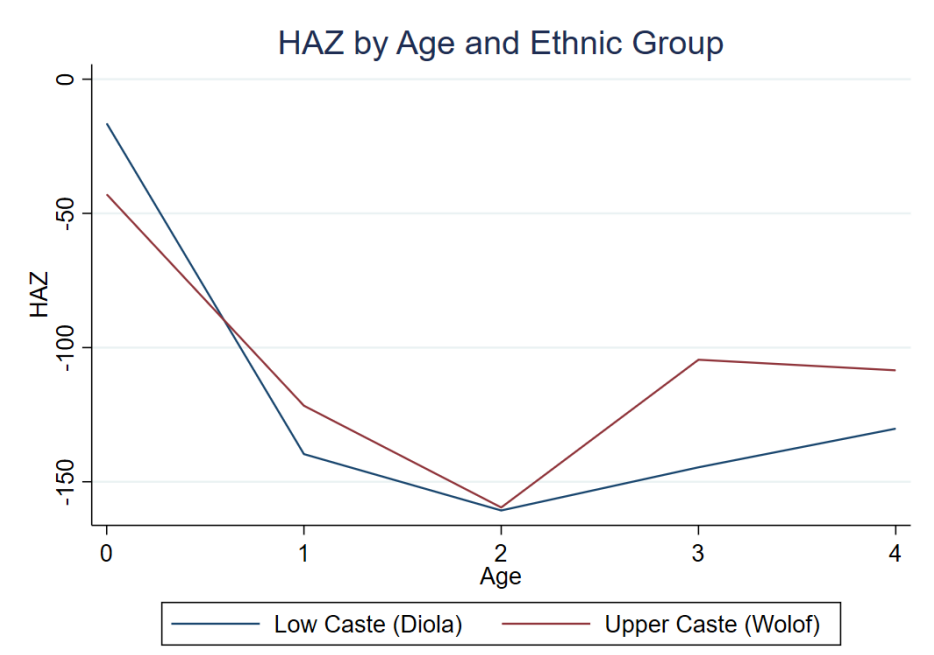
Eugene Baldonado
ECON193A
Professor Joseph Cummins
SID: 862070316

HAZ Difference in Disadvantaged/Advantaged Groups: Diola and Wolof tribes in Senegal

The differences in socioeconomic status and between the Senegal tribes of the Diola (majority Ethnic group) and Wolof (minority Ethnic group) should seemingly produce differences in the biological age-dynamics in child growth patterns between the two groups; however in statistical analysis, this difference was not as prominent as expected. In the 2010 Senegal IPUMS-DHS dataset collected, the two social groups selected were the Diola (Jola) and Wolof ethnic groups, as there were key differences between the two groups that could make one group (the Diola) more disadvantaged than the latter, and were thus tested for any differences.

The statistical analysis performed on these two groups regressed HAZ (Height adjusted Z-score from the median provided by the World Health Organization) unconditionally on a child's current age between the two groups (Diola and Wolof); as well as a conditional regression analysis that included controls for various endowment and investment variables. In these analyses, sample weight (pweight) is accounted for accordingly to factor in differences in samples to represent their population accordingly. Aside from these conditional and unconditional regressions, an oaxaca decomposition analysis was also performed to find how much the differences between the two groups' HAZ scores were explained by either endowment or investment variables. The results varied, **shown below is the HAZ differences between the Diola and Wolof**, which showed the Diola having larger babies initially, and then the Wolof having larger children in ages 3-4, on average.



The results from the unconditional and conditional regressions, revealed that there were already minimum group differences in the unconditional regression, which was noticed since the differences between the two groups contained 0 in each ages' confidence intervals when accounting for standard errors. Additionally, controlling for conditional covariates yielded varying results; such as controls producing larger group differences instead of smaller (closer to zero) for some ages, in comparison to the unconditional regressions. These results definitely weren't to be expected, but there could have been omitted variable biases from the regressions, which can be backed by literary review that can explain why the differences in HAZ-score may be smaller between groups.

From the two social groups selected in the Wolof and Diola, the Wolof being a primary ethnic group in Senegal, this group was chosen as the advantaged group. The Wolof, although primarily more rural, contained more urban demography than the Diola and were clustered into smaller villages headed by chiefs (Appendix) and contained higher maternal height and weight per child, on average. In contrast, the Diola were much more urban and were Senegal's main producers of the country's staple goods' rice, "in addition to other crops for domestic consumption and export" (Diola). As for why they were chosen for the disadvantaged group, the Diolas of Casamence (Southern Ethnic groups of Senegal) during this time period of 2010, experienced prevailing civil wars, and were often called pejorative terms from Northern Senegalese tribes (*specifically* ethnic Wolofs) (Diola(Jola)). This meant that Northern Senegalese tribes always referred to the Southern Senegalese Ethnic groups as lesser, with one term "naq" literally meaning a "primitive person from the bush" (Appendix). The Diola also had interethnic tensions between other southern ethnic groups that played into a "long-running Casamence rebellion that was characterized by grievous human rights abuses' " (Bureau). The combination of interethnic tensions causing civil wars between other Southern Senegalese groups, socio-inferior status from Northern Senegalese groups, and their rebellions against the Wolof made this minority group seem relatively disadvantaged to the Wolof group. Despite these shortcomings, it is important to relay that the Diola's more rural demography and growth of crops could have caused the differences in HAZ scores in children in comparison to the Wolof to be significantly *less* considering the Diola were a main exporter of Senegal's staple crops. The Diola ethnic group provided Senegal most of their staple goods (either for domestic consumption or export) which is a possible investment variable that was not included in the IPUMS-DHS dataset. If this variable could have been controlled for, it could have made the group differences closer to 0, or more in favor of the Wolof.

In measuring human capital in the following regression analyses, HAZ score can be a viable measure of health capital as it can represent a child's starting point and growth over time with differing endowment and investment variables. In this case, at the child's starting point it can be expected that the large portion of differences is explained by endowment variables, and over time investment variables can alter children from relatively close starting points to larger group differences between them. Although papers regarding height and growth of children put most of their importance on endowment variables, research suggested by Professor Joseph

Cummins (Neha Agarwal, et. al) displayed that when accounting for children in different caste groups in India, the difference in HAZ score could be explained by both endowment and investment variables almost equally. Considering the IPUMS-DHS dataset on Senegal didn't have more specific data on the specific diets the children were receiving throughout their growth period, this could have been a significant investment variable when accounting for HAZ score differences between the Wolof and the Diola- especially due to the Diola harvesting a large portion of Senegal's crops either for domestic uses or for exports. Nevertheless, HAZ score proves to be a reliable indicator of health capital when including covariates that measure the differences between two groups with both endowment and investment directly correlating to the outcome.

The collection of data from the IPUMS-DHS website included the original data from the Senegal 2010 dataset which included 12,326 total points of data. From these points, 3,837 were from the Wolof ethnic group, and 464 were from the Diola ethnic group. After cleaning for all the variables and covariates used, there were a total of 1,432 points, of which 1,269 were from the Wolof ethnic group and 133 from the Diola ethnic group. As said previously, regression analysis included per weight recognition, which took sample weight into account. This way, each ethnic group samples were able to accurately represent their population. In analyzing these two groups, it can be important to take a look at the **sample summary statistics (below)** to find differences between the two groups in various endowment and investment variables relating to health capital.

Sample Summary Statistics by Group

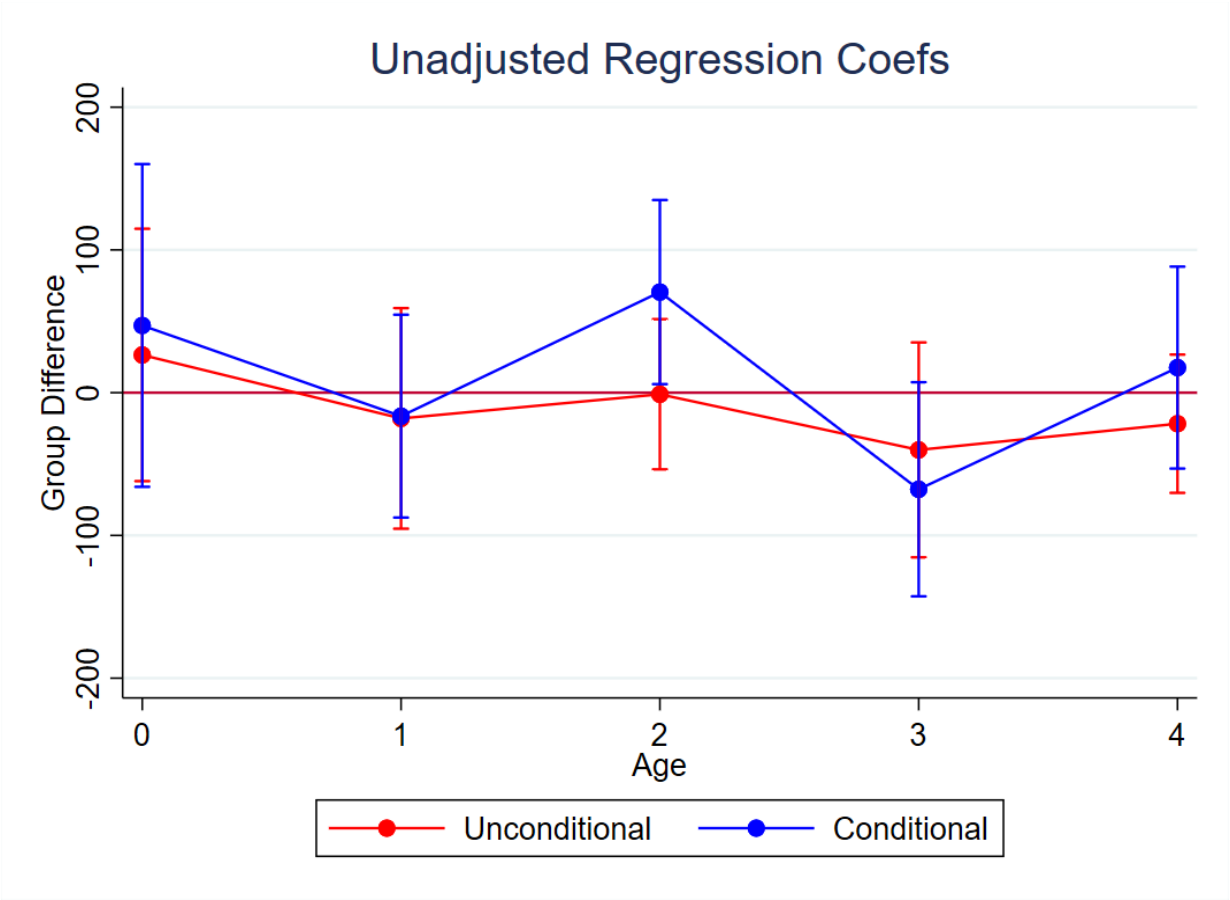
	Wolof mean	Diola mean
height for age standard deviations from median (who)	-114.8	-106.1
current age of child in years	1.84	1.53
urban-rural status	0.29	0.43
highest educational level	0.26	0.92
age	29.5	30.5
household wealth index in quintiles	2.80	2.76
weight of woman (kilos)	600.7	579.7
height of woman in centimeters	1643.8	1592.5
has electricity	0.66	0.65
barrier to woman's health care: getting permission	0.21	0.15
barrier to woman's health care: lack money for treatment	0.58	0.53
barrier to woman's health care: distance to facility	0.37	0.40
barrier to woman's health care: not want to go alone	0.16	0.11
Observations	1269	133

It can be seen that the Wolof contains a larger portion of urban civilization, a slightly higher wealth index, and mothers who are slightly heavier and slightly taller than those of the Diola. Overall, they seem to have similar barriers to health care across the board. As shown in the summary statistics, it seems certain endowment variables in weight/height reveal that mothers of the Wolof tribe are slightly larger than those of the Diola. Although similar in many aspects, differences in the Wolof and Diola's endowment and investment variables may explain their group differences in HAZ-score, to which statistical analysis can be performed on both conditional and unconditional regressions to analyze.

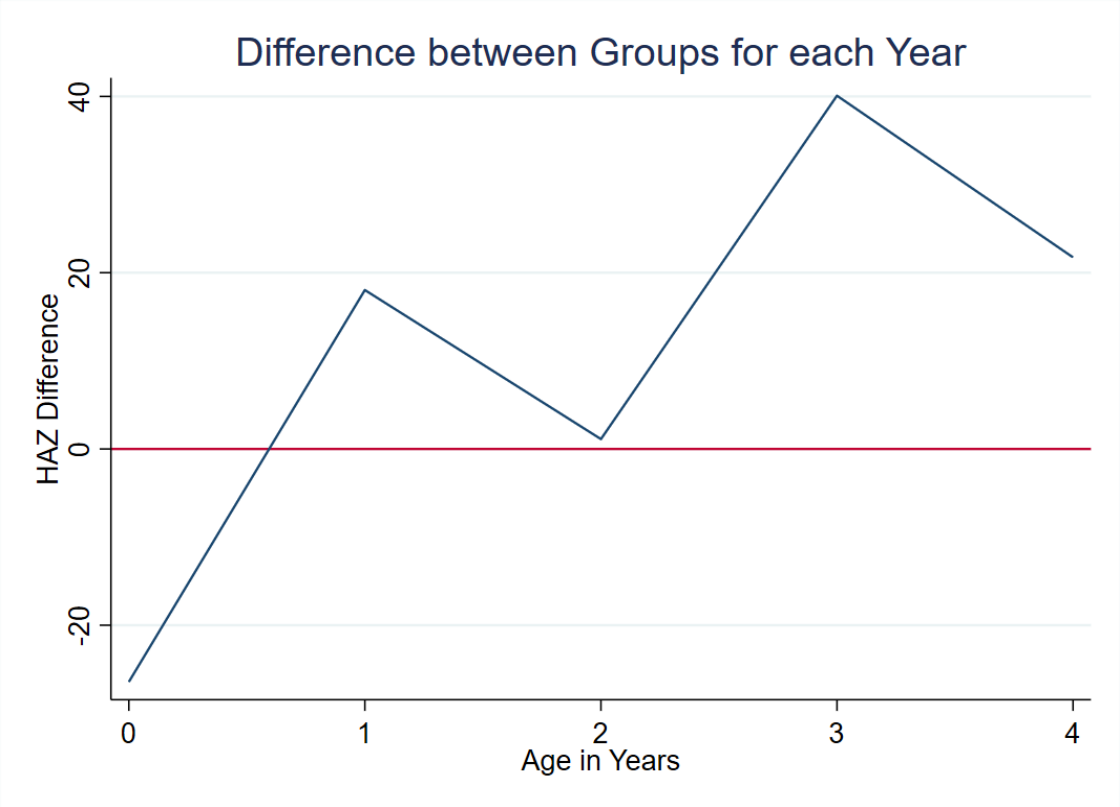
$$Y_{ig} = \beta_0 + \Delta * Diola_{ig} + \beta_1 * X_{ig} + \boldsymbol{\varepsilon}_{ig}$$

*Highlighted is included in Conditional Regression, not in Unconditional

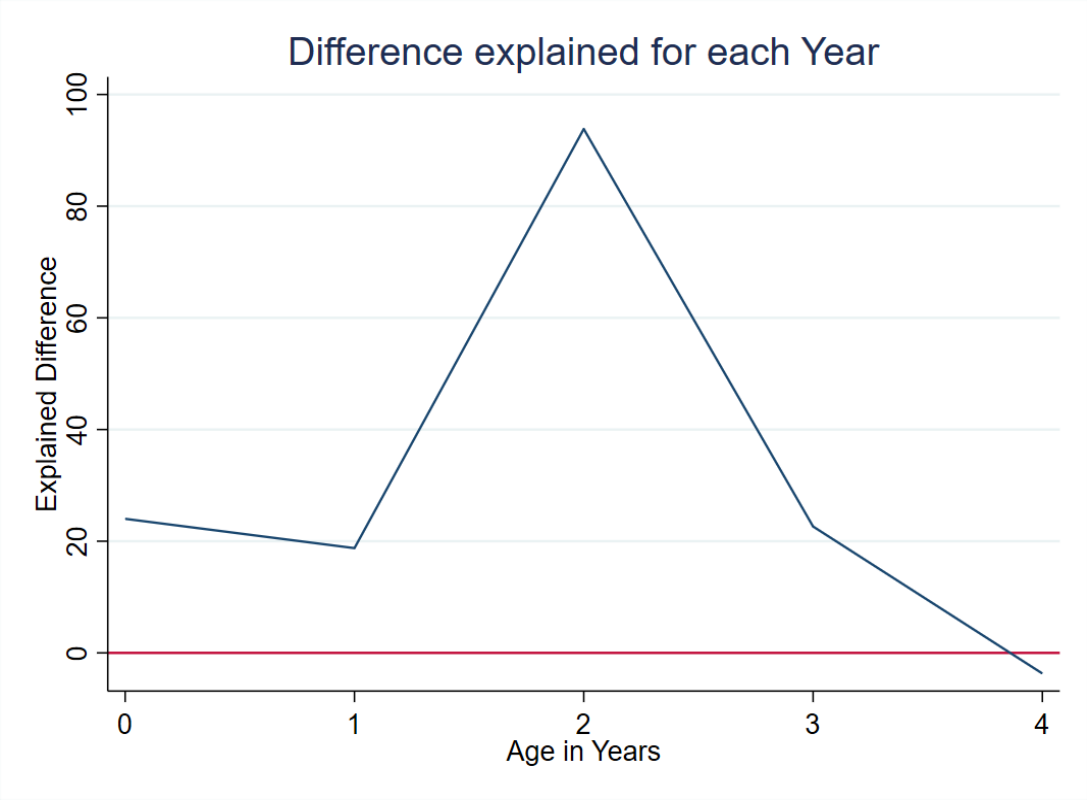
Statistical analysis was performed for the Wolof/Diola HAZ-score difference by regressing HAZ score on a dummy variable representing whether the individual child **i** (subscript) was either in the Diola **g** (subscript) socioeconomic/ethnic group. This regression (not including $\beta_1 * X_{ig}$) is the unconditional regression and represents the average difference between the two groups by the coefficient Δ. In this case, the coefficient Δ represents the average HAZ score difference if an individual child is from the Diola ethnic group. The coefficient β_0 represents the average HAZ score of a child from the Wolof ethnic group (when the Diola dummy variable is 0) and the coefficient $\boldsymbol{\varepsilon}$ represents the error term, which includes any omitted variables or variability not included in the regression model. As for the conditional regression, the covariates include varying endowment and investment variables to be controlled for. The conditional regression includes variables controlling for weight/height/age/education level of the mother, the family's wealth quartile index, the delivery place the baby was born, whether they treat their water, if they have electricity, and barriers to health care. These covariate controls are included in separate β's but are simplified in the regression equation example. From the unconditional and conditional regressions, the standard error was multiplied by 1.96 (5% critical value) and added and subtracted to the group differences to find the confidence intervals for each age from 0-4 at the 5% level. **Results from these regressions are shown below.**



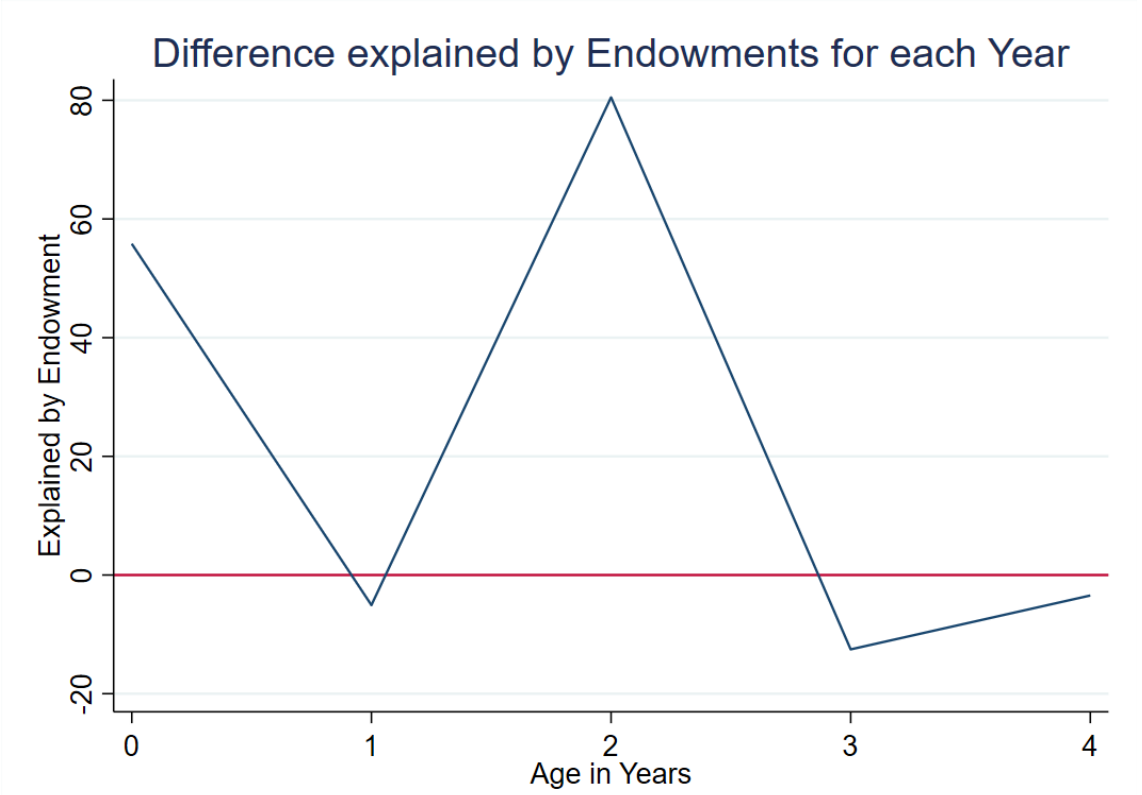
The results find the unconditional group differences for Child HAZ-age score contains 0 in each of the group difference confidence intervals from ages 0-4, which means at the 5% level of significance, we cannot reject the null hypothesis that the group difference between the Wolof and Jiola tribe isn't zero. In trying to control for the covariates to explain the differences between the Wolof and Diola HAZ score differences, there are varying results. The HAZ difference actually increases for the Diola (disadvantaged group) on average for ages 0, 2 and 4, predicting that the Diola actually would contain higher HAZ-scores than the Wolof under the regression with these controls, but stays about the same for age 1, and predicts even larger differences between the Wolof and Diola at age 3 (with Wolof holding higher HAZ-scores), on average. In controlling for these covariates, the differences between the Wolof and Diola groups still contained 0 in all ages except for 2 for individual children's HAZ score, on average; similarly to our unconditional regression, we cannot reject the possibility that the differences between the Wolof and Diola's HAZ-score are 0 at the 5% level of significance.

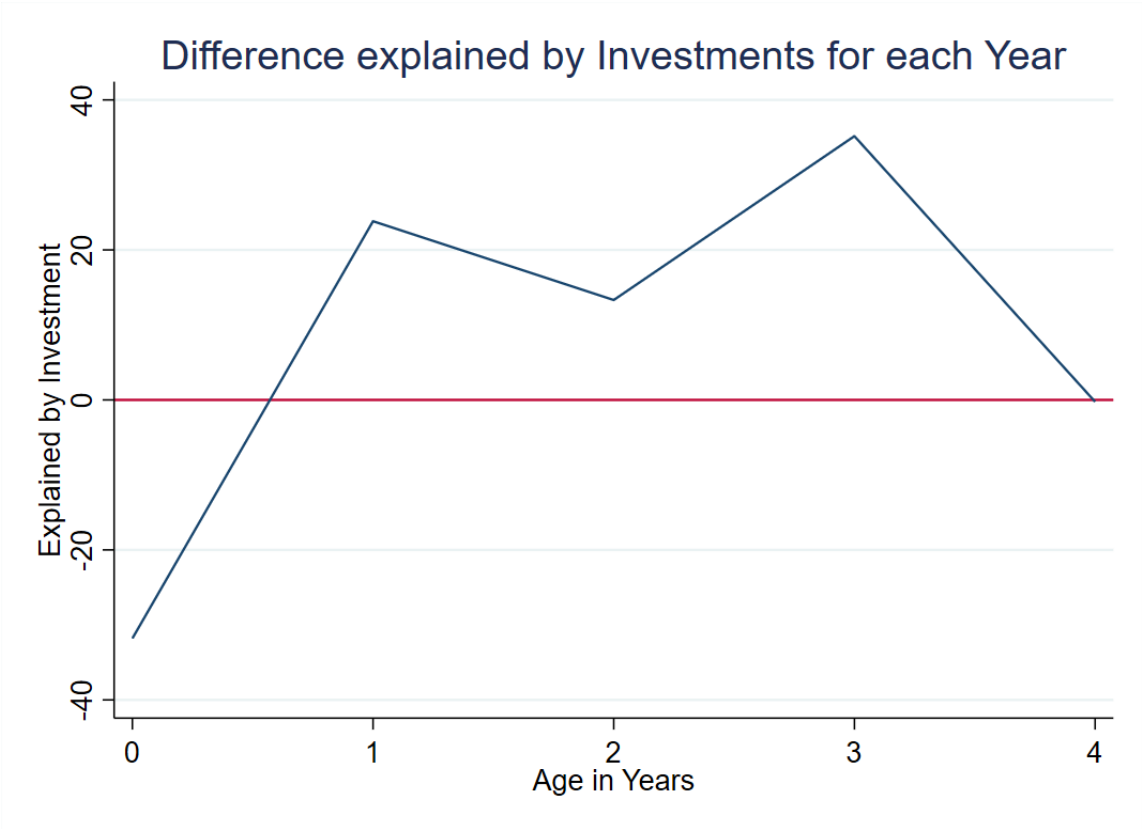


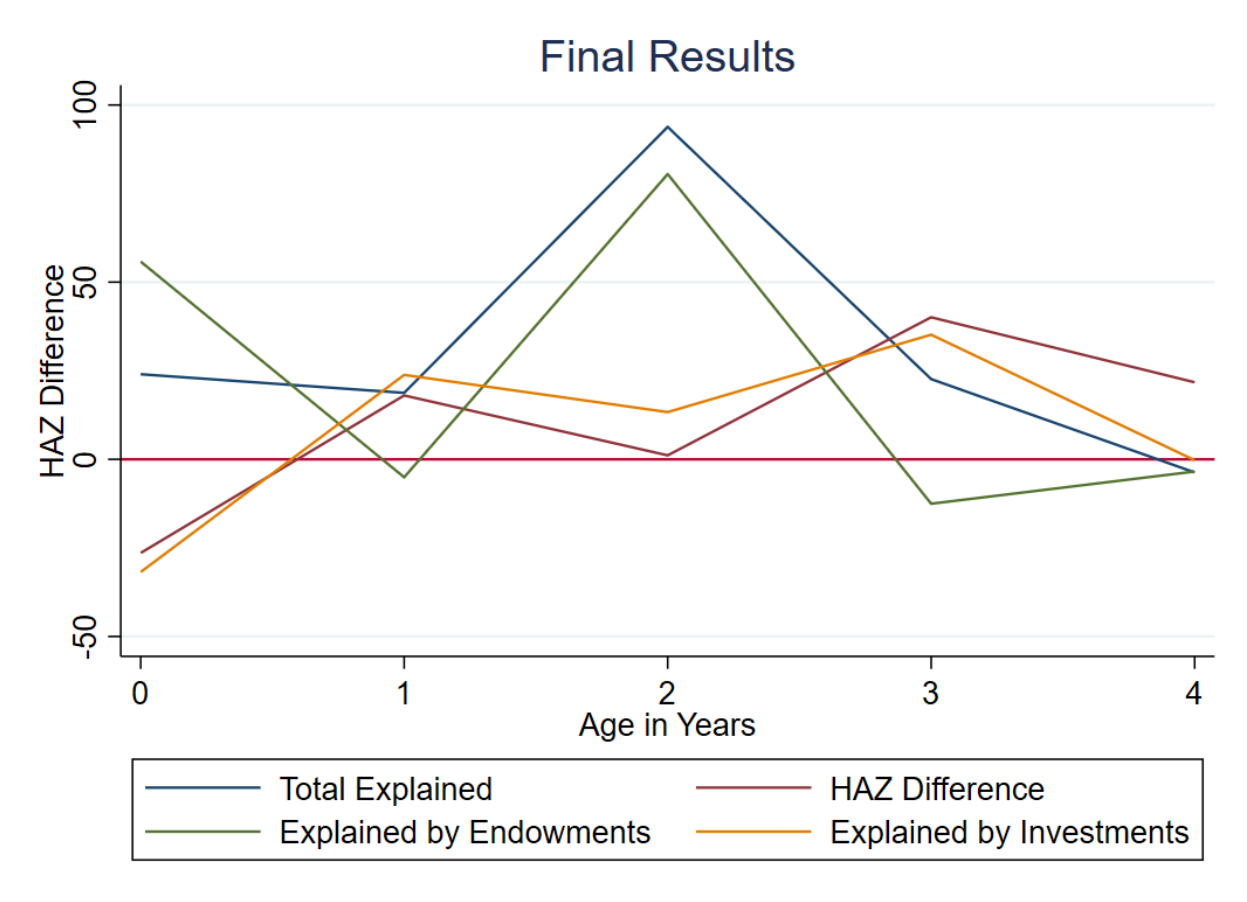
Above is the HAZ score difference between the Wolof and the Diola in a singular graph. Revisualizing this graph will help explain the oaxaca-blinder decomposition results from the dataset. From this dataset, the oaxaca-blinder decomposition will describe the group differences in how much of the group differences were explained by either endowment or investment variables and how much of that explained total was from endowment or investment. In this case, the HAZ score difference between the Wolof and the Diola will be described by the endowment variables: mother's age/weight/height, and delivery place, and investment variables: wealth quartile range, access to electricity, educational level, and barriers to health care. From the results, it can be found that the largest (unfortunately, smallest difference in HAZ score between groups) portion explained by endowment/investment variables were in children age 2, while children aged 0, 1, 3 and 4 all were largely left to unexplained differences not in the variables included in the endowment or investment variables. **This can be seen in the graph below which displays the total difference explained for each year.**



In addition to the total difference explained by either investment or endowment variables, the oaxaca-blinder decomposition reveals how much is explained by investment or endowment, which found interesting results. Endowment explained most of the miniature difference in age 2 between the HAZ-scores of the Wolof and Diola, and generally decreased after the child was born. This means that the endowment variables were explaining most of the difference in HAZ score when the child was born, and as the child grew older (aside from age 2) explained less of the difference from the HAZ-scores in the Diola and Wolof ethnic groups. In contrast, the results from the investment variables seemed to explain more of the HAZ-score differences past the child's birth. This can be seen in the results, as the explained difference by investments starts negative, and increases (spiking at age 3) which shows that over the child's growth more of the difference in HAZ-score can be represented through the investment variables of the mother. This decomposition concludes that the differences between the Wolof and Diola have portions explained by both endowment and investment variables, with their effects varying across ages. **Below are the graphs displaying the group differences in HAZ scores explained by Investment or Endowment variables.**


Difference explained by Endowments for each Year
Explained by Endowment
80
60
40
20
0
-20
0
1
2
3
4
Age in Years


Difference explained by Investments for each Year
Explained by Investment
40
20
0
-20
-40
0
1
2
3
4
Age in Years



Above are each of the four graphs from the oaxaca-decomposition in a single graph.

In conclusion, although showing some differences in the HAZ-age profile graph, it cannot be concluded that there were differences in the HAZ scores of the Diola and Wolof tribe. In the HAZ-age profile graph, it could be seen that the Wolof contained a higher HAZ-score from ages 1 to age 4, with the Diola only containing a higher HAZ at birth, on average. These results, when compared to the regression analysis performed unconditionally and conditionally (controlling for covariates) still found that the confidence intervals contained 0 across each age, which means that it cannot be concluded that these group differences are not zero. Although the Diola had a lesser perceived socioeconomic status by the Wolof, had civil wars, and interethnic issues disadvantaging them in comparison to the Wolof, they also produced some of Senegal's key crops, which could have been an investment variable (not included in the dataset) that provided equally high HAZ-scores as the advantageous Wolof ethnic group. Due to the possibilities of omitted variable bias, and lack of statistical significance, it cannot be concluded that the Wolof and Diola had varying HAZ-scores across the ages of children.

Works Cited

Agarwal, Neha, et al. “Caste Differences in Child Growth: Disentangling Endowment and Investment Effects.” *UCR*, 2022, 0 - 29.

Appendix 1: Demographics of Senegal: Ethnicity and Religion (By Region and Department in %). https://link.springer.com/content/pdf/bbm%3A978-0-230-61112-2%2F1.pdf.

Bureau of Democracy Human Rights and Labor. “2010 Human Rights Report: Senegal.” *Justice.gov*, U.S. Department of State Diplomacy in Action, 2010, https://www.justice.gov/sites/default/files/eoir/legacy/2013/06/10/senegal_1.pdf.

“Diola (Jola).” *Minority Rights Group*, MRG, 20 Jan. 2021, https://minorityrights.org/minorities/diola-jola/.