Maternal Literacy and Numeracy Skills and Child Health in Ghana†

Niels-Hugo Blunch

Department of Economics

The George Washington University

Washington, D.C., USA

nblunch@gwu.edu

PRELIMINARY RESULTS
July 10, 2004

Preliminary results for paper to be presented at *Northeast Universities Development Consortium Conference*, HEC Montréal, October 1-3, 2004.

.

Abstract:
This paper examines the impact of mothers' skills and schooling on the production of children's health in Ghana. The analysis considers intermediate outcomes including pre- and post-natal care and vaccinations, and final outcomes, including illnesses and mortality. Previous studies of the determinants of child health have mostly been limited to investigating the impact of maternal schooling only and, as a consequence, largely have not considered skills and also have ignored alternative routes to acquiring skills, such as adult literacy programs. Analyzing a recent household survey for Ghana, this paper addresses both of these issues. Preliminary results for a specification where all regressors are treated as predetermined indicate that skills are largely not important once education is controlled for but at the same time also indicate a positive association between adult literacy course participation and child health. The latter points towards the potentially important role of adult literacy programs in promoting child health, something which has previously received little to no attention in the economics literature.

† I am grateful to Donald Parsons and David Ribar for helpful comments and suggestions. Remaining errors and omissions are my own. The data were kindly provided by the Ghana Statistical Service. The findings and interpretations, however, are those of the author and should not be attributed to the Ghana Statistical Service.

"The education of parents, notably that of the mother, appears to be an omnibus. It affects the choice of mates in marriage. It may affect the parents' preferences for children. It assuredly affects the earnings of women who enter the labor force. It evidently affects the productivity of mothers in the work they perform in the household, including the rearing of their children. It probably affects the incidence of child mortality, and it undoubtedly affects the ability of parents to control the number of births."

Theodore W. Schultz (Schultz, 1973)

1. Introduction

One of the strongest and most consistent findings in development, health and labor economics is the positive relationship between schooling and well-being. This empirical relationship has been confirmed in numerous studies across different time periods, countries and measures of well-being. These studies generally treat education as a "black box", however. What is measured is not what a person has learned in terms of skills such as for example literacy and numeracy but rather what level or grade has been completed. Two main issues are involved here. First, the link between schooling and well-being really goes from schooling to skills to productivity to well-being. As the link between schooling and skills is more tenuous in developing countries due to often poor school quality, it is imperative that this part of the process receive particular attention in empirical analyses in this context. Second, policies focusing on education rather than on skills might be misdirected. With multiple paths to achieving skills (including formal education and adult literacy programs) and with limited public budgets, cost-effectiveness of programs is essential.

In response to these issues, I suggest that literacy, numeracy and other skills be viewed as intermediate outputs in a production process where the main inputs are formal (child) schooling and non-formal (adult) literacy course attendance. Subsequently, literacy, numeracy and other skills enter as inputs in a production process to generate the final outputs of well-being, including health outcomes.

Building on the above sketched two-pronged production process, this paper examines the relationship between mothers' literacy and numeracy skills, formal education and adult literacy course participation and child health in Ghana. The health measures examined include child vaccinations, child mortality, pre- and postnatal care and child morbidity. The contribution of this paper includes (1) analyzing the impact on child health from skills, including reading and writing skills and that for both English and indigenous languages, as well as numeracy and other skills and (2) including adult literacy course participation as a pathway of achieving skills, two issues which have not been addressed in the previous literature. It considers how these skills affect the production of health, including mothers' pre-natal care and children's vaccinations and post-natal care, and how they affect outputs, including children's illnesses and mortality.

The rest of this paper is structured as follows. The next section presents the conceptual framework of this paper, while section three discusses estimation strategies and related issues. Section four presents the data, discusses sample restrictions and also provides preliminary, descriptive analyses of the interlinkages of mother's literacy and numeracy skills, formal schooling and adult literacy course participation in Ghana. The multivariate econometric analyses follow in section five, while section six concludes and provides directions for further research.

2. Conceptual Framework

The inter-linkages between skills and child health are examined in the context of Grossman's (1972) health production model. In the original model, an individual maximizes utility with respect to his/her own health and consumption. I extend the model by letting the mother also obtain utility from child health and by allowing the human capital effects to come from a set of individual skills, rather than from education per se. Further, the skills effects run from the mother's skills to the child's health. While this model might be posed entirely in terms of a verbal description, a mathematical representation helps highlight some important issues and the latter will therefore be pursued in the following.

Specifically, I consider a two-person household consisting of a mother and a child in which the mother has preferences over the child's health (Z_1) and other commodities (Z_2). The utility of final goods is affected by three types of preference shifters: human capital skills, S, observed family background including needs or fertility, ethnicity/tribal association, B, and unobserved characteristics including tastes, δ, giving rise to the following utility function:

$$U = u(Z_1, Z_2; S, B, \delta) \qquad (2.1)$$

The utility function is assumed to be well-behaved that is, the marginal utility is strictly increasing at a decreasing rate for both goods, so that $u_{Z_1} > 0, u_{Z_2} > 0, u_{Z_1Z_1} < 0, u_{Z_2Z_2} < 0$.

The household's utility maximization is subject to three types of constraints: technological, budget and time constraints. First, the technological constrains are given by the two production functions f_1 and f_2, which give output of child health and all other goods as functions of their respective inputs of a market good (X) and mothers' time (T):

$$Z_1 = f_1(X_1, T_1; S, \eta, C) \qquad (2.2)$$

$$Z_2 = f_2(X_2, T_2; S) \qquad (2.3)$$

In (2.2.) and (2.3) mothers' human capital skills act as a conditioning variable by increasing output for given inputs[1]. In the case of child health, there are two additional conditioning

[1] To simplify the discussion, skills are not modeled explicitly here. Following Blunch (2004), one might imagine skills being produced from time in child schooling, T_1, time attending adult literacy classes, T_2, the quality of these two types of education, Q_1 and Q_2, conditioned by a taste shifter, φ, capturing different tastes for education due to for example religion, ethnicity and/or cultural norms and traditions in the community: $S = s(T_1, T_2, Q_1, Q_2, \varphi)$.

variables: η, which is the (unobserved) initial child health endowment and C, which includes health infrastructure, treatment practices, and disease environment of the community.

The household's budget constraint defines the consumption frontier of the household as a function of its potential income sources. Specifically, the household may obtain income from engaging in labor activities, supplying H amounts of labor at the rate W, which is affected by the vector of human capital skills, S:

$$W(S)H \geq P_1X_1 + P_2X_2 \tag{2.4}$$

Lastly, the maximization of (2.1) is also subject to a time-constraint:

$$T_1 + T_2 + H = K \tag{2.5}$$

K in (2.6) is the maximum time available for home-production and market work after accounting for time to eat and sleep, say, 16 hours a day (alternatively, it could be normalized to one).

The problem of the mother, therefore, is to maximize (2.1) with respect to T_1, T_2, X_1, X_2 and H subject to the constraints (2.2)-(2.5), that is to decide the amount of time and goods inputs in the production child health and other commodities and the amount of time devoted to market work so as to maximize utility subject to the set of constraints.[2] Solving the model yields a series of market goods demands and production time supply functions. The child health input demand function has our main interest:

$$X_1^* = x_1(W(S), P_1, P_2, S, B, \delta, \eta, D) \tag{2.6}$$

Substituting X_1^* from (2.6) into (2.2) yields the reduced form child health production function:

$$Z_1^* = z_1(W(S), T_1, P_1, P_2, S, B, \delta, \eta, D) \tag{2.7}$$

The child health input demand function (2.6) and the child health production function (2.7) are what will be estimated in the empirical analyses.

[2] The amount of time devoted to market work and one of the market good inputs are redundant due to the linear dependence between these variables.

From this discussion several issues with implications for the empirical analyses come out. First, the reduced form child health input demand and child health production functions (2.6) and (2.7) hint at which variables are important determinants conceptually and therefore at which variables should be included in the empirical analyses. These variables include the mother's skills level, her wage rates as well as prices of health care and time use of health care services/child health production, needs and tastes. Second, the model yield several testable implications (I will return to this after extending the conceptual framework a bit). Third, the reduced form (input) demand for child health market goods (2.6) and the reduced form child health production function (2.7) make clear that human capital skills have both direct and indirect effects on child health input demands and child health production. May it be possible to disentangle the direct and indirect effects empirically? (2.6) and (2.7) hint that it is: inclusion of skills will capture the direct effects, while inclusion of wages will control for indirect effects.

Lastly, the conceptual model outlined above also highlights the importance of unobserved heterogeneity and endogeneity for the subsequent analyses. From the presence of δ (unobserved family characteristics, including tastes) and η (unobserved child health endowment) in both (2.6) and (2.7), unobserved heterogeneity is seen to affect both child health input demand and the production of child health. Further, the issue of endogeneity or simultaneity involved in an examination of determinants of child health input demands and child health outcomes are also apparent from (2.6) and (2.7): child health input demand and child health outcomes, which are chosen by the mother, both depend on skills and wages, which are themselves chosen by the mother, also. In turn, these twin issues of unobserved heterogeneity and endogeneity/simultaneity highlighted by this model are something, which needs to be dealt with in the empirical analyses.

The treatment of human capital as one generic skill, *S*, simplified the presentation for the mathematical model above. At the same time, however, this simplified model helped bring out several key points related to both the theoretical interlinkages of child health inputs and outcomes, mothers' skills and other variables and the subsequent empirical analyses. Conceptually, however, a few extensions seem warranted. First, rather than skills (implicitly) being obtainable via only one route, it is more realistic to consider several routes for achieving skills. In particular, I suggest that skills may be obtained either from formal schooling during childhood or from participation in adult literacy programs during youth or adulthood. Second, rather than just the one composite factor, *S*, human capital skills would seem to consist of several individual components (each of which are obtainable through one or both of the two alternative routes of achieving skills). Since the data further allows discriminating between at least some of these components, it seems fruitful to discuss in more detail the exact channels through which the different components of *S* affect child health outcomes. In so doing I will distinguish between direct effects on the mother's home productivity (working through her production function (2.2)) and other indirect skills effects.

Starting with the direct skills effects, there are several reasons why skills might affect the mother's home productivity of child health. First, the production of child health depends crucially on literacy and numeracy skills—being able to read and accurately follow prescriptions, for example. Second, health issues play a major role in education, particularly in adult literacy programs (where 10 out of the total 28 topics taught in addition to literacy and numeracy skills include health related issues, for example "Family Planning", "Immunization", "Safe Motherhood and Child Care" and "Safe Drinking Water"). Whereas literacy and numeracy skills may be viewed as more generally applicable skills, the latter may be viewed as a more specialized skill, which mainly affects the home productivity of child health. So how might increased efficiency in the (home) production of child health work its way through this modified Grossman health production model? Initially, (home) production will shift towards the production of child health, assuming that this is a normal good and that it is not relatively "much" more time-intensive than other commodities. At the same time, however, there will be more time available for market work, which will enable the individual to purchase more of the market-good input for production of child health and other commodities. While this effect therefore depends on the relative time and goods intensities of the various commodities, the net effect on the production of child health is most likely positive.

The indirect effects work mainly through the household's consumption possibilities. Most importantly, an individual's wages may increase from participation in schooling activities. This could be due to a direct productivity effect from literacy and numeracy skills or from socialization or discipline skills obtained from childhood schooling. Alternatively, earnings capacity may increase either from credentialism or signaling (Spence, 1973) obtained from childhood schooling or from participation in adult literacy programs, where participants learn about income generating activities and frequently engage in them directly under the direction of the teacher. The increased income potential of the household reduces the need to depend on children as a source of income, thus decreasing child labor. In turn, this will positively affect child health. Again, both substitution and income effects may be operating here—the net effect from these indirect effects, however, is likely to be positive. In addition to affecting the household's consumption possibilities, participation in schooling activities may also affect needs or tastes for child health. First, parents may become aware of the harmful effects of child labor on child schooling and child health. Second, the composition of the consumption basket may shift from predominance of food to include more non-food items, including child health related items.

A few potentially important determinants are not included in the conceptual framework. For example, savings are not allowed. Spill-over effects from having other literates and/or literacy course participants in the household (and/or in the community) are also not allowed. Focusing on the latter, I conjecture that the household decision-maker(s) might invest in education for the oldest daughter, say, in order to release the mother for market work or working at the farm or other household enterprise, while the oldest daughter takes care of younger siblings. A way to incorporate this formally into the model would be to simply include the skills of the oldest daughter in the child health production function, (2.2).

Based on this expanded conceptual framework I will examine the following research questions. First, do literacy and numeracy increase demand for child health inputs and improve final child health outcomes. Second, if so what is the relative efficiency of the different "literacies"? Third, do the literacy/numeracy effect work through wages and fertility and, if so, what is the relative magnitude of the impacts? Fourth, has education any impact on intermediate and final child health outcomes once the impact from literacy and numeracy has been controlled for? Fifth, are there indirect effects on child health outcomes from having literates and/or literacy course participants in the household (and/or in the community)? Sixth, are there differences in skills efficiency related to location?[3]

3. Estimation Strategies and Issues

[3] The motivation behind this research question is the fact that huge posters inform about for example water safety, how to be protected against HIV/AIDS, particularly in urban areas – in English. Hence, it may be conjectured that English is relatively more important (efficient) for transmitting health knowledge in urban areas, while indigenous languages may be as or possibly even more important (efficient) for transmitting health knowledge in rural areas.

From the previous section skills directly affected child health outcomes by increasing the efficiency of household productivity and indirectly affected these outcomes by increasing consumption possibilities and changing tastes. The empirical analysis will rely on linear specifications of the optimal intermediate and final child health outcome equations. These equations are written:

$$X_i = \alpha + \beta_1 S_i + \beta_2 E_{1i} + \beta_3 E_{2i} + \beta_4 W_i + \beta_5 F_i + \beta_6 A_i + \beta_7 P_i + \beta_8 C_i + \varepsilon_i \qquad (3.1)$$

$$Z_i = \alpha + \beta_1 S_i + \beta_2 E_{1i} + \beta_3 E_{2i} + \beta_4 W_i + \beta_5 F_i + \beta_6 A_i + \beta_7 P_i + \beta_8 C_i + \nu_i \qquad (3.2)$$

where X_i is intermediate child health output; Z_i is final child health output; S_i is literacy and numeracy skills; E_{1i} is childhood schooling; E_{2i} is adult literacy course participation; W_i is wages or earnings; F_i is fertility; A_i is access to health facilities (for example health center, midwife, doctor) in the community; P_i is prices on health related goods and services; C_i is a vector of other controls, including age of the mother (to capture experience and time having been of child-bearing age) and the child (to capture child needs and age specific productivity effects in health production), geographical location, and ethnicity and ε_i is an error-term capturing unobservables. (3.1), therefore, is a commodity production function, while (3.2) is a factor or intermediate output demand function.

4. Data and Descriptive Analyses

The Ghana Living Standards Survey (GLSS) is a nationally representative, stratified multi-purpose household survey, carried out in 1987/88, 1988/89, 1991/92 and 1998/99 as four independent cross-section surveys. The most recent round of these (GLSS 4) is used for the analyses in this paper. The household survey contains information on educational attainment, participation in adult literacy courses, literacy and numeracy, as well as information on background variables such as age, gender, tribal association/ethnicity and region, which are also important factors in analyses of human capital processes. In addition to the household survey, each round also includes a community and a price questionnaire. The community questionnaire contains information on access to facilities, including schools, hospitals, markets, roads, public transportation and adult literacy programs. Due to the difficulties involved in defining communities in urban areas the community questionnaire was only administered to rural areas. The most important variables will now be described in turn, starting with the dependent variables.

Child health outcomes

Five different child health outcomes from the GLSS 4 are examined in this paper. They may be classified into intermediate child health outcomes (vaccinations, pre-and post-natal care) and final child health outcomes (child morbidity and mortality). The information is provided by the mother, except for the information on morbidity, which was provided by the head of household or other adult member (which may or may not have been the mother). Since these variables are crucial to subsequent analyses, I will got through them in some detail, starting with the intermediate child health outcomes.

The intermediate child health outcomes are all based on information on whether the service in question was ever received by the child (vaccinations and postanatal care) or mother (prenatal care). The samples subjected to the questions on vaccination and postnatal are children 7 years and below and 5 years and below, respectively, while the question on prenatal care was only given to females, who either were pregnant currently or had been pregnant within the past 12 months (and were between 15 and 49 years old).

One problem with these measures is that it is not known when the service was provided. For vaccinations, for example, it is not known when a vaccination was provided, which type(s) and whether the full series of vaccinations for a given type was given (for example, polio and dpt vaccinations, to be fully effective, each require three consecutive vaccinations). Further, it may be that the child will or will not receive a (the) vaccination(s) in the future. One way to address the latter, though, would be to chose a lower cut-off high enough that it would seem that children not vaccinated at this age would likely never be vaccinated, say 3 years of age. Since this causes a large drop in observations this strategy is not pursued here, however. Similar issues exist for postnatal care, except that the timing problem is reversed: postnatal care seems to be most relevant for younger children, so that here the issue is to chose an appropriate upper cut-off, say, two or three years. While timing problems do not appear substantial for the prenatal care measure, the exclusion of mothers who are not currently pregnant or have not been pregnant within the past 12 months naturally limits the sample size considerably. Additionally, the prenatal and postnatal care measures are potentially riddled with unobserved heterogeneity (or self-selection): mothers who have experienced complications either with their current pregnancy or past pregnancies would seem to be more likely to seek prenatal care for themselves and postnatal care for their children. Vaccinations do not appear to be prone to these issues to the same degree, implying that this is more objective measure of child well-being as related to the child's health.

Moving to the final child health outcomes, the information on morbidity is available for all household members and reveals whether an individual has suffered from an illness or an injury or both during the past 2 weeks. While it therefore confounds factors that may be systematically related to mother's education and literacy and numeracy skills (such as malnutrition or lack of preventive care, such as vaccinations and prenatal care—which are *not* due to lack of resources) with factors that are not (such as accidents), in practice very few (less than one percent) answered "both". The child mortality measures are constructed from the fertility module, which includes information on the number of children ever born and ever died to a woman (15 to 49 years old) but not when, which is unfortunate: it would have been useful to be able to combine the information on child mortality with birth-spacing, since the latter may be an important determinant of child mortality. Additionally, since it is possible that the child(ren) died far back in the past, the explanatory variables, which are current, may be poor predictors as a result. If a mother has recently participated in an adult literacy program, for example, this of course has no impact on the past deaths. Similarly, high child mortality might have induced the mother to participate in the programs in order to be able to prevent future deaths. With all its timing problems, however, it still appears fruitful to exploit this information. I therefore construct three measures of mortality based on this information: the total number of children ever died (preferred measure), as well as a binary measure for whether any children have died to a women and a measure of the share of children died out of the children ever born. The two latter measures are included so as to evaluate the robustness of the results for the preferred measure.

As was also the case for the intermediate child health outcomes, the final child health outcomes are potentially prone to issues of unobserved heterogeneity (self-selection), as well. Again, this may be not be equally relevant for all final child health outcome measures. Child mortality certainly leaves minimal room for subjective assessment, while this is not the case for child illnesses. Complicating the issue of subjectivity related to child illnesses, however, is the possibility that it may go either way: more educated mothers, while being more likely to be able to prevent their children's illnesses, at the same time also would seem to be more likely to be able to properly diagnose their children's symptoms as indeed a disease rather than, say "tiredness" or "laziness".

Literacy and numeracy

The information on literacy skills from the GLSS 4 include Ghanaian reading and writing proficiency and English reading and writing proficiency, while numeracy measures the ability to do written calculations. The question on English reading (writing) skills is: "Can (NAME) read (write) a letter in English?", while the question on Ghanaian reading (writing) skills is: "In what Ghanaian language can (NAME) write a letter?" (stating the one in which (NAME) is most proficient). The question on written calculations is: "Can (NAME) do written calculations?" The respondent to these as to most of the other questions in the survey is "preferably the head of household, if not available, any adult member of the household who is able to give information on the other household members". While this may be an issue for concern, it is hard to correct. Since this is "the" way these types of surveys are typically done, I will therefore follow standard practice and assume that the individual who has answered the questions has sufficient knowledge of household members, i.e. that the data are reliable. Another concern, which may be examined a bit further, is the subjective nature of the literacy and numeracy measures. In order to gain additional insights into this issue and the extent to which it poses any problems in practice educational attainment and skills proficiency for adults is tabulated in Table 5.1.

Table 5.1 Distribution of Self-reported Skills Across Highest Educational Level Completed

	Ghanaian reading	Ghanaian writing	English reading	English writing	Written Calculations
Full sample	0.435	0.400	0.498	0.519	0.609
None	0.060	0.046	0.038	0.033	0.132
Primary school	0.435	0.381	0.462	0.422	0.758

Middle school	0.741	0.689	0.875	0.846	0.960
Junior Secondary school	0.700	0.656	0.882	0.861	0.968
Secondary and above	0.874	0.841	0.999	0.998	0.996
Vocational	0.735	0.651	0.997	0.991	1.000
Other	0.937	0.885	1.000	1.000	1.000

Notes: Sample is individuals 15-65 years of age who have answered whether they have attended an adult literacy course, yielding a total of 13,403 observations.

Three findings from Table 5.1 indicate that such concerns may be unwarranted. First, the skill incidence does not appear heavily inflated. Second, literacy and numeracy rates increase with the level of education completed. Third, few literates have not attended school (some of these may be genuine, though, resulting from home-schooling or participation in adult literacy programs).

A related but somewhat different issue is the potentially high correlation among the four literacy measures and the numeracy measure empirically. The main issue here is that while conceptually the four literacy measures and the numeracy measures span five distinctly different dimension in the "skills space", as it were, empirically the high correlation among the five measures may cause "funny" results in terms of inference: statistically insignificant results of one or more of the measures and/or results of opposite directions of effects between sets of these variables, say between English reading and writing skills.[4] The reason such correlation may come about is that when an individual can write, she or he will also be able to read and also tend to be able to do written calculations, although the association between writing and reading skills would seem to be somewhat stronger than that between writing (and reading) skills and numeracy skills. To examine this issue in a bit more detail, consider the following correlation matrix for the estimation sample for the morbidity analyses (chosen since this is the largest of the five sub-analyses estimation samples and therefore likely to shed relatively most light on this

[4] Such results were actually obtained for preliminary analyses where all five skills measures were included simultaneously.

issue):

Table 5.2 Correlation Matrix for Full Set of Literacy and Numeracy Skills Variables

	Ghanaian reading	Ghanaian writing	English reading	English writing	Written calculations
Ghanaian reading	1.000	--	--	--	--
Ghanaian writing	0.896	1.000	--	--	--
English reading	0.687	0.679	1.000	--	--
English writing	0.680	0.679	0.950	1.000	--
Written calculation	0.642	0.604	0.711	0.688	1.000

Notes: Sample consists of the 4754 mothers from the morbidity analysis sub-sample. All correlation coefficients are statistically significant from zero at 1 percent.

The results from Table 4.2 confirm that there is a high correlation between Ghanaian reading and writing skills on the one hand and English reading and writing skills on the other, while the correlation between either of the four literacy measures and numeracy is somewhat smaller.

As a result of the high correlation empirically among the four literacy measures and the numeracy measure, I will include only the two writing skills measures and the numeracy skills measure in the multivariate analyses. The motivation for this is that writing skills may be interpreted as the higher standard, relative to reading skills: if an individual can write, she can also read, while the opposite is not necessarily the case. While the correlation between writing skills and numeracy skills is also high, it is not as high as that between reading and writing skills and also may be argued to encompass skills, which conceptually are quite distinct from either reading or writing skills. It is therefore included in the multivariate analyses despite the sizable positive correlation with the writing skills measures.

Education variables

Educational attainment is measured as the highest level completed, ranging from “none” through “university” and also includes vocational and technical training. Two sets of educational attainment variables are used: one in which the different levels are included as dummy variables and one in which the levels have been converted into years of schooling. Adult literacy course participation is a binary measure, stating whether an individual has ever attended an adult literacy course program. A problem with this, of course, is that the time of participation is unknown. An individual may just have started attending a class, for example, in which case the impact from the program will not have fully kicked in yet. Any impact from adult literacy course participation, therefore, is likely to be downwards biased. Further, schools and adult literacy programs are not homogenous. Specifically, school quality and content may be very different across areas or across time[5] (for example, in 1987 major schooling reforms aiming at improving the efficiency, quality and relevance of Ghanaian education were undertaken). Remote areas are also more likely to have fewer resources and therefore lower quality of schools and instruction. Additionally, private schools may be more effective than public schools in generating skills. Quality and content of adult literacy programs may vary across time or across areas, as well, since these programs are—and for a long time have been—offered by many different providers, including several different NGOs and the government. There is only information on whether or not an individual participated, however, and not on who the provider was.

Economic variables

The information on wages and earnings in the GLSS 4 is riddled with many zeros. For example,

[5] These factors are partially controlled for, however, by inclusion of age, rural-urban location and region of residence in the empirical analyses, as these variables capture components of school quality differences related to cohort and geographical location. Note that the confounding of school quality and age and cohort effects among these variables is not a concern here, since the goal is merely to control for these factors so as to obtain valid inference from our main variables of interest, mother's literacy and numeracy skills and formal schooling and adult literacy course participation.

of the 4,406 mothers in the sub-sample for the morbidity analyses (the largest of the five sub-samples) almost 30 percent report zero earnings. Examining this a bit further by cross-validating with these mothers' responses to the labor module of the survey, these zeros are genuine in the sense that the overwhelming majority of these mothers, 99.6 percent, report not having done work in the past 12 months for which a wage or any other payment was received. Of the remaining 0.4 percent, half reveal themselves as economically inactive when answering the question on the main occupation during the past 12 months. In conclusion, 99.8 of the reported zero earnings are genuine. Even though the zero earnings responses are genuine it still seems problematic to include them in the estimations, since the earnings distribution is so heavily skewed towards zero. I therefore run a Heckman selection model for (the log of) mother's (daily) earnings, using as identifying instruments marital status and the number of children born. For mothers where the observed wage rate is zero, the predicted wage rate is then imputed. Presumably, if a mother faces a higher (potential) wage rate, she will have relatively more bargaining power within the household (this is of course somewhat confounded with the skills and schooling of the mother, as well).

Additionally, geographical variables (rural-urban location and region of residence) capture economic conditions specific to the area (as well as everything else related to rural-urban residence or the region in question).

Access to facilities and prices of health services

While these variables are mainly included so as to ensure valid inference on impacts from the variables of primary interest, namely mothers' skills, schooling and adult literacy course participation, their construction is a bit tricky and therefore require somewhat detailed explanations as compared to what one might first think is warranted.

The natural first point when discussing the construction of community level health variables would seem to be the community questionnaire. However, although the community questionnaire contains fairly detailed information on availability of health personnel and facilities in the community,[6] there are several reasons to turn to the household survey for community level health information. First, in the community questionnaire there is no information on prices of services (and in the price questionnaire there is only limited information on health related items, such as aspirin, paracetamol and penicillin), so that the household part of the GLSS 4 would have to be consulted for that information, anyway. Second, due to the difficulty of defining communities in urban areas, the community questionnaire was only administered in rural areas. Using the information from the community questionnaire, therefore, automatically decreases the effective estimation samples substantially. I therefore construct community health information on facility availability and prices of services using the information on actual usage of health services from the household survey.

Specifically, questions asked include whether an individual consulted a health practitioner or dentist or visited a health center or consulted a traditional healer during the past two weeks, type of health practitioner (traditional healer, doctor, dentist, nurse, and so on) and type of facility (hospital, dispensary, pharmacy, clinic, and so on). On the cost side, there is information on the fee paid for the last vaccination given to a child. For children, who has been taken for postnatal care during the past 12 months there is information on whether or not there was a fee and, if so, how much is usually paid for one consultation. Additionally, for women who were pregnant

[6] This includes information on which types of facilities and health personnel are available in the community as well as the distance both in terms of physical distance and travel time. Facilities include hospitals, drugstore/chemical store, pharmacy, maternity home, clinic or health post and family planning clinic, while health personnel include doctor, nurse, pharmacist, trained midwife, family planning worker, community health worker, traditional birth attendant, traditional healer and medical assistant.

during the past 12 months and received any pre-natal care during this pregnancy, there is information on how much was paid for the first consultation.

This information may be used to create variables for access to and prices of health care services in the community in a way such that these variables, by construction, are exogenous to the individual even though they are based on information of actual usage. Specifically, based on the responses from individuals who has consulted a health practitioner within the past two weeks, I create a binary access variable defined as one if at least one individual in the community (child or adult) has attended a health practitioner at a hospital and zero otherwise. Similarly, consultation cost may be constructed as the average cost for the given type of service. Missing observations quickly become an issue here, however, so I focus only at a few prices: cost of pre- and postnatal care consultations and vaccinations. To ensure exogeneity of the access and price variables I calculate these variables for each household separately, leaving out the contribution of the individual household from the calculations.

An important issue in these calculations is the level of aggregation. The sample contains 300 enumeration areas (clusters), covering 101 of the 110 districts in Ghana, both of which may be further divided into either one of the three ecological zones (Coastal, Forest and Savannah) or rural-urban location. In principle, community averages could be calculated at a level of aggregation according to either of these variables (or combinations of these variables). However, the calculations face an obvious trade-off between including more observations in the analysis and losing variation in the calculated community access or price variable. Since these variables are only of secondary importance and so as to include as many observations as possible, thereby possibly increasing the precision in the estimates of the parameters of particular interest, namely mother’s skills, schooling and adult literacy participation, I chose the district level as the level of aggregation in the construction of these variables. In sum, the community level information constructed includes availability of hospital and cost of vaccinations, pre- and postnatal care. The resulting variables are based on an average number of observations per district of 10.4 (access to hospital), 13.5 (vaccination cost), 15.1 (cost of postnatal consultation) and 9.3 (cost of prenatal consultation) but with a somewhat wide range. For example, the minimum number of observations on postnatal consultation cost per district is 1, while the maximum number of observations is 100. While the range of the number of observations for the other measures is not quite as extreme, they all have at least one district with only one observation. The typical number of observations per district across the measures is about 8 or 9. Again, while these measures are clearly prone to criticism in terms of their precision (or lack thereof), they are not essential to the subsequent analysis but mainly included as additional controls so as to increase the validity of the inference from the skills and schooling variables.

Other variables related to child health inputs and outcomes

Other variables related to child health inputs and outcomes not already captured by the previous groupings include the age of the mother and the child, the number of other adults and children in the household, fertility, water source and type of sanitation of the household.

The age of the mother proxies the potential general experience of the mother, while the age of the child proxies the needs of the child (the latter is only available for the vaccinations, postnatal care and morbidity samples, however, since the mortality and prenatal care samples are at the level of the mother). For example, vaccinations are more needed at the earlier ages (and also mostly administered to young children, say, below age three). Both of these are entered with a linear and a quadratic term to allow for non-linearities. The number of other adults (than the mother) and other children (than the child in question) indicates availability of time resources for child care, while the latter at the same time also indicates needs in terms of other children in the household. Fertility may affect the health status and health investments in a child either positively or negatively. Since fertility is a measure of specific child experience, a mother with a higher fertility may have obtained more health knowledge than one with a lower fertility, for example as related to the benefits of vaccinations and pre- and postnatal care. On the other hand, a higher fertility also means that there will be less resources for the new child, both economically and in terms of child care time. Which of these effects exerts the strongest effect on child health inputs and outcomes is an empirical question, however. I include the number of other children born to the woman as a measure of fertility. A problem with this measure is that not all women in the sample have completed their fertility cycle, so that the fertility measure will be downwards biased. Any measured impact will therefore provide a conservative estimate of the fertility effect.

In addition to these other variables related to the health inputs and outcomes the household survey includes information on the source of drinking water (indoor plumbing, public standpipe, rainwater, and so on) and type of toilet (flush toilet, pit latrine, and so on). This information is particularly relevant when examining morbidity and mortality determinants. For example, diarrhea, which has been estimated by the World Health Organization to kill about 2.2. million people each year, mainly children in developing countries,[7] is thought to be caused mainly by contaminated water. The contamination may be caused by human faeces from municipal sewage, septic tanks and latrines, for example. It is more common when there is a shortage of clean water for drinking, cooking and cleaning. To capture these factors, I create a binary variable for whether the household has access to piped water and whether the household has a

[7] http://www.who.int/water_sanitation_health/diseases/diarrhoea/en/

flush toilet.

Lastly, ethnicity/tribal association, rural-urban location and region of residence may capture the taste for child health within the culture of origin and/or the culture of the local community. At the same time, however, these variables confound cultural factors with the economic conditions and experiences of different ethnicities and between different geographical locations. Again, since I am mainly interested in ensuring that valid estimates may be obtained for the skills and schooling variables and not in distinguishing between the relative impacts of cultural and economic factors per se, the issue of confoundedness as it relates to these variables is not critical to the analyses in this paper.

Sample restrictions

As was also apparent from the previous discussion on child health outcome variables, the way the questions pertaining to child health outcomes were administered in the survey implicitly gives some of the sample restrictions, while others are to be chosen. When analyzing the determinants of ever being vaccinated and ever received post-natal care, the samples are therefore restricted to children in the relevant age ranges (vaccinations: 7 years of age or younger; post-natal care: 5 years of age or younger) for which information on mothers' literacy and numeracy skills, formal schooling and adult literacy program participation is available. When analyzing the determinants of pre-natal care, the sample is restricted to women between the ages of 15 and 49 who were pregnant within the past 12 months, while the determinants of child mortality is examined for all women between the ages of 15 and 49 years of age. To ensure consistency between the different sub-samples, the estimations involving child morbidity, vaccinations and post-natal care are restricted to children for whom the mother is between the ages of 15 and 49.

Moving to the explanatory variables, mothers should have had a chance to complete primary schooling, while at the same time being eligible for participation in adult literacy programs (the lower limit). Also, individuals should not be "too old", since then measurement issues start to kick in more (upper limit). Restricting the sample to women between the ages of 15 and 49 therefore remains a reasonable strategy. Lastly, some explanatory variables are missing for some observations, which causes a further drop in the sample sizes, though in most cases not substantial. Table 4.3 summarizes the sample restrictions and the impact on the estimation sample sizes for the various analyses.

While the drop in observations from the initial to the estimation samples is not alarming, I nonetheless examine the impact of the sample restrictions on sample selection. This is done by tabulating the variable means for the initial and sample means and testing for statistical different differences between the two (see Appendix A, tables A1 and A2). Again, while there are statistically significant differences between the initial and estimation samples in some cases, the dropout in absolute terms is rather low.

Table 4.3 Sample Restrictions and Impact on Estimation Sample Sizes

Sub-analysis	Sample	Initial sample	Estimation sample
(1) Ever vaccinated	Children 0-7 years	4791	4406
(2) Mortality	Females, 15-49 years	4167	4144
(3) Received prenatal care	Females, 15-49 years, who were pregnant in past 12 months	1142	1038
(4) Ever received postnatal care	Children 0-5 years	3588	3539
(5) Morbidity	Children 0-7 years	4821	4754

Notes: Initial sample is sample, for which information on the dependent variable is available and the mother is between 15 and 49 years old. Estimation sample is sample where all explanatory variables (gender and age of child, age, literacy and numeracy skills, formal education, adult literacy course participation and earnings of mother, rural-urban location, region of residence, ethnicity/tribal association and access to and prices of health related services in the community, water access and sanitation of household) are available.

Preliminary analyses of the determinants of child health outcomes

In order to provide preliminary analyses of child health outcomes, these are tabulated across mothers' literacy and numeracy skills, school attendance and adult literacy course participation in Table 4.4. Again, while the reading skills variables are excluded form the multivariate analyses, it still seems worthwhile to examine the bivariate associations between these measures and the child health measures.

Judging from this evidence, there are major differences in child health outcomes across the mother's skills and schooling. First, mothers proficient in Ghanaian reading or writing, English reading or writing or written calculations experience lower child mortality and are also more likely to vaccinate their children than both the "average mother" (as measured by the sample means) and mothers who are not proficient in these skills. Second, mothers proficient in English reading or writing or written calculations are more likely to seek post-natal care for their child than are both the "average mother" (as measured by the sample means) and mothers who are not proficient in these skills. Third, mothers who have ever attended school are more likely to vaccinate their children, seek pre-natal care for themselves and post-natal care for their children and also experience lower child mortality both the "average mother" (as measured by the sample means) and mothers who never attended school.

I will now examine these child health outcomes gaps (child health determinants) further using multivariate statistical techniques.

Table 4.4 Children's Vaccinations, Mortality, Pre- and Post-natal Care and Illnesses Across Maternal Literacy and Numeracy Skills, Schooling and Literacy Course Participation

	Full Sample Mean	Ghanaian Writing	English Writing	Written Calculations	Attended School	Adult Literacy Course Participation
Ever vacc.	0.919	0.963***	0.968***	0.961***	0.948***	0.964***
Number	0.536	0.323***	0.270***	0.316***	0.394***	0.669

died						
Pre-natal	0.828	0.883*	0.879*	0.876**	0.873***	0.778
Post-natal	0.394	0.409	0.443**	0.433**	0.419*	0.449*
Ill	0.319	0.305	0.308	0.315	0.327	0.375*

Notes: Samples are final estimation samples as shown in Tables 5.1 and 5.2 and contain children 7 years old or younger, except for (1) post-natal care, which is for children 5 years or younger, (2) pre-natal care, which is only measured for women who were pregnant within the past 12 months and (3) mortality, which is measured for women between 15 and 49 years of age. Sample for literacy course participants is individuals who completed primary school or less. For presentation purposes the individual cell sizes for health variables have been omitted; they are available upon request. ***: statistically significant from the reference category (not proficient in Ghanaian writing, English writing and so on) at 10 percent, **: statistically significant from the reference category at 5 percent, *: statistically significant from the reference category at 1 percent.

5. Multivariate Analyses

The multivariate analyses contain three specifications for each of the five child health measures, the last of which is the "preferred" specification and the first two are included for sensitivity analyses (Table 5.1). The first specification includes mothers' literacy and numeracy skills, formal education, adult literacy course participation and strictly exogenous variables such as gender, age and age squared of the child and the age and age squared of the mother. The second specification adds weakly exogenous variables such as health facility access, regional variables and cost of health services. The last specification adds potentially endogenous variables, including mothers' earnings, water and sanitation of the household and fertility variables.

Judging on the evidence from the preferred specifications in Table 5.2, there seems to be some support of literacy and numeracy skills effects independent of the effects from formal education and adult literacy course participation for some of the child health measures. In particular, English writing skills has a negative and statistically significant impact on child mortality. This points towards English writing skills being the "higher standard" in terms of skills as far as child mortality is concerned. Note, however, that this does not mean that English reading skills are not important: individuals who writes must be able to read, as well, and therefore that skill is implicitly included in English writing skills. For the unrestricted illness measure the English reading and writing skills variables are both statistically significant but with opposite signs, indicating the existence of collinearity between the two variables. For vaccinations and number of children ever died written calculations have a statistically negative and positive impact, respectively, for the initial specification but it disappears after introducing more controls.

Table 5.1 Overview of Variables Included in the Different Specifications

(1) For all equations:	Model 1	Model 2	Model 3
	Contains mother's skills, education and strictly exogenous controls:	Adds weakly exogenous controls:	Adds potentially endogenous variables:
Mother's literacy and numeracy skills:			
Ghanaian writing	X	X	X
English writing	X	X	X
Written calculations	X	X	X
Mother's education:			
Literacy course participation	X	X	X
Primary	X	X	X
Middle/JSS	X	X	X
Secondary and above	X	X	X
Vocational and other	X	X	X
Economic variables:			
Mother's (potential) wage rate			X
Income of other HH members			X
Access to facilities:			
Access to hospital in community		X	X
Other variables related to child health inputs and outcomes:			
Age of mother	X	X	X
Age of mother squared	X	X	X
Ethnicity/tribe of mother	X	X	X
Number of other adults in HH			X
Number of other children in HH			X
Access to piped water			
Urban-rural location		X	X
Regional dummies		X	X
(2) For vaccinations, postnatal care and morbidity equations, only:			
Age of child	X	X	X
Age of child squared	X	X	X
Gender of child	X	X	X
(3) For vaccinations, morbidity and mortality equations, only:			
Vaccination cost (cluster average)		X	X
(4) For prenatal care equation, only:			
Prenatal consultation cost (cluster average)		X	X
(5) For postnatal care equation, only:			
Postnatal consultation cost (cluster average)		X	X

In accordance with the previous literature, formal educational attainment has strong and statistically significant effects in several cases, although the introduction of literacy and numeracy skills and adult literacy course participation seems to pick up some of the impact from formal education so that there is not as many statistically significant formal education variables as one might otherwise have anticipated.

Table 5.2 Results for Preferred Specifications

	Vaccinations	Number of dead children	Prenatal care	Postnatal care	Illness
Ghanaian reading	0.211	-0.015	-0.059	0.072	-0.009
	[0.264]	[0.133]	[0.226]	[0.201]	[0.112]
Ghanaian writing	-0.137	0.056	-0.039	-0.153	-0.059
	[0.289]	[0.135]	[0.280]	[0.190]	[0.116]
English reading	0.1	0.159	0.313	0.157	-0.403**
	[0.335]	[0.146]	[0.355]	[0.204]	[0.158]
English writing	-0.027	-0.292**	-0.006	0.018	0.390**
	[0.323]	[0.147]	[0.341]	[0.200]	[0.163]
Written calculations	0.097	-0.059	-0.161	0.061	-0.01
	[0.145]	[0.088]	[0.186]	[0.098]	[0.089]
Literacy course	0.269**	-0.037	0.212	0.227**	0.155*
	[0.132]	[0.094]	[0.184]	[0.098]	[0.086]
Primary	0.064	-0.049	0.400**	0.004	-0.034
	[0.125]	[0.084]	[0.173]	[0.105]	[0.077]
Middle/JSS	0.323*	0.044	0.157	-0.101	0.04
	[0.190]	[0.116]	[0.193]	[0.129]	[0.099]
Secondary and above	NA	0.117	0.142	0.407*	0.105
	NA	[0.214]	[0.312]	[0.221]	[0.231]
Vocational	NA	0.32	-0.042	-0.123	-0.239
	NA	[0.208]	[0.624]	[0.256]	[0.221]
Other education	NA	-5.893***	NA	-0.283	0.146
	NA	[0.223]	NA	[0.385]	[0.401]
Specification tests:					
(1) Hausman tests:					
Model 2 vs. 1	F(16, 270) = 0.90	F(16, 279) = 17.39***	F(15, 262) = 0.60	F(19, 274) = 0.58	F(19, 275) = 1.53*
Model 3 vs. 2	F(28, 258) = 0.75	F(28, 267) = 18.09***	F(27, 250) = 0.74	F(31, 262) = 0.55	F(31, 263) = 0.48
Model 3 vs. 1	F(16, 270) = 1.75**	F(16, 279) = 28.28***	F(15, 262) = 1.17	F(19, 274) = 0.67	F(19 ,275) = 1.23
(2) Joint significance of additional variables:					
Model 2 vs. 1	F(12, 274) = 2.34***	F(12, 283) = 4.97***	F(12, 265) = 1.22	F(12, 281) = 1.31	F(12, 282)= 3.48***
Model 3 vs. 2	F(2, 284) = 4.70***	F(2, 293) = 255.94***	F(2, 275) = 5.53***	F(2, 291) = 4.48**	F(2, 292) = 2.85*
Model 3 vs. 1	F(14, 272) = 2.14***	F(14, 281) = 46.02***	F(14, 263) = 1.98**	F(14, 279) = 1.68*	F(14, 280) = 2.98***
Number of observations	4439	4144	1130	3569	4789

Notes: "NA" indicates that variable dropped out due to being a perfect predictor, whereby the corresponding observation is dropped from the estimation, as well. Robust Huber-White Sandwich (Huber, 1967; White, 1980) standard errors in brackets. *: statistically significant at 10 percent; ***: statistically significant at 5 percent; ****: statistically significant at 1 percent. Since the asymptotic assumptions of the Hausman (1978) test are not satisfied when incorporating survey weights in the estimations, the Generalized Hausman test is applied here.

Adult literacy course participation is positively statistically significant for vaccinations and postnatal care, indicating the impact of the health knowledge, which is an integral part of the adult literacy course curriculum in Ghana (Blunch and Pörtner, 2004) on individual health seeking behavior. Again, had only formal educational attainment been included in the analyses, this important channel of health knowledge diffusion would have been overlooked. Somewhat puzzling at first, the impact from adult literacy participation on child illnesses is statistically significant and positive. One explanation for this is that adult literacy course participants may better correctly diagnose diseases, which would seem to inflate the self-reported child morbidity measure. This is supported by the results from the specification using the restricted child morbidity measure: when children had to stop their usual activities, the impact of adult literacy course participation is halved and is no longer statistically significantly different from zero. The intuition behind this is that when the illness is so serious that the children has to stop their usual activities, the health knowledge obtained from adult literacy courses is not "necessary" in terms of assessing whether the child is ill or not.

Again, remembering the weaknesses of the various child health measures as discussed earlier, these results should be interpreted with care, especially for the cases of child illnesses, and pre- and post-natal care, whereas the results for child vaccinations and child mortality seems to hold up a bit stronger. Altogether, however, these results point towards the importance of literacy and numeracy skills and adult literacy programs for child health, two issues which have not received much attention in the previous literature of child health determinants.

6. Conclusion

Viewing the accumulation of human capital as a two-pronged production process, where schooling produces skills, which subsequently enters as inputs (or intermediate outputs) to generate the final outcome of child health, this paper examines the relationship between mothers' literacy and numeracy skills, formal education and adult literacy course participation and child health in Ghana. The health measures examined include child vaccinations, child mortality, pre- and postnatal care and child morbidity. The contribution of this paper includes (1) analyzing the impact on child health from skills, including reading and writing skills and that for both English and indigenous languages, as well as numeracy and other skills and (2) including adult literacy course participation as a pathway of achieving skills, two issues which have not been addressed in the previous literature. It considers how these skills affect the production of health, including mothers' pre-natal care and children's vaccinations and post-natal care, and how they affect outputs, including children's illnesses and mortality.

The results indicate some support for skills affecting child health outcomes independently of formal educational attainment. Most significantly, the results indicate that the health knowledge skills obtained through adult literacy course participation significantly improves children's health in terms of vaccinations and postnatal care.

More research is required, however, in order to address whether the link is causal, rather than merely indicating correlation. In particular, the current analyses implicitly assume that all explanatory variables are exogenous to child health outcomes. While that may not be too far off for some of the variables (gender, rural-urban location and region), for others this assumption is more problematic (mother's schooling, skills, earnings, water and sanitation). The many endogenous variables in this study and the resulting requirement for instruments render instrumental variables methods infeasible as a means of addressing endogeneity. Future research will therefore address these issues by introducing—and hence controlling for—unobserved heterogeneity. This will be done by introducing a latent factor capturing unobservables such as "type" or preferences, either assuming a specific functional form for the distribution of the factor or using the discrete factor approximation approach first suggested by Heckman and Singer (1984) and generalized by Mroz and Guilkey (1992). The latter is somewhat less restrictive but also entails a greater burden computationally.

References:

Asenso-Okyere, W.K., F.A. Asante and M. Nubé (1997) "Understanding the Health and Nutritional Status of Children in Ghana", *Agricultural Economics*, 17: 59-74.

Becker, Gary S. (1965) "A Theory of the Allocation of Time", *Economic Journal*, 75: 493-517.

Blunch, Niels-Hugo (2004) "Literacy, Numeracy and Proximate Determinants", Mimeo, Washington, DC: The George Washington University.

Blunch, Niels-Hugo and Claus Pörtner (2004) "Adult Literacy Programs in Ghana: An Evaluation", Mimeo, Human Development Network Education Department, Washington, DC: World Bank.

Glewwe, Paul and Jaikishan Desai (1999) "Child Health and Mother's Schooling in Ghana", in Paul Glewwe (ed) *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*, London: Macmillan.

Grossman, Michael (1972) "On the Concept of Health Capital and the Demand for Health", *Journal of Political Economy*, 80 (2): 223-55.

Heckman, James and B. Singer (1984) "A Method for Minimizing the Impact of Distributional Assumptions on Econometric Models for Duration Data", *Econometrica*, 52: 271-320.

Huber, P. J. (1967) "The Behavior of Maximum Likelihood Estimates under Nonstandard

Conditions", In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* Vol. 1, Berkeley, CA: University of California Press.

Mroz, Thomas and David K. Guilkey (1992) "Discrete Factor Approximations for Use in Simultaneous Equation Models With Both Continuous And Discrete Endogenous Variables", Mimeo, Department of Economics, University of North Carolina, Chapel Hill.

Schultz, Theodore W. (1973) "The Value of Children: An Economic Perspective", *Journal of Political Economy*, 81 (2): 2-13.

Spence, Michael A. (1973) "Job Market Signaling", *Quarterly Journal of Economics*, 87: 355-74.

White, H. (1980) "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity", *Econometrica*, 48: 817–830.

APPENDIX A: Comparisons of Means of Initial and Estimation Samples

Table A1 Comparisons of Means of Initial and Estimation Samples: Vaccinations and Child Mortality

Sample/Dependent Variable	Vaccine			Mortality (all measures)		
	Estimation sample	Excluded Observations	Difference	Estimation sample	Excluded Observations	Difference
Female	0.503	0.555	-0.052	NA	NA	NA
Age, child	3.493	3.204	0.290	NA	NA	NA
Age squared/100, child	0.174	0.163	0.011	NA	NA	NA
Age, mother	31.468	28.460	3.008**	33.243	31.020	2.223
Age squared/100, mother	10.341	8.439	1.902**	11.676	10.171	1.505
Akan	0.472	0.541	-0.069	0.485	0.192	0.293***
Ewe	0.126	0.205	-0.080	0.133	0.127	0.005
Ga-Adangbe	0.088	0.065	0.022	0.093	0.117	-0.024
Other ethnicity	0.315	0.189	0.126	0.290	0.564	-0.274**
Number other children 0-7	1.137	0.912	0.224	NA	NA	NA
Number born	NA	NA	NA	3.978	3.190	0.788
Access to hospital	0.779	0.774	0.004	0.805	0.729	0.076
Piped water	0.276	0.230	0.046	0.337	0.118	0.219***
Flush toilet	0.035	0.027	0.008	0.056	0.027	0.029
Urban	0.273	0.182	0.092	0.339	0.318	0.021
Western	0.124	0.130	-0.006	0.114	0.046	0.069*
Central	0.097	0.109	-0.012	0.092	0.014	0.078***
Greater Accra	0.090	0.125	-0.035	0.120	0.115	0.006
Eastern	0.103	0.162	-0.059	0.110	0.276	-0.166
Volta	0.123	0.023	0.100***	0.120	0.025	0.096***
Ashanti	0.160	0.255	-0.095	0.170	0.098	0.072
Brong-Ohofa	0.094	0.000	0.094***	0.091	0.224	-0.132
Northern	0.133	0.114	0.019	0.108	0.162	-0.054
Upper West	0.029	0.000	0.029**	0.029	0.000	0.029*
Upper East	0.047	0.082	-0.035	0.045	0.042	0.003
Number of observations	4760	31		4144	23	

Notes: ***: statistically significant at 10 percent, **: statistically significant from the reference category at 5 percent, *: statistically significant from the reference category at 1 percent.

Table A2 Comparisons of Means of Initial and Estimation Samples: Pre- and Postnatal Care and Illnesses

Sample/Dependent Variable	Prenatal care			Postnatal care			Illnesses (both measures)		
	Estimation sample	Excluded Observations	Difference	Estimation sample	Excluded Observations	Difference	Estimation sample	Excluded Observations	Difference
Female	NA	NA	NA	0.510	0.466	0.044	0.501	0.534	-0.033
Age, child	NA	NA	NA	2.467	2.129	0.339	3.510	3.233	0.277
Age squared/100, child	NA	NA	NA	0.090	0.077	0.013	0.176	0.163	0.013
Age, mother	29.818	22.652	7.167	30.674	27.864	2.810*	31.477	29.114	2.363
Age squared/100, mother	9.300	5.367	3.933	9.849	8.216	1.633	10.349	8.913	1.436
Akan	0.501	0.339	0.163	0.473	0.719	-0.246**	0.472	0.521	-0.048
Ewe	0.124	0.121	0.003***	0.124	0.113	0.010	0.126	0.235	-0.109
Ga-Adangbe	0.094	0.143	-0.049	0.088	0.115	-0.027	0.088	0.063	0.025
Other ethnicity	0.286	0.191	0.095	0.316	0.053	0.263***	0.314	0.182	0.133
Number other children 0-5	NA	NA	NA	0.778	0.525	0.253	1.262	1.343	0.080
Number other children 0-7	NA	NA	NA	NA	NA	NA	1.140	0.878	0.261
Access to hospital	0.783	0.822	-0.039	0.775	0.882	-0.106	0.778	0.783	-0.005
Piped water	0.285	0.308	-0.023*	0.263	0.314	-0.051	0.278	0.222	0.056
Flush toilet	0.035	0.075	-0.040***	0.031	0.047	-0.017	0.035	0.026	0.009
Urban	0.284	0.348	-0.064	0.262	0.295	-0.032	0.275	0.175	0.100
Western	0.122	0.131	-0.009***	0.124	0.118	0.006	0.123	0.125	-0.002
Central	0.104	0.063	0.041	0.098	0.230	-0.132	0.096	0.105	-0.008
Greater Accra	0.086	0.197	-0.111	0.088	0.158	-0.070	0.091	0.121	-0.030
Eastern	0.104	0.130	-0.025	0.105	0.121	-0.017	0.104	0.193	-0.089
Volta	0.133	0.115	0.018	0.123	0.000	0.123***	0.123	0.022	0.101***
Ashanti	0.158	0.277	-0.119	0.162	0.306	-0.144	0.161	0.245	-0.085
Brong-Ohofa	0.115	0.039	0.076***	0.092	0.000	0.092***	0.093	0.000	0.093***
Northern	0.116	0.073	0.044	0.133	0.067	0.066	0.133	0.110	0.023
Upper West	0.028	0.007	0.021	0.028	0.000	0.028**	0.028	0.000	0.028**
Upper East	0.037	0.024	0.013	0.047	0.000	0.047***	0.048	0.079	-0.031
Number of observations	1132	10		3569	19		4789	32	

Notes: ***: statistically significant at 10 percent, **: statistically significant from the reference category at 5 percent, *: statistically significant from the reference category at 1 percent.

APPENDIX B: Sensitivity Analyses and Specification Tests

Table B1 Results from "Naïve" Estimation (Not Taking Endogeneity into Account): Vaccinations and Mortality

Dependent variable:	Vaccinations			Number of children ever died			Any children ever died			Share of children ever died		
Estimation method:	Probit			Ordered probit			Probit			OLS		
Model:	1	2	3	1	2	3	1	2	3	1	2	3
Ghanaian reading	0.168	0.184	0.211	0.017	0.014	-0.015	0.061	0.056	0.032	-0.007	-0.005	-0.006
	[0.251]	[0.262]	[0.264]	[0.151]	[0.144]	[0.133]	[0.170]	[0.165]	[0.156]	[0.017]	[0.017]	[0.016]
Ghanaian writing	-0.18	-0.137	-0.137	0.051	0.045	0.056	0.025	0.007	0.004	0.017	0.013	0.013
	[0.274]	[0.291]	[0.289]	[0.162]	[0.150]	[0.135]	[0.182]	[0.172]	[0.159]	[0.017]	[0.016]	[0.016]
English reading	-0.018	0.133	0.1	-0.07	-0.007	0.159	0.004	0.071	0.209	0.004	0.007	0.014
	[0.304]	[0.335]	[0.335]	[0.111]	[0.119]	[0.146]	[0.123]	[0.132]	[0.155]	[0.016]	[0.017]	[0.018]
English writing	0.096	-0.034	-0.027	-0.210*	-0.239*	-0.292**	-0.282**	-0.305**	-0.338**	-0.026	-0.026	-0.026
	[0.305]	[0.327]	[0.323]	[0.118]	[0.128]	[0.147]	[0.133]	[0.145]	[0.160]	[0.017]	[0.018]	[0.019]
Written calculations	0.247*	0.155	0.097	-0.160**	-0.126	-0.059	-0.103	-0.064	0.009	-0.012	-0.008	-0.003

	[0.145]	[0.144]	[0.145]	[0.074]	[0.082]	[0.088]	[0.091]	[0.098]	[0.103]	[0.014]	[0.013]	[0.013]
Literacy course	0.296**	0.279**	0.269**	0.028	-0.034	-0.037	0.145	0.083	0.089	0.007	0.000	0.001
	[0.129]	[0.132]	[0.132]	[0.074]	[0.076]	[0.094]	[0.098]	[0.100]	[0.117]	[0.014]	[0.014]	[0.015]
Primary	0.143	0.094	0.064	-0.133	-0.111	-0.049	-0.187**	-0.167*	-0.122	-0.018	-0.013	-0.009
	[0.116]	[0.119]	[0.125]	[0.084]	[0.086]	[0.084]	[0.095]	[0.097]	[0.096]	[0.013]	[0.013]	[0.012]
Middle/JSS	0.412**	0.336*	0.323*	-0.188*	-0.138	0.044	-0.224**	-0.182	-0.038	-0.019	-0.012	-0.001
	[0.179]	[0.184]	[0.190]	[0.106]	[0.108]	[0.116]	[0.109]	[0.111]	[0.119]	[0.017]	[0.017]	[0.017]
Secondary and above	NA	NA	NA	-0.502***	-0.377**	0.117	-0.565***	-0.454**	-0.039	-0.024	-0.012	0.018
	NA	NA	NA	[0.174]	[0.180]	[0.214]	[0.186]	[0.188]	[0.215]	[0.025]	[0.026]	[0.028]
Vocational	NA	NA	NA	-0.255	-0.045	0.32	-0.225	-0.022	0.305	-0.001	0.018	0.041
	NA	NA	NA	[0.199]	[0.212]	[0.208]	[0.219]	[0.234]	[0.229]	[0.029]	[0.030]	[0.029]
Other education	NA	NA	NA	-6.431***	-6.985***	-5.893***	NA	NA	NA	-0.073***	-0.050**	-0.011
	NA	NA	NA	[0.186]	[0.195]	[0.223]	NA	NA	NA	[0.021]	[0.020]	[0.024]
Specification tests:												
(1) Hausman tests:												
Model 2 vs. 1	F(16, 270) = 0.90			F(16, 279) = 17.39***			F(15, 280) = 2.78***			F(16, 279) = 2.93***		
Model 3 vs. 2	F(28, 258) = 0.75			F(28, 267) = 18.09***			F(27, 268) = 12.85***			F(28, 267) = 2.63***		
Model 3 vs. 1	F(16, 270) = 1.75**			F(16, 279) = 28.28***			F(15, 280) = 21.86***			F(16, 279) = 6.40***		
(2) Joint significance of additional variables:												
Model 2 vs. 1	F(12, 274) = 2.34***			F(12, 283) = 4.97***			F(12, 283) = 4.07***			F(12, 283) = 3.64***		
Model 3 vs. 2	F(2, 284) = 4.70***			F(2, 293) = 255.94***			F(2, 293) = 170.15***			F(2, 293) = 35.73***		
Model 3 vs. 1	F(14, 272) = 2.14***			F(14, 281) = 46.02***			F(14, 281) = 28.54***			F(14, 281) = 8.48***		
No. of observations		4439			4144			4142			4144	
R^2	NA	NA	NA	NA	NA	NA	NA	NA	NA	0.03	0.04	0.06

Notes: "NA" indicates that variable dropped out due to being a perfect predictor, whereby the corresponding observation is dropped from the estimation, as well. Robust Huber-White Sandwich (Huber, 1967; White, 1980) standard errors in brackets. *: statistically significant at 10 percent; ***: statistically significant at 5 percent; ****: statistically significant at 1 percent. Since the asymptotic assumptions of the Hausman (1978) test are not satisfied when incorporating survey weights in the estimations, the Generalized Hausman test is applied here.

Table B2 Results from "Naïve" Estimation (Not Taking Endogeneity into Account): Ilnesses, Pre- and Post-natal Care.

Dependent variable:	Prenatal care			Postnatal care			Illness			Ill, stopped usual activities		
Estimation method:	Probit			Probit			Probit			Probit		
Model:	1	2	3	1	2	3	1	2	3	1	2	3
Ghanaian reading	-0.024	-0.062	-0.059	0.032	0.075	0.072	-0.05	-0.036	-0.009	0.06	0.079	0.093
	[0.224]	[0.224]	[0.226]	[0.199]	[0.199]	[0.201]	[0.115]	[0.115]	[0.112]	[0.125]	[0.130]	[0.130]
Ghanaian writing	-0.098	-0.062	-0.039	-0.129	-0.176	-0.153	-0.012	-0.041	-0.059	-0.077	-0.101	-0.113
	[0.271]	[0.274]	[0.280]	[0.188]	[0.189]	[0.190]	[0.118]	[0.120]	[0.116]	[0.123]	[0.127]	[0.124]
English reading	0.29	0.296	0.313	0.135	0.168	0.157	-0.456***	-0.397**	-0.403**	-0.341*	-0.283	-0.285
	[0.341]	[0.355]	[0.355]	[0.221]	[0.205]	[0.204]	[0.156]	[0.159]	[0.158]	[0.177]	[0.184]	[0.183]
English writing	-0.005	-0.025	-0.006	0.04	0.034	0.018	0.433***	0.385**	0.390**	0.26	0.197	0.2
	[0.310]	[0.334]	[0.341]	[0.206]	[0.200]	[0.200]	[0.160]	[0.162]	[0.163]	[0.175]	[0.181]	[0.181]
Written calculations	-0.201	-0.216	-0.161	0.129	0.105	0.061	0.003	0.008	-0.01	0.036	0.033	0.019
	[0.192]	[0.187]	[0.186]	[0.100]	[0.097]	[0.098]	[0.088]	[0.088]	[0.089]	[0.098]	[0.093]	[0.095]
Literacy course	0.159	0.2	0.212	0.250**	0.237**	0.227**	0.169*	0.166*	0.155*	0.079	0.099	0.085
	[0.188]	[0.184]	[0.184]	[0.108]	[0.099]	[0.098]	[0.087]	[0.085]	[0.086]	[0.093]	[0.087]	[0.087]
Primary	0.430**	0.378**	0.400**	0.04	0.012	0.004	-0.039	-0.025	-0.034	-0.019	-0.002	-0.005
	[0.171]	[0.173]	[0.173]	[0.115]	[0.107]	[0.105]	[0.078]	[0.078]	[0.077]	[0.083]	[0.083]	[0.083]
Middle/JSS	0.282	0.202	0.157	-0.046	-0.089	-0.101	0.031	0.046	0.04	0.017	0.032	0.031

	[0.194]	[0.193]	[0.193]	[0.132]	[0.130]	[0.129]	[0.099]	[0.100]	[0.099]	[0.109]	[0.112]	[0.111]
Secondary and above	0.251	0.175	0.142	0.568**	0.494**	0.407*	0.023	0.062	0.105	0.11	0.121	0.154
	[0.310]	[0.313]	[0.312]	[0.236]	[0.226]	[0.221]	[0.208]	[0.215]	[0.231]	[0.232]	[0.241]	[0.258]
Vocational	0.202	0.096	-0.042	0.021	-0.066	-0.123	-0.293	-0.223	-0.239	-0.346	-0.31	-0.305
	[0.585]	[0.598]	[0.624]	[0.262]	[0.252]	[0.256]	[0.211]	[0.218]	[0.221]	[0.250]	[0.261]	[0.263]
Other education	NA	NA	NA	-0.036	-0.159	-0.283	-0.03	0.136	0.146	0.17	0.305	0.315
	NA	NA	NA	[0.409]	[0.446]	[0.385]	[0.475]	[0.465]	[0.401]	[0.501]	[0.496]	[0.462]
Specification tests:												
(1) Hausman tests:												
Model 2 vs. 1	F(15, 262) = 0.60			F(19, 274) = 0.58			F(19, 275) = 1.53*			F(19, 275) = 1.05		
Model 3 vs. 2	F(27, 250) = 0.74			F(31, 262) = 0.55			F(31, 263) = 0.48			F(31, 263) = 0.28		
Model 3 vs. 1	F(15, 262) = 1.17			F(19, 274) = 0.67			F(19 ,275) = 1.23			F(19, 275) = 1.11		
(2) Joint significance of additional variables:												
Model 2 vs. 1	F(12, 265) = 1.22			F(12, 281) = 1.31			F(12, 282)= 3.48***			F(12, 282) =3.51***		
Model 3 vs. 2	F(2, 275) = 5.53***			F(2, 291) = 4.48**			F(2, 292) = 2.85*			F(2, 292) = 3.42**		
Model 3 vs. 1	F(14, 263) = 1.98**			F(14, 279) = 1.68*			F(14, 280) = 2.98***			F(14, 280) = 3.47***		
No. of observations		1130			3569			4789			4789	

Notes: "NA" indicates that variable dropped out due to being a perfect predictor, whereby the corresponding observation is dropped from the estimation, as well. Robust Huber-White Sandwich (Huber, 1967; White, 1980) standard errors in brackets. *: statistically significant at 10 percent; ***: statistically significant at 5 percent; ****: statistically significant at 1 percent. Since the asymptotic assumptions of the Hausman (1978) test are not satisfied when incorporating survey weights in the estimations, the Generalized Hausman test is applied here.