**Exercise 5**

The binary variable to be explained is approve, which is equal to one if a mortgage loan to an individual was approved. The key explanatory variable is white, a dummy variable equal to one if the applicant was white. The other applicants in the data set are black and Hispanic.

To test for discrimination in the mortgage loan market, a linear probability model can be used:

1. If there is discrimination against minorities, and the appropriate factors have been controlled for, what is the sign of ?
2. Use data *loanapp.gdt* to regress approve on white. Interpret the coefficient on white. Is it statistically significant? Is it practically large?
3. As controls, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat1*, *mortlat2*, and *vr*. What happens to the coefficient on white? Is there still evidence of discrimination against nonwhites?
4. Now, allow the effect of race to interact with the variable measuring other obligations as a percentage of income (*obrat*). Is the interaction term significant?
5. Estimate a Probit model of approve on white. Find the estimated probability of loan approval for both whites and nonwhites. How do these compare with the linear probability estimates?
6. Now, add the variables *hrat*, *obrat*, *loanprc*, *unem*, *male*, *married*, *dep*, *sch*, *cosign*, *chist*, *pubrec*, *mortlat1*, *mortlat2*, and *vr* to the Probit model. Is there statistically significant evidence of discrimination against nonwhites?
7. Estimate the model from part (f) by Logit. Compare he coefficient on white to the Probit estimate.