic regression to obtain estimates for propensity scores, with the variables place of residence and provinces. Pseudo R squared measures for the model ranged from .182 (Cox and Snell R square) to .250 (Nagelkerke R square), which were surprisingly high. (The results of matching to be added).

Summary and Conclusion

Demographic research indicates a strong statistical association between maternal education and infant mortality. An important question is whether the association also indicates the existence of a strong causal relationship or that it is a spurious association. Fixed effect models show that the association attenuates considerably for some countries and disappears altogether for others (Desai and Alva 1995).

In this paper, I use data from Iran Demographic and Health Survey (DHS) and try to estimate the causal effect of maternal education on infant mortality within a counterfactual framework. In order to obtain a causal estimate the condition for ignorabilty should hold. I use logistic regression and interpret the estimated coefficient within the framework of counterfactual causal analysis. Invoking a weak assumption of ignorability, I suggest that the coefficient of mothers' education represents the effect of education on the infant mortality of the educated mothers. The use of a dummy for the variable education helps to justify this assumption. I use place of residence, region (provinces), and household characteristics as proxies for access to health services, and socioeconomic status of mothers. The distribution of health services in Iran provides an opportunity to control for this important community characteristic affecting child health. The results of analysis indicate that the introduction of variables representing socioeconomic and community characteristics attenuates the effect of maternal education considerably. In fact, in the data set for first births, the coefficient of maternal education drops considerably and approaches insignificance. The estimated coefficients represent only the partial effect of mother's education on their infant mortality. The effect through the use of health services and increased income has been, to some extent, excluded. It appears that the partial effect of education on infant mortality of educated mothers is small. However, the effect of education through other pathways (use of health services and increased income) could be large. Given that I have used proxies for the socioeconomic and community characteristics, the true effects should be less than those estimated here. If the condition of ignorability does not hold, however, then the estimated coefficients underestimate the effect of education on the infant mortality of educated mothers.

In addition, I use matching on propensity scores in order to estimate the effect of maternal education on infant mortality. In order to estimate propensity scores, I use place and region of residence as proxies for the socioeconomic status of mother's parents. (Results not available yet).

References

Bicego, G.T. and J.T. Boerma. 1991. Maternal education and child survival: A comparative analysis of DHS data, in *Proceedings of the Demographic and Health Surveys World Congress*, Vol. 1, IRD/ Macro International, Inc., Columbia, MD, pp. 177-204.

Caldwell, J.C. 1979. Education as a factor in mortality decline: An examination of Nigerian data. *Population Studies* 33: 395-413.

Caldwell J.C. 1994. How is greater maternal education translated into lower child mortality? *Health Transition Review* 4: 224-29.

Caldwell J.C. and P. Caldwell. 1993. Women's position and child morbidity and mortality in less developed countries, in N. Federici, K.O.Mason, and S. Sogner (eds.), *Women's Position and Demographic Change*, Clarendon Press, Oxford, pp. 122-39.

Cleland, J. 1990. Maternal education and child survival: Further evidence and explanations, in J.C. Caldwell, S. Findley, P. Caldwell, G. Santow, J. Braid, and D. Broers-Freeman (eds.), *What We Know about the Health Transition: The Cultural, Social and Behavioral Determinants of Health*, Health Transition Center, Australian National University, Canberra, pp. 399-419.

Cleland, J.C. and J.K. van Ginneken. 1988. Maternal education and child survival in developing countries: The search for pathways of influence. *Social Science and Medicine* 27: 1357-1368.

Desai, S. and S. Alva. 1998. Maternal education and child health: Is there a strong causal relationship? *Demography* 35: 71-81.

Heckman, J. J. 1992. Randomization and social policy evaluation. In Manski, C.F. and I Garfinkel, eds., *Evaluating Welfare and Training Programs*, Cambridge, MA: Harvard University Press, 1992, pp. 201-30.

Hobcraft, J. 1993. Women's education, child welfare and child survival: A review of the evidence. *Health Transition Review* 3: 159-175.

Hobcraft, J.N., J.W. McDonald, and S.O. Rutstein. 1984. Socioeconomic factors in infant and child mortality: A cross-national comparison. *Population Studies* 38: 193-223.

Holland, P.W.1986. Statistics and causal inference. *Journal of American Statistical Association*, 81: 945-70.

Kalton, G. 1983. *Introduction to Survey Sampling*. Sage Publications, Newbury Park and London.

Kravdal, O. 2004. Child mortality in India: The community-level effect of education. *Population Studies*, 58: 177-192.

Mensch, B., H. Lentzner, and S.H. Preston. 1985. *Socioeconomic Differentials in Child Mortality in Developing Countries*. Department of International Economic and Social Affairs, United Nations, New York.

Ministry of Health, Treatment and Medical Education. 2000. *Seemaye Jamiyat va Salamat dar jomhorie Islamie Iran, Mehr 1379 (Iran Demographic and Health Survey, DHS 2000)*. Ministry of Health, Treatment and Medical Education, Tehran [in Persian].

Naghavi, M., N. Jafari, E. Jamshid Beiggi, S.H. Vasegh, A.M. Azad, and M.E. Akbari. 2005. *Degargoonie Symaie Salamat Dar Roosta Neshinan Iran, 1372-1382 [The Transformation of the Health profile of Iranian Rural Population, 1993-2003]*. Ministry of Health, Treatment and Medical Education, Tehran [in Persian].

Palloni, A. 1981. Mortality in Latin America: Emerging patterns. *Population and Development Review* 7: 623-49.

Rosenbaum, P.R. and D.B. Rubin. 1983. The central role of propensity score in observational studies for causal effects. *Biometrika*, 76:41-55.

Rosenzweig, M.R. and T.P. Schultz. 1982. Child mortality and fertility in Columbia: individual and Community effects. *Health Policy and Education* 2: 305-348.

Shadpour, K. 2001. Primary health care system of the Islamic Republic of Iran, in A.H. Mehryar, *Proceedings of the First International Workshop on Integrated Approach to Reproductive Health and Family Planning in the Islamic Republic of Iran, Isfahan, 31 April-13 May, 2000*, Institute for Research on Planning and Development, Tehran, pp. 239-269.

Statistical Center of Iran. 2004. *Iran Statistical Yearbook 1382 [March 2003 – March 2004]*. Statistical Center of Iran, Tehran.

Winship, C. and M. Sobel. 2004. Causal inference in sociological studies, in M. Hardy and A. Bryman (eds.), *Handbook of Data Analysis*, Sage Publications, London.