Empirical Assignment

Niklas Leander Kampe - 16-611-618

December 17, 2020

1 Descriptive statistics

In order to receive a summarized look on the data set of Uganda, the first chapter takes a closer look on the most important individual characteristics by the surveyed as an approximation for the whole population.

1.1 Average age of population & age distribution

The average age of the population in Uganda based on the National Panel Survey between 2013-2014, adjusted by the survey-specific weights, is **22.5 years** (indicated by the red line in Figure 1). Furthermore, the age distribution across the surveyed looks as follows:
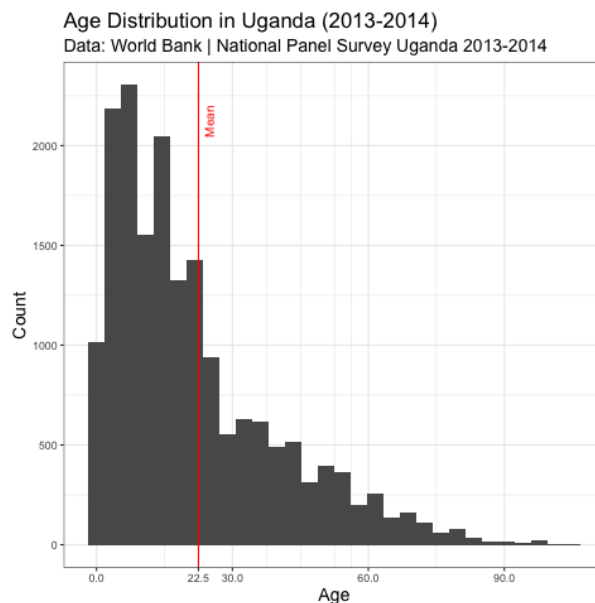

Figure 1: Age Distribution in Uganda (2013-2014)

In contrast to Western countries, in which the age is distributed in a honeycomb-like manner, with the middle-aged (40-60 years) making up the largest proportion of the population, Uganda has a flattening age distribution with the highest proportion of young people. Furthermore, there is a significantly lower proportion of the old population, which suggests a lower life expectancy (higher mortality) as well as a higher fertility which are a sign of a not yet fully completed demographic transition in Uganda compared to high-income countries.

1.2 Average years of schooling & education gap

The average years of schooling in Uganda based on the National Panel Survey between 2013-2014 is **6.73 years**. Furthermore, the education gap between males and females, calculated as the weighted mean of gender-adjusted years of schooling, results in **0.594 years**. In conclusion, the average male citizens thus has around **7 months more** of schooling compared the average female citizen.

1.3 Literacy rates (by age groups)

The share of people with the ability to read and write in Uganda, defined as the literacy rate, based on the National Panel Survey between 2013-2014 is **62.96%**. It's important to note that the ability to only write or only read is not included in the literacy rate. Split up by the age groups of 15-25, 25-35, 35-45 and 45-55 years-olds, the distribution of the literacy rates looks as follows:
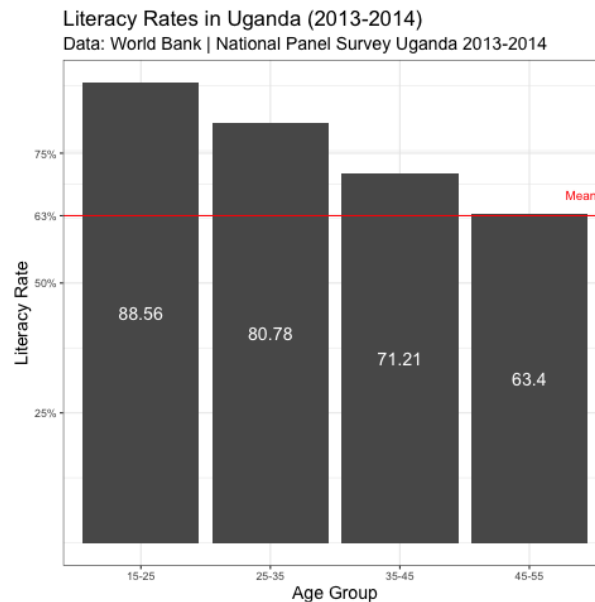

Figure 2: Literacy Rates by Age Groups in Uganda (2013-2014)

Compared to the mean of 62.95% over the whole population, all included age groups exhibit literacy rates above the mean. Furthermore, strict decreasing rates in increase of the age can be determined. While the mean is below the literacy rates of all included age groups, one can comprehend that a further decreasing ability to read and write in even older age groups is to be expected. A potential reason for that negative correlation between the age and the literacy can be the improving education system in the last decades which led to steadily increasing literacy rates across the population. While the ability to read and write is mostly taught in the young ages, the decreasing rates in ages are expected to stay, but at higher levels over future middle- and long-time horizons as long as the increase and improvements in the education system remain.

1.4 Education levels (by age groups)

The shares of population with a primary, secondary and tertiary education within the working age population of 15-55 years-olds in Uganda, defined as the highest completed education levels, based on the National Panel Survey between 2013-2014 are **41.49%, 8.45% and 7.93% respectively**. Split up by the age groups of 15-25, 25-35, 35-45 and 45-55 years-olds, one get the following education levels:

(a) Education Level Rates
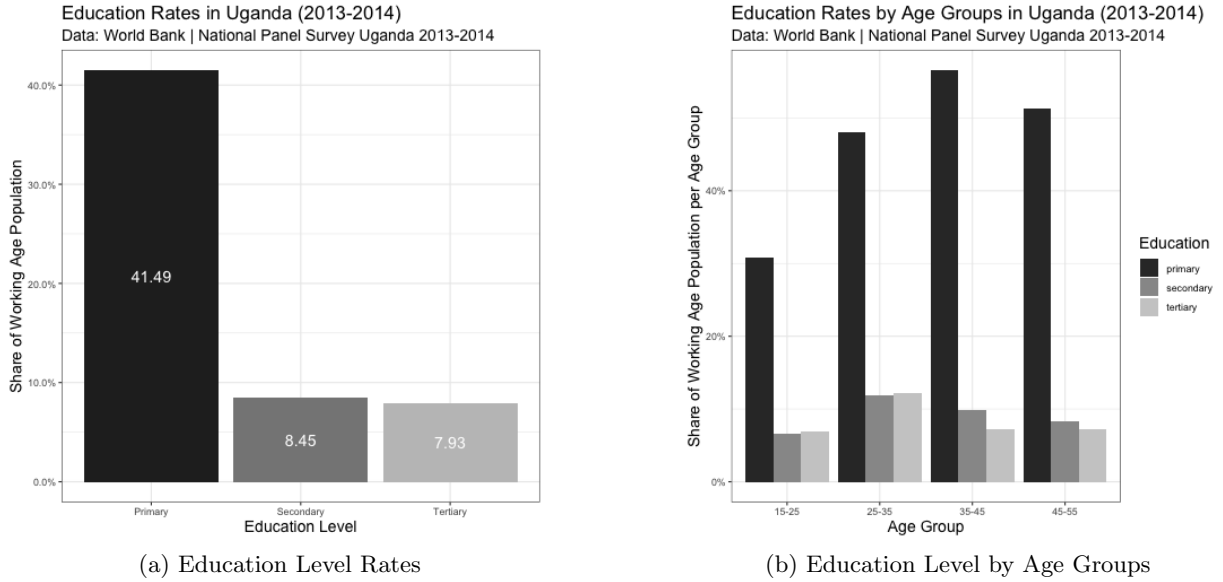
(b) Education Level by Age Groups

Figure 3: Education Level Rates of Working Age Population in Uganda (2013-2014)

In accordance with Figure 3, it must be noted that the data is computed based on the highest completed degree according to primary, secondary and tertiary education. Therefore, it's possible that a higher share of the working age population per age group have completed a tertiary compared to the share completed a secondary education as the education levels are not aggregated. Across all age groups, it can be seen that completion of primary education represents by far the highest proportion per age group. Due to the comparatively low proportion in the 15-25 age group, a relatively low value can also be identified across all age groups, since at the same time a relatively high proportion of the entire aged population is made up (see Figure 1). With regard to the education level distributions within the age groups, the images are similar, although the low proportions of secondary and tertiary education are obvious. However, secondary and tertiary education are at very similar levels across age groups. Compared to Western countries, it can be said that the education rates across age groups and education levels are significantly lower, which is related to many social, political and legal issues, such as non-compulsory school attendance across Uganda. Potential policies to further incentive education in the population in order to ultimately affect the income levels are compulsory schooling, the reduction of schooling costs split up into time and expenditure costs or subsidizing schooling investments which may not be optimal due to credit constraints, but also the steering of demographic development with lower fertility (to increase education investments per child) and lower mortality (to increase the profitability of education due to a longer working horizon).

2 Labor markets in low-income countries

In order to get an insight about the labor markets in low-income countries, on the example of Uganda, the following chapter will take a closer look on the structure of the labor market, the amount of hours worked, the wage distribution as well as the returns to education measured as the increase in wage due to years of schooling.

2.1 Labor market status

The first section of the labor market analysis states the labor market status in Uganda in the form of the share of the working age population and the employment rates.

2.1.1 Share of working age population

The share of the working age population, measured as the rate of 15-55 years-olds to the whole population, based on the National Panel Survey between 2013-2014 is **47.30%**. Therefore, 52.70% of the whole population, approximated by the surveyed, are either under 15 or over 55 years old and not defined as being in the working age.

2.1.2 Employment rate (wage vs self-employed)

The employment rate of the working age population (definition as in chapter 2.1.1), measured as people that either (i) work for any payment, (ii) run a business, (iii) help in any business without any kind of payment, (iv) work as an apprentice or (v) work on the owns household's farm based on the National Panel Survey between 2013-2014 is **70.59%**. Furthermore, the share of the working age population working as a wage worker is **19.36%** and working as a self-employed, either employing or not employing any other people, is **15.50%**. However, the highest share is represented by unpaid workers.

2.2 Wage measurement

The second section of the labor market analysis deals with the wage measurement in Uganda, in particular the hourly wages across the population, for which also the weekly hours worked have to be computed.

2.2.1 Weekly hours worked

Based on the National Panel Survey 2013-2014 in Uganda, one can compute the weekly hours worked by wage workers, who represent around 19% of the working age population (see chapter 2.1.1) in the last 7 days from the date of the personal interview. The resulting distribution looks as follows:
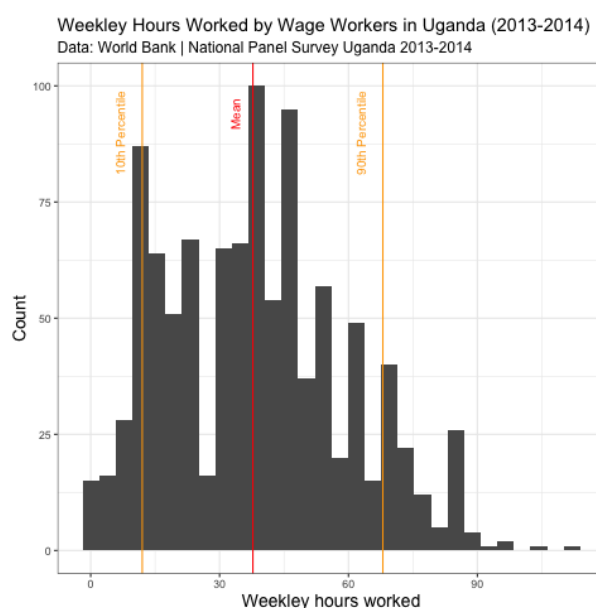
Figure 4: Weekly Hours Worked in Uganda (2013-2014)

Based on the distribution of the weekly hours worked, it can be determined that the weighted mean is **37.75 hours**. Furthermore, the 10th and the 90th percentile are **12 and 68 hours** respectively. Furthermore, according to the graph, one can see that there is also quite a major share which works a lot more than 40-50 hours which defines a full-time equivalent in the Western countries.

2.2.2 Hourly wages

The hourly wages, computed out of the last wage income received and the weekly hours worked, based on the National Panel Survey 2013-2014, are distributed as follows:
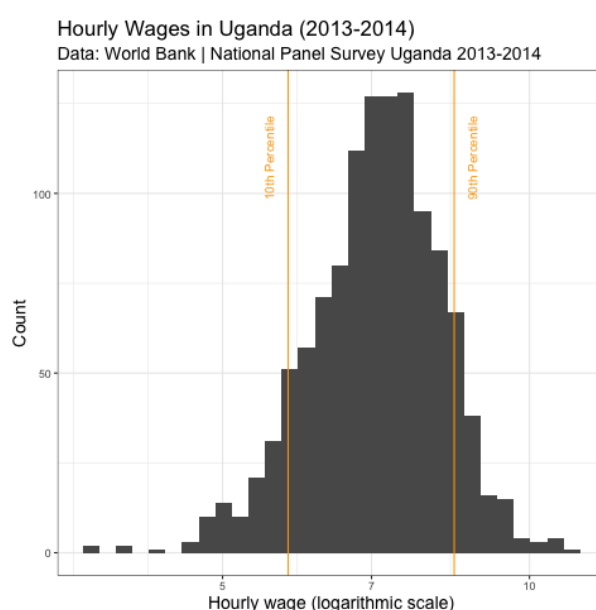

Figure 5: Hourly Wages in Uganda (2013-2014)

The distribution of the hourly wages is based on the adjustment of a logarithmic scale in order to be able to show the large variance within the wage levels more clearly. In accordance to the figure and the computive statistics, it can be determined that the wages are primarily distributed **between 330 and 4,500 USh** (back-transformed from logarithmic scale) which define the 10th and 90th percentile of the data.

2.3 Returns to education

In the third section of the labor market analysis, we will take a look at the the return to education measured by the marginal increase in hourly wages, as well as including control variables based on individual characteristics.

2.3.1 Returns to education (excl. control variables)

In order to measure the correlation between the hourly wage and the years of schooling, following regression has been run:

$$\log(w_i) = \alpha + \beta_1 s_i + \beta_2 s_i^2 + \epsilon_i$$

where w_i donates the monthly wage, α the intercept, s_i the number of school years, β_1 and β_2 the corresponding coefficients of the polynomial regression and ϵ_i the unobserved error term. i represents the notation per individual. Based on this regression model, the logarithmic monthly wage is regressed on the explanatory variable of the years of schooling using the data from the National Survey Panel 2013-2014. The following results are achieved:

	Model 1
(Intercept)	6.80***
	(0.19)
Education Years	−0.07
	(0.05)
Education Years ^{2}	0.01***
	(0.00)
R^2	0.20
Adj. R^2	0.20
Num. obs.	561

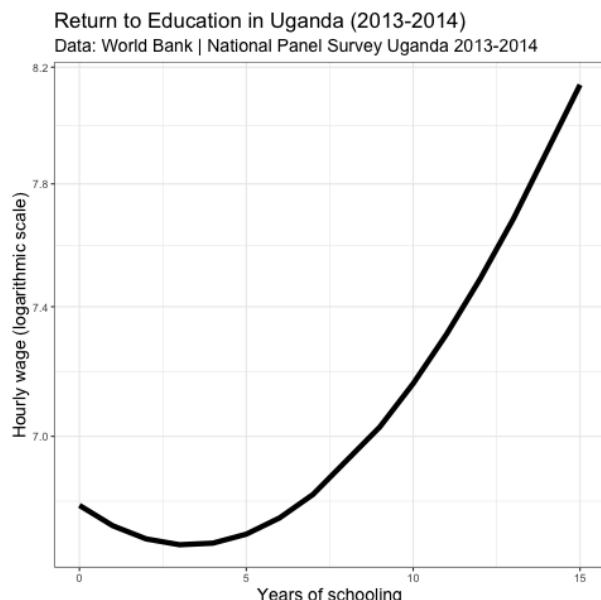***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Table 1: Regression Output and Plot (excl. control variables)

Based on the regression run, the following hourly wages are predicted using the education years between 0 and 15 which represents the schooling and education system in Uganda:

	Education Years	log(Predicted Hourly Wage)	Predicted Hourly Wage
1	0.00	6.80	894.22
2	1.00	6.74	842.88
3	2.00	6.70	811.54
4	3.00	6.68	798.13
5	4.00	6.69	801.78
6	5.00	6.71	822.74
7	6.00	6.76	862.36
8	7.00	6.83	923.29
9	9.00	7.03	1127.97
10	10.00	7.16	1287.09
11	11.00	7.31	1500.17
12	12.00	7.49	1786.05
13	13.00	7.68	2172.04
14	15.00	8.14	3423.59

Table 2: Predicted hourly wages (excl. control variables)

The choice of the polynominal structured regression is usually based on the assumption that an increase in the explanatory variable is resulting in a even relatively further increase in the dependent variable which leads to a non-linear relationship, and so it is assumed in this specific case. Furthermore, the explanatory variables education years and education years squared are treated as orthogonal parameters to avoid collinearity which could lead to misleading estimations. Finally, the logarithmic scale on the y-axis in the dependent variable is used to estimate the percentage increase based on a change in the explanatory variable. Looking back on the regression results in Table 1, the intercept (accordingly α in the underlying formula) defines the expected monthly wage in logarithmic terms when years of schooling are zero $E[log(y)|x = 0]$. Since a polynomial regression of the second degree is used in this case, the interpretation must be observed taking into account the respective other degree and does not correspond to the same interpretation as that of the linear regression. The graph, the coefficients and the estimation results show that there is a percentage reduction in the hourly wage up to the fourth year of education and that a percentage increase can only be observed from the fifth educational year onwards. The negative value of the first and the positive value of the second β-coefficient could explain that the wage growth correlates negatively with the education years until a specific threshold value, which is defined as the fourth education year in this case, and positively upwards that threshold. This observation can have several reasons, such as the comparatively high level of hourly wages for individuals without or really small levels of education, the small sample size and the associated relatively high weighting of individual observations, as well as the fact that fully completed degrees of education define threshold values for the formation of expectations within the job market and employees. However, it is difficult to draw concrete conclusions from the data and the results. Furthermore, the variable R^2 indicates the share of the explained variation in the dependent variable by the explanatory variables, in this case the share of the variation in hourly wages that can be explained by the number of years of education. At first glance, the share of 20% seems rather small, but it is already a remarkable result

if one keeps in mind that the hourly wage probably depends on a number of other variables relating to an individual. It can be concluded that the value of a completed school education (primary, secondary or tertiary) leads to higher hourly wages in percentage terms and thus puts the individual in a relatively better position. However, it must be individually weighted whether the financial costs as well as the opportunity costs during the education period can be offset or exceeded by the comparative advantages in the form of the higher hourly wage over the period of employment. Ultimately, other explanatory variables must also be included in order to increase the degree of certainty (R^2) and to explain the variation in hourly wages more precisely.

2.3.2 Returns to education (incl. control variables)

On the basis of the results discussed in the previous chapter, further control variables are now included in order to further increase the explanatory power of the regression. The following adapted regression is performed:

$$\log(w_i) = \alpha + \beta_1 e_i + \beta_2 e_i^2 + \gamma_1 \text{gender}_i + \gamma_2 \text{location}_i + \gamma_3 \text{age}_i + \epsilon_i$$

where w_i donates the monthly wage, α the intercept, e_i the number of school years, β_1 and β_2 the corresponding coefficients of the polynomial regression and ϵ_i the unobserved error term. Furthermore, the control variables $gender_i$, $location_i$ (both dummy variables) and age_i are introduced which determine if the person is female (value = 0) or male (1), if it lives in a rural (0) or urban (1) location and how old it is. i represents the notation per individual. Based on this regression model, the logarithmic monthly wage is regressed on the explanatory variable of the years of schooling as well as the characteristics gender, location and age using the data from the National Survey Panel 2013-2014. The following results are achieved:

	Model 2
(Intercept)	5.90***
	(0.22)
Education Years	−0.06
	(0.05)
Education Years ^{2}	0.01***
	(0.00)
Gender	0.34***
	(0.08)
Location	0.10
	(0.08)
Age	0.02***
	(0.00)
R^2	0.27
Adj. R^2	0.27
Num. obs.	561

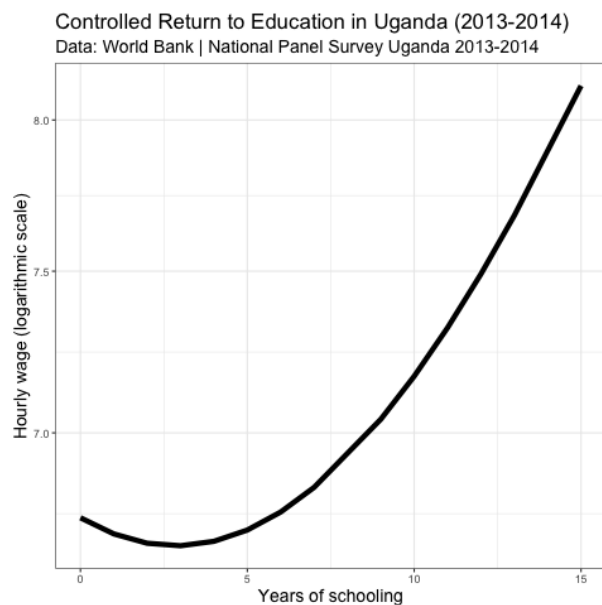***$p < 0.001$; **$p < 0.01$; *$p < 0.05$

Table 3: Regression Output and Plot (excl. control variables)

Based on the regression run, the following hourly wages are predicted using the education years between 0 and 15 controlled by the individual characteristics gender, location and age:

	Education Years	log(Predicted Hourly Wage)	Predicted Hourly Wage	Δ to Model 1
1	0.00	6.75	854.76	-39.45
2	1.00	6.70	816.05	-26.83
3	2.00	6.68	794.54	-16.99
4	3.00	6.67	788.95	-9.18
5	4.00	6.68	798.93	-2.85
6	5.00	6.72	825.08	2.34
7	6.00	6.77	868.99	6.63
8	7.00	6.84	933.39	10.10
9	9.00	7.04	1142.21	14.24
10	10.00	7.17	1301.31	14.22
11	11.00	7.32	1511.97	11.80
12	12.00	7.49	1791.59	5.54
13	13.00	7.68	2165.02	-7.02
14	15.00	8.12	3353.49	-70.10

Table 4: Predicted hourly wages (incl. control variables)

As already emerged from the model in Chapter 2.3.1, based on the coefficient of determination of 20%, the hourly wage depends not only on the number of years of education, but also on other variables such as individual characteristics. Therefore, it makes sense to adapt the model with the addition of control variables in order to be able to explain a higher proportion of the variation in the dependent variable. Based on the regression results in Table 3 it can be seen that the intercept drops from 6.80 to 5.90 and the first coefficient of the polynomial regressor of the education years slightly drops from -0.07 to -0.06, but the second regressor remains at the level of 0.01. Furthermore, the control variables had coefficients of 0.34 (gender), 0.10 (location) and 0.02 (age). Finally, a higher degree of determination can be identified defined by a higher level of R^2, which means that, when the control variables are included, a proportion of now 27% (compared to 20% in Model 1) of the variation in the dependent variable can be explained by the independent variables. Thus, one gets a higher goodness of fit of the model. From these values it can be deduced that the expected value of the logarithmically scaled hourly wage, assuming no school education $E[log(y)|x = 0]$, is now lower (in absolute values approx. 40 USh). Consideration of the coefficients of the control variables, including the previously introduced definition of the dummy variables, shows that the growth in hourly wages is higher for male individuals than for female individuals. If one now looks at the gender difference in education from Chapter 1.2, it can be concluded that there is also a gender difference in the pay for wage labor in Uganda. Furthermore, it can be said that individuals in urban regions can expect higher hourly wages. Ultimately, the hourly wage also depends on the age of the employee, with the wage tending to rise slightly with age. In addition, Table 4 shows that the estimates up to the fifth and from the thirteenth year of education are lower compared to the model in Chapter 2.3.1 and are higher between the sixth and twelfth year of education. In conclusion, it can be said that the estimate of the hourly wage or the percentage change in

the hourly wage due to completed years of education with the addition of control variables, in this case individual characteristics, produces better estimation results and explains a larger proportion of the dependent variable. However, since the value R^2 at 27% still allows for optimization potential, further adjustment to include additional control variables would make sense in order to obtain an even more detailed insight into the development of hourly wages in Uganda.

2.4 Who is self-employed? (BONUS)

In order to get a better and deeper understanding on why people in Uganda are either wage workers or self-employed, which defines one of the major differences between low- and high-income countries, the following statistics are computed and visualized based on the National Panel Survey 2013-2014:
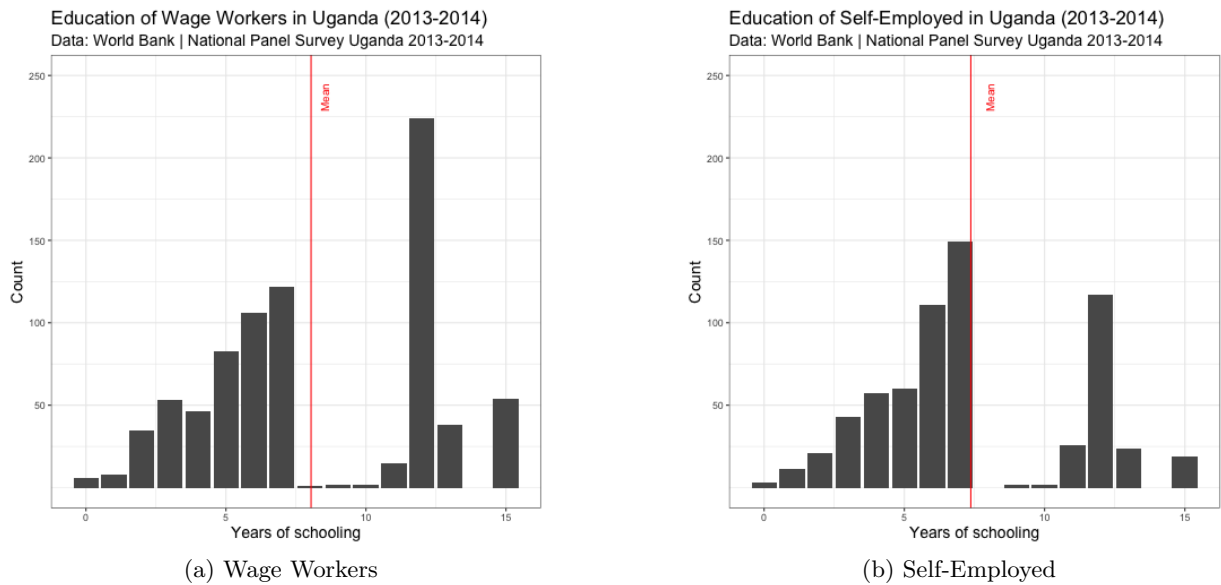

(a) Wage Workers (b) Self-Employed

Figure 6: Years of Schooling of Wage Workers and Self-Employed in Uganda (2013-2014)

Wage Workers	Wage Workers	Self-Employed	Self-Employed
Male	Female	Male	Female
68.36	31.64	50.08	49.92

Table 5: Gender Gaps per Employment Type in Uganda (2013-2014)

Wage Workers	Wage Workers	Self-Employed	Self-Employed
Urban	Rural	Urban	Rural
37.27	62.73	37.39	62.61

Table 6: Location Differences per Employment Type in Uganda (2013-2014)

Mean Age Wage Workers	Mean Age Self-Employed
35.01	38.59

Table 7: Mean Ages per Employment Type in Uganda (2013-2014)

Mean Weekly Hours Worked Wage Workers	Mean Weekly Hours Worked Self-Employed
38.59	38.65

Table 8: Mean Weekly Hours Worked per Employment Type in Uganda (2013-2014)

Based on the statistics in Figure 6 as well as in the Tables 5-8, the difference and sources of wage workers and self-employed can be better developed. The main distinguishing features between wage earners and the self-employed are, on the one hand, the number of years of education, which is on average 8.04 for wage workers and 7.37 for the self-employed and thus 0.67 lower. The distribution also shows that a significantly larger proportion of wage workers have completed second education, while the self-employed mainly have primary education (similar to the school system in Uganda). On the other hand, there are also differences in the characteristics of individuals within the types of employment. The proportion of men among wage workers is almost 68%, with this proportion among the self-employed being around 50% and thus balanced. Therefore, it can be said that men tend to do more paid work than women. Furthermore, if the regions are included, it can be seen that statistics are balanced between wage workers and the self-employed with regard to urban and rural distribution. About 63% of both types of employment are in rural and 37% in urban areas. However, it must be noted here that it can be assumed that the sectors of work are not identical. For example, it can be assumed that in the context of self-employment, agriculture plays a central role in rural areas, although agriculture is less likely to be an issue in urban areas. However, this cannot be further elaborated and analyzed using the National Panel Survey. When looking at the average age of wage workers and the self-employed, there is also a difference, with wage workers being on average 35 years old and self-employed almost 38.6 years old. This could be due to the fact that wage workers retire earlier and / or have the opportunity to quit work earlier due to the higher average additional earnings. Finally, with a look at the last statistic in Table 8, it can be shown that the average weekly working hours for both wage workers and the self-employed is around 38.6 hours and no main differences can be identified.

In addition to the statistics already discussed above, which show the difference between wage and self-employment in Uganda, there are other important characteristics that further explain the rate of wage and self-employment. The demographic as well as the economic development level of a country plays an important component here. As already explained in the previous chapters, the transition from agriculture to the manufacturing sector represents an important development in the context of an industrializing economy. This means that the population has largely switched from independent agricultural work (for self-sufficiency and to feed the families) to wage employment in the manufacturing sector. As can be shown on the basis of current empirical data, developed countries have significantly higher wage employment rates. From this it can be concluded that the development towards wage employment in the course

of the switch to manufacturing or to the service sector can bring a higher level of output and income and thus represents one of the central challenges of developing countries in order to be comparatively similar and competitive with other countries.

Now that the descriptive statistics of the differences between wage workers and the self-employed have been further elaborated, the last section is devoted to a concrete analysis of the extent to which the type of employment differs in rural and urban areas and what possible reasons lead to this. The following figure supplements the analysis with the question of the distribution of types of employment in the regions in Uganda:
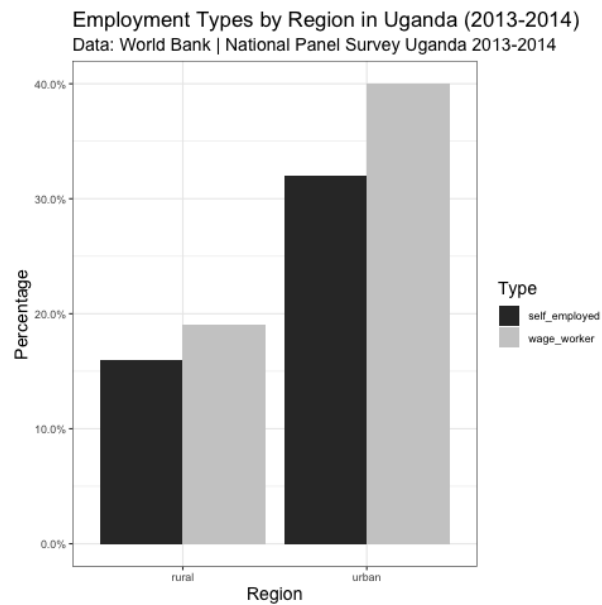

Figure 7: Employment Type by Region in Uganda (2013-2014)

Based on Figure 7 it can be said that the proportion of self-employed is comparatively higher than the proportion of wage workers in rural areas. Furthermore, it can be concluded that the proportion of non-employed individuals in rural areas is also higher, as the cumulative proportion in rural areas is lower and as the classification of employed persons is divided into (i) wage workers, (ii) self-employed and (iii) unpaid workers. This can be connected with the fact that the rural population often pursues self-sufficiency in order to feed the family with the help of family farms and to ensure survival. Furthermore, the cost of living in rural areas is presumably lower, as there is no right to property in the sub-Saharan countries and thus makes life in the city more expensive, which in most cases requires a secure basic income. As the previous section suggested, gainful employment is an important part of rural areas in Uganda. Thus, as the previous section suggested, self-employment is an important part of rural areas in Uganda.