Demographic and Cultural Factors in African Women’s Education- Education and Economic Development

By Julie Norman

# Abstract

Many factors are involved in a group of women receiving any amount of education. Particularly, in developing countries where legal requirements often aren’t applicable. Discovering what household characteristics best allow for female household members to receive more formal education can inform policy regarding development and female empowerment. Using data from 11 African countries and 8500 randomly selected households and farming cooperatives, I’ve created a multi-variable regression to model what characteristics have an impact on female education. I found that there’s a significant difference in average household education between countries and religious groups; additionally, the presence of electricity has a positive relationship with my education variable. I also found that an increase in the female to male gender ratio, the education of the household’s men, and the quantity of the harvest for the farm are all associated with a higher female education average.

(JEL, I25)

Table of Contents

[Abstract 1](#_Toc500417676)

[Introduction and Literature Review 2](#_Toc500417677)

[Data Descriptions and Visualization 5](#_Toc500417678)

[Regression Summary 6](#_Toc500417679)

[Diagnostic Plots with Outliers 6](#_Toc500417680)

[Diagnostic Plots without Outliers 7](#_Toc500417681)

[Variable Scatterplots 9](#_Toc500417682)

[Average Female Education in a Household, by Country 10](#_Toc500417683)

[RESET Test 10](#_Toc500417684)

[Method and Model Specification 11](#_Toc500417685)

[Empirical Analysis 14](#_Toc500417686)

[Estimation and Inference Results 14](#_Toc500417687)

[Statistical Assumptions and Robustness 16](#_Toc500417688)

[Conclusions and Applications 18](#_Toc500417689)

# Introduction and Literature Review

The goal of this research project is to discover what elements of a farming household or cooperative best allows the women and girls in that household to achieve higher levels of formal education. Women’s education vastly impacts a region’s economic growth and development. According to the World Bank[[1]](#footnote-1), female education has a positive association with not just development but also general productivity, income, and quality of life.

However, improving access to schools is not enough. The issue with improving female education in developing countries is demand for female education. The private return on education for women is significantly lower than the investment costs, especially since boys and men tend to be a higher priority culturally, which puts them first in line for a formal education. Discovering what household characteristics best allow that group of women and girls to receive their education opens doors for policy makers and humanitarians to better target their work to help these women and their communities. I found that specific household characteristics such as country, household gender ratio, average education of household men, presence of electricity, harvest quantity, and religion have statistically significant relationships to the education of that household’s women and girls.

The data set I’m utilizing is from the Center for Environmental Economics and Policy in Africa and has primarily been used to describe the impact of climate change on African farmers across the continent. The data set includes nearly 9,500 samples of households across various African countries including Burkina Faso, Cameroon, Ghana, Niger and Senegal, Egypt, Ethiopia, Kenya, South Africa, Zambia and Zimbabwe.

Much of the research surrounding women’s education in Africa specifically focuses on the positive benefits of women’s education for the society. One of the largest benefits is HIV prevention. Through a literature review and a randomized control trial Jukes, Mathew, and Simmons conclude that there is a growing body of evidence that keeping girls in school helps lower their risk of HIV contraction.[[2]](#footnote-2) Within the African continent there have been 2.4 million deaths and 3.2 million incidences of infection as of 2006.[[3]](#footnote-3) Another large impact on the development of African countries is the presence of high population growth. Work by Neeru Gupta and Mary Mahy from the World Health Organization and the United Nations Children’s Fund (UNICEF) respectively, conclude that a girl’s education from the secondary level and beyond indicates a lower probability of giving birth within adolescence.[[4]](#footnote-4) These are merely two examples, out of a vast literature, that support the importance of education for African women and girls.

Additionally, research from Esther Duflo’s shows how household conditions affect a female household member’s education. Her research on Old- Age Pensions in South Africa concludes that pensions received by women had a large impact on anthropometric status (weight and height by age) of girls with little comparative improvement for boys. No such relationship was found with male recipients.[[5]](#footnote-5) These results indicate that the gender for the head of household, as well has the number of adult women present, may have a positive impact on the wellbeing of younger female family members, including their ability to receive an education. Further research by Cynthia B. Lloyd and Ann K. Blanc shows that higher education levels for head of households and female headed households tend to better encourage education for girls whereas the biological relationship between children and adults as well as large bias against girls in male headed households is weaker than one might suspect.[[6]](#footnote-6) Clearly, household conditions impact women’s education. My research provides impactful information on African households and the elements involved in female education.

# Data Descriptions and Visualization

### Regression Summary

Residuals:

Min 1Q Median 3Q Max

-13.2378 -1.3523 -0.0372 1.3443 13.6305

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -2.733e-01 1.129e-01 -2.421 0.015509 \*

countryEgypt 1.650e+00 1.717e-01 9.608 < 2e-16 \*\*\*

countryEthiopia 1.787e+00 3.582e-01 4.988 6.23e-07 \*\*\*

countryGhana 8.307e-01 1.519e-01 5.470 4.62e-08 \*\*\*

countryNiger 4.719e-01 1.328e-01 3.553 0.000383 \*\*\*

countrySenegal 5.486e-01 1.303e-01 4.209 2.59e-05 \*\*\*

countrySouth\_Africa 2.866e+00 2.071e-01 13.840 < 2e-16 \*\*\*

countryZambia 1.273e+00 1.516e-01 8.396 < 2e-16 \*\*\*

countryCameroon 2.371e+00 1.680e-01 14.113 < 2e-16 \*\*\*

countryKenya 3.822e+00 1.686e-01 22.664 < 2e-16 \*\*\*

countryZimbabwe -2.711e-01 1.697e-01 -1.597 0.110289

hhg\_ratio 4.979e-01 8.123e-02 6.129 9.24e-10 \*\*\*

I(hhg\_ratio^2) -1.227e-01 1.630e-02 -7.532 5.49e-14 \*\*\*

educ\_mean\_male 2.964e-01 2.225e-02 13.321 < 2e-16 \*\*\*

I(educ\_mean\_male^2) 6.921e-03 1.501e-03 4.611 4.06e-06 \*\*\*

hhelectricyes 7.603e-01 9.794e-02 7.762 9.30e-15 \*\*\*

qharv\_total 1.030e-06 1.755e-07 5.871 4.49e-09 \*\*\*

I(qharv\_total^2) -5.974e-14 1.450e-14 -4.121 3.81e-05 \*\*\*

hhreligAfrican Traditional 3.600e-01 2.044e-01 1.761 0.078284 .

hhreligChristianity 3.413e+00 2.789e+00 1.224 0.221031

hhreligIslam 9.224e-01 9.349e-01 0.987 0.323827

hhrelignonreligious -9.763e+00 2.793e+00 -3.496 0.000475 \*\*\*

hhreligother 5.540e-01 1.112e-01 4.984 6.35e-07 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

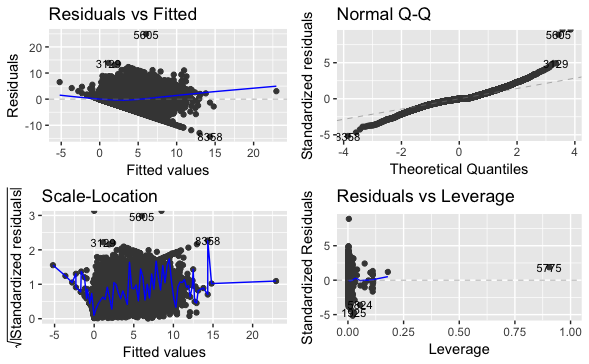
Residual standard error: 2.785 on 8490 degrees of freedom

Multiple R-squared: 0.5119, Adjusted R-squared: 0.5107

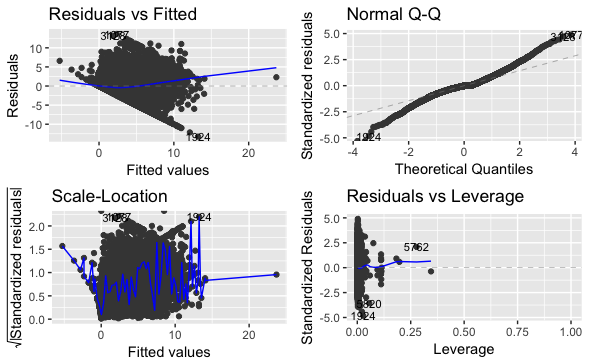
F-statistic: 404.8 on 22 and 8490 DF, p-value: < 2.2e-16

Interpretation of variables and coefficients start on page 11

### Diagnostic Plots with Outliers

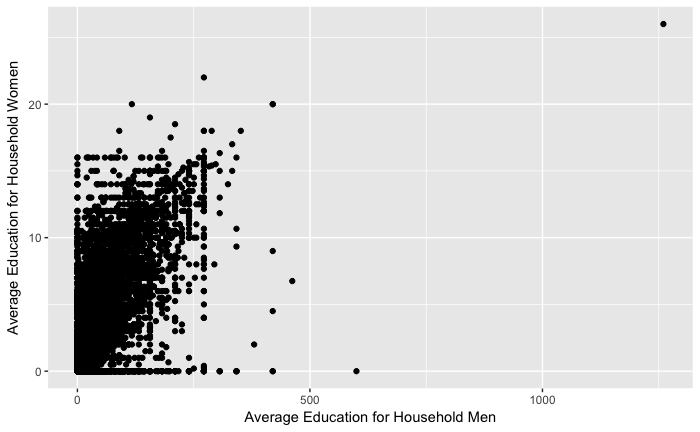


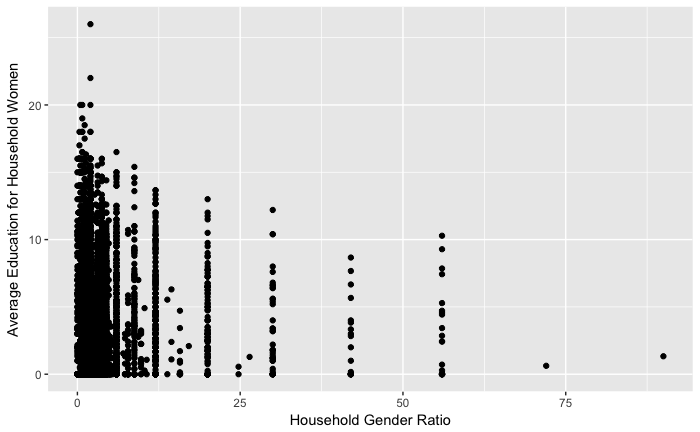
### Diagnostic Plots without Outliers

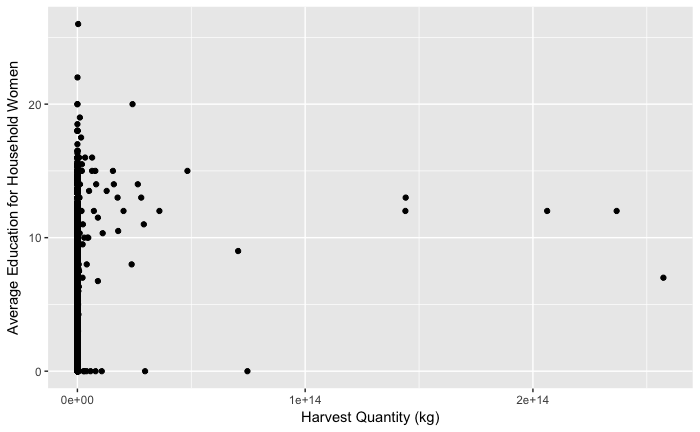


The plots above show the data set with and without five different outliers from various countries. The residuals versus fitted plot shows the residuals of my regression with the appropriate fitted line. While the residuals are clustered, within the clustered area they appear to relatively random. There appears to be a strong negative linear association in the lower left bound residuals. This is likely due to the presence of missing values within the dataset. Since the Q-Q plot is nearly linear the parameters are approximately normally distributed. The Scale- Location graphic has a line that is randomly plotted but generally heads in a horizontal direction so I have sufficient evidence for homoscedasticity. The final plot, Residuals vs Leverage, shows extreme clustering toward the lower values x values, therefore I do have influential points within my data set. I’ve decided to keep all outliers aside from those deleted earlier since they fit in with the general trend of the data, as shown on page nine.

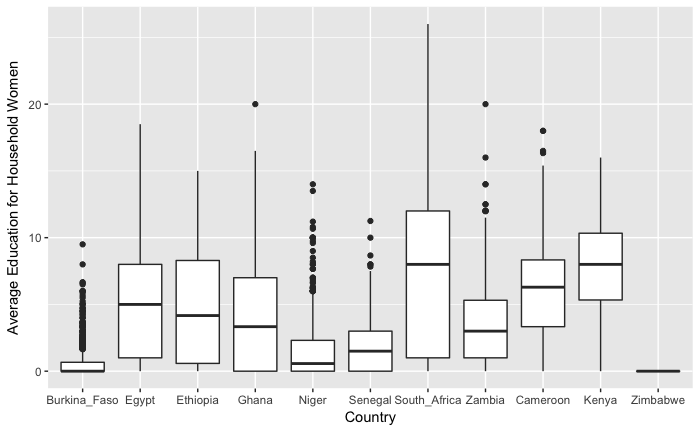
### Variable Scatterplots







### Average Female Education in a Household, by Country



In my analysis of country variation, I used Burkina Faso as the baseline for comparison. All countries appear to have the same minimum at zero years while the means vary within ten years of each other and maximums across 25 years. I included country location within my regression primarily as a control variable. There are too many factors and differences between countries beyond policy and culture that a significant conclusion as to the nature of the variation isn’t valid.

### RESET Test

data: reg1

RESET = 25.312, df1 = 2, df2 = 8488, p-value = 1.096e-11

Since the result of the RESET Test is signification at the 0.1% level, I have sufficient evidence to support including quadratic relationships in my model.

# Method and Model Specification

Regression

To predict household female education, I calculated the average education of women within the household to use as my response variable. My method of estimation is a linear model with two different sets of dummy variables, country and religion, with a quadratic association for the rest of my variables, see page six for coefficient statistical significance measures. My analysis focuses on farming cooperatives primarily around 4- 8 people, though some are as high as 38 people. Throughout my analysis, I will refer to these groups as “households”.

Each variable in my regression considers a different aspect of the household. The country variable aims to control for the vast variations in culture, policy, and geography that exist between countries. The country variable helps achieve a *cetris paribus* analysis for the other variables. I also included religion as a categorical variable. Due to limitations within the data, the counts are only high enough for “African Traditional” and “Other” to be valid. Therefore my analysis will focus on those religious categories only. The comparison of the two categories capture the impact of religion, and culture, on household female education. My final categorical variable is the presence of electricity, simply as a yes or no response. The presence of electricity has been economically linked with fertility[[7]](#footnote-7) and it acts as an indicator for poverty level of the household.

I’ve also included the variables household gender ratio to see if a larger female presence has an impact on household female education. I also included a variable for average education of men within the household. I want to see how living around more educated men impacted the women and girls in the household. The final variable I’m testing for, also with a quadratic relationship, is the farm’s harvest quantity in kilograms. This variable acts as an indicator for both farm size and farm success.

Data collection at the scale needed for my analysis often comes with imprecisions. I would prefer to included variables to account for marriage, and more variables for relative wealth. However, that information was not sufficiently available. I also considered including a variable for the quantity of various crops. While some of the crops were significant, all the crops together account for less than two percent of the variation in average female education so I opted to omit those variables.

# Empirical Analysis

## Estimation and Inference Results

0.1128 0.1717 0.3582 0.1519

0.1328 0.1303 0.2071

0.153 0.1695 0.1701

0.1516 0.08123

0.0163

0.02225

0.001501

0.09794 1.755 10-7

1.45 10-14

0.2044 2.789

0.9349 2.793 0.1112

Within my country calculations, with Burkina Faso as the baseline, South Africa had the highest average while Zimbabwe had the lowest. Each country was statistically significant at the 0.1% level, accept Zimbabwe at a 10% level. Due to the infinite variation factors between countries, I consider this variable a control variable. The rest of the variables are focused on individual household characteristics. I found that the ratio of females to males in the household has a predicted quadratic relationship to the average female education level at a 0.1% level. I tested this variable to see if having less women would mean more focus on education or if more women in the household meant female education was a higher priority. The quadratic has its predicted maximum where the ratio is two and 2.8 months of additional education. The average education of the men in the household has a statistically significant impact the 0.1% level as well. The quadratic predicted that as men’s education increases, the average female education increases exponentially. This indicates that, boys likely aren’t receiving an education over girls, it’s a culture of support and desire for education that leads to more years.

Harvest quantity is also quadratic where the predicted maximum is 4,163,088 kg harvested with three additional years of education, with a significance level at 0.1%. This kilogram calculation is lies within the “large” category of farm meaning that my model predicts that as farms increase in size, female household members become more educated on average. The next household variables are categorical. I tested the presence of electricity in the household as a binary variable and found that at a 0.1% level of significance where electricity is associated with nine more months of mean education relative to a household with no electricity. I also compared the religion of the household head. Due to data limitations in counts, the only categories are *African Traditional* at 10% significance, and *Other* which is significant at the 0.1% level. Both are relative to the random collection of household heads with no response for the religion question. Between the two, *Other Religion* had a mean education higher by two months compared to *African Traditional* indicating that our model shows that religion does play a factor but it’s not very high relative to other factors. The predicted impact of each variable is when all other variables are held constant.

## Statistical Assumptions and Robustness

For my model to accurately reflect my parameters it needs to meet the Gauss- Markcov assumptions. The first assumption is that there is a correct model specification. Using the RESET test, I found that all my quantitative variables were best modeled as a quadratic so this assumption is met, see page 10 for RESET test results. The next assumption requires that the data sample have both sufficiently large counts and be randomly sampled. I have approximately 8,500 randomly sampled households. Since my counts are larger than the traditional minimum of 30 – 50 this assumption is met. The third assumption requires that there is no relationship between the variables and the error term. The best way to prove this is through a plot of the residuals. The upper left plot on page seven shows a plot of the residuals including a fitted line. While the residuals are clustered in what appears to be a negative and linear, the residuals within that cluster appear to be random. The condition isn’t met with the certainty that is preferred, but for the purposes of this paper they are met. The next assumption is a requirement of no perfect collinearity. Since none of the variables in my regression are derived from each other this condition is met. The final condition requires homoscedasticity in the regression. On page six the plot on the bottom left shows the spread of the residuals along a range of predictors. Since the residual points are random and the line is in a horizontal direction the condition is met.

Given my R2 above fifty percent, I consider my parameters appropriately robust estimators of the true relationship between education and my other variables. While my parameters do not explain all the variation, given the breadth of my topic and the accessibility of relevant information my analysis is appropriate. I also used a data count well above 8,000 which further supports the reliability of my conclusions. I also compared my regression with and without outliers, as discussed on page 7-8, and found that there were only minute chances with and without the outliers

# Conclusions and Applications

Women’s education is key to the progress of a society. Unfortunately, education women and girl’s is easier said than done. There are infinitely many factors involved, factors that span genetics and luck to culture and birthplace. Previous studies have found that external stimulus, such as increasing education access, doesn’t suffice. We need to do more. My research focuses on what conditions are most associated with a higher level of female education within an individual African household. Given my results, each variable I tested for has worthwhile insight into how we can educate women better.

The first variable I tested was country location. During testing I found a wide variation in my education variable. Due to the number of characteristics that vary between countries, I can’t conclude anything specific as to the origin of this variation. Further research should occur on this topic, ideally in the form of a small-scale comparison between similar villages in different countries. I also tested variations in religion. The only two religions I can compare with any accuracy are *African Traditional* and *Other.* While the difference between the two only amounts to a few months of education difference, well within the margin of error, the coefficients provide worthwhile information. Since both coefficients are positive, relative to a random selection of nonresponsive households, religion has a positive impact on household female education. Much of the concern surrounding female education involves the patriarchal cultures featured in many of these households. The assumption is that this culture of patriarchy, often perpetuated by religion, is holding these women back. My results contradict this. They indicate that the practice of local religion by the head of household, most often a man, is associated with higher levels of education. Meaning female education is not as associated with religion and culture as previous research may suspect.

I also tested characteristics of the farm itself, including output and the presence of electricity. I found that farms with more kilograms of output tended to have more educated females on that farm. This results indicate that a possible explanation for this trend is infrastructure access. Larger farms likely have better facilities and more people to help if the farm needs more attention. A farm with higher yields also tends to be more productive, meaning that ease of farming may also allow some of the women’s labor to be spared in favor of educational attainment. Additionally, the presence of electricity had a positive impact on education. This supports my theory of farm infrastructure as well as ease of educational studying. Access to heat and light extend that amount of practical working time for an individual and may give them leisure hours to devote towards education.

Additionally, I tested the impact of the household’s men on the women through investigating a gender ratio as well as the impact of the surround men’s education. I found that up until a 2:1 ratio of females to males in a household, women’s education increases as more females live in the household. I also found that the education level of the household’s women increase with that of the men. These results indicate that female education is less a matter of resource allocation and more a matter of individual advocacy. If the household had limited resources and choose between educated males and females we would have the opposite results. Shifting from monetary investment to female advocacy will provide better outcomes for both women and their economies.

Supporting women in developing countries and their education is a multifaceted issue. My research and the body of information from around the world, can improve the lives of these women and their community for the better.

1. King, E M. “Educating girls and women: investing in development.” *World Bank*, no. 1990, 1990, p. 17. *Popline*, United States Agency for International Development (USAID, www.popline.org/node/379153. [↑](#footnote-ref-1)
2. 2 Women’s education prevents HIV: Jukes, Matthew & Simmons, Stephanie & Bundy, Donald. (2008). Education and Vulnerability: the role of schools in protecting young women and girls from HIV in Southern Africa. AIDS (London, England). 22 Suppl 4. S41-56. 10.1097/01.aids.0000341776.71253.04. [↑](#footnote-ref-2)
3. Bunnell R, Mermin J, De Cock KM. HIV Prevention for a Threatened Continent Implementing Positive Prevention in Africa. JAMA. 2006;296(7):855–858. doi:10.1001/jama.296.7.855 [↑](#footnote-ref-3)
4. Girls education in impact on childbearing: Gupta, Neeru, and Mary Mahy. "Adolescent childbearing in sub-Saharan Africa." *Demographic Research*8 (2003): 93-106. doi:10.4054/demres.2003.8.4. [↑](#footnote-ref-4)
5. Gender of pension: Esther Duflo; Grandmothers and Granddaughters: Old‐Age Pensions and Intrahousehold Allocation in South Africa, *The World Bank Economic Review*, Volume 17, Issue 1, 1 June 2003, Pages 1–25, <https://doi.org/10.1093/wber/lhg013> [↑](#footnote-ref-5)
6. Lloyd, Cynthia B., and Ann K. Blanc. "Children's Schooling in Sub-Saharan Africa: The Role of Fathers, Mothers, and Others." *Population and Development Review* 22, no. 2 (1996): 265-98. doi:10.2307/2137435. [↑](#footnote-ref-6)
7. Grimm, Michael, et al. “Does Electrification Spur the Fertility Transition? Evidence From Indonesia.” *Demography*, vol. 52, no. 5, 2015, pp. 1773–1796., doi:10.1007/s13524-015-0420-3.

   [↑](#footnote-ref-7)