University of Cape Town

School of Economics

Gendered Differences in Hours Spent on Paid Work in South Africa: A Lesson from the Covid-19 Pandemic

ECO3021S Project

Group 90

Submission date: 4 October 2022

ABSTRACT

The COVID-19 pandemic has had a profound effect on hours spent on paid work by females, as they work fewer hours than men on average. In studying these differences, this analysis controls for age, race, education, sector, geographical location, and hours spent on childcare. Contrary to most literature, this paper finds that race and hours spent on childcare as well as geographical location are not significant determinants of hours worked by gender whereas sector and education influence hours spent on paid work by genders. With little policy on hours worked in South Africa, the paper recommends that policy targeting to incentivize women to work longer hours be introduced.

TABLE OF CONTENTS

1. INTRODUCTION

Having a legacy of gender and racial discrimination and being characterized as the most unequal country in the world, South Africa has witnessed an increase in the inequality gap along racial and gender margins as a result of the COVID-19 pandemic (World Bank, 2022). In the labour market, for example, women are more likely to work from home than men whereas along racial lines, uneducated and poor Africans are less able to work from home compared to other races (Nwosu, Kollamparambil, & Oyenubi, 2021). Among these labour-market inequalities lies the inequality in hours worked by males and females.

This paper studies closely the differences in hours spent on paid work by males and females, partly as a result of working from home due to the pandemic. The data used for the study comes from NIDS-CRAM Wave 5, which focuses on two weeks of March 2021.

In carrying out this study, the rest of the paper is structured as follows: the next section looks at available literature on the study of hours worked by males and females, with emphasis on factors affecting hours worked. Section 3 defends the methodology used for this study and the empirical model is specified. Next, Section 4 discusses the data used and the preliminary analysis, and Section 5 discusses the results obtained from the empirical model. Finally, the paper concludes and makes policy recommendations.

2. LITERATURE REVIEW

With gender inequalities being a centre of many debates in the world, the effect of gender differences on working hours has been studied extensively over the years across many parts of the world. In India for example, Krishnakumar and Visnawathan (2021) study the factors affecting women labour-participation based on data from a survey. They find that in traditional households, women tend to work fewer hours on paid work than men, as they have to make enough time to take care of children while doing household chores. Kishnakumar and Visnawathan (2020) also find that educated Indian women prefer to work longer hours on paid work than uneducated ones, and these educated high-earning women tend to hire housekeepers and caretakers to take care of the household work instead.

Using survey data from the United States, Weinberg, Reagan, and Yankow (2004) find that neighbourhood location is a significant factor in determining employee working hours. They estimate that a 28 percent increase in annual working hours would result from a one standard deviation increase in neighbourhood employment. It is also found that moving to a neighbourhood with a higher level of job density is associated with an increase in annual working hours two years after the move (Weinberg, Reagan, & Yankow, 2004).

Literature on hours worked by different genders in South Africa, as well as on hours worked in general, is limited. One South African study is that of Casale and Posel (2020) who find that South African women work 35 less hours per week compared to men. This difference in working hours, according to Casale and Posel (2020), is due to women being primary caregivers to children than men. Rubiano-Matulevich and Viollaz (2019) confirms this in their study when they find that women work more on unpaid work and childcare than men. Both studies suggest that hours spent on childcare is a significant determinant of hours worked by men and women.

Working on NIDS-CRAM Waves 1 – 5 panel data, Casale and Shepherd (2021) find that during the COVID-19 pandemic, employed women saw a 35% decrease in mean working hours between February and April 2020, falling from 35.3 to 23 hours per week, while employed men only saw a 26% (38.8 to 28.8 hours) decrease per week. These differences are also due to women spending more time on childcare than men, as well as due to differences in sectors that are male- and female-dominated (Cale & Shepherd, 2021). Other South African studies suggest that race may be a key determinant in hours spent on paid work, especially considering South Africa's legacy of perpetual racial inequality.

Benhura and Magejo (2021) – also working on NIDS-CRAM Waves 1 – 5 panel data – find that socioeconomic differences in the ability to work from home between Africans and non-Africans increased during the COVID-19 pandemic, with Africans being mostly unable to work from home. This resulted in a decrease in paid hours worked due to stringent lockdown levels (Benhura & Magejo , 2021). Still on the South African context, some studies consider age to be a determinant of hours spent on paid work. For example, Espi, Ranchhod, and Leibrandt (2021) find that younger people are more likely to work fewer hours than older people during the COVID-19 pandemic, especially since the youth unemployment rate continues to increase in the country.

Doan *et al*. (2022) find that older workers between the ages of 50 and 70 can typically work full-time without putting their health at risk. On the other hand, they find that older workers who were in poor health work 7 to 11 fewer hours per week than younger workers. Doan *et al.* (2022) also consider racial and sectorial differences when studying factors affecting hours spent on paid work. They find that older white-collar workers work an average of 7-9 hours more per week than their counterparts in blue- and pink-collar occupations, probably due to having better jobs and thus deriving greater utility from these jobs (Doan *et al.*, 2022).

Despite the preceding research, there are still some gaps in the available literature. In the South African context, for instance, there is insufficient research into the ways in which the working hours of employed people differ, especially with focus on gender differences. Using NIDS-CRAM Wave 5 cross-sectional data, this paper attempts to fill this gap by examining how gender differences influence hours spent on paid work. To do this, this paper uses the preceding literature as guidance, and so, controls for age, race, education levels, employment sector, geographical area, and hours spent on childcare as determinants explaining disparities in hours worked by males and females. The methodology used in studying these effects is discussed in the next section, followed by the defence of the model used for the study.

3. METHODOLOGY AND MODEL SPECIFICATION

The methodology used for this empirical analysis is a *Multiple Linear Regression*. A Multiple Linear Regression is a statistical technique used to study the relationship between an outcome variable and an explanatory variable using a straight line called the line of best fit (Wooldridge, 2012). The line of best fit, and hence the parameters of the regression model, is estimated using the ordinary least squares method of parameter estimation. This method is appropriate for the analysis conducted in this paper because the paper hypothesises a linear relationship between the outcome variable and all the explanatory variables. In this empirical analysis, the outcome variable of interest is the hours spent on paid work per day, and the key explanatory effect is gender. That is, this paper is interested in explaining differences in hours spent on paid work by different genders during the COVID-19 pandemic in South Africa.

Hours spent on paid work are expected to be different between males and females, with males expected to work more hours on paid work than females. This hypothesis is consistent with Bick's, Fuchs-Schündeln's, and Lagakos' (2017) findings. To test this hypothesis, controls for male and female differences in age, race, education levels, employment sector, hours spent on childcare, and geographical location are made.

Age is expected to have a negative relationship with hours worked. Adults are expected to work less hours than young people due to having numerous other responsibilities in their societies and families. A negative correlation between hours spent on childcare and hours spent on paid work is expected as there is a trade-off between the two (Bick, Fuchs-Schündeln, & Lagakos, 2017). Finally, geographical location is controlled for as people in rural areas are expected to work fewer hours on paid work than people in urban areas (Krishnakumar & Viswanathan, 2021). This is mainly because of little to no availability of paid work in rural areas, especially in the South African context (Wilson, 2011). To test these hypotheses, simplifying assumptions are made in the next section.

3.1. Model Assumptions

In estimating the regression model, this paper makes use of the Gauss-Markov assumptions of Multiple Linear Regression, as presented by Wooldridge (2012). The first assumption made is that the population model is linear in parameters. Since this paper uses the multiple *linear* regression model, this assumption thus holds. Secondly, the data used in this research is a random sample drawn from the population. This assumption holds, as the data used in this paper was collected using a random procedure (Ingle, Brophy, & Daniels, 2021). The paper then assumes that there are no perfect collinear relationships between explanatory variables. To test this assumption, **Table I** in the Appendix shows the correlation matrix between all variables considered in the model. As evident in the table, the third assumption holds as none of the variables are perfectly correlated. The fourth assumption made is that the mean of the error terms in the model are independent of the explanatory effects. This assumption is tested later in the analysis. Finally, the paper assumes that the variation in the latent factors, given any value of the explanatory effects, is constant. Using these Gauss-Markov assumptions, the regression models can be specified.

3.2. Empirical Specifications

There are four analyses considered in this paper. The first analysis (sub-analysis) is the simple linear regression between hours spent on paid work and gender. The sub-analysis explores the relationship between the two factors. To further study this relationship, the second analysis is the multiple regression containing all explanatory effects and the third and fourth analyses are the subsample regressions on the female and male subsamples, respectively.

The sub-analysis focuses on the effect of gender on the hours spent on paid work. To estimate this effect, a Simple Linear Regression is used. Similar assumptions to the Gauss-Markov's are made, with differences in the third assumption, which for the simple linear regression case, states that there must be variation in the explanatory variable. **Figure I** of the Appendix shows the distribution of responses by gender, and as evident in the figure, there seems to be variation in gender. The Simple Linear Regression model is defined as follows:

$$Workinghrs_{paid} = \beta_0 + \beta_{female} \cdot Female + u$$

In the main analysis, controls are then made to the simple analysis in order to further examine the determinants of hours worked on paid work. The main analysis is specified as follows:

$$Workinghrs_{paid} = \beta_0 + \beta_{female} \cdot Female + \beta_{age} \cdot Age + \beta_{race} \cdot Race + \beta_{edu} \cdot Education$$

$$+\beta_{sector} \cdot Employment\ Sector + \beta_{childcare} \cdot Childcare +$$

$$\beta_{rural} Rural + u$$

Where *$Workinghrs_{paid}$* is the hours spent on paid work, the outcome variable. Hours spent on paid work is a continuous variable measured in hours per day, ranging between 0 and 24 hours. *Female* is a categorical (binary) variable with a binary value of zero indicating that the respondent is male and a value of 1 indicating female. *Age* is a continuous variable, which measures a respondent's age in years. *Race* is a categorical variable with four levels, namely, African/Black, White, Indian/Asian, and Coloured; with White being the reference category for the analysis. On average, White people are expected to work the least hours than non-Whites, as they have more economic resources (Wilson, 2011). *Education* is a categorical variable with four levels: no schooling; primary education; secondary education; and tertiary education, with no schooling being the base category. This is because people with no education are expected to work fewer hours due to relative lower earnings as well as lack of necessity to work from home during the pandemic (Keswell & Poswell, 2004). *Employment sector* is a categorical variable, with 10 different categories of sectors, with domestic work used as the base sector. *Childcare* is a continuous variable measuring the hours spent on childcare per day of a working week. *Rural* is a binary variable, with a value of one being rural and zero being urban areas.

Of interest to this study is differences in hours spent on paid work by persons of different genders. To compare between genders, subsample regressions on females and males are made. The first subsample regression is on females, and is specified in the same manner as the main analysis, except with observations only from females. Similarly, the subsample regression on males is the main analysis, with responses only from males. These subsample regressions allow for cross-gender comparisons.

4. DATA AND PRELIMINARY ANALYSIS

4.1. Data Description

The data used for this research comes from Wave 5 of the National Income Dynamics Survey – Coronavirus Rapid Mobile Survey (NIDS-CRAM), which is a rapid household survey that investigates the socioeconomic impacts of the South African national lockdown during the covid-19 pandemic (Ingle, Brophy, & Daniels, 2021). The data collected by NIDS-CRAM in Wave 5 – and prior Waves – is panel data, comprising of continuing and temporary sample members aged 18 years or older (Ingle, Brophy, & Daniels, 2021). The members of the panel come from a broadly representative and random sample of South African adults from 2017, who were re-interviewed in 2020 and 2021 for NIDS-CRAM (Ingle, Brophy, & Daniels, 2021). This analysis, however, is restricted to the cross-sectional data from Wave 5 and ignores prior waves.

The Wave 5 cross-sectional data was collected between 6 April 2021 and 11 May 2021 and was collected through the use of Computer Assisted Telephone Interviewing, in which respondents were contacted via telephone to be interviewed (Ingle, Brophy, & Daniels, 2021). A sample of 5862 respondents was successfully collected for Wave 5 (Ingle, Brophy, & Daniels, 2021). In this analysis, after cleaning the dataset, there are 2931 observations. This has important implications on the results.

4.2. Data Challenges

The small sample size implies that there may be little variation in the responses, thus affecting the accuracy and unbiasedness of the results. Because the cross-sectional data comes from a panel data, true patterns of observations are lost. The true pattern is especially lost in the Wave 5 cross-sectional data as the survey focuses only on two weeks of the month of March 2021. This short time is not representative of the long-term effects of the Covid-19 pandemic effects on hours worked by males and females, and so, just in the case of a similar study conducted by Epsi, Leibbrandt, and Ranchhod (2021), the results obtained from this analysis may necessarily be exploratory, but shed light on the effects of the Covid-19 pandemic on hours spent on paid work, nonetheless.

4.3. Preliminary Analysis

4.3.1. Data Cleaning

Other challenges in the data include responses of people who indicated that they did not know an answer to a question or refused to answer the question at all. Where such a challenge is met in the analysis, the observation is considered to be missing and is dropped from the dataset in the analysis as they hold no material value in the analysis. Another key challenge in the data is that of outliers.

Where an observation is an outlier and does not make economic sense, it is often dropped from the analysis. For example, in the outcome variable, three respondents indicated to have worked more than 15 hours a day. Considering the hours needed for personal care, it is unrealistic for one to spend 17 hours on paid work. Thus, the observations are dropped from the analysis.

4.3.2. Descriptive Statistics

This analysis uses ASDOC to create descriptive statistics tables. ASDOC is a Stata program written by Shah (2018). Table 1 below describes the hours spent on paid work by persons of different age and racial groups, education levels, and geographical areas.

Table 1: Hours Worked on Different Attributes

Attribute	Mean	Std. Dev.	min	max	N
Gender					
Female	7.794	1.919	1.000	15	553
Male	8.496	2.039	1.000	15	488
Age					
Youth	8.154	2.013	2.000	15	455
Adult	8.099	2.002	1.000	14	586
Race					
Whites	8.136	2.075	2.000	14	44
Blacks	8.114	2.054	1.000	15	876
Coloureds	8.17	1.654	4.000	15	112

Indians/Asians	8.333	.707	8.000	10	9
Education					
No Schooling	6.417	2.712	1.000	9	12
Primary	7.627	1.961	2.000	12	83
Secondary	8.189	2.114	1.000	15	562
Tertiary	8.202	1.785	2.000	14	377
Geographical Area					
Rural	7.91	2.206	1.000	15	256
Urban	8.206	1.929	2.000	15	724

(Source: NIDS-CRAM, 2021)

As evident in Table 1, males spend 0.702 more hours on paid than females on average. This is consistent with Bick's, Fuchs-Schündeln's, and Lagakos' (2017) findings. Also in line with this paper's expectations, youths work longer hours on paid work than adults. Along racial lines, Indians/Asians work more hours on average than the rest of the racial groups. However, this may be misleading as there are only 9 observations for this racial group. Black, Coloured, and White people seem to work similar hours on average, with White people working slightly fewer hours than other racial groups. This meets the paper's expectations.

There are greater disparities in hours spent on paid work between individuals of different levels of education. Individuals with no schooling, although having the smallest number of observations, work the least hours and those with a tertiary education seem to work more hours on average. This finding is consistent with those of Keswell and Poswell (2004). Whereas, there are smaller differences in hours worked by people in different geographical areas. Those who live in rural areas work slightly fewer hours on average compared to urban residents, however, the difference in average hours worked is not as enormous as expected by the paper. This is possibly due to working classes having moved back to rural areas to work online while they spend time with family during the pandemic.

While there may be differences in the average hours spent on paid work by different attributes, some differences are not statistically significant and may not necessarily explain disparities in hours spent on paid work. To test whether the differences in means are significant, this analysis makes use of T-tests. Table 2 below shows the T-tests for each category, with corresponding p-values.

Table 2: T-tests for Significance

	obs1	obs2	Mean1	Mean2	difference	Std error	t value	p value	significance
Male/female	488	553	8.496	7.794	.702	.122	5.7	0	***
youths	586	455	8.099	8.154	-.055	.126	-0.45	0.661	
Whites	997	44	8.123	8.136	-.014	.309	-0.05	0.964	
Blacks	165	876	8.169	8.114	.056	.171	0.35	0.745	
Indian/Asians	1032	9	8.121	8.334	-.212	.672	-0.3	0.752	
Coloureds	929	112	8.117	8.169	-.052	.201	-0.25	0.794	
rural	724	256	8.206	7.91	.295	.146	2.05	0.043	**

Note: *** p<0.01, ** p<0.05, * p<0.1 **(Source: NIDS-CRAM, 2021)**

The t-tests in table 2 tests the null hypothesis that there is no difference between the observed means against the alternative hypothesis that there is a difference between the means. Along racial margins, the average hours worked by Whites is not different from hours spent by non-Whites, with a large p-value close to 1. This contradicts the paper's expectations in racial differences. The differences in hours worked by people in different geographical areas is statistically significant at the 0.043 level of significance. Finally, the differences in the average hours spent on paid work between males and females is statistically significant, with a probability of making a type I error (p-value) of 0. Other important factors that could explain this significant difference are the hours spent on childcare and the sector in which both genders work in, as sectors define hours worked.

Figure 1 below compares the average hours spent by males and females on paid work and childcare.

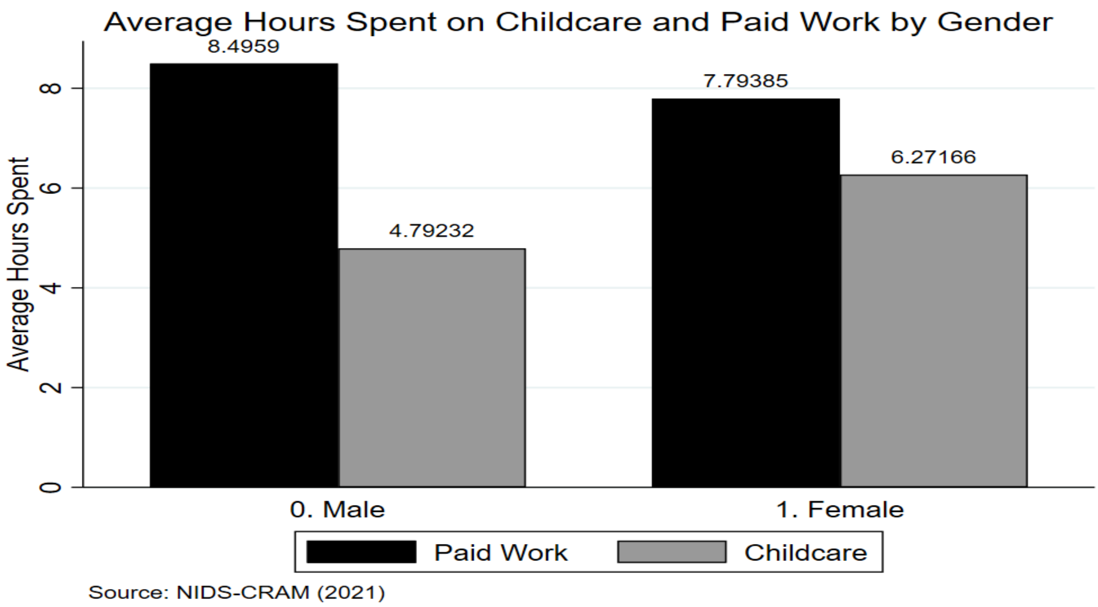


FIGURE 1

As evident in the figure immediately above, females spend more hours on childcare than males on average. This is expected, as during the pandemic, women's hours spent on childcare increased substantially due to school closure, thus limiting hours they spend on paid work (Casale & Sheperd, 2021).

Figure 2 below shows the distribution of the average hours spent on paid work by individuals across different employment sectors. While people working in majority of the sectors work 8 hours per day on average, the mining and financial sectors have the longest working hours of 9.5 per day. This is expected even during the COVID-19 pandemic, as the mining and financial sectors are considered essential sectors even under strict lockdown regulations (Department of Trade, 2020).

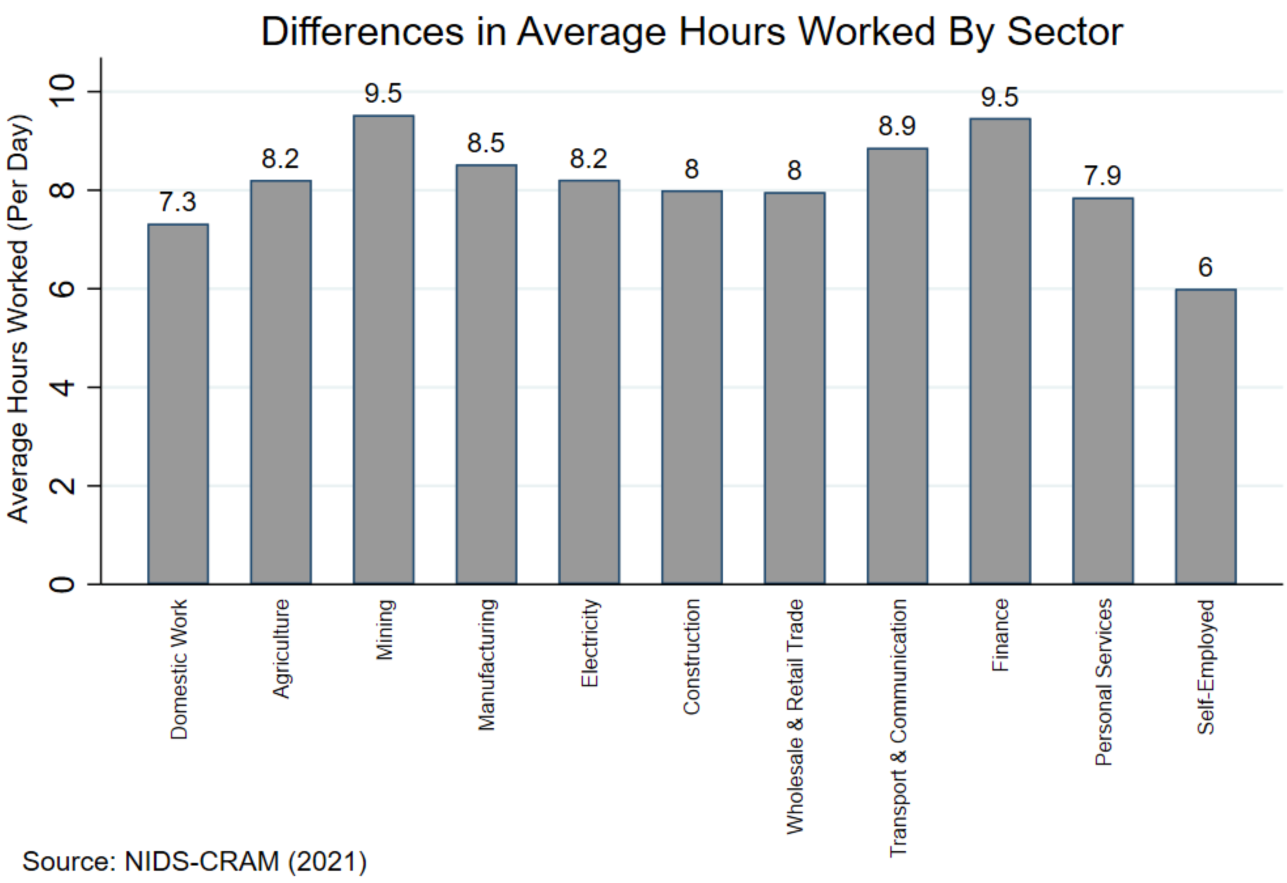


FIGURE 2

Complementary to the figure above is figure 3 shown below, which shows the distribution of males and females across sectors. Contrary to the hypothesis of this analysis, of the total labour force, 46% are males and 54% are females. At face value, this could mean that more females spend time on paid work than males, and could potentially speak to the evolving South African society in which more females work for pay. However, figure 3 shows that most of these females are in sectors which do not require above-average working hours such as personal services and domestic work. In sectors that require the longest hours spent on paid work – the financial and mining sectors, that is – there are more males in both sectors than females, with 28% and 74% more males than females in the financial and mining sectors respectively. Since the financial and mining sectors have the longest working hours than the rest of the sectors, and since there are more males than females in both sectors, it is thus suggested that *ceteris paribus* males will naturally spend more hours on paid work than females, despite more females participating in the labour force than males.

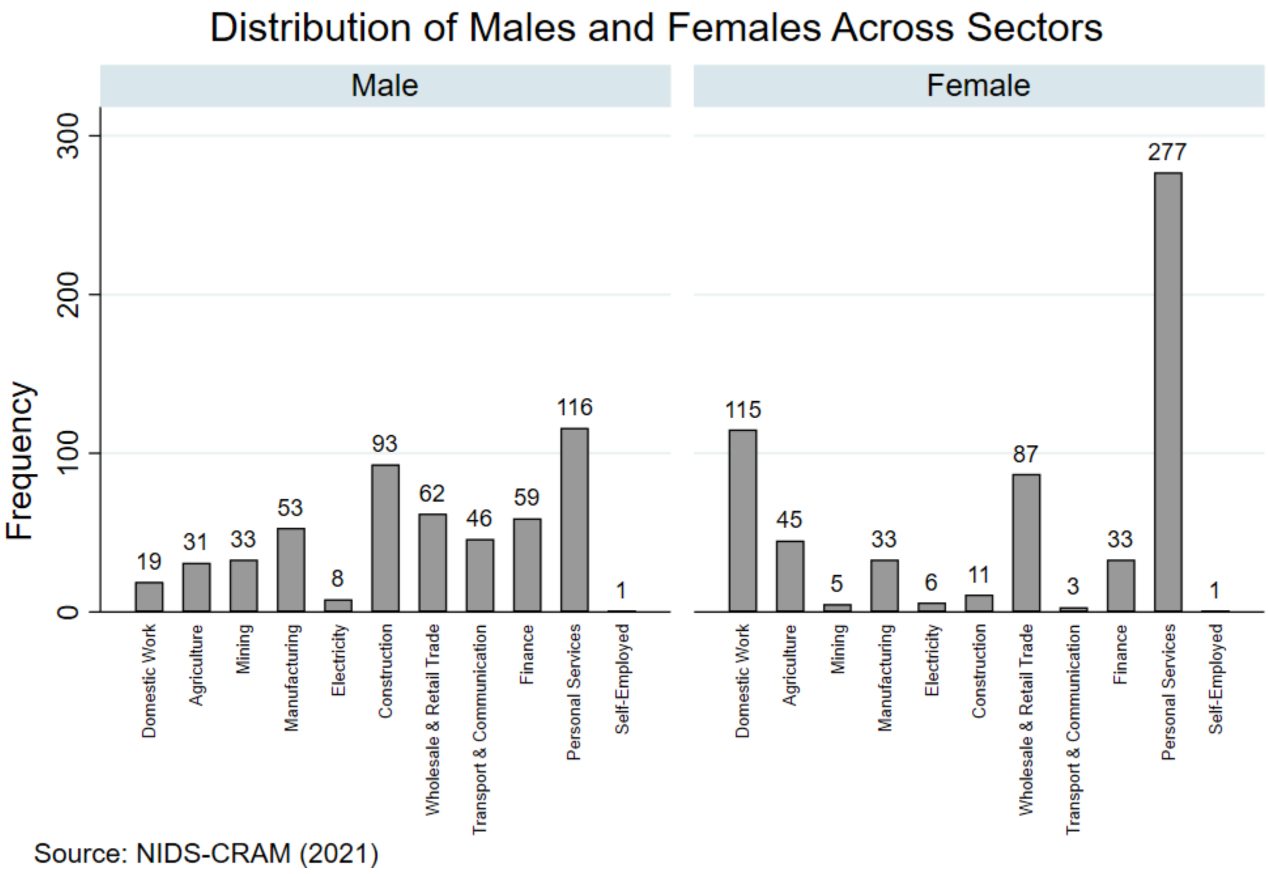


FIGURE 3

Next, this paper models the factors influencing hours spent on paid work (main analysis), as well as the relationship between hours spent on paid work and gender; and further models the differences in hours spent on paid work by males and females, with controls on all explanatory effects (subsample regressions).

5. RESULTS AND INTERPRETATIONS

Table 3: OLS Estimates

VARIABLES	(1) Simple Model	(2) Full Model - Based	(3) Female	(4) Male
Youth		-0.326**	-0.231	-0.346
		(0.160)	(0.226)	(0.245)
Female	-0.702***	-0.544***		
	(0.123)	(0.168)		
Race				
African/Black		0.869	0.648	0.793
		(0.539)	(0.740)	(0.703)
Coloured		0.884	0.648	0.909
		(0.555)	(0.754)	(0.740)
Asian/Indian		0.543	1.058	0.046
		(0.618)	(0.838)	(0.809)
Education				
Primary		1.229	2.175**	-0.754
		(0.853)	(1.091)	(1.002)
Secondary		1.862**	2.656**	0.265
		(0.833)	(1.069)	(0.943)
Tertiary		1.926**	2.904***	0.097
		(0.841)	(1.074)	(0.969)
Rural		-0.096	-0.165	0.071
		(0.179)	(0.232)	(0.292)
Sector				
Agriculture		1.076***	0.950**	1.852***
		(0.337)	(0.428)	(0.620)
Mining		2.259***	1.416***	3.249***
		(0.499)	(0.424)	(0.761)
Manufacturing		1.271***	0.860	2.228***
		(0.409)	(0.558)	(0.701)
Electricity		1.008**	0.664*	1.945**
		(0.414)	(0.379)	(0.945)
Construction		0.276	0.261	0.930
		(0.388)	(0.946)	(0.613)
Wholesale & Retail Trade		0.764**	0.317	1.939***
		(0.361)	(0.431)	(0.672)
Transport & Communication		1.108*	2.182***	1.875**
		(0.657)	(0.312)	(0.867)
Finance		1.998***	1.053*	3.222***
		(0.431)	(0.638)	(0.680)
Personal Services		0.401	0.245	1.064*
		(0.289)	(0.334)	(0.592)
Childcare		0.014	-0.014	0.049*
		(0.018)	(0.023)	(0.028)
Constant	8.496***	5.130***	4.293***	5.877***
	(0.089)	(0.978)	(1.265)	(1.153)
Observations	1,041	660	403	257
F-statistic	32.73	5.23	-	4.04
P-value	0.000***	0.000***	-	0.000***
R-squared	0.031	0.131	0.068	0.202

NOTE: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1 (Source: NIDS-CRAM, 2021)

5.1. Interpretation of Results

Table 3 above shows the OLS estimates predicted by the regression models. Column (1) shows the OLS estimates from the simple regression model. Column (2) shows the estimates from the main analysis, whereas columns (3) and (4) depict the OLS estimates obtained from the regressions on the female and male subsamples, respectively. Herein this section, analysis on the simple model is made, followed by the full model, and then the subsample regressions.

The outcome variable in Model (1) is the hours worked and the explanatory effect is female, with male being the gender of reference. Significant at the 1% level, the OLS estimate shows that *ceteris paribus*, females work 0.7 hours less than males. This finding meets this paper's expectations and is consistent with Bick's, Fuchs-Schündeln's, and Lagakos' (2017) findings. In Model (2), various effects are controlled for in order to further study this relationship.

Of the explanatory effects in Model (2), youth, female, education, and sector are the only ones statistically significant. Contrary to this paper's expectations, young people seem to work fewer hours than adults on average, with youths working 0.326 less hours on average, *ceteris paribus*. As expected, females spend less time on paid work than males, with females working 0.54 less hours.

The results also confirm that education affects hours spent on paid work. As expectated, persons with an education work longer hours than those without education. Individuals with a primary-level of education spend 1.23 hours more on paid work than those without schooling, however, this finding is not statistically significant. This is not surprising, as people with a primary education could arguably be in the same class as those with no education at all. Significant at the 5% level, persons with secondary and tertiary education work 1.86 and 1.93 hours more than those without education, *ceteris paribus*. These findings are in line with this paper's expectations, and confirm both results seen in Table 1, as well as Keswell's and Poswell's (2004) findings.

The mining and financial sectors, as well as the manufacturing sector are all significant at the 1% level. As shown earlier in figure 3, these sectors are predominately male-composed, and explain disparities in hours worked better as evidenced hereabove. Contrary to this paper's expectations, race and childcare are not statistically significant effects on hours spent on paid work. This finding may be attributed to the small sample size whence from these results are drawn. To deepen the analysis on differences in hours worked between males and females, subsample regressions are estimated.

In Model (3), when the main regression is only fitted on the female subsample, it is evident that differences in education levels affect the hours spent on females more, with tertiary education being the most significant determinant. These females with tertiary education work 2.9 hours more than those without schooling, *ceteris paribus*. This finding is significant at the 1% level, and confirms that tertiary-educated women mostly prefer to work for pay than those without school, as they are most likely to be on senior positions or better jobs (Ehrenberg, Smith, & Hallock, 2021). In the male subsample regression in Model (4), however, education is not a significant determinant of hours worked. This result is contrary to the expected relationship between education and hours spent on paid work, and so may necessarily warrant further research.

Hours spent on childcare is still not significant in the female subsample regression. This is unexpected, as a negative relationship between hours spent on childcare and hours spent on paid work by females is hypothesised by this paper. For the male subsample, hours spent on childcare is significant at the 10% significance level – suggesting that for an additional hour spent on childcare, hours spent on paid work increases by 0.05 hours on average, *ceteris paribus*, for males. This result is not practically significant, however, as there is a trade-off between spending time on childcare and paid work. The results suggest that hours spent on childcare may not necessarily be a determinant of hours spent on paid work for both genders. The uncertainty around this finding also warrants further research.

5.2. Overall Model

Table 3 further shows that the models specified in the analysis are all significant at the 1% level – that is, the models are able to explain disparities in hours worked overall. **Figure II** in the Appendix shows the residuals plotted against the predicted values in the main analysis.

As evident in the figure, heteroskedasticity is evident in the distribution of the residuals. This implies that the assumption made by this paper that there is homoskedasticity in the error terms does not hold, and so the estimates, although unbiased due to assumptions 1 – 4 holding, are not necessarily the best linear unbiased estimators. This is mainly due to the small sample size on which this analysis was performed.

6. CONCLUSION

This paper studies factors contributing to disparities in hours worked by males and females. Using Regression Analysis, this paper finds that hours worked by males and females on paid work are explained mainly by differences in sectors to which they belong, with male-dominated sectors working more hours on average than female-dominated sectors. Females generally work fewer hours on paid work than males, although this difference cannot be attributed to racial differences as well as age, geographical area, as well as hours spent on childcare. Even so, there is a need for policy change in South Africa to address the differences in hours worked by gender.

7. RECOMMENDATIONS

The only policy that exists on hours worked by females in South Africa is the Basic Conditions of Employment Act (BCEA), which mandates that the maximum typical working duration is 45 hours per week. This policy does not speak to gender-differences in hours worked, and thus this indicates little effort in closing gender inequality gap in South Africa.

As such, this paper recommends that legislators and government employ policies that are centred around balancing working hours between men and women. For example, a policy that can be adopted is that of childcare grant system in the USA, where working mothers are incentivised to keep working by giving them childcare grant.

8. REFERENCES

Bick, A., Fuchs-Schündeln, N., & Lagakos, D. (2017). How Do Hours Worked Vary with Income? Cross-Country Evidence and Implications. *National Bureau of Economic Research*. Working Paper No. 21874.

Casale, D., Posel, D., & Mosimo, J.(2020). Gender and Work in South Africa. *Research Gate*. Doi:10.13140.

Casale, D., & Shepherd, D. (2021). The gendered effects of the Covid-19 crisis in South Africa: Evidence from NIDS-CRAM waves 1–5. *Development Southern Africa*, *39*(5), 644-663. Doi: 10.1080/0376835x.2022.2036105.

Doan, T., LaBond, C., Banwell, C., Timmins, P., Butterworth, P. and Strazdins, L., 2022. Unencumbered and still unequal? Work hour - Health tipping points and gender inequality among older, employed Australian couples. *SSM - Population Health*, 18, p.101121.

Espi-Sanchis, G., Leibbrandt, M. and Ranchhod, V., 2021. Age, employment and labour force participation outcomes in COVID-era South Africa. *Development Southern Africa*, 39(5), pp.664-688.

Department of Trade, Industry and Competition. (2020). *Government Gazzet No. 43308*. Available: https://www.gov.za

Doan, T., LaBond, C., Banwell, C., Timmins, P., Butterworth, P. and Strazdins, L., 2022. Unencumbered and still unequal? Work hour - Health tipping points and gender inequality among older, employed Australian couples. *SSM - Population Health*, 18, p.101121.

Espi-Sanchis, G., Leibbrandt, M. and Ranchhod, V., 2021. Age, employment and labour force participation outcomes in COVID-era South Africa. *Development Southern Africa*, 39(5), pp.664-688.

Francis,D, Ramburuth-Hurt,K and Valodia,I. (2020). A gender-responsive employment recovery: Building back fairer. Retrieved from: https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_751785.pdf

Keswell, M. & Poswell, L. (2004). Returns to Education in South Africa: A Retrospective Analysis Sensitivity Analysis of the Available Evidence. *South African Journal of Economics*, *72(4), 834 – 860*. ISSN 0038-2280

Krishnakumar, J., & Viswanathan, B. (2021). Role of social and institutional factors in Indian women's labour force participation and hours worked. *Journal of the Asia Pacific Economy*, *26(2), 230 – 251*.DOI: 10.1080/13547860.2021.1917095.

Nwosu, C.O., Kollamparambil, U., & Oyenubi, A. (2021). Socioeconomic Inequalities in Ability to Work from Home During the Coronavirus Pandemic: The Case of South Africa. *NIDS-CRAM Wave 5*. Available: https://cramsurvey.org/reports/.

Rubiano Matulevich, E., & Viollaz, M. (2019). *Gender Differences in Time Use: Allocating Time between the Market and the Household*. Papers.ssrn.com. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3437824.

Salisbury, T. (2016). Education and inequality in South Africa: Returns to schooling in the post-apartheid era. *International Journal Of Educational Development*, *46*, 43-52. https://doi.org/10.1016/j.ijedudev.2015.07.004

Shah, A. (2018). ASDOC: Stata module to create high-quality tables in MS Word from Stata output. Statistical Software Components S458466, Boston College Department of Economics.

Weinberg,B, Reagan,P and Yankow,J. (2004). Do Neighborhoods Affect Hours Worked? Evidence from Longitudinal Data. Retrieved from https://www.researchgate.net

Wilson, F. (2011). Historical Roots of Inequality in South Africa. *Economic History of Developing Regions*, 26(1), 1 – 15. DOI: 10.1080/20780389.2011.583026.

Wooldridge, J.M. (2012). *Introductory Econometrics: A Morden Approach*. 5th ed. Cambridge, MIT Press.

World Bank. (2022). *New World Bank Report Assesses Sources of Inequality in Five Countries in Southern Africa*. Available: https://www.worldbank.org/en/news/press-release/2022/03/09/new-world-bank-report-assesses-sources-of-inequality-in-five-countries-in-southern-africa.

9. APPENDIX

TABLE I: CORRELATION MATRIX OF ALL VARIABLES

	hours worked	youth	female	race	education	rural	sector	childcare
hours worked	1	0.01	0.06	0.04	-0.06	0.15	-0.03	-0.02
youth	0.01	1	-0.001	0.31	0.07	0.01	0.06	0.03
Female	-0.17	-0.10	1	-0.10	-0.21	0.05	0.0028	-0.06
Race	0.01	-0.13	-0.03	1	014	-0.05	0.30	0.08
Education	0.08	0.30	-0.05	0.14	1	-0.03	-0.13	0.01
Rural	-0.06	0.0028	0.05	-0.21	-0.10	1	-0.10	-0.17
Sector	0.03	0.06	0.01	0.07	0.31	-0.001	1	0.01
childcare	-0.02	-0.03	0.15	-0.06	0.04	0.06	0.01	1

(SOURCE: NIDS-CRAM, 2021)

FIGURE I: DISTRIBUTION OF HOURS WORKED BY GENDER

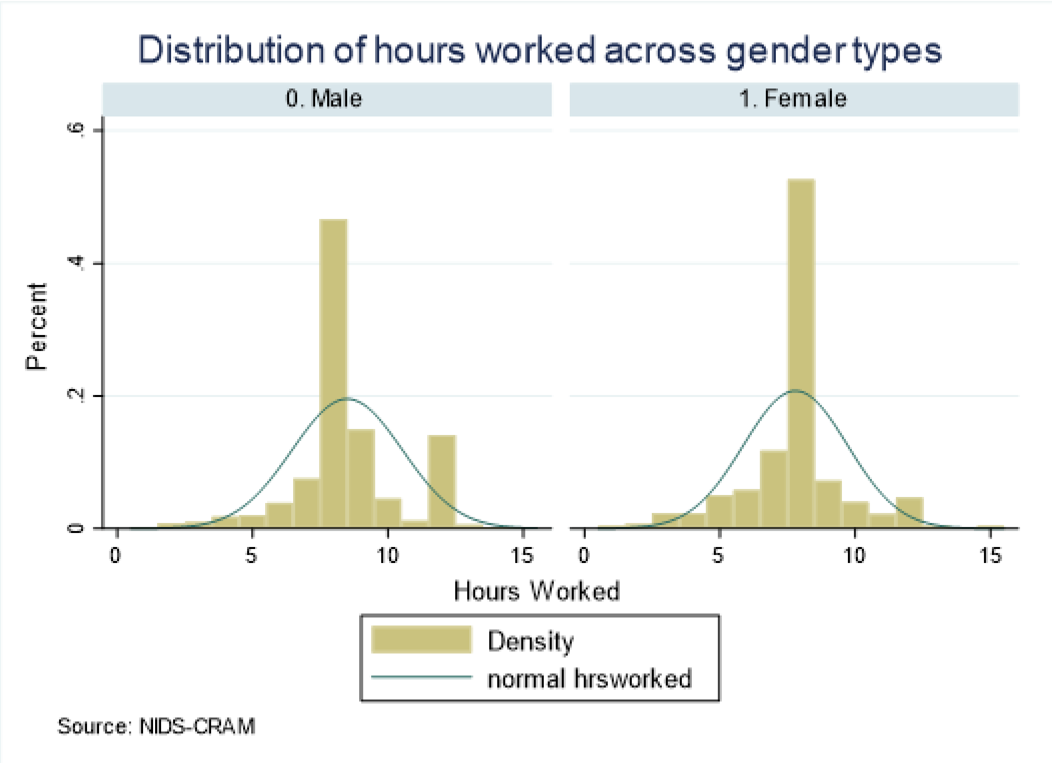
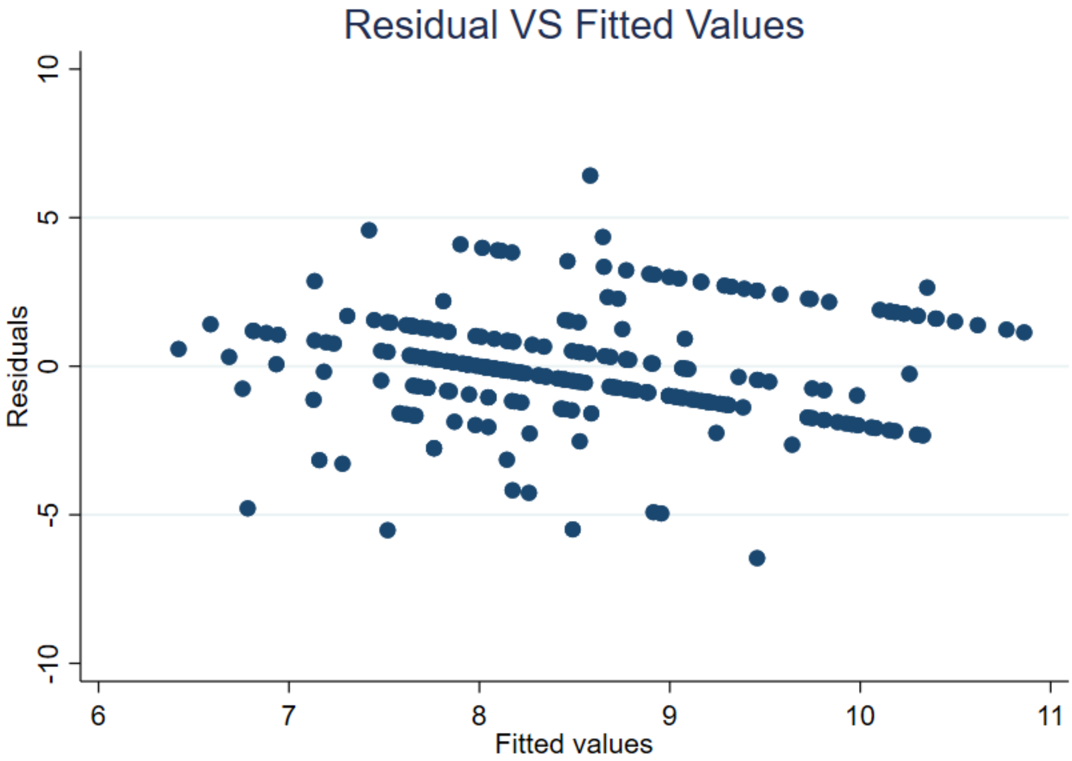


FIGURE II: RESIDUAL PLOT



ATTACHMENT: STATA CODE

```
clear all
set more off
numlabel, add

cd "C:\Users\ROCKETTE NGOEPE\Desktop\SCHOOL FOLDER\2022\ECO3021S\Project"

log using Project2022, replace text

use dataset2022.dta, clear

ssc install pshare
ssc install fre
*-----------------------------------------------------------------------------
                                        ***1. Cleaning and Recoding Variables***
*-----------------------------------------------------------------------------

*-----------------------------------------------------------------------------
            ***1. Preliminary Analysis******
*-----------------------------------------------------------------------------



*Dependent/Outcome Variable
codebook w5_nc_emhrs_mar
tab w5_nc_emhrs_mar
ren w5_nc_emhrs_mar hrsworked
recode hrsworked (-3 -9=.)
tab hrsworked
drop if hrsworked<0
drop if hrsworked >15 & hrsworked<.

tab hrsworked
sum hrsworked,d
pshare estimate hrsworked, nq(4)

histogram hrsworked, percent


* Explanatory Variables
*-----------------------------------------------------------------------------
*       Youth
*-----------------------------------------------------------------------------

lookfor age

su  w5_nc_best_age_yrs
des w5_nc_best_age_yrs
tab w5_nc_best_age_yrs , m
recode w5_nc_best_age_yrs (65/100 = .), gen(age)
sum age
tab age
```

```
* Changing the variable to a categorigal variable
gen youth=.
replace youth=1 if age<=35
replace youth=0 if age>35
label var youth "Growth Stage"
label val youth youthv1
label define youthv1 0"0. Adults" 1"1. Youth"

tab youth
by youth, sort : sum hrsworked
ttest hrsworked, by(youth)

histogram hrsworked, discrete normal ytitle(Percent) xtitle(Hours Worked) by(, title(Distribution of hours worked across age groups) note(Source: NIDS-CRAM)) by(youth)
*------------------------------------------------------------------------------
*           Gender
*------------------------------------------------------------------------------

lookfor gender

des w5_nc_best_gen
codebook  w5_nc_best_gen
su          w5_nc_best_gen
tab w5_nc_best_gen , m

* What do the following codes do?
recode w5_nc_best_gen (1 = 0) (2 = 1), gen(female)
label var female "Gender"
label val female genderv1
label define genderv1 0"0. Male" 1"1. Female"
tab female
fre female
by female, sort : sum hrsworked
ttest hrsworked, by(female)

histogram hrsworked, discrete normal ytitle(Percent) xtitle(Hours Worked) by(, title(Distribution of hours worked across gender types) note(Source: NIDS-CRAM)) by(female)
*------------------------------------------------------------------------------
*           Race
*------------------------------------------------------------------------------

tab w5_nc_best_race
sum w5_nc_best_race
des w5_nc_best_race
codebook w5_nc_best_race



//  Now recode for white and indian/asian and label them.
```

```
recode w5_nc_best_race (1 2 3 = 0) (4 = 1), gen(white)
tab white w5_nc_best_race
label var white "White"
tab white w5_nc_best_race, m //cross tab
tab white
label val white white_lbl
label define white_lbl 0"Non-White" 1"White"
tab white

//	the same intuition applies to the command below for indian/asian dummy
//	variable

recode w5_nc_best_race (1 2 4 = 0 "Non-Indian/Asian") (3 = 1 "Indian/Asian"), gen(indas)
label var indas "Indian/Asian"
tab w5_nc_best_race indas
label val indas indas_lbl
label define indas_lbl 0"Non-Indian/Asian" 1"Indian/Asian"
tab indas

//Black as the base variable
recode w5_nc_best_race (2 3 4 = 0 "Non-Black") (1 = 1 "African/Black"), gen(black)
label var black "African/Black"
tab w5_nc_best_race black
label val black black_lbl
label define black_lbl 0"Non-Black" 1"Africa/Black"
tab black

//Coloured as base variable
recode w5_nc_best_race (1 3 4 = 0 "Non-Coloured") (2 = 1 "Coloured"), gen(coloured)
label var coloured "Coloured"
tab w5_nc_best_race coloured
label val black coloured_lbl
label define coloured_lbl 0"Non-Coloured" 1"Coloured"
tab coloured

ren w5_nc_best_race race //rename race

recode race (4 =0) //white as base in regression model

by race, sort : sum hrsworked


histogram hrsworked, discrete percent normal ytitle(Percent) xtitle(Hours Worked) by(, title(Distribution of hours worked across race groups) note(Source: NIDS-CRAM)) by(race)


*-----------------------------------------------------------------------------
*	Education
*-----------------------------------------------------------------------------
```

```
fre w5_nc_edschgrd
ren w5_nc_edschgrd secondary
recode secondary (-9 -3 13/18 20/22 =.) (19=0)
tab secondary
fre w5_nc_edter
gen education=.
replace education=0 if secondary==0
replace education=1 if secondary>0 & secondary<=7
replace education=2 if secondary>7 & secondary<=12
replace education=3 if w5_nc_edter==1
label var education "Highest education completed"
label val education educationv1
label define educationv1 0"0.No Schooling" 1"1.Primary" 2"2.Secondary" 3"3.Tertiary"
tab education
by education, sort : sum hrsworked

histogram hrsworked, discrete percent normal ytitle(Percent) xtitle(Hours Worked) by(, title(Distribution of hours
worked across education levels) note(Source: NIDS-CRAM)) by(education)

*------------------------------------------------------------------------------
*Area
*------------------------------------------------------------------------------

ren w5_nc_geo2011 location
tab location
recode location (-3 -9=.)

tab location
gen rural =.
replace rural = 1 if location == 1 | location == 3
replace rural = 0 if location == 2
tab rural
label var rural "Geographical Locations"
label val rural ruralvl
label define ruralvl 0"0. Urban" 1"1. Rural"
tab rural


by rural, sort : sum hrsworked
ttest hrsworked, by(rural)

histogram hrsworked, discrete percent normal ytitle(Percent) xtitle(Hours Worked) by(, title(Distribution of hours
worked across area) note(Source: NIDS-CRAM)) by(rural)

*------------------------------------------------------------------------------
* Sector
*------------------------------------------------------------------------------

codebook w5_nc_emsect_c
tab w5_nc_emsect_c
su w5_nc_emsect_c
```

```
des w5_nc_emsect_c
ren w5_nc_emsect_c  sector
label var sector "Employment Sector"
tab sector
recode sector (-3 -8 -9=.) (-2 = 10)
label val sector sector1
label define sector1 0"Domestic Work" 1"Agriculture" 2"Mining" 3"Manufacturing" 4"Electricity" 5"Construction"
6"Wholesale & Retail Trade" 7"Transport & Communication" 8"Finance" 9"Personal Services" 10"Self-Employed"
tab sector
su sector




*------------------------------------------------------------------------------
                ***Child Care***
*------------------------------------------------------------------------------

codebook w5_nc_chldcar_mar
tab w5_nc_chldcar_mar
ren  w5_nc_chldcar_mar childcare
recode childcare (-8=.)
tab childcare

graph bar (mean) hrsworked (mean) childcare, over(female) blabel(bar) ytitle(Average Hours Spent) title(Average Hours
Spent on Childcare and Paid Work by Gender) note(Source: NIDS-CRAM (2021))


*------------------------------------------------------------------------------
**Descriptive Data Tables**
*------------------------------------------------------------------------------
ssc install asdoc, replace
help asdoc

*Descriptive Statistics Table 1
//this table gives the summary stats for average hrs worked by persons of different attributes (gender, age, race,
educational levels, geographical location)
*Variables by hrs worked
asdoc sum hrsworked if female == 1, stat(mean sd min max N) by female append, save(table1.doc) title(Table 1: Hours
Worked on Different Attributes) //female
asdoc sum hrsworked if female == 0, stat(mean sd min max N) rowappend //males
asdoc sum hrsworked if youth == 1, stat(mean sd min max N) rowappend //youth
asdoc sum hrsworked if youth == 0, stat(mean sd min max N) rowappend //adult
asdoc sum hrsworked if white == 1, stat(mean sd min max N) rowappend //whites
asdoc sum hrsworked if black == 1, stat(mean sd min max N) rowappend //blacks
asdoc sum hrsworked if coloured == 1, stat(mean sd min max N) rowappend //coloureds
asdoc sum hrsworked if indas == 1, stat(mean sd min max N) rowappend //indian/asian
asdoc sum hrsworked if education ==0, stat(mean sd min max N) rowappend //no schooling
asdoc sum hrsworked if education == 1, stat(mean sd min max N) rowappend //primary education
asdoc sum hrsworked if education == 2, stat(mean sd min max N) rowappend //secondary education
asdoc sum hrsworked if education == 3, stat(mean sd min max N) rowappend //tertiary education
```

```
asdoc sum hrsworked if rural == 1, stat(mean sd min max N) rowappend //rural
asdoc sum hrsworked if rural == 0, stat(mean sd min max N) rowappend //urban

*Descriptive Statistics Table 1b
*T-test Table

//differences in avg hrs worked by gender (male vs female)
asdoc ttest hrsworked, by(female) append save(table2.doc) title(Table 2: T-Tests)

//differences in average hours worked by age groups (young vs old)
asdoc ttest hrsworked, by(youth) rowappend

//differences in avg hrs worked by race: whites vs non-whites
asdoc ttest hrsworked, by(white) rowappend

//differences in avg hrs worked by race: blacks vs non-blacks
asdoc ttest hrsworked, by(black) rowappend

//differences in avg hrs worked by race: indas vs non-indas
asdoc ttest hrsworked, by(indas) rowappend

//differences in avg hrs worked by race: coloured vs coloured
asdoc ttest hrsworked, by(coloured) rowappend

//differences in avg hrs worked by geographical locations (rural vs urban)
asdoc ttest hrsworked, by(rural) rowappend

/*Use the
pvalues to guide you with inputting the stars that show significance:
 *** if p-value < 0.01
 ** if 0.01 < p-value < 0.05
 * if 0.05 < p-value < 0.1*/

*-----------------------------------------------------------------------------
                                                                    ***Regressions***
*-----------------------------------------------------------------------------

ssc install outreg2
help outreg2

*-----------------------------------------------------------------------------
                                                    *Simple Linear regression
*-----------------------------------------------------------------------------

//regress hrsworked with gender
reg hrsworked i.female

//output simple regression in Excel as a table
outreg2 using regression.xls, bdec(3) excel replace ctitle(Simple Model) label

*-----------------------------------------------------------------------------
```

```
                                        *Multiple regression
*------------------------------------------------------------------------------
*1. Full model with control variables

//regress hrsworked with all explanatory
reg hrsworked i.youth i.female i.race i.education i.rural i.sector childcare, robust

//output the regression in Excel as a table
outreg2 using regression.xls, bdec(3) excel append ctitle (Full Model - Based) label

*2.Full Regression with controls done on a particular sub sample
//Determinants of hours worked amongst Female vs Male

//regressing hrsworked with all variables to test for females
reg hrsworked i.youth i.female i.race i.education i.rural i.sector childcare if female==1, robust
outreg2 using regression.xls, bdec(3) excel append ctitle (Female) label

//regressing hrsworked with all variables to test for males
reg hrsworked i.youth i.female i.race i.education i.rural i.sector childcare if female==0, robust
outreg2 using regression.xls, bdec(3) excel append ctitle (Male) label

*------------------------------------------------------------------------------
                                        *ADDITIONALS TO THE APPENDIX*
*------------------------------------------------------------------------------

*Correlation Matrix
pwcorr hrsworked youth female race education rural sector childcare

*Residual Distributions
predict uhat
histogram uhat //distribution of residuals
rvfplot

cap log close
```