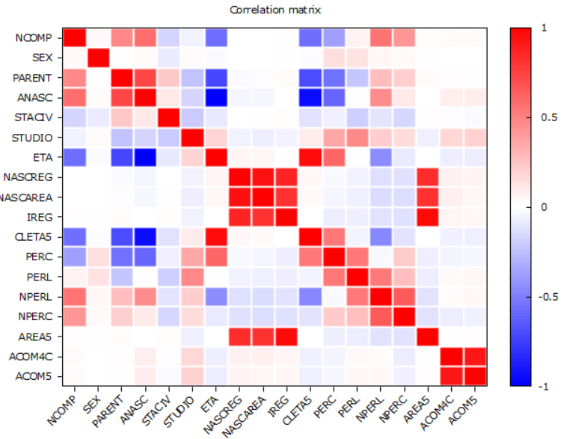
TAKE HOME (ALBERTO SARTINI)

I built a logistic regression model for LFP as dependent variable, that’s because LFP is a dichotomous variable.

I dropped all the variables with too many Nas, then I plotted the matrix of correlations and decided to drop the most correlated variables.



Subsequently I transformed the variable STATCIV into dummy variables where:

1=MARRIED

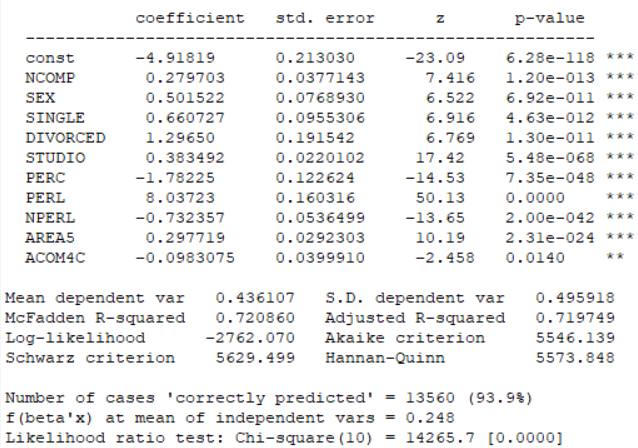
2=SINGLE

3=DIVORCED

Forcing me to drop 4=WIDOW

The current list of explanatory variables would be: NCOMP, MARRIED, SINGLE, DIVORCED, STUDIO, ETA, PERC, PERL, NPERC, NPERL, AREA5, ACOM4C.

The following is the stepwise logistic regression of the full sample

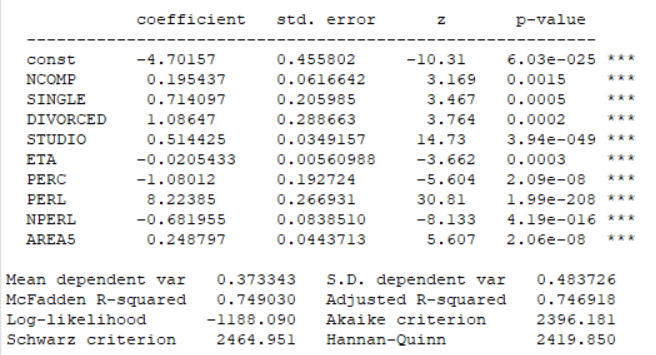
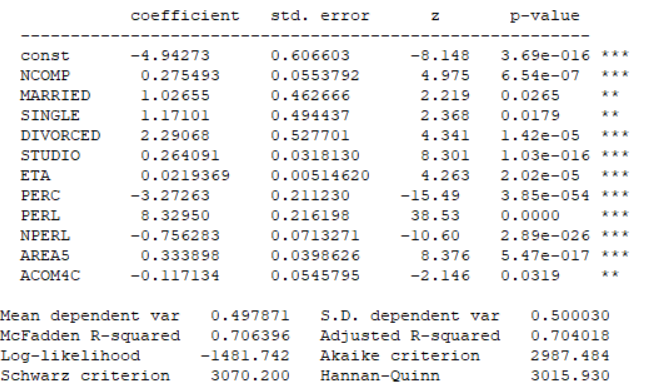


As we can see the McFadden R-squared is 0.72, this can lead us to say that the model is decent.

PERL has a value of 8, this means that if a person has work income than the probability to participate to the labour force increases by a factor of e^8 which is a lot and that makes sense.

If we look at SEX we can see that if you are a male it’s likely that you participate to the labour force; also STUDIO has a positive coefficient, the higher is your educational level the higher are the chances

Now let’s see the two models for the male (left) and female (right) subsets

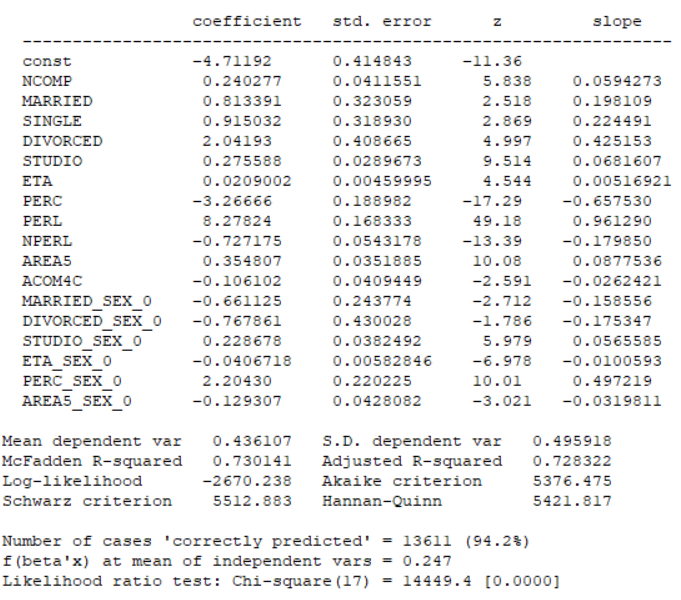


Both of them do a great job in prediction, it’s interesting that in both models PERL maintains the same “big” magnitude and ETA doesn’t seem to play a big role, that’s because the LFP ranges from 15 y.o. to 65 y.o. PERC has a negative coefficient in both samples, that’ because if you receive a subsidy you don’t need to work.

The stepwise regression maintains the MARRIED variable for the male subset but drops it for the female, I think that this happens because marriage plays a more important role in participating in the labour force when you are a man, what I mean is that if a man is married it is highly likely that he works in order to be able to support the family but the same thing is usually not true for the woman ( I hope my comment is not misinterpreted, I do not want to sound sexist at all).

It’s interesting to notice that STUDIO coefficient value in the female sample is twice as big as the one in the male sample and the p-value is way smaller; this could mean that education plays a more important role in labour participation for women, in other words it may seem that it would be harder for a woman to have a job if she doesn’t reach a certain educational level.

Lastly, let’s take a look to the interaction model.



We can confirm what I said above, It could be likely that a married woman doesn’t work and it’s also likely that a divorced woman doesn’t work because she may receive an income from her ex-husband but, again I am generalizing. STUDIO maintains a positive effect while AREA5 has a negative coefficient, this means that a woman living in the south has less chance to work.