**EXERCISE 2**

**2a)**

Focusing on the variable *nettfa*, it is very evident that most individuals do not have considerable financial assets. Even though the mean stands at $13 594,98 , the standard deviation is enormous, showing huge dispersion in the distribution of said financial wealth. Also, if we combine this information with the (min, max) range, the following is clear: most individuals actually have negative financial assets – they are in debt.

This is one example of a variable where the “mean” is a bad way to describe the distribution – because there is huge variation across individuals.

**2b)**

Given the model:

: An increase of $1000 in annual income of an individual leads on average to an increase in the net total financial assets of $782,59, all else constant.

and : *age* and *agesq* must be interpreted jointly. As such, taking the derivative of *nettfa* in order to *age*, we get the following result: A variation of one year in age of the individual leads on average to a variation in the net total financial assets of $[(-1,568 + 0,0568\**age*)\*1000], all else constant - depending thus on the age level.

: An individual that is eligible for a 401(k) pension fund through her employer has on average more 683,66$ in net total financial assets than an individual with the same characteristics but not eligible for the pension fund.

Age leads to an increase in net total financial assets from age 28 onwards (technically 27,6 – but we rounded it upwards because age is usually represented only as integer). The derivation to conclude this is the following:

**2c)**

We test for the presence of heteroskedasticity in the model using White’s test. The test hypothesis:

No heteroskedasticity

: Heteroskedasticity

The test-statistic:

The decision rule: Reject if: > , with -p- being the number of regressors of White’s auxiliar regression, in this case 12.

We conclude for the presence of heteroskedasticity, due to the p-value of 0.0009 obtained through STATA.

**2d)**

First, let us state that since the Standard Errors are estimated resorting to Bootstrap calculations, they will most likely be different for each time anyone runs the code. However, the coefficients remain the same across estimations, and as we only need to interpret coefficients, this is not a big problem.

For , the 10th percentile, an increase in annual income by $1000 leads to a decrease in net total financial assets of $ -17,91.

For , the 25th percentile, an increase in annual income by $1000 leads to an increase in net total financial assets of $71,29.

For , the 50th percentile, an increase in annual income by $1000 leads to an increase in net total financial assets of $323,93.

For , the 75th percentile, an increase in annual income by $1000 leads to an increase in net total financial assets of $797,72.

For , the 90th percentile, an increase in annual income by $1000 leads to an increase in net total financial assets of $1291,106.

Apart from this, we can say that the coefficient for income increases with each increase of the percentile – the top quantiles increase their net total financial assets more than the bottom quantiles. Meaning that there is a very spread-out unequal distribution of *nettfa* – within-group inequality. There is more net total financial assets increases for the top quantiles that have more income than for the bottom quantiles with lower incomes.

The only comparable to OLS estimate is the one for the 50th percentile. For OLS we got an increase of $782,59 on average, and with Quantile Regression we get an increase of $323,93. This gap between estimates shows the upper outliers in the data, and really show the difference between average and median. In OLS the average is way beyond the 50th percentile value, because the upper percentile values drag the average up. In QR, we can get a fairer picture of the impact of income on net total financial assets.

All of this also points towards the heteroskedasticity in the OLS model that we determined previously – when you compare OLS with QR, it makes perfect sense. The variation (or squared S.E.) is higher for higher percentiles.

**2e**)

Endogeneity through omitted variable bias. Income affects both the probability of an individual being eligible for a (401 k) pension and its net total financial assets. As such, this regression is inconsistent.