

1) The  $r^2$  tells you that 56% of the variability in the  $y$ 's is explained by your  $x$ 's. The degrees of freedom are given by  $n-k$ , no. of observations (4360) – no. of parameters to be estimated (alpha, beta, gamma: 3) = 4357

We also have the SE of the model which is 1.47. The smaller the better, but here, we don't have anything to compare it with (it is not a ratio). If we had another model, as we will, in the following question, then we can compare and select models using either  $r^2$  or SE of model.

Global significance (or F-test). It measures the significance of the model, that is, of ALL coefficients taken together. We have to look for the two last rows in the Eviews output. In fact, with only reading the last row (if you know how) it will be enough. The closer to absolute zero the more strongly you reject  $H_0$ .  $H_0$  here is  $\alpha=\beta=\gamma=0$ ;  $H_1$  is all different from zero. We want to reject  $H_0$ , so as to have all coefficients being globally significant.

Individual significance (t-test). It measures the individual significance of the coefficients in the model. You can have a model which is globally significant but at the same time has one or some coefficients being non-significant. Again, we should check the Prob. column and reason as before. The closer to zero the more strongly we reject  $H_0$ . In this case  $H_0$  is, e.g.,  $\beta = 0$ ; whereas  $H_1$  is  $\beta$  different from zero. Again, we want to reject  $H_0$ .

How to read the output of coefficients? Let's suppose that 100 people get an additional year of education (EDU). The effect will be that they will have, on average, XX children less; all other things the same. The same reasoning applies to the other variable/coefficient: let's suppose that 100 people grow older by one additional year. This will have a positive impact on the amount of children that they want to have, equal to 17 (on average).

The intercept (C) does not have, as in other many cases, any economic meaning in our example. Its interpretation can be easily dismissed here.

The four columns: coefficient, std. error, t-stat and Prob. are related. Notice that  $\text{coef./error} = \text{t-stat}$ . And prob. gives you the answer in terms of individual significance. Remember that the conventional levels of significance are 1, 5, and 10%. That means, that you will make mistakes, 1, 5, and 10% of the times, that is, you will reject  $H_0$  when  $H_0$  is true. In our example we strongly reject  $H_0$  in all cases, both in the individual test and in the global test. All can be said to be rejected at the 1% level. Remember also that there's no such thing as rejecting at the 0% (this is not an exact science, we always have errors).

The std. error gives you an idea of the degree of variability of the estimated coefficient from sample to sample. Assuming that we can repeat the sampling as many times as possible, then (according to GM), we will have the least variance/variability. This implies efficiency/accuracy of the estimated coefficients.

2) One stark difference b/ the 2 models is this new added variable  $\text{AGE}^2$ . This new variable tries to account for the fact that as people grow older, even if they want to have more babies (see the + sign in the AGE variable), they want to do that (or they can do that) at a decreasing rate. That is, a 5 year old child cannot have children, at 15 you can, at 20-30-40 you can/want,

at 50/60 you might want but you most probably can't. Bottom line: you grow older then you will have more probability of having babies, but beyond certain point (which is different from person to person) you will have fewer children, not more. Such variable as AGE<sup>2</sup> can be also applied to a human capital model, where experience affects incomes positively but at a decreasing rate (older people, even if they are more experienced, can get out of date much easily than younger workers; e.g. the use of computers, languages, etc.). The  $r^2$  is now higher, implying a better fit of the model, because we get to explain a larger part of the variability in the  $y$ 's (now  $r^2$  is almost 57%). However, we should not forget about the adjusted  $r^2$  (we will see this soon). The SE is now smaller (as compared with the previous model), which is a good thing. Moreover, the model is globally significant, as before. Notice that, among all these indicators, the F-stat is not useful to compare models (it is if one of them is not able to reject  $H_0$ ). The F-stat is like a necessary condition; you have to pass (that is, you have to reject  $H_0$ ). Akaike and Schwarz are useful for comparing among competing models. The smaller these indicators the better.