Does the Family Dynamics Affect Education?

By Hazel, Cory, Allen and Albert
Linear Algebra Poster Project
Spring 2015, CSULB

Introduction

Previous studies have shown that parent's attitude and behaviors have a significant impact on how well their children perform in school and work environment. Children of well-educated and well-aware parents tend to be more responsible in their life and therefore do relatively better.

In this study, we will first explore the correlation between **mother's education** and the **respondent's education**. However, it is important to note that mother's education, although important, is not the only factor that affects the respondent's education. Ignoring other important factors can lead the results to suffer from omitted variable bias. Hence, to overcome such omitted variable bias we will include factors like **father's education** and **the number of siblings** to see how that affects the respondent's education. These new variables will also help isolate the effect of mother's education on the respondent's education.

Data

A sample data of 936 observations was used to explore the question on how the family dynamics affects the respondent's education. The respondent's were asked to share information on their number of years of education and the parents', and the total number of siblings. However, there were some respondents that did not answer all questions. After dropping the missing observations, the final data set comprised of a **total of 722 observations.**

Method

We used **The Least Squares Model** to analyze and understand this relationship between the respondent's education and other factors like mother's education, father's education and the number of siblings. The least squares line $y=\beta_0+\beta_1 x$ is called line of regression of y on x, where, y = dependent variable, x = independent variable and β_0 &β_1 are regression coefficients.

We can also write this system as:

$$X_{p \times (n+1)} \beta_{(n+1) \times 1} = y_{p \times 1}$$

where p = *# of obs. = 722* and n = *# of independent variables (1-3)*

Question: Is the system $\mathbf{X\beta = y}$ ***consistent for all y***?
Since n << p, $\mathbf{X\beta = y}$ does not have a unique solution.
This is an over-determined system.

Therefore we solve $\mathbf{X\beta = y}$ via the Normal Equation:

$$X^T(\mathbf{X\beta = y})$$

$$(X^TX)\beta = (X^Ty) \text{ (Now, } (X^TX) = \textit{full rank})$$

$$\textit{Therefore, we obtain } \beta = (X^TX)^{-1}(X^Ty)$$

MATLAB is utilized to solve βvia the Normal Equation as well as the QR Factorization. The codes are shown below for reference:

```
X = csvread('Datafor247.csv');
Y = X(:,1);
col_ones = ones(length(Y),1);
X1 = [col_ones X(:,3)];
X2 = [col_ones X(:,2:4)];
[Q1,R1] = qr(X1,0);
[Q2,R2] = qr(X2,0);

%% via the Normal Equation
beta1 = inv(X1'*X1)*(X1'*Y);
beta2 = inv(X2'*X2)*(X2'*Y);

%% via the QR-factorization
beta1_qr = inv(R1)*Q1'*Y;
beta2_qr = inv(R2)*Q2'*Y;
```

Results

Model 1: *education=10.48 + 0.29 mother's education*

β_0 of 10.48 suggests that on an average the respondent has about 10 years of education. β_1 of 0.29 suggests that there is a positive correlation between the mother's education and the respondent's education. The results infer that children with more educated mothers tend to also get more education.

Model 2: *education=10.36 + 0.21 father's education + 0.13 mother's education - 0.09 numb. of siblings*

β_0 of 10.36, similar to Model 1, also suggests that on an average the respondent has about 10-11years of education. Positive coefficients for both father's education and mother's education suggest that there is a positive correlation between the parent's education and the respondent's education. We also see evidence that there is an inverse relationship between the number of siblings and education.

Conclusion

In conclusion, family dynamics do have an effect on the respondent's education**. The results infer that children of educated and smaller families tend to get more education**. It is important to note that coefficient on mother's education shifted from 0.29 to 0.13 after adding new variables to the 2nd model. This magnitude change explains how 2nd model fixed the omitted variable bias problem in Model 1. This model could further be improved to obtain more accurate results by adding more factors like the family income, working or stay-at-home mothers, ethnicity, the country they reside in, etc.

Acknowledgements: Thanking Dr. Jen-Mei Chang for all her guidance and help with the models and MATLAB coding.
Citation: Crosnoe, R., Mistry, R. S. and Elder, G. H. (2002), *Economic Disadvantage, Family Dynamics, and Adolescent Enrollment in Higher Education.* Journal of Marriage and Family, 64: 690–702.
Lay, David. (2011) *Linear Algebra and It's Applications*