I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10663

# A Teenager in Love: Multidimensional Human Capital and Teenage Pregnancy in Ghana

Niels-Hugo Blunch

MARCH 2017

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 10663

# A Teenager in Love: Multidimensional Human Capital and Teenage Pregnancy in Ghana

**Niels-Hugo Blunch**
*Washington and Lee University*
*and IZA*

MARCH 2017



**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## A Teenager in Love: Multidimensional Human Capital and Teenage Pregnancy in Ghana*

I examine teenage pregnancy in Ghana, focusing on the role and interplay of Ghanaian and English reading skills, formal educational attainment, and adult literacy program participation. Pursuing several alternative identification strategies three main results are established. First, I confirm the finding from previous studies that educational attainment is negatively related to teenage pregnancy. Second, however, once Ghanaian and English reading skills are introduced, the association between educational attainment and teenage pregnancy decreases or disappears altogether. Third, for the girls who have not completed primary school, adult literacy program participation is associated with a much lower probability of experiencing a teenage pregnancy.



**JEL Classification:** I21, J24

**Keywords:** teenage pregnancy, human capital, literacy, adult literacy programs, Ghana



**Corresponding author:**
Niels-Hugo Blunch
Department of Economics
Washington and Lee University
Lexington, VA 24450
USA

E-mail: blunchn@wlu.edu



* I thank David Bishai, Marco Gonzalez-Navarro, Anthony Keats, Melanie Khamis, Nishith Prakash, Claus Pörtner, Howard White, and participants at the Danish Academic Economists in North America Annual Meetings; the Economic Development in Africa conference (Centre for the Study of African Economies, Oxford), the Midwest International Economic Development Conference, the Population Association of America Annual Meetings, Southern Economic Association Annual Meetings, and seminar participants at Aarhus University, Ohio University, Wesleyan University, and Seattle University for helpful comments and suggestions. I gratefully acknowledge financial support from Washington and Lee University's Lenfest Summer Research Grant. Part of this research was carried out while visiting Princeton University's Economics Department; the hospitality of the Department is gratefully acknowledged. 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 data is proprietary to the Ghana Statistical Service but can be obtained from the GSS website at http://www.statsghana.gov.gh/DataRequest.html. The Stata code used for the analysis in this paper is available upon request.

*A bright future is the best contraceptive.*

–Marian Wright Edelman

**1. Introduction.**

In days past, with low average life expectancy many women gave birth—and therefore, perhaps rationally so—early on in life. Today, however, birth at an early age is usually considered an adverse outcome—even in developing countries—due, among other things, to increasing life expectancies over most of the latter half of the previous century. Teenage pregnancy arguably is a case in point here. Not only will a teenage mother be less able to take care of the child at birth due to her often low socioeconomic status but future care will be affected as well, since the accumulation of human capital will be at the expense of her earlier caretaking, all else equal.

In line with this discussion, recent years have seen increased attention to the relationship between human capital and teenage pregnancy/childbearing (Heckman, Stixrud, and Urzua, 2006; Levine and Painter, 2003; Ribar, 1993, 1994), especially for the US. Many of these studies consider the direction going from pregnancy to individual human capital outcomes and also mostly focus at developed countries.

This study extends previous research in several directions. First, again, while previous studies mostly have examined the direction going from pregnancy to individual human capital outcomes, I consider the possibility that basic skills and schooling, some of which are obtained during the early years, have a potential role in avoiding "too early" (certainly in terms of possibly affecting subsequent additional human capital accumulation) pregnancies. Second, I consider adult literacy programs as a potential alternative to the formal education system in terms of acquiring life skills, including health knowledge. This is a channel that seems especially promising for developing countries where substantial parts of the population have acquired only low levels of human capital, including basic skills such as literacy and numeracy—but at the same time this has received only limited attention in previous research. Third, I allow for possible separate effects from two alternative forms of reading skills, namely native reading skills and the skills of the official government language (which in this case also is the language of the previous colonial power). Fourth, in exploring these multidimensional human capital-teenage pregnancy linkages the effects of the individual human capital components are allowed to be endogenous, where many previous studies have taken human capital to be predetermined, if

not exogenous.[1] Lastly, this study considers a developing country, where the bulk of previous research has focused on developed countries, especially the US.[2]

The empirical application for this analysis is the West African country of Ghana. Ghana provides an exciting context for examining the issue of teenage pregnancy, especially in terms of the possible linkages to adult literacy programs and literacy.

First, while teenage pregnancy is a problem in many developing as well as developed countries, teenage pregnancy has increased substantially in Ghana in recent years, almost doubling between 2009 (43,465 pregnancies) and 2013 (83,917 pregnancies) among those aged 15 to 19 years, so that these pregnancies now account for about 12.3 percent of all pregnancies in Ghana (GHS, 2013).[3] Since pregnancies for adolescents are both associated with potential health problems—both for the mother and the child—as well as with increased risk of the teenage mother dropping out from school (if enrolled, in the first place), this massive recent increase in teenage pregnancies in Ghana would seem to be an important policy concern; and warranting also increased attention from researchers.

Second, however, adult literacy programs—which have a long history in Ghana[4]—seem to provide a potential vehicle in the fight against teenage pregnancy. Alas, these programs have been scaled down—if not practically abandoned—in recent years, particularly due to their seemingly dismal track record in creating literacy and numeracy skills for participants[5]. This is in line with similar developments across the developing world, including also major players such as the World Bank (Chowhury, 1995). Yet, there is a case to be made that even if the programs may not be successful in creating their stated objectives of literacy and numeracy, due to their nature of being multiplex programs—containing also a Health[6], an Income-Generation/

[1] Although it turns out that ultimately this does not appear necessary for the application considered here.
[2] A caveat here is that a substantial literature exists on the related notions of "age at first birth" or "first-birth timing," though this literature seems to be predominantly Asian (e.g., Hirschman and Rindfuss, 1980; Bloom and Reddy, 1986; and Basu, 1993)—so that explicitly examining teenage pregnancy, per se, and especially doing so for the Sub-Saharan African country of Ghana, is still a contribution to the literature in this dimension.
[3] Similarly, several key reproductive health statistics in Ghana are comparatively low—including median age at first intercourse (18.4 years for women and 20.0 years for men), at first birth (20.7 years), and at first marriage (19.8 years for women and 25.9 for men) (GSS et al., 2009).
[4] See Blunch and Pörtner (2005), Appendix B (whereupon much of the following discussion is also based) for additional details.
[5] See, e.g., Abadzi (1994) and Ortega and Rodríguez (2008).
[6] Topics include family planning, teenage pregnancy, environmental hygiene, immunization, HIV/AIDS, safe motherhood and child care, drug abuse, traditional medicine, and safe drinking water.

Occupational Skills[7], and a Civic Education[8] component—they may still be beneficial to participants in terms of improving their livelihoods in these other dimensions. Here, previous evidence seems to suggest a substantial impact on health outcomes especially from the health component of the program (Blunch, 2013). Since these programs are relatively inexpensive on both the supply and the demand side—by typically using voluntary teachers (frequently day-time teachers from the formal education system) and meeting just a few hours a few times a week, in the evening (thus allowing learners to work during the day)—even modest returns in terms of improved outcomes of participants are likely to be cost-effective.[9]

Third, public health campaigns in Ghana (for vaccination-drives, HIV/AIDS information, etc.) frequently incorporate banners in the public sphere—frequently in English, the official language of Ghana. There is therefore a case to be made that English reading skills[10] might be more important for affecting health outcomes, including avoidance of a teenage pregnancy, than Ghanaian reading skills.

In response to these issues, the analysis in this paper focuses on the role and interplay of four types of human capital as potential determinants of teenage pregnancy in Ghana: Ghanaian and English reading skills, formal educational attainment, and adult literacy program participation. Allowing for human capital to potentially be endogenous, the analysis first pursues an instrumental variables strategy, using several alternative sets of identifying instruments. As this indicates that skills and schooling may be treated as predetermined to whether a teenage girl has experienced a pregnancy, the main analysis continues with OLS/linear probability models, and also offers matching as an alternative estimation method to corroborate the main results.

## 2. Data

The Ghana Living Standards Survey is a nationally representative multi-purpose household survey, the fifth round of which (GLSS 5, carried out in 2005/06) is used for the analyses in this paper. The household survey contains information on fertility, formal educational attainment,

[7] Topics include cocoa farming, maize cultivation, dry season farming, basket weaving, animal husbandry, bee-keeping, oil palm cultivation, borrowing money for work, hygienic way of preserving and selling fish, farm extension services, pottery, and soap making.
[8] Topics include taxation, bushfires, interstate succession law, child labour, chieftaincy, community empowerment, and expensive funerals.
[9] See Blunch (2013) for a cost-benefit analysis of adult literacy programs in Ghana in the context of child mortality outcomes.
[10] Again, most likely obtained from the formal education system and only to a lesser degree from participating in adult literacy programs.

participation in adult literacy programs, reading skills, as well as information on background variables such as age, parental education, marital status, and tribal association/ethnicity, which are also important factors in analyses of human capital processes.[11]

Teenage pregnancy, the dependent variable for the analysis here, is constructed from the fertility module. This module includes information on reproductive outcomes including whether or not a woman has ever experienced a pregnancy and the number of children ever born and ever deceased to a woman (12 to 49 years old) but not the dates of these events. However, since the estimation sample is restricted to teenage girls, only, the event will not have happened very far back in the past, though it is still possible that the explanatory variables, some of which are current, may be poor predictors. If a teenage girl has recently participated in an adult literacy program, for example, this of course has no impact on the past pregnancies. Similarly, an earlier pregnancy might have induced the girl to participate in the program in order to be able to prevent future pregnancies. In turn, all of these are additional reasons to pursue the IV strategy discussed in the following section. I construct a binary measure which takes the value one if the teenage girl has ever been pregnant and zero otherwise.

Starting with the focal explanatory variables, formal educational attainment is constructed as a set of three binary variables, corresponding to the completion of primary school, middle/junior secondary school, and secondary and above[12] (with the reference category consisting of individuals who never attended school or have primary incomplete). Adult literacy program participation is a binary measure, stating whether an individual has ever attended an adult literacy program. A problem with this 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 taken full effect. This would lead to a downward bias in the estimated impact. Similarly, the adult literacy program could have been completed long ago, possibly leading to depreciation of any skills or knowledge obtained from program participation. Also, the quality and content of adult literacy programs may vary across time or across areas. Unfortunately, the GLSS 5 does not have additional information about timing of participation or about quality or content of the program (though the use of geographic fixed effects in the estimations goes some way towards controlling for program quality and/or content). As these measurement issues may all be considered specific types of measurement error, these issues in

[11] Descriptive statistics for the analysis samples are reported in Table 1.
[12] Eight girls report having completed "other education." These are dropped since it is not clear what "other education" is.

and by themselves point towards the use of instrumental variables methods for this application (as measurement error is one of the possible sources of the larger issue of endogeneity—which the use of instrumental variables is one way of remedying).

Ghanaian and English reading proficiency is constructed from the education module. The question on English reading skills is: "Can (NAME) read a letter in English?" while the question on Ghanaian reading skills is: "In what Ghanaian language can (NAME) read a letter?" The subjective nature of these questions and the use of indirect reports are sources of concern—consistency checks elsewhere, however, give some measure of confidence in this information (Blunch, 2014).

Additional controls include region of birth,[13] region of residence, and cluster fixed effects capture economic conditions specific to the area (as well as everything else related to the region or community in question). The inclusion of current residence fixed-effects (in two alternative flavours, including either regional or cluster fixed effects) help control for possible non-random program placement on the part of government officials or other decision makers in charge of allocating adult literacy programs across Ghana.[14] For example, the North and Upper East and Upper West regions suffer from both less availability and lower quality of educational facilities than the rest of Ghana, as do rural areas versus urban areas.

Socioeconomic background, perhaps most importantly parental education, is also important. For example, more educated parents are both a source of health/contraceptive knowledge and also more likely to instil in their children relatively stronger preferences for avoiding teenage pregnancies than are less educated parents. Additional factors include age, economic factors such as labour market and general economic conditions in the area of residence and ethnicity/tribal association, which is included to capture cultural differences in attributes regarding teenage pregnancy. Cultural and contextual factors are also captured by the region and cluster fixed-effects mentioned earlier.

From the previous discussion, since the survey module pertaining to reproductive behaviour was administered to women between 12 and 49 years of age, and since the analysis here considers teenage girls, this gives the first implicit sample restriction—namely to girls 13-19 years of age (both included). Together with the fact that information on whether or not ever

[13] In addition to having been born in one of the ten different regions in Ghana, the variable also allows for being born abroad (three categories: Other ECOWAS, Other African (than ECOWAS), Outside Africa).
[14] Though it is of course possible that individuals have attended the adult literacy program in another region or community—though the inclusion of region of birth fixed effects in addition to either region of residence or cluster of residence fixed effects seem to address this to some extent, at least at the regional level.

having experienced a pregnancy also must be available, this yields an initial sample of 2,757 observations. Some explanatory variables are missing for some observations, which causes a further drop in the sample size in arriving at the final, effective analysis sample of 2,746 teenage girls for the full sample. From the resulting table of descriptive statistics (see Table 1, below) teenage pregnancy is more common in rural areas, where about 10.6 percent of teenagers have experienced a pregnancy—as compared to about 8.5 percent for the full sample. Similarly, at 9.4 percent, having experienced a pregnancy is more common among teenage girls with less than primary completed than among teenage girls in general (full sample: 8.5 percent). The former of these findings are in line with the fact that rural girls also have far less education on average; for example, the group of girls who has completed less than primary accounts for only about 40.6 percent of the full sample but about 51.5 percent of the rural girls. The finding that teenage girls with less than primary completed in rural areas have a very similar incidence of ever having experienced a pregnancy (about 10.5 percent) to rural teenage girls in general (about 10.6 percent) is both an indication of the poor quality of formal education in rural areas and the relative effectiveness of adult literacy programs in helping avert teenage pregnancies through the reproductive messages taught therein (since the adult literacy program participation rate among the latter group—which is also, indeed, the target group of these programs—is about 4.1 percent, as compared to only 2.5 percent of rural girls overall). Again, these are merely descriptive statistics—to validly assess the relative effectiveness of adult literacy programs in averting teenage pregnancies requires a multivariate analysis, the framework of which is discussed in the next section.

[Table 1 about here]

## 3. Methodology

This section reviews the methodology applied in this paper. First, a simple framework in the human capital tradition is outlined; this is then followed by a discussion of identification strategies and other issues related to the empirical analysis.

### *Conceptual Framework*

The traditional economic approach to fertility views the fertility decision as a rational choice, essentially being a matter of how the demand for children is affected by a host of other factors.

These factors include the relative price of children or, similarly, the opportunity cost of child rearing, the cost of fertility regulation (including psychic costs as well as cost involved with finding and using a given method), as well as the expected future economic opportunities.[15] For instance, if the demand for female labour in the labour market increases, so that female wages increase, women's demand for children is expected to decline. Similarly, if economic prospects are improving more generally, investment in education is expected to increase—thus again decreasing the demand for children and therefore also decreasing teenage fertility. Relatedly, Grossman developed his health human capital model to examine the relationship between the education and the health status of an individual (Grossman, 1972). In the original Grossman model an individual obtains utility from one's own health and the education affects run from own education (as a whole that is, not considering the components of education, such as reading skills) to own health. Here, one might consider teenage pregnancy a specific type of (averse) health (outcome), with especially high risk involved for younger girls. Similar to the rational choice fertility framework this framework once again assumes certainty (perfect foresight) and perfect information.

There are several reasons why the rational choice fertility or the health human capital models may yield overly simplified (or even somewhat unrealistic) conceptual frameworks in their pure forms, especially when the specific type of fertility examined is teenage fertility. First, teenagers may not be all that well informed, so that the notion of fertility as a rational choice may seem a bit of a stretch when considering this particular group—especially if thinking about potentially important factors such as future labour market conditions, including wages. Second, the traditional economic view that fertility behaviour is simply a consequence of rational choice is also at odds with the concept of 'natural fertility' in the demography literature—with related practices, which additionally may be associated with social, economic, and/or cultural conditions in a given society (Easterlin, 1986: 517-18). Third, rather than considering "education" per se, it would seem useful to consider also its component parts, such as reading skills, which may be particularly important for the case of teenage pregnancy as a specific health outcome of interest.

Based on this discussion a combined and augmented rational choice fertility health human capital framework, specifically adapted to the Ghanaian context, may serve as a more useful conceptual framework to better understand the factors governing teenage pregnancy in Ghana than either of the traditional, pure rational choice or the health human capital frameworks.

[15] One of the seminal papers here is Becker (1960).

Some of the factors in this combined, augmented framework are special, if not unique, to the Ghanaian context—perhaps most importantly, human capital.

To better understand how human capital—including formal and non-formal education, and reading skills and health knowledge—affect teenage pregnancy specifically in the Ghanaian context, while at the same time tying this in with previous related research, it is useful to first recall that previous research has distinguished between two main roles of education for affecting health. The first of these has been termed "allocative efficiency"—denoting the fact that educated individuals combine health inputs more efficiently and therefore are able to produce more and/or better health than uneducated individuals (Rosenzweig, 1995).[16] Here, through their increased health knowledge, more educated teenagers are better able to combine health inputs related to pregnancy avoidance (such as the use of several different types of contraceptives and/or practices simultaneously—for example combining the use of condoms with the use of anti-spermicidal lubricant) to achieve even lower risk of experiencing a teenage pregnancy than if using only one input.

They may also consult the doctor and other health personnel more frequently about the use of contraception and related practices, thereby lowering the risk of experiencing a pregnancy. The second main role of education for affecting the production of health has been termed "productive efficiency," which denotes the fact that the health productivity of an individual—here, a teenage girl—is higher the more education they have (and—since skills result from education—therefore also the more and better skills, they have). More generally, educated individuals can therefore produce more health output—including averting a teenage pregnancy—for a given amount of inputs (Grossman, 1972). For example, teenagers with more education are likely to be better able to produce a higher decrease in the risk of experiencing a pregnancy out of a given set of contraceptive inputs and are also better able to read and accurately follow prescriptions and instructions, for example—thus ensuring more effective contraceptive use.

Similarly, skills work through the individual's consumption possibilities and preferences—including the consumption of contraception and preferences for fertility and/or contraceptive use—in several different ways. Most importantly, an individual's income may increase from participation in schooling activities. This could be due to a direct productivity effect from literacy and numeracy or from socialization or discipline skills obtained from

[16] Specifically examining the case of mother's education and child health—though the concept translates to any type of health related behaviour.

schooling. Alternatively, earnings capacity may increase either from credentialism/diploma effects or signalling (Spence, 1973) obtained from formal education. In addition to affecting the individual's consumption possibilities, participation in schooling activities may also affect the knowledge of—and needs or tastes for—contraception and fertility. Indeed, if an individual were not aware of contraception practices and/or their usefulness, why would s/he demand them in the first place?

Additionally, in terms of skills specifically in the Ghanaian context, there are at least three possible reasons/mechanisms why having more skills are related to a decreased risk of experiencing a pregnancy as a teenager. First, girls who can read are more likely to benefit from public health campaigns, by being able to read flyers, banners and signs with public health messages—many of which pertain to HIV/AIDS, and, therefore, at least indirectly, to undesired pregnancies, as well.[17] This is especially true for English reading skills, since most of these campaigns are done using English (the official/government language). Again, the relevance of this channel does not necessarily require that reading these messages is where the health knowledge is initially created—alternatively, this channel could instead be considered a reinforcement of already acquired health knowledge. Second, in terms of schooling, girls with higher levels of formal schooling are more likely to have experienced sex education in school, as are also participants in adult literacy programs (as discussed in the Introduction, several of the topics here, in addition to basic literacy and numeracy skills, explicitly deal with family planning and related issues). Importantly, even if these programs seem to have obtained fairly dismal outcomes in terms of literacy and numeracy outcomes (Blunch, 2013; Blunch and Pörtner, 2011), participants may still obtain this reproductive health knowledge—so that this effectively becomes "knowledge in time" (as opposed to the health knowledge from childhood schooling, which may have deteriorated to some extent). Third, both girls with (more) formal schooling and adult literacy program participants are likely to both have higher (expected/future) wages, as well as being more likely to currently be working (the so-called "incarceration effect")—and can therefore also be expected to be less likely to experience a teenage pregnancy.

From this discussion, it is clear that skills, formal educational attainment and adult literacy program participation all are potentially endogenous to whether a teenage girl experiences a pregnancy, conceptually (though it is still possible—even likely, due to the nature of the data—that the reading skills, as well as the schooling and adult literacy program

[17] I have witnessed this myself, on numerous occasions in both urban and rural areas in Ghana.

participation may have been obtained prior to becoming pregnant, so that they are at least predetermined, if not exogenous). This will therefore need to be addressed in the empirical strategy (see below).

Socioeconomic background, perhaps most importantly parental education, is also potentially important. For example, more educated parents are both a source of health/contraceptive knowledge and also more likely to instil in their children relatively stronger preferences for avoiding teenage pregnancies than are less educated parents. Additional factors include age, economic factors such as labour market and general economic conditions in the area of residence.

*Estimation Strategy and Issues*

The conceptual framework discussed in the previous subsection suggests that individual skills and schooling can directly affect whether a teenage girl experiences a pregnancy through the acquisition of health knowledge, including contraceptive knowledge, and the relatively higher opportunity costs for girls with more human capital and also suggest additional factors that are potentially important for experiencing a teenage pregnancy and therefore should be included in the empirical specifications.[18] The empirical analysis will examine this relationship, using linear approximations of the optimal teenage pregnancy equation, where human capital is measured by dummy variables for Ghanaian and English reading skills, level of formal schooling completed, and participation in adult literacy programs.

The natural starting point is estimating a regression of having experienced a teenage pregnancy on the explanatory variables discussed in the previous section by OLS[19] (i.e. as a Linear Probability Model, LPM). Further, so as to allow for arbitrary heteroskedasticity, the estimation will be carried out using Huber-White standard errors (Huber, 1967; White, 1980). To allow for the possibility that observations are correlated within communities the standard errors are also adjusted for within-cluster correlation (Wooldridge, 2010).

[18] At a minimum, if these factors are not included, one may systematically over- or underestimate the strength of the human capital-teenage pregnancy relationship.

[19] As is well known, there may be some concern about using OLS, or, in effect, the linear probability model (LPM), when the dependent variable is binary. For example, predicted probabilities may fall outside the (0,1)-range and heteroskedasticity also is present by default. However, it can be argued that the LPM approximates the response probability well, especially if (1) the main purpose is to estimate the partial effect of a given regressor on the response probability, averaged across the distribution of the other regressors, (2) most of the regressors are discrete and take on only a few values and/or (3) heteroskedasticity-robust standard errors are used in place of regular standard errors (Wooldridge, 2010). All three factors seem to work in favour of the LPM for the purposes of the application here.

Since formal educational attainment, adult literacy, and Ghanaian and English reading skills all reflect different dimensions of an individual's human capital, it is possible that this induces multicollinearity in the estimations—with possible implications for the interpretation of the results for these individual dimensions. To explore whether multicollinearity is present I will pursue the two main methods available, namely estimate the pairwise simple correlations and the variance inflation factor (VIF) scores.

After thus presenting the benchmark estimation method, several issues pertaining to the estimation of the human capital-teenage pregnancy relationship need to be addressed—where the two most important issues arguably are estimation samples (including the issue of relevant counterfactuals for program participants (non-participants) and the possible endogeneity of the human capital variables.

Regarding estimation samples, in addition to estimating the human capital-teenage pregnancy relationship for the full sample, there is a compelling case to be made for carrying out the estimations across several subsamples, also. First, girls with particularly low—or no—completed education seem particularly vulnerable. Also, with one focus of this paper being on adult literacy programs and since, methodologically, the correct counterfactual for teenage girls who have attended adult literacy programs is uneducated teenage girls who would potentially enrol in these programs, there is a case to be made that only girls who have not completed primary education (or higher) should be included in the estimation sample when one is explicitly interested in the effect of adult literacy program participation on teenage pregnancy. Further, from a policy perspective, these girls would also seem to be more at a disadvantage in terms of both human capital and providing for their children and should therefore be of special interest to policy makers. A separate analysis for teenage girls with less than primary education completed therefore seems warranted. Second, rural areas suffer more from lack of adequate provision of educational facilities than urban areas. Girls from rural areas are also more likely to suffer from social exclusion following a pregnancy as teenagers, due to the more traditional social norms in rural areas. A separate analysis for rural areas therefore seems warranted, as well. In addition to these two sub-samples, it would seem potentially useful to combine the two, i.e. to examine girls with no formal education from rural areas in a separate analysis—these girls may be particularly vulnerable, both in terms of a potential pregnancy and having particularly low human capital.

One potentially important econometric issue pertaining to these estimations (which was also emphasised in the conceptual framework) is that reading skills, formal educational

attainment and/or adult literacy program participation all may potentially be endogenous for teenage pregnancy. In response to this, I initially pursue an instrumental variable (IV) / Two-Stage Least Squares (2SLS) strategy using two of the main IV strategies that have been used in recent years: either using as IVs (1) various combinations of time of year, birth cohort, and/or geographical area of birth dummies to capture variation in institutional factors relevant for human capital accumulations such as compulsory schooling laws or expansion of educational programs (Angrist and Krueger, 1991; Duflo, 2001) or (2) variables for proximity or exposure to educational institutions in the local area (Card, 2001). This strategy has previously been applied to studies for Ghana (Blunch, 2013; Blunch and Pörtner, 2011), where more information about these instruments can be found.[20]

While the empirical strategy outlined here allows educational attainment to be endogenous, it is still possible that an estimation strategy that treats educational attainment as exogenous is ultimately preferred. There are two main reasons why this might be the case: either the instruments are not necessary for a given application (even if they are valid) or they are invalid—or both. First, as was noted above, conceptually educational attainment could potentially be regarded as predetermined (if not exogenous) for this application—if this is indeed the case and the two other potential sources of endogeneity bias, namely omitted variables and measurement error are also not too strong, it is possible that the Wu–Hausman test for endogeneity will pass (i.e. cannot reject exogeneity). If this is the case it would seem prudent to use ordinary least squares instead, since this would then be efficient relative to IV/2SLS. In this case, effectively, IV/2SLS is not called for, i.e. "the medicine would be worse than the disease." Additionally, however, it is useful to combine the test for endogeneity with examining the validity of the instruments—for at least two reasons. First, even in the presence of endogeneity (as determined by the Wu-Hausman test) the OLS results may still be preferred if the instruments are not valid. Second, if the Wu-Hausman test fails to reject exogeneity, it is still useful to determine whether that is merely due to the instruments being invalid to begin with.

The validity of the instruments, in turn, has two dimensions. First, if the instruments in the first stage are weak, the use of IV/2SLS is questionable to begin with (Bound, Jaeger, and Baker 1995). This is typically tested as a joint (F-) test of statistical significance of the identifying/excluded instruments (that is, excluded from the second-stage regression) in the first-

[20] Due to space constraints, as well as since this strategy ultimately turns out not to be warranted for the analysis here (as revealed by the results from IV/2SLS specification tests), the details are omitted here (but are available upon request).

stage regression of the endogenous variable(s). In particular, based on the work in Staiger & Stock (1997), the "Rule of 10" has found wide application, meaning that for a given instrument to be strong, it should yield an F-statistic of 10 or more (in practice, one would ideally like an even higher F-statistic, especially when there is more than one endogenous variable/instrument). Second, if there are more instruments than there are potentially endogenous variables, it is possible to test for the validity of the over-identifying instruments using the Hansen–Sargent test for over-identification. If they do not pass this test, the instruments are not valid, either. This can be tested as a chi-square test of N times the $R^2$ from a regression of the IV residuals on the full set of instruments (Wooldridge, 2010: 134–136). So, passing both the tests for weak instruments and for over-identification (*if* there are more instruments than potentially endogenous variables) are necessary but not sufficient conditions for the instruments to be valid.

Even if the tests for endogeneity would seem to suggest that the IV estimation strategy is not empirically relevant for this application, it still seems prudent to extend the analysis with additional alternative estimation strategies to help support the (potentially biased) OLS/LPM results, if feasible. For the application here, additionally using matching on observables—whereby the treatment effect from a binary treatment can be estimated—seems particularly useful as an additional, alternative estimation strategy.[21] The treatment effect of formal educational attainment will be estimated for the full and rural samples, whereas the effect of adult literacy program participation is estimated for the sample of teenage girls with less than primary education completed. In both cases I use Mahalanobis covariate matching[22], implementing the robust analytical standard errors proposed by Abadie and Imbens (2006) to correct for possible heteroskedasticity and impose common support (as also suggested by Rosenbaum and Rubin, 1983).

## 4. Results

This section reviews the results from the multivariate models, focusing at the results for educational attainment, adult literacy program participation and Ghanaian and English reading

[21] The variables used for the matching include all explanatory variables from the previous regressions except the possibly endogenous explanatory variables (formal educational attainment, adult literacy program participation, and (English and Ghanaian) reading skills).
[22] For details on Mahalanobis matching, see for example Rosenbaum and Rubin (1985).

skills.[23] I will start by first determining the preferred estimation method—where the "candidates" are OLS and 2SLS—and then discuss the results.

The results from specification tests indicate that the use of OLS, i.e. the Linear Probability Model (LPM) is preferable for this application overall (Appendix, Tables A1-A3). First, the results from Wu-Hausman tests indicate that reading skills, schooling, and adult literacy program participation are not endogenous to experiencing a teenage pregnancy (Appendix, Tables A1-A3). It therefore seems prudent to use OLS, since this will yield more efficient estimates if IV is not called for. But is this merely due to having "bad" instruments? The results from the F-tests of the joint significance of the identifying instruments from the first stage of the 2SLS procedure indicate that the identifying instruments are certainly not strong across all the potentially endogenous variables and different specification, though they are sometimes quite strong—with significance levels of 1 percent or better in many cases and also quite high F-statistics in a few cases (Appendix, Tables A1-A3)—with several of the first stage F-statistics exceeding the required threshold of 10 (the "Rule of 10") suggested in Staiger and Stock (1997).

Second, the identifying instruments also pass Hansen's (1982) J-test for over-identification thus supporting their validity in terms of the second stage results, as well (Appendix, Tables A3-A5).[24] Together, this gives some credence to the finding of lack of endogeneity not being entirely due to the instruments merely not being valid in the first place—though they certainly could be (even) stronger predictors of the potentially endogenous variables.

Especially the Access X Individual Characteristics IV specification, and here again especially for the specification with cluster fixed effects, for the target group of women with less than primary completed—which is probably the preferred specification among all the specifications—comes out particularly strong in terms of the predictive power of the instruments from the first stage (and also passes the over-identification test). But some of the other specifications certainly suffer from weak instruments.

The reading skills, schooling, and adult literacy program participation coefficients from the OLS ever pregnant regressions are presented in Table 2 for the four estimation samples: full sample, girls with less than primary completed, rural sub-sample, and girls with less than primary completed from rural areas. For each of these estimation samples, two sets of results are

[23] The full set of results is shown in Tables A4 and A5 in the Appendix.
[24] It should be remembered that if the instruments are valid then these tests will pass with high probability; therefore, the tests are necessary but not sufficient for the instruments to be valid.

presented (with two different specifications of geographical fixed effects, one with regional fixed effects and one with cluster fixed effects—where the latter can be considered "cleaner," though at the expense of less precision). The first set of results includes all variables, except reading skills, and the second set of results adds Ghanaian and English reading skills. Since the main focus of this paper is on the impact of skills and schooling, the results from the additional controls (as described earlier) are omitted in this table for brevity (but are included in the Appendix, Tables A4 and A5).

[Table 2 about here]

Starting with the results that can be compared with the previous literature—namely the results for formal education—it is clear from the first column for both the full and the rural samples in Table 2 that formal educational attainment is highly negatively associated with experiencing a teenage pregnancy in both substantive and statistical terms. For the full sample, for example, having completed middle school or junior secondary is associated with about a minus five percentage-points lower probability of experiencing a teenage pregnancy (relative to the reference group, less than primary completed). The results for having completed secondary and above is even more pronounced, at about minus 12 percentage-points. Both of these estimated coefficients are also quite precisely measured and therefore highly statistically significant (at a one percent level of statistical significance).

This is in line with previous findings from the related literature, both for Sub-Saharan Africa (Palermo and Peterman, 2009; Gupta and Mahy, 2003), as well as for (the much larger literature on) Asia (Hirschman and Rindfuss, 1980; Bloom and Reddy, 1986; and Basu, 1993). Gupta and Mahy (2003), for example, examine (among other things) the determinants of pregnancy among adolescent girls (18-24 years of age) before age 18 for eight different Sub-Saharan African countries and find that in all countries having eight or more years of education is associated with being between about 50 and 90 percent less likely to experience a pregnancy before age 18 than if having no education completed (the reference category). For Ghana, in particular, it is found that having 1-7 years of education and having eight or more years of education, respectively, are associated with being between about 25 and 73 percent less likely to experience a pregnancy before age 18 than if having no education completed (the reference category). This is consistent with the results found in Table 2, given the differences in the reference categories, the definition of the educational categories, and the dependent variables (i.e. actual birth versus pregnancy).

Once reading skills are introduced, however, this effect from formal educational attainment decreases or even disappears, in some cases—as can be seen from the second column for both the full and the rural samples in Table 2. In turn, reading skills seem to matter both in statistical and substantive terms (except for the rural subsample and for girls with no formal education from rural areas for the specification with cluster fixed effects). For the full sample, for example (specification with regional FEs) at about minus 5.5 percentage-points, the estimate for English reading is about double that of Ghanaian reading, at 1.7 percentage-points. This is consistent with the observation that most public health campaigns in Ghana are carried out in English (which is also the "official"/government language more generally). The finding that reading skills matter, in turn, highlights both the importance of skills (and output from education) as opposed to merely school participation, per se, and therefore also the importance of school quality: if quality is low, actually obtaining these reading skills is not a given even if one attends school. This finding therefore also highlights part of the contribution of this study, namely by shifting focus from merely education to the skills, including reading skills, obtained from that education.

The estimates for adult literacy program participation are virtually zero, in statistical as well as substantive terms, for the full sample, as well as the rural subsample. Again, this is not really surprising as it may be claimed that the correct estimation sample here should focus exclusively on girls with no formal education (since these are the relevant potential participants of adult literacy programs in the first place). Alas, when restricting the estimation sample accordingly, there are now substantively large, negative and statistically significant effects on teenage pregnancy—at about minus 9 percentage-points for the specification using regional FEs, while somewhat higher (but also less precisely measured) for the specification using cluster fixed effects, at about 13-14 percentage-points (and still statistically significant, though only at 10 percent).

While the results for the full specification with the full set of education and reading skills variables—where the estimated coefficients are (or become) small(er) speak strongly to a weak(er) relationship with teenage pregnancy in *substantive* terms—it is useful to explore whether multicollinearity can help explain the lack of *statistical* significance occurring in some cases (sometimes occurring after the reading skills are included). While it is not clear exactly "how high is high," though a VIF of 10 sometimes is chosen as the (arbitrary) threshold (Wooldridge, 2009: 96-99), having simple correlations ranging from about 0.10 to about 0.36 in

absolute value for the main simple correlations of interest here (namely those between, on the one hand, educational attainment and, on the other, the outcomes of this educational attainment, namely Ghanaian and English reading skills) and all individual VIF scores below 4 (and the average VIF scores below 3) multicollinearity does not seem to be a serious issue here.[25]

To support the OLS results from the main analysis I also estimate the impact of formal educational attainment and adult literacy program participation on teenage pregnancy using matching on observables as an alternative identification strategy. Tables A6 and A7 in the Appendix present the results from matching on observables, using Mahalanobis matching and incorporating the heteroskedasticity-consistent analytical standard errors proposed by Abadie and Imbens (2006), for a range of treatments and estimation samples. For formal educational attainment a set of three different treatments are considered: ever having attended school, completing primary and above, and completing junior secondary and above (Table A6), while for adult literacy program participation one treatment is considered: having ever attended an adult literacy program (Table A7).[26]

From the tables it appears that the matching results are consistent with the previous (regression/OLS) results overall. There are some differences in magnitude and statistical significance, of course. At about 9-11 percentage-points, the estimates for adult literacy program participation, for example, are both slightly higher and much more precisely measured than was the case for OLS using regional FEs (between 7-9 percentage-points), though still not as high as the OLS results using cluster FEs (about 13-14 percentage-points).

To examine the robustness of the main results I also perform several sensitivity analyses.[27] First, it would seem useful to examine in (even) more detail the possible heterogeneity with respect to the adult literacy variable—for example, along the lines of Blunch and Pörtner (2011), which finds that the effect of adult literacy program participation on household economic welfare decreases as the education level of the household goes up. It turns out, however, that no girls with junior secondary school or higher has also attended an adult literacy program. The only group that can be examined in addition to the "less than primary completed" group from the main analysis, therefore, is girls with primary completed. From these

[25] Due to space constraints the results tables are not shown here—but they are available upon request.
[26] For both types of education treatments the variables used for the matching include all explanatory variables from the previous regressions except the possibly endogenous explanatory variables (formal educational attainment, adult literacy program participation, and (English and Ghanaian) reading skills).
[27] Due to space constraints the results tables for the sensitivity analyses are not shown here—but they are available upon request.

results the estimated coefficient for adult literacy program participation is statistically indistinguishable from zero, supporting the conjecture in Blunch and Pörtner (2011) of adult literacy programs and formal education being highly substitutable. Second, it is possible that the married women in this sample face very different exposures and risks, so that one might argue that these women should be excluded from the estimation sample. To examine whether inclusion of the married women affect the results, I re-estimated the models without these (84) women in the estimation sample. The results were very similar overall, though the results for adult literacy program participation were slightly strengthened. Most notably here the results for the women from the full sample with no formal education completed using cluster fixed effects saw an increase in the coefficient estimate of about 3 percentage points, to about (minus) 16 percentage points, in total—which further was statistically significant at a 10 percent level of significance.

In summary, the main impression from the results obtained here is that formal educational attainment and adult literacy program participation both have substantial (negative) effects on the probability of experiencing a teenage pregnancy, though the latter is particularly pronounced among girls who have not completed primary school. Additionally, there seems to be a separate, distinct effect from reading skills, especially English reading skills.

Notably, had one relied only on the full sample results, the conclusion would have been that adult literacy programs are ineffective in decreasing teenage pregnancy. In turn, this also highlights the importance of selecting appropriate estimation samples in applied econometrics analysis more generally.

## 5. Conclusion

This paper examines the association between the human capital of Ghanaian teenage girls and their probability of having ever experienced a pregnancy. The relationship between human capital was explored along several dimensions of human capital, including formal educational attainment, as well as adult literacy program participation and English and Ghanaian reading skills.

Altogether, though as always there necessarily are caveats involved when conducting empirical research—especially when attempting to estimate causal effects—as was also discussed at length above, all the results taken together, as a whole, indicate a substantial effect of reading skills, especially English reading skills, and of adult literacy program participation on

reducing teenage pregnancy in Ghana. In turn, this adds to the previous literature on the relationship between formal educational attainment and teenage pregnancies. Arguably, health knowledge, especially related to family planning and contraceptive use—which, as previously discussed, is an integral part of the adult literacy program curriculum in Ghana—explains part of the effect of adult literacy program participation on child mortality established here.

To put these results into perspective, primary education in Ghana is six years, while adult literacy programs are 21 months (in the case of the most widespread program, supplied by the Non-formal Education Department under the Ministry of Education) and with less frequent attendance, say, two or three evenings a week for a few hours. Hence, the effective time spent undertaking an adult literacy program is far less than completing primary school and yet appears to have a substantial effect on teenage pregnancy (if not on literacy, Blunch, 2013; Blunch and Pörtner, 2011).

Where does this effect from adult literacy program participation on teenage pregnancy come from? Arguably, health knowledge—which, as previously discussed, is an integral part of the adult literacy course curriculum in Ghana—explains part of the effect of adult literacy program participation on teenage pregancy established here. This is especially true for the full specifications estimated previously, since reading skills are controlled for. In turn, this supports findings from previous research that health knowledge is particularly important for health outcomes, and more so than literacy and numeracy (see, e.g., Glewwe, 1999 for a study of maternal skills and schooling and child health in Morocco).

On the other hand, previous research has found that adult literacy programs have not been all that successful in creating literacy (and numeracy) skills—though this is their stated objective—among participants (Abadzi, 1994; Ortega and Rodríguez, 2008), which also seems to be one of the main reasons why adult literacy programs appears to have been abandoned to a large extent in recent years across the developing world, including by major players such as the World Bank (Chowdhury, 1995). Yet, the results here indicate that adult literacy programs may still have something to contribute in terms of improving livelihoods among participants. Importantly, this is true even if improving livelihoods is not in terms of the officially stated program objectives of providing literacy and numeracy skills but rather in terms of enabling girls to avoid experiencing a teenage pregnancy. And thereby enabling them to build a brighter future for themselves and their *future* children.

**References:**


Abadie A. and Imbens, G., 2006, 'Large Sample Properties of Matching Estimators for Average Treatment Effects', *Econometrica*, 74, 235-267.

Abadzi, Helen, 1994, 'What We Know about Acquisition of Adult Literacy: Is There Hope?' World Bank Discussion Paper no. 245, World Bank, Washington, DC.

Angrist, Joshua D. and Alan B. Krueger, 1991, 'Does Compulsory School Attendance Affect Schooling and Earnings?' *Quarterly Journal of Economics*, 106, 979-1014.

Basu, Alaka Malwade, 1993, 'Cultural Influences on the Timing of First Births in India: Large Differences That Add up to Little Difference', *Population Studies*, 47, 85-95.

Becker, Gary S., 1960, 'An Economic Analysis of Fertility', in: George B. Roberts (Ed.) *Demographic and Economic Change in Developed Countries*, Princeton: National Bureau of Economic Research.

Bloom, David E. and Reddy, P. H., 1986, 'Age Patterns of Women at Marriage, Cohabitation, and First Birth in India', *Demography*, 23: 509-23.

Blunch, Niels-Hugo, 2013, 'Staying Alive: Adult Literacy Programs and Child Mortality in Rural Ghana', *World Development*, 42, 114–126.

Blunch, Niels-Hugo, 2014, 'Literacy and Numeracy Production and Education Sector Reform: Evidence from Ghana', *Education Economics*, 22, 209-235.

Blunch, Niels-Hugo and Claus Pörtner, 2005, 'Literacy, Skills and Welfare: Effects of Participation in an Adult Literacy Program', University of Washington Economics Department Working Paper Series, UWEC-2005-23, University of Washington, Seattle.

Blunch, Niels-Hugo and Claus Pörtner, 2011, Literacy, Skills and Welfare: Effects of Participation in Adult Literacy Programs', *Economic Development and Cultural Change*, 60, 17-66.

Card, David, 2001, 'Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems', *Econometrica*, 69, 1127-1160.

Chowdhury, K. P., 1995, 'Literacy and Primary Education', Discussion Paper no. 50, Human Capital and Development Operations Policy, World Bank, Washington, DC.

Duflo, Esther, 2001, 'Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment', *American Economic Review*, 91, 795-813.

Easterlin, Richard A., 1986, 'Economic Preconceptions and Demographic Research: A Comment', *Population and Development Review*, 12, 517-528.

GHS (Ghana Health Service), 2013, '2013 Annual Reproductive and Child Health Report', Accra: Ghana Health Service. http://www.ghanahealthservice.org/downloads/GHS-REproductive_and-Child-Health-Annual-Report-2013.pdf → Accessed August 17, 2016.

Ghana Statistical Service (GSS), Ghana Health Service (GHS), and ICF Macro, 2009, 'Ghana

Demographic and Health Survey 2008: Key Findings', Calverton, Maryland, USA: GSS, GHS, and ICF Macro.

Glewwe, Paul, 1999, 'Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco', *Journal of Human Resources*, 34, 124-159.

Gupta, Neeru and Mary Mahy, 2003, 'Adolescent childbearing in sub-Saharan Africa: Can increased schooling alone raise ages at first birth?' *Demographic Research* Volume 8, Article 4, 93-106.

Heckman, James J., Jora Stixrud, and Sergio Urzua, 2006, 'The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior', *Journal of Labor Economics*, 24, 411-482.

Hirschman, C. and Rindfuss, R., 1980, 'Social, Cultural, and Economic Determinants of Age at Birth of First Child in Peninsular Malaysia', *Population Studies*, 34, 507-18.

Huber, P. J., 1967, 'The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions', in: Lucien M. Le Cam and Jerzy Neyman (Eds.) *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, *Vol. 1*, Berkeley, CA: University of California Press.

Imbens, Guido W., 2004, 'Nonparametric Estimation of Average Treatement Effects Under Exogeneity: A Review', *The Review of Economics and Statistics*, 86, 4-29.

Levine, David I. and Gary Painter, 2003, 'The Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a Within-School Propensity Score Matching Estimator', *The Review of Economics and Statistics*, 85, 884-900.

Ortega, Danile and Francisco Rodríguez, 2008, 'Freed from Illiteracy? A Closer Look at Venezuela's Misión Robinson Literacy Campaign' *Economic Development and Cultural Change*, 57, 1-30.

Palermo, Tia and Amber Peterman, 2009, 'Are Female Orphans at Risk for Early Marriage, Early Sexual Debut, and Teen Pregnancy: Evidence from Sub-Saharan Africa', *Studies in Family Planning*, 40, 101-12.

Ribar, D.C., 1993, 'A Multinomial Logit Analysis of Teenage Fertility and High School Completion', *Economics of Education Review*, 12, 153–164.

Ribar, David, 1994, 'Teenage Fertility and High School Completion', *The Review of Economics and Statistics*, 76, 413–424.

Rindfuss, R.R, S.P. Morgan and K. Offutt, 1996, 'Education and the Changing Age Pattern of American Fertility', *Demography*, 33, 277-290.

Rosenbaum, P., and D. Rubin, 1983, 'The Central Role of the Propensity Score in Observational Studies for Causal Effects', *Biometrika*, 70, 41-55.

Rosenbaum, P., and D. Rubin, 1985, 'Constructing a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity', *American Statistician*, 39, 33-38.

Staiger, D. and J.H. Stock, 1997, 'Instrumental Variables Regression with Weak Instruments', *Econometrica*, 65, 557–586.

Upchurch, Dawn M., Lee A. Lillard, and Constantijn W.A. Panis, 2002, 'Nonmarital

Childbearing: Influences of Education, Marriage, and Fertility', *Demography*, 39, 311-329.

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

Wooldridge, Jeffrey M., 2009, *Introductory Econometrics*, Fourth Edition, Mason, Ohio: The South-Western CENGAGE Learning.

Wooldridge, Jeffrey M., 2010, *Econometric Analysis of Cross-Section and Panel Data*, Second Edition, Cambridge, Massachusetts: The MIT Press.

**Table 1. Descriptive Statistics for Estimation Samples**

| | *Rural + urban:* | | | | *Rural, only:* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Full sample:* | | *Less than Primary Completed:* | | *Full sample:* | | *Less than Primary Completed:* | |
| | Mean: | Std Dev: | Mean: | Std Dev: | Mean: | Std Dev: | Mean: | Std Dev: |
| **Dependent variable:** | | | | | | | | |
| Ever pregnant | 0.085 | 0.278 | 0.094 | 0.291 | 0.106 | 0.308 | 0.105 | 0.307 |
| **Educational attainment:** | | | | | | | | |
| Adult literacy program participation | 0.015 | 0.120 | 0.029 | 0.169 | 0.025 | 0.156 | 0.041 | 0.197 |
| Never attended school or some primary | 0.406 | 0.491 | 1.000 | 0.000 | 0.515 | 0.500 | 1.000 | 0.000 |
| Primary | 0.336 | 0.473 | 0.000 | 0.000 | 0.333 | 0.472 | 0.000 | 0.000 |
| Junior secondary | 0.227 | 0.419 | 0.000 | 0.000 | 0.144 | 0.351 | 0.000 | 0.000 |
| Secondary and above | 0.030 | 0.172 | 0.000 | 0.000 | 0.008 | 0.090 | 0.000 | 0.000 |
| **Skills:** | | | | | | | | |
| Reads Ghanaian | 0.467 | 0.499 | 0.196 | 0.397 | 0.374 | 0.484 | 0.168 | 0.374 |
| Reads English | 0.705 | 0.456 | 0.391 | 0.488 | 0.600 | 0.490 | 0.337 | 0.473 |
| **Age cohort:** | | | | | | | | |
| Age 13 | 0.168 | 0.374 | 0.278 | 0.448 | 0.181 | 0.385 | 0.267 | 0.443 |
| Age 14 | 0.145 | 0.352 | 0.194 | 0.395 | 0.153 | 0.360 | 0.196 | 0.397 |
| Age 15 | 0.165 | 0.371 | 0.183 | 0.387 | 0.168 | 0.374 | 0.191 | 0.394 |
| Age 16 | 0.152 | 0.359 | 0.115 | 0.320 | 0.159 | 0.366 | 0.114 | 0.318 |
| Age 17 | 0.113 | 0.316 | 0.067 | 0.250 | 0.108 | 0.311 | 0.070 | 0.255 |
| Age 18 | 0.148 | 0.355 | 0.091 | 0.288 | 0.134 | 0.341 | 0.096 | 0.295 |
| Age 19 | 0.110 | 0.313 | 0.072 | 0.259 | 0.097 | 0.296 | 0.065 | 0.247 |
| **Parental education:** | | | | | | | | |
| Mother none | 0.578 | 0.494 | 0.738 | 0.440 | 0.696 | 0.460 | 0.787 | 0.409 |
| Mother primary | 0.121 | 0.326 | 0.106 | 0.308 | 0.118 | 0.323 | 0.100 | 0.300 |
| Mother above primary | 0.279 | 0.449 | 0.134 | 0.341 | 0.176 | 0.381 | 0.100 | 0.300 |
| Mother don't know/missing | 0.022 | 0.146 | 0.022 | 0.148 | 0.010 | 0.101 | 0.013 | 0.112 |
| Father none | 0.374 | 0.484 | 0.553 | 0.497 | 0.511 | 0.500 | 0.630 | 0.483 |
| Father primary | 0.073 | 0.260 | 0.072 | 0.258 | 0.070 | 0.256 | 0.061 | 0.240 |
| Father above primary | 0.514 | 0.500 | 0.329 | 0.470 | 0.395 | 0.489 | 0.280 | 0.449 |
| Father don't know/missing | 0.040 | 0.196 | 0.046 | 0.210 | 0.024 | 0.152 | 0.029 | 0.168 |
| **Marital status:** | | | | | | | | |
| Married or informal union | 0.050 | 0.218 | 0.064 | 0.245 | 0.067 | 0.250 | 0.072 | 0.259 |
| **Geographical info:** | | | | | | | | |
| Urban (residence) | 0.424 | 0.494 | 0.278 | 0.448 | 0.000 | 0.000 | 0.000 | 0.000 |
| Western (residence) | 0.096 | 0.294 | 0.096 | 0.295 | 0.106 | 0.307 | 0.105 | 0.307 |
| Central (residence) | 0.092 | 0.289 | 0.075 | 0.264 | 0.098 | 0.297 | 0.075 | 0.264 |
| Greater Accra (residence) | 0.146 | 0.353 | 0.089 | 0.284 | 0.029 | 0.168 | 0.031 | 0.174 |
| Eastern (residence) | 0.145 | 0.352 | 0.118 | 0.323 | 0.169 | 0.375 | 0.128 | 0.334 |
| Volta (residence) | 0.079 | 0.270 | 0.085 | 0.279 | 0.092 | 0.289 | 0.077 | 0.267 |
| Ashanti (residence) | 0.173 | 0.378 | 0.135 | 0.341 | 0.161 | 0.368 | 0.121 | 0.326 |
| Brong-Ahafo (residence) | 0.104 | 0.305 | 0.117 | 0.322 | 0.111 | 0.314 | 0.104 | 0.306 |
| Northern (residence) | 0.087 | 0.281 | 0.154 | 0.361 | 0.110 | 0.313 | 0.187 | 0.391 |
| Upper West (residence) | 0.034 | 0.182 | 0.054 | 0.226 | 0.052 | 0.223 | 0.066 | 0.248 |
| Upper East (residence) | 0.044 | 0.206 | 0.076 | 0.265 | 0.072 | 0.259 | 0.104 | 0.306 |
| Western (birth) | 0.094 | 0.292 | 0.090 | 0.286 | 0.097 | 0.296 | 0.097 | 0.296 |
| Central (birth) | 0.096 | 0.295 | 0.085 | 0.278 | 0.103 | 0.304 | 0.086 | 0.281 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Greater Accra (birth) | 0.096 | 0.295 | 0.060 | 0.238 | 0.033 | 0.178 | 0.030 | 0.170 |
| Eastern (birth) | 0.154 | 0.361 | 0.118 | 0.323 | 0.164 | 0.370 | 0.119 | 0.324 |
| Volta (birth) | 0.094 | 0.292 | 0.097 | 0.296 | 0.095 | 0.294 | 0.087 | 0.281 |
| Ashanti (birth) | 0.170 | 0.375 | 0.123 | 0.328 | 0.151 | 0.358 | 0.108 | 0.311 |
| Brong-Ahafo (birth) | 0.098 | 0.297 | 0.102 | 0.303 | 0.101 | 0.301 | 0.085 | 0.279 |
| Northern (birth) | 0.093 | 0.290 | 0.163 | 0.370 | 0.116 | 0.320 | 0.195 | 0.397 |
| Upper West (birth) | 0.041 | 0.198 | 0.064 | 0.246 | 0.055 | 0.227 | 0.072 | 0.259 |
| Upper East (birth) | 0.051 | 0.219 | 0.086 | 0.281 | 0.078 | 0.269 | 0.116 | 0.320 |
| Other ECOWAS (birth) | 0.010 | 0.097 | 0.005 | 0.074 | 0.006 | 0.075 | 0.005 | 0.070 |
| Other Africa (birth) | 0.002 | 0.048 | 0.001 | 0.028 | 0.002 | 0.050 | 0.000 | 0.000 |
| Outside Africa (birth) | 0.002 | 0.048 | 0.005 | 0.071 | 0.000 | 0.000 | 0.000 | 0.000 |
| N | 2,746 | | 1,192 | | 1,474 | | 813 | |

*Notes:* Calculations incorporate sampling weights and also adjust for within-community correlation/clustering (Wooldridge, 2010).
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table 2. Skills and Schooling Coefficients from OLS Ever Pregnant Regressions**

| | *Rural + urban:* | | | | *Rural, only:* | | | |
|---|---|---|---|---|---|---|---|---|
| | *Full sample:* | | *Less than Primary Completed:* | | *Full sample:* | | *Less than Primary Completed:* | |
| | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* |
| **Regional FEs:** | | | | | | | | |
| Adult literacy program participation | -0.058 | -0.058 | -0.087*** | -0.087*** | -0.07 | -0.070* | -0.089*** | -0.090*** |
| | [0.038] | [0.037] | [0.029] | [0.028] | [0.044] | [0.042] | [0.032] | [0.032] |
| Primary | -0.013 | 0.017 | | | -0.022 | 0.012 | | |
| | [0.011] | [0.012] | | | [0.017] | [0.019] | | |
| Middle/JSS | -0.048*** | -0.01 | | | -0.068** | -0.021 | | |
| | [0.017] | [0.016] | | | [0.029] | [0.030] | | |
| Secondary and above | -0.123*** | -0.083*** | | | -0.185*** | -0.133*** | | |
| | [0.029] | [0.028] | | | [0.035] | [0.033] | | |
| Ghanaian Reading | | -0.017* | | -0.002 | | -0.021 | | 0.006 |
| | | [0.010] | | [0.012] | | [0.015] | | [0.021] |
| English reading | | -0.055*** | | -0.038*** | | -0.058*** | | -0.039** |
| | | [0.013] | | [0.012] | | [0.017] | | [0.016] |
| $R^2$ | 0.501 | 0.508 | 0.516 | 0.519 | 0.521 | 0.529 | 0.508 | 0.511 |
| N | 2,746 | 2,746 | 1,192 | 1,192 | 1,474 | 1,474 | 813 | 813 |
| **Cluster FEs:** | | | | | | | | |
| Adult literacy program participation | -0.062 | -0.064 | -0.127 | -0.128 | -0.074 | -0.079 | -0.136* | -0.137* |
| | [0.065] | [0.063] | [0.083] | [0.082] | [0.076] | [0.074] | [0.081] | [0.080] |
| Primary | -0.005 | 0.026 | | | -0.015 | 0.017 | | |
| | [0.015] | [0.016] | | | [0.023] | [0.024] | | |
| Middle/JSS | -0.047** | -0.008 | | | -0.064 | -0.02 | | |
| | [0.021] | [0.023] | | | [0.039] | [0.041] | | |
| Secondary and above | -0.123*** | -0.080** | | | -0.165*** | -0.121*** | | |
| | [0.036] | [0.038] | | | [0.046] | [0.046] | | |
| Ghanaian Reading | | -0.011 | | -0.006 | | -0.015 | | -0.012 |
| | | [0.014] | | [0.033] | | [0.021] | | [0.049] |
| English reading | | -0.066*** | | -0.022 | | -0.064** | | -0.021 |
| | | [0.019] | | [0.028] | | [0.026] | | [0.036] |
| $R^2$ | 0.6 | 0.606 | 0.665 | 0.666 | 0.621 | 0.626 | 0.644 | 0.644 |
| N | 2,746 | 2,746 | 1,192 | 1,192 | 1,474 | 1,474 | 813 | 813 |

*Notes:* Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-community correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. *: statistically significant at 10 percent; **: statistically significant at 5 percent; ***: statistically significant at 1 percent. Additional controls include remaining variables from Table 1, including variables for birth cohort, urban residence, region of birth, region of residence, and parental education.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

## Appendix

**Table A1. Specification Tests for 2SLS Ever Pregnant Regressions: Predictive Power of Identifying Instruments (First Stage), Endogeneity, and Overidentification (Second Stage), Identification Strategy I: Access to Educational Facilities (Rural Sample, only)**

| | *Rural, only:* | | | |
|---|---|---|---|---|
| | *Access* | | *Access X Individual Characteristics* | |
| | *Full Sample:* | *Less than Primary:* | *Full Sample:* | *Less than Primary:* |
| **(i) Using region FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | 1.25 [0.266] | 1.18 [0.310] | 41.18 [0.000] | 10.75 [0.000] |
| Primary | 2.65 [0.005] | | 10.37 [0.000] | |
| Middle/Junior Secondary | 1.74 [0.080] | | 3.47 [0.000] | |
| Secondary or higher | 0.84 [0.575] | | 0.27 [1.000] | |
| English reading | 5.87 [0.000] | 3.56 [0.000] | 5.43 [0.000] | 8.11 [0.000] |
| Ghanaian reading | 2.53 [0.008] | 2.07 [0.033] | 17.82 [0.000] | 36.76 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | 1.30 [0.257] | 0.09 [0.966] | 1.30 [0.257] | 0.47 [0.702] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | 1.16 [0.762] | 4.64 [0.590] | 1.16 [0.762] | 30.01 [0.995] |
| **(ii) Using cluster FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | NA | NA | 65.66 [0.000] | 89.93 [0.000] |
| Primary | NA | | 11.96 [0.000] | |
| Middle/Junior Secondary | NA | | 7.80 [0.000] | |
| Secondary or higher | NA | | 1.19 [0.182] | |
| English reading | NA | NA | 4.28 [0.000] | 526.51 [0.000] |
| Ghanaian reading | NA | NA | 25.46 [0.000] | 984.94 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | NA | NA | 1.32 [0.246] | 1.98 [0.117] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | NA | NA | 39.17 [0.905] | 43.67 [0.787] |
| N | 2,746 | 1,192 | 1,474 | 813 |

*Notes:* Terms in brackets are the p-values of the corresponding test-statistic. The tests incorporate sampling weights and robust Huber-White (Huber, 1967; White, 1980) standard errors and also adjust for within-community

correlation/clustering (Wooldridge, 2010). The explanatory variables in estimations are the same used for the estimations in Table 1. NA: "Not Available" → Since this specification cannot be estimated, due to the instruments being perfectly correlated with the cluster FEs for this specification.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A2. Specification Tests for 2SLS Ever Pregnant Regressions: Predictive Power of Identifying Instruments (First Stage), Endogeneity, and Overidentification (Second Stage), Identification Strategy II: Exposure to Educational Facilities (Rural Sample, only)**

| | *Rural, only:* | | | |
|---|---|---|---|---|
| | *Exposure* | | *Exposure X Individual Characteristics* | |
| | *Full Sample:* | *Less than Primary:* | *Full Sample:* | *Less than Primary:* |
| **(i) Using region FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | 2.11 [0.051] | 3.43 [0.017] | 1.11 [0.339] | 0.97 [0.493] |
| Primary | 1.74 [0.110] | | 6.26 [0.000] | |
| Middle/Junior Secondary | 1.35 [0.235] | | 4.66 [0.000] | |
| Secondary or higher | 1.31 [0.253] | | 0.67 [0.833] | |
| English reading | 7.03 [0.000] | 3.28 [0.021] | 2.26 [0.003] | 2.05 [0.009] |
| Ghanaian reading | 2.58 [0.018] | 1.23 [0.299] | 1.48 [0.100] | 3.09 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | 1.43 [0.201] | 1.20 [0.310] | 1.12 [0.350] | 2.02 [0.111] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | NA (exactly identified) | NA (exactly identified) | 12.23 [0.346] | 8.21 [0.877] |
| **(ii) Using cluster FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | NA | NA | 0.56 [0.919] | 0.78 [0.719] |
| Primary | NA | | 6.07 [0.000] | |
| Middle/Junior Secondary | NA | | 3.61 [0.000] | |
| Secondary or higher | NA | | 1.00 [0.459] | |
| English reading | NA | NA | 0.86 [0.621] | 1.04 [0.411] |
| Ghanaian reading | NA | NA | 0.98 [0.477] | 3.30 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | NA | NA | 1.12 [0.348] | 0.73 [0.535] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | NA | NA | 5.42 [0.908] | 12.54 [0.562] |
| N | 2,746 | 1,192 | 1,474 | 813 |

*Notes:* Terms in brackets are the p-values of the corresponding test-statistic. The tests incorporate sampling weights and robust Huber-White (Huber, 1967; White, 1980) standard errors and also adjust for within-community correlation/clustering (Wooldridge, 2010). The explanatory variables in estimations are the same used for the estimations in Table 1. NA: "Not Available" → Since this specification cannot be estimated, due to the instruments being perfectly correlated with the cluster FEs for this specification.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A3. Specification Tests for 2SLS Ever Pregnant Regressions: Predictive Power of Identifying Instruments (First Stage), Endogeneity, and Overidentification (Second Stage), Identification Strategy III: Region of Birth X Birth Cohort**

| | *Rural + Urban:* | | *Rural, only:* | |
|---|---|---|---|---|
| | *Full Sample:* | *Less than Primary:* | *Full Sample:* | *Less than Primary:* |
| **(i) Using region FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | 0.75 [0.928] | 0.91 [0.656] | 0.82 [0.817] | 1.08 [0.345] |
| Primary | 18.52 [0.000] | | 8.70 [0.000] | |
| Middle/Junior Secondary | 7.53 [0.000] | | 5.33 [0.000] | |
| Secondary or higher | 1.45 [0.015] | | 0.80 [0.842] | |
| English reading | 1.63 [0.002] | 2.38 [0.000] | 4.95 [0.000] | 7.65 [0.000] |
| Ghanaian reading | 5.00 [0.000] | 3.16 [0.000] | 12.42 [0.000] | 19.72 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | 1.18 [0.317] | 1.54 [0.204] | 1.49 [0.182] | 0.72 [0.542] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | 58.74 [0.411] | 46.90 [0.742] | 46.90 [0.674] | 42.77 [0.787] |
| **(ii) Using cluster FE:** | | | | |
| **IV Predictive Power/First-stage joint F-test of identifying IVs:** | | | | |
| Adult literacy program participation | 0.57 [0.997] | 0.53 [0.998] | 0.68 [0.961] | 2.49 [0.000] |
| Primary | 21.54 [0.000] | | 4.34 [0.000] | |
| Middle/Junior Secondary | 4.22 [0.000] | | 3.74 [0.000] | |
| Secondary or higher | 1.05 [0.369] | | 5.04 [0.000] | |
| English reading | 3.17 [0.000] | 2.57 [0.000] | 2.18 [0.000] | 3.24 [0.000] |
| Ghanaian reading | 3.15 [0.000] | 6.17 [0.000] | 13.74 [0.000] | 29.16 [0.000] |
| **Test for Endogeneity of Schooling & Skills:** | | | | |
| Wu (1973)-Hausman (1978) endogeneity test | 0.47 [0.834] | 0.39 [0.762] | 0.75 [0.612] | 0.52 [0.667] |
| **Second-stage test for instrument validity:** | | | | |
| Hansen (1982) J-test for overidentification | 55.56 [0.529] | 44.24 [0.798] | 57.08 [0.292] | 50.17 [0.506] |
| N | 2,746 | 1,192 | 1,474 | 813 |

*Notes:* Terms in brackets are the p-values of the corresponding test-statistic. The tests incorporate sampling weights and robust Huber-White (Huber, 1967; White, 1980) standard errors and also adjust for within-community correlation/clustering (Wooldridge, 2010). The explanatory variables in estimations are the same used for the estimations in Table 1.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A4. Skills and Schooling Coefficients from OLS Ever Pregnant Regressions: Full Results (Regional FEs)**

| | *(1) Rural + urban:* | | | | *(2) Rural, only:* | | | |
|---|---|---|---|---|---|---|---|---|
| | *(i) Full sample:* | | *(ii) No formal edu:* | | *(i) Full sample:* | | *(ii) No formal edu:* | |
| | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* |
| Primary | -0.013 | 0.017 | | | -0.022 | 0.012 | | |
| | [0.011] | [0.012] | | | [0.017] | [0.019] | | |
| Middle/JSS | -0.048*** | -0.01 | | | -0.068** | -0.021 | | |
| | [0.017] | [0.016] | | | [0.029] | [0.030] | | |
| Secondary and above | -0.123*** | -0.083*** | | | -0.185*** | -0.133*** | | |
| | [0.029] | [0.028] | | | [0.035] | [0.033] | | |
| Adult literacy program participation | -0.058 | -0.058 | -0.087*** | -0.087*** | -0.07 | -0.070* | -0.089*** | -0.090*** |
| | [0.038] | [0.037] | [0.029] | [0.028] | [0.044] | [0.042] | [0.032] | [0.032] |
| English reading | | -0.055*** | | -0.038*** | | -0.058*** | | -0.039** |
| | | [0.013] | | [0.012] | | [0.017] | | [0.016] |
| Ghanaian Reading | | -0.017* | | -0.002 | | -0.021 | | 0.006 |
| | | [0.010] | | [0.012] | | [0.015] | | [0.021] |
| Married or informal union | 0.790*** | 0.776*** | 0.663*** | 0.658*** | 0.756*** | 0.742*** | 0.649*** | 0.644*** |
| | [0.029] | [0.029] | [0.058] | [0.058] | [0.035] | [0.036] | [0.062] | [0.062] |
| Age 14 | 0.004 | 0 | 0.002 | -0.002 | 0.001 | 0 | 0.001 | -0.002 |
| | [0.006] | [0.006] | [0.009] | [0.009] | [0.009] | [0.010] | [0.012] | [0.012] |
| Age 15 | 0.016** | 0.01 | 0.003 | -0.001 | 0.013 | 0.009 | 0.008 | 0.005 |
| | [0.008] | [0.008] | [0.012] | [0.011] | [0.011] | [0.011] | [0.016] | [0.015] |
| Age 16 | 0.036*** | 0.029** | 0.066*** | 0.060** | 0.052*** | 0.046** | 0.090** | 0.085** |
| | [0.012] | [0.012] | [0.025] | [0.023] | [0.019] | [0.019] | [0.036] | [0.034] |
| Age 17 | 0.082*** | 0.072*** | 0.060** | 0.050* | 0.096*** | 0.085*** | 0.074** | 0.067** |
| | [0.017] | [0.016] | [0.028] | [0.027] | [0.026] | [0.025] | [0.034] | [0.032] |
| Age 18 | 0.144*** | 0.135*** | 0.193*** | 0.184*** | 0.176*** | 0.169*** | 0.238*** | 0.231*** |
| | [0.026] | [0.026] | [0.063] | [0.063] | [0.044] | [0.044] | [0.083] | [0.084] |
| Age 19 | 0.142*** | 0.128*** | 0.202*** | 0.189*** | 0.188*** | 0.175*** | 0.231*** | 0.221*** |
| | [0.026] | [0.024] | [0.047] | [0.047] | [0.040] | [0.038] | [0.065] | [0.065] |
| Mother primary | -0.001 | 0.002 | -0.029** | -0.024* | 0.007 | 0.011 | -0.043** | -0.036* |
| | [0.014] | [0.014] | [0.013] | [0.013] | [0.021] | [0.020] | [0.021] | [0.020] |
| Mother above primary | -0.003 | 0.001 | 0.007 | 0.015 | 0 | 0.005 | 0.014 | 0.022 |
| | [0.011] | [0.011] | [0.019] | [0.019] | [0.017] | [0.017] | [0.031] | [0.032] |
| Mother don't know/missing | 0.002 | 0.005 | 0.089 | 0.089 | -0.015 | -0.013 | 0.007 | 0.011 |
| | [0.028] | [0.027] | [0.064] | [0.063] | [0.023] | [0.024] | [0.034] | [0.035] |
| Father primary | -0.012 | -0.004 | 0.004 | 0.012 | -0.011 | -0.006 | -0.006 | -0.002 |
| | [0.015] | [0.014] | [0.022] | [0.023] | [0.024] | [0.023] | [0.035] | [0.035] |
| Father above primary | 0.017 | 0.024** | 0.029* | 0.036** | 0.019 | 0.027* | 0.038* | 0.043** |
| | [0.012] | [0.012] | [0.015] | [0.016] | [0.015] | [0.015] | [0.021] | [0.022] |
| Father don't know/missing | 0.032 | 0.027 | 0.045 | 0.039 | 0.023 | 0.009 | -0.008 | -0.016 |
| | [0.022] | [0.021] | [0.035] | [0.035] | [0.033] | [0.031] | [0.026] | [0.026] |
| Urban (residence) | -0.020** | -0.018** | -0.040*** | -0.037*** | | | | |
| | [0.009] | [0.008] | [0.013] | [0.012] | | | | |
| Western (residence) | 0.006 | 0.01 | -0.004 | 0 | -0.068 | -0.077 | -0.061 | -0.067 |
| | [0.034] | [0.033] | [0.071] | [0.070] | [0.058] | [0.060] | [0.100] | [0.101] |
| Central (residence) | 0.014 | 0.014 | 0.064 | 0.064 | 0.035 | 0.022 | 0.069 | 0.064 |
| | [0.017] | [0.017] | [0.048] | [0.048] | [0.032] | [0.036] | [0.074] | [0.076] |
| Eastern (residence) | 0.053** | 0.056** | 0.003 | 0.004 | 0.067 | 0.058 | 0.001 | 0.001 |
| | [0.026] | [0.026] | [0.032] | [0.032] | [0.063] | [0.062] | [0.068] | [0.069] |
| Volta (residence) | 0.015 | 0.01 | -0.053 | -0.062 | -0.013 | -0.026 | -0.054 | -0.067 |
| | [0.021] | [0.021] | [0.045] | [0.045] | [0.035] | [0.037] | [0.073] | [0.074] |
| Ashanti (residence) | -0.007 | -0.004 | 0.002 | 0.004 | -0.007 | -0.02 | 0.033 | 0.024 |
| | [0.021] | [0.020] | [0.052] | [0.051] | [0.044] | [0.047] | [0.081] | [0.082] |
| Brong Ahafo (residence) | 0.02 | 0.021 | 0.036 | 0.037 | 0.043 | 0.036 | 0.056 | 0.052 |
| | [0.033] | [0.033] | [0.074] | [0.074] | [0.050] | [0.054] | [0.093] | [0.095] |
| Northern (residence) | 0.06 | 0.053 | 0.096 | 0.095 | 0.139 | 0.126 | 0.168 | 0.162 |
| | [0.048] | [0.049] | [0.069] | [0.068] | [0.090] | [0.091] | [0.115] | [0.116] |
| Upper West (residence) | -0.058 | -0.051 | -0.008 | -0.004 | 0.091 | 0.092 | 0.159 | 0.159 |
| | [0.044] | [0.044] | [0.077] | [0.076] | [0.083] | [0.086] | [0.107] | [0.109] |
| Upper East (residence) | 0 | 0.002 | 0.024 | 0.027 | 0.009 | -0.002 | 0.039 | 0.032 |
| | [0.028] | [0.029] | [0.064] | [0.065] | [0.052] | [0.058] | [0.095] | [0.098] |
| Western (birth) | 0.022 | 0.018 | 0.048 | 0.038 | 0.153** | 0.152** | 0.147 | 0.149* |
| | [0.033] | [0.033] | [0.067] | [0.067] | [0.060] | [0.059] | [0.090] | [0.090] |
| Central (birth) | -0.019 | -0.022 | -0.012 | -0.024 | 0.014 | 0.012 | 0.028 | 0.025 |
| | [0.016] | [0.015] | [0.046] | [0.046] | [0.026] | [0.028] | [0.060] | [0.060] |
| Eastern (birth) | -0.029 | -0.036 | 0.031 | 0.02 | 0.007 | -0.002 | 0.094 | 0.088 |
| | [0.022] | [0.022] | [0.033] | [0.033] | [0.053] | [0.053] | [0.060] | [0.058] |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Volta (birth) | -0.013 | -0.016 | 0.051 | 0.046 | 0.067** | 0.060* | 0.107* | 0.107* |
| | [0.019] | [0.019] | [0.045] | [0.044] | [0.033] | [0.034] | [0.059] | [0.057] |
| Ashanti (birth) | 0.015 | 0.007 | 0.039 | 0.025 | 0.065 | 0.06 | 0.057 | 0.057 |
| | [0.022] | [0.022] | [0.051] | [0.050] | [0.045] | [0.044] | [0.072] | [0.071] |
| Brong Ahafo (birth) | -0.003 | -0.006 | 0.03 | 0.02 | 0.035 | 0.029 | 0.077 | 0.077 |
| | [0.033] | [0.033] | [0.080] | [0.079] | [0.048] | [0.049] | [0.087] | [0.087] |
| Northern (birth) | -0.057 | -0.06 | -0.05 | -0.062 | -0.066 | -0.077 | -0.064 | -0.069 |
| | [0.046] | [0.047] | [0.068] | [0.068] | [0.087] | [0.088] | [0.110] | [0.110] |
| Upper West (birth) | 0.019 | 0.016 | 0.001 | -0.007 | -0.096 | -0.104 | -0.134 | -0.137 |
| | [0.044] | [0.044] | [0.076] | [0.076] | [0.082] | [0.084] | [0.099] | [0.099] |
| Upper East (birth) | -0.036 | -0.048* | -0.02 | -0.035 | 0.005 | -0.008 | 0.011 | 0.01 |
| | [0.025] | [0.026] | [0.064] | [0.066] | [0.049] | [0.053] | [0.086] | [0.087] |
| Other ECOWAS (birth) | 0.058 | 0.047 | 0.276* | 0.266* | 0.022 | -0.001 | 0.109 | 0.103 |
| | [0.055] | [0.054] | [0.160] | [0.158] | [0.030] | [0.031] | [0.075] | [0.074] |
| Other Africa (birth) | -0.069 | -0.059 | 0.039 | 0.052 | -0.098 | -0.097 | | |
| | [0.044] | [0.048] | [0.049] | [0.049] | [0.070] | [0.075] | | |
| Outside Africa (birth) | -0.009 | 0.001 | -0.021 | -0.022 | | | | |
| | [0.025] | [0.024] | [0.057] | [0.053] | | | | |
| Constant | 0.005 | 0.036** | -0.024 | 0.001 | -0.050*** | -0.008 | -0.080** | -0.06 |
| | [0.015] | [0.016] | [0.022] | [0.023] | [0.016] | [0.020] | [0.034] | [0.039] |
| $R^2$ | 0.501 | 0.508 | 0.516 | 0.519 | 0.521 | 0.529 | 0.508 | 0.511 |
| N | 2,746 | 2,746 | 1,192 | 1,192 | 1,474 | 1,474 | 813 | 813 |

*Notes:* Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-community correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. *: statistically significant at 10 percent; **: statistically significant at 5 percent; ***: statistically significant at 1 percent.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A5. Skills and Schooling Coefficients from OLS Ever Pregnant Regressions: Full Results (Cluster FEs)**

| | *(1) Rural + urban:* | | | | *(2) Rural, only:* | | | |
|---|---|---|---|---|---|---|---|---|
| | *(i) Full sample:* | | *(ii) No formal edu:* | | *(i) Full sample:* | | *(ii) No formal edu:* | |
| | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* | *Only education* | *Adding skills* |
| Primary | -0.005 | 0.026 | | | -0.015 | 0.017 | | |
| | [0.015] | [0.016] | | | [0.023] | [0.024] | | |
| Middle/JSS | -0.047** | -0.008 | | | -0.064 | -0.02 | | |
| | [0.021] | [0.023] | | | [0.039] | [0.041] | | |
| Secondary and above | -0.123*** | -0.080** | | | -0.165*** | -0.121*** | | |
| | [0.036] | [0.038] | | | [0.046] | [0.046] | | |
| Adult literacy program participation | -0.062 | -0.064 | -0.127 | -0.128 | -0.074 | -0.079 | -0.136* | -0.137* |
| | [0.065] | [0.063] | [0.083] | [0.082] | [0.076] | [0.074] | [0.081] | [0.080] |
| English reading | | -0.066*** | | -0.022 | | -0.064** | | -0.021 |
| | | [0.019] | | [0.028] | | [0.026] | | [0.036] |
| Ghanaian Reading | | -0.011 | | -0.006 | | -0.015 | | -0.012 |
| | | [0.014] | | [0.033] | | [0.021] | | [0.049] |
| Married or informal union | 0.785*** | 0.768*** | 0.616*** | 0.612*** | 0.754*** | 0.738*** | 0.641*** | 0.637*** |
| | [0.037] | [0.038] | [0.096] | [0.096] | [0.046] | [0.046] | [0.088] | [0.087] |
| Age 14 | 0 | -0.004 | 0.004 | 0.002 | 0.011 | 0.009 | 0.01 | 0.008 |
| | [0.011] | [0.011] | [0.019] | [0.020] | [0.018] | [0.018] | [0.023] | [0.024] |
| Age 15 | 0.017 | 0.012 | -0.009 | -0.012 | 0.015 | 0.012 | -0.012 | -0.015 |
| | [0.012] | [0.011] | [0.023] | [0.023] | [0.018] | [0.018] | [0.027] | [0.026] |
| Age 16 | 0.037** | 0.030* | 0.085** | 0.081** | 0.065** | 0.060** | 0.110** | 0.107** |
| | [0.017] | [0.017] | [0.041] | [0.039] | [0.026] | [0.025] | [0.053] | [0.052] |
| Age 17 | 0.074*** | 0.065*** | 0.034 | 0.03 | 0.086** | 0.078** | 0.045 | 0.042 |
| | [0.022] | [0.021] | [0.049] | [0.047] | [0.035] | [0.033] | [0.054] | [0.052] |
| Age 18 | 0.148*** | 0.138*** | 0.245** | 0.240** | 0.185*** | 0.178*** | 0.270** | 0.266** |
| | [0.035] | [0.035] | [0.108] | [0.110] | [0.058] | [0.058] | [0.125] | [0.128] |
| Age 19 | 0.147*** | 0.132*** | 0.213*** | 0.206*** | 0.202*** | 0.191*** | 0.258*** | 0.251*** |
| | [0.030] | [0.029] | [0.073] | [0.073] | [0.049] | [0.048] | [0.091] | [0.091] |
| Mother primary | -0.007 | -0.005 | 0.004 | 0.006 | -0.005 | 0.001 | 0.01 | 0.011 |
| | [0.020] | [0.019] | [0.023] | [0.023] | [0.027] | [0.027] | [0.031] | [0.032] |
| Mother above primary | 0.001 | 0.004 | 0.031 | 0.034 | 0.018 | 0.021 | 0.048 | 0.05 |
| | [0.014] | [0.014] | [0.034] | [0.034] | [0.025] | [0.024] | [0.046] | [0.046] |
| Mother don't know/missing | -0.015 | -0.014 | 0.07 | 0.069 | 0.03 | 0.027 | 0.088 | 0.084 |
| | [0.044] | [0.043] | [0.047] | [0.048] | [0.029] | [0.031] | [0.072] | [0.076] |
| Father primary | -0.012 | -0.005 | 0.015 | 0.02 | -0.01 | -0.007 | 0.029 | 0.033 |
| | [0.021] | [0.021] | [0.049] | [0.050] | [0.034] | [0.034] | [0.056] | [0.058] |
| Father above primary | 0.014 | 0.021 | 0.042 | 0.047 | 0.015 | 0.022 | 0.058* | 0.063* |
| | [0.015] | [0.015] | [0.028] | [0.030] | [0.020] | [0.020] | [0.034] | [0.036] |
| Father don't know/missing | 0.018 | 0.012 | 0.108* | 0.104* | -0.015 | -0.028 | 0.085* | 0.08 |
| | [0.025] | [0.025] | [0.057] | [0.057] | [0.033] | [0.031] | [0.050] | [0.052] |
| Urban | -0.057 | -0.089 | 0.123 | 0.101 | | | | |
| | [0.065] | [0.066] | [0.095] | [0.099] | | | | |
| Western (birth) | -0.008 | -0.01 | 0.041 | 0.036 | 0.127* | 0.128* | 0.236* | 0.238* |
| | [0.036] | [0.036] | [0.081] | [0.081] | [0.067] | [0.067] | [0.124] | [0.124] |
| Central (birth) | -0.038** | -0.041** | -0.045 | -0.052 | 0.018 | 0.015 | 0.088 | 0.083 |
| | [0.017] | [0.017] | [0.060] | [0.062] | [0.039] | [0.040] | [0.084] | [0.084] |
| Eastern (birth) | -0.051* | -0.058* | 0.071 | 0.064 | 0.022 | 0.008 | 0.195* | 0.193* |
| | [0.030] | [0.030] | [0.065] | [0.065] | [0.048] | [0.049] | [0.115] | [0.116] |
| Volta (birth) | -0.028 | -0.028 | 0.052 | 0.044 | 0.077* | 0.069 | 0.151** | 0.145** |
| | [0.027] | [0.027] | [0.075] | [0.075] | [0.045] | [0.044] | [0.075] | [0.072] |
| Ashanti (birth) | -0.006 | -0.012 | 0.02 | 0.012 | 0.044 | 0.039 | 0.129 | 0.125 |
| | [0.020] | [0.020] | [0.053] | [0.054] | [0.045] | [0.046] | [0.095] | [0.094] |
| Brong Ahafo (birth) | -0.011 | -0.019 | -0.015 | -0.02 | 0.025 | 0.011 | 0.126 | 0.126 |
| | [0.041] | [0.041] | [0.080] | [0.081] | [0.071] | [0.071] | [0.139] | [0.137] |
| Northern (birth) | -0.066 | -0.071 | -0.066 | -0.072 | -0.09 | -0.106 | 0.002 | -0.002 |
| | [0.060] | [0.062] | [0.088] | [0.088] | [0.121] | [0.123] | [0.152] | [0.153] |
| Upper West (birth) | -0.039 | -0.04 | -0.078 | -0.085 | -0.137 | -0.146 | -0.119 | -0.125 |
| | [0.058] | [0.059] | [0.126] | [0.127] | [0.127] | [0.126] | [0.168] | [0.167] |
| Upper East (birth) | -0.022 | -0.032 | 0.064 | 0.059 | 0.077 | 0.068 | 0.204 | 0.205 |
| | [0.042] | [0.043] | [0.095] | [0.096] | [0.080] | [0.081] | [0.136] | [0.135] |
| Other ECOWAS (birth) | 0.045 | 0.037 | 0.307 | 0.302 | -0.014 | -0.027 | 0.214* | 0.215* |
| | [0.074] | [0.075] | [0.226] | [0.225] | [0.061] | [0.065] | [0.125] | [0.126] |
| Other Africa (birth) | -0.094 | -0.064 | 0.021 | 0.026 | -0.093 | -0.052 | | |
| | [0.061] | [0.061] | [0.088] | [0.088] | [0.118] | [0.125] | | |
| Outside Africa (birth) | -0.13 | -0.127 | -0.242** | -0.228** | | | | |
| | [0.176] | [0.167] | [0.103] | [0.102] | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Constant | 0.039 | 0.106* | -0.078 | -0.049 | -0.016 | 0.06 | 0.119 | 0.145 |
| | [0.058] | [0.062] | [0.097] | [0.105] | [0.081] | [0.089] | [0.168] | [0.176] |
| $R^2$ | 0.6 | 0.606 | 0.665 | 0.666 | 0.621 | 0.626 | 0.644 | 0.644 |
| N | 2,746 | 2,746 | 1,192 | 1,192 | 1,474 | 1,474 | 813 | 813 |

*Notes:* The results for the cluster FEs have been omitted for brevity (available upon request). Robust Huber-White (Huber, 1967; White, 1980) standard errors, adjusted for within-community correlation/clustering (Wooldridge, 2010), in brackets under parameter estimates. *: statistically significant at 10 percent; **: statistically significant at 5 percent; ***: statistically significant at 1 percent.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A6. Formal Schooling Average Treatment Estimates for Ever Pregnant Outcome Using Mahalanobis Matching**

| Outcome: | Ever attended school: | | | Primary and above: | | | Junior secondary and above: | | |
|---|---|---|---|---|---|---|---|---|---|
| Number of neighbors: | 1 | 3 | 5 | 1 | 3 | 5 | 1 | 3 | 5 |
| **Full sample:** | | | | | | | | | |
| (i) Using region FE | -0.021 | -0.045* | -0.043* | -0.026** | -0.030*** | -0.028*** | -0.029** | -0.023* | -0.025** |
| | [0.025] | [0.024] | [0.022] | [0.012] | [0.011] | [0.010] | [0.014] | [0.012] | [0.012] |
| N | 2,738 | 2,738 | 2,738 | 2,746 | 2,746 | 2,746 | 2,743 | 2,743 | 2,743 |
| (ii) Using cluster FE | -0.086*** | -0.058*** | -0.044*** | -0.003 | 0.001 | 0.002 | 0.013 | -0.006 | -0.016++ |
| | 0.022 | 0.015 | 0.013 | 0.012 | 0.010 | 0.009 | 0.016 | 0.011 | 0.010 |
| N | 1,081 | 1,081 | 1,081 | 1,968 | 1,968 | 1,968 | 1,720 | 1,720 | 1,720 |
| **Rural sample:** | | | | | | | | | |
| (i) Using region FE | -0.010 | -0.030+ | -0.024 | -0.030++ | -0.021+ | -0.023* | -0.035+ | -0.012 | -0.018 |
| | [0.019] | [0.022] | [0.020] | [0.019] | [0.016] | [0.014] | [0.025] | [0.034] | [0.025] |
| N | 1,458 | 1,458 | 1,458 | 1,472 | 1,472 | 1,472 | 1,458 | 1,458 | 1,458 |
| (ii) Using cluster FE | -0.079*** | -0.070*** | -0.057** | -0.004 | -0.009 | -0.011 | 0.025 | 0.011 | 0.015 |
| | 0.028 | 0.021 | 0.024 | 0.018 | 0.015 | 0.013 | 0.029 | 0.039 | 0.057 |
| N | 683 | 683 | 683 | 1,031 | 1,031 | 1,031 | 687 | 687 | 687 |

*Notes:* Additional control variables include the explanatory variables from the ever pregnant OLS regressions in Table 1, including variables for parental education, birth cohort, region of birth, region of current residence, urban residence (the last two variables not included in the specification with cluster FEs due to perfect collinearity). Estimations incorporate the heteroskedasticity-consistent analytical standard errors proposed by Abadie and Imbens (2006) (in brackets under parameter estimates). The lower numbers of observations for the specifications with cluster FEs are due to some of the communities having no variation in treatment status. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent; +: statistically significant at 20 percent.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).

**Table A7. Adult Literacy Program Participation Average Treatment Estimates for Ever Pregnant Outcome Using Mahalanobis Matching (Sample with Less than Primary Completed)**

| Estimation sample: | *Rural + urban:* | | | *Rural, only:* | | |
|---|---|---|---|---|---|---|
| Number of neighbors: | 1 | 3 | 5 | 1 | 3 | 5 |
| (i) Using region FE | -0.107*** | -0.112*** | -0.113*** | -0.094*** | -0.107*** | -0.110*** |
| | [0.016] | [0.015] | [0.013] | [0.016] | [0.013] | [0.013] |
| N | 960 | 960 | 960 | 624 | 624 | 624 |
| (ii) Using cluster FE | -0.103** | -0.094* | -0.091+ | -0.066 | -0.098++ | -0.093++ |
| | 0.052 | 0.051 | 0.066 | 0.073 | 0.060 | 0.060 |
| N | 148 | 148 | 148 | 139 | 139 | 139 |

*Notes:* Additional control variables include the explanatory variables from the ever pregnant OLS regressions in Table 1, including variables for parental education, birth cohort, region of birth, region of current residence, urban residence (the last two variables not included in the specification with cluster FEs due to perfect collinearity). Estimations incorporate the heteroskedasticity-consistent analytical standard errors proposed by Abadie and Imbens (2006) (in brackets under parameter estimates). The lower numbers of observations for the specifications with cluster FEs are due to some of the communities having no variation in treatment status. ***: statistically significant at 1 percent; **: statistically significant at 5 percent; *: statistically significant at 10 percent; ++: statistically significant at 15 percent; +: statistically significant at 20 percent.
*Source:* Ghana Living Standards Survey (Round 5, 2005/06).