Empirical

In 1993, South Africa extended its pension system to black citizens, whereas it had been reserved only for whites. It provides cash transfers to any woman aged 60 or older and to any man aged 65 or older.

We are seeking your expertise in health economics to study the effect of this policy on the health of children aged 6 months to 5 years, particularly their weight and height. In order to study the effect of this policy *ceteris paribus*, you need to consider the presence of parents, rural/urban/metropolitan, household size, number of people in the household in each of the age categories 0-5, 6-14, 15-24, and 25-49, and household income.

Using data from the 1993 South African National Household Survey available on Brightspace, you are invited to:

1. Combine the sub-databases (there are at least 5 sub-files to combine) and then construct all the variables needed for the analysis before proceeding to step 3. Make sure you have created all variables at the household level so that each child has his or her household information on the same line as their individual information. Some households' information are not automatically transcribed for children when you combine the data.

2. In order to help you in using the right data, you need to submit the list of sub-files you intend to use and your do file for combining the data **preferably by June 3rd**. This part of the final exam is worth 2 marks out of 20 marks. If you get right you will automatically secure 2 marks. If you get it wrong we will give you the correct files to use and the way to combine them. Please note that in order to create your variables at the household level, you may refer to Resources for Empirical Applications available in Brightspace. You have all the necessary information there.

3. Restrict the sample to the population suitable for the study by referring to the problem statement.

4. Identify dependent and control variables.

5. Produce the following table of descriptive statistics (mean and standard deviation):

	Eligibility for pension		
	Woman	Man	None
Household characteristics			
Rural residence			
Presence of the father			
Presence of the mother			
household size			
Income and pension			
Man receiving the pension			
Woman receiving the pension			
Non-pension income			
Pension income			
Per capita income			
Anthropometric data			
Size-by-age z-score			
Weight-by-size z-score			
Observations			

6. Estimate the impact of the reception of the pension in general (i.e., any person receives the pension), women pension reception, and men pension reception on girls and boys separately. Interpret and comment your results.

7. Estimate the impact of pension eligibility in general (i.e., any person eligible for the pension), women eligibility for pension, and men eligibility for pension on girls and boys separately. Interpret the results.

8. Suppose now that there is a reason to believe that these cash transfers (i.e., pension reception) are not random, which could bias our estimate. Propose an alternative approach that might be effective to solve this endogeneity problem. Specify the identifying assumptions needed to identify the effect of this policy and specify the empirical approach you will apply.

9. Run the regressions specified in 8 (for girls and boys) and interpret your results.

10. The impact of receiving the pension might be different for younger children under two years and older children (over two years). Propose an appropriate approach to study the effect of eligibility considering the difference on the young and old children.

11. Run the linear regressions specified in 10 (for girls and boys). Remember to create the necessary variables for each model.

12. **Bonus: Propose a third approach by exploiting the eligibility cutoff point to study this problem.**

Submission format:

1. A pdf file in which you write your answers as well as your results of the empirical applications presented in a table with a format similar to articles published in economic journals;
2. Your do file for all your codes including all the steps of data preparation.
3. A log file that includes your results.