Marno Verbeek-A Guide to Modern Econometrics (2004) Unemplyment Benefit pg 198

## Marno Verbeek-A Guide to Modern Econometrics (2004) Unemplyment Benefit pg

## Logit Probit Regression - Marno Verbeek-A Guide to Modern Econometrics (2004) pg 198

### 7.1.6 Illustration: the Impact of Unemployment Benefits on Recipiency

As an illustration we consider a sample3 of 4877 blue collar workers who lost their jobs in the US between 1982 and 1991, taken from a study by McCall (1995). Not all unemployed workers eligible for unemployment insurance (UI) benefits apply for it, probably due to the associated pecuniary and psychological costs. The percentage of eligible unemployed blue collar workers that actually applies for UI benefits is called the take-up rate, and it was only 68% in the available sample. It is therefore interesting to investigate what makes people decide not to apply.

The amount of UI benefits a person can receive depends upon the state of residence, the year of becoming unemployed, and his or her previous earnings. The replacement rate, defined as the ratio of weekly UI benefits to previous weekly earnings, varies from 33% to 54% with a sample average of 44%, and is potentially an important factor for an unemployed worker’s choice to apply for unemployment benefits. Of course, other variables may influence the take-up rate as well. Due to personal characteristics, some people are more able than others to find a new job in a short period of time and will therefore not apply for UI benefits. Indicators of such personal characteristics are schooling, age, and, due to potential (positive or negative) discrimination in the labour market, racial and gender dummies. In addition, preferences and budgetary reasons, as reflected in the family situation, may be of importance. Due to the important differences in the state unemployment rates, the probability of finding a new job varies across states and we will therefore include the state unemployment rate in the analysis. The last type of variables that could be relevant relates to the reason why the job was lost. In the analysis we will include dummy variables for the reasons: slack work, position abolished, and end of seasonal work.

We estimate three different models, the results of which are presented in Table 7.2. The *linear probability model is estimated by ordinary least squares*, so no corrections **for heteroskedasticity are made and no attempt is made to keep the implied probabilities between 0 and 1**. The **logit** and **probit model** are both estimated by maximum likelihood. Because the logistic distribution has a variance of ??^2/3, the estimates of ?? obtained from the logit model are roughly a factor ??/???3 larger than those obtained from the probit model, acknowledging the small differences in the shape of the distributions. Similarly, the estimates for the linear probability model are quite different in magnitude and approximately four times as small as those for the logit model (except for the intercept term). Looking at the results in Table 7.2, we see that the signs of the coefficients are identical across the different specifications, while the statistical significance of the explanatory variables is also comparable. This is not an unusual finding. Qualitatively, the different models typically do not provide different answers.

# Set working directory:

setwd("C:/Users/Willians/Desktop/Mestrado/Econometria - Schonerwaldt") # note / instead of \ in windows

# Store the data from benefits.dta in an object (database):

database <- read.dta("C:/Users/Willians/Desktop/Mestrado/Econometria - Schonerwaldt/R Scripts/Data/unemployment\_benefits/benefits.dta", convert.factors=FALSE)

# Explore the data features (columns) and 5 Number Summary Stats

head(database)  
summary(database)

# Create a vector with the packages to be installed:

list.of.packages <- c("AER", "sandwich", "lmtest", "car", "dplyr", "stargazer", "ggplot2", "foreign",  
 "openintro","OIdata", "gdata", "doBy","ivpack", "psych","plm", "readxl")  
  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]  
  
if(length(new.packages)) install.packages(new.packages)  
  
lapply(list.of.packages, require, character.only = TRUE)

# Logit Regression

logit <- glm(y ~ rr + rr2 + age + age2 + tenure + slack + abol + seasonal + head + married + dkids + dykids + smsa + nwhite + yrdispl + school12 + male + statemb + stateur, family = binomial(link = "logit"),   
 data = database)  
summary (logit)

# Probit Regression

probit <- glm(y ~ rr + rr2 + age + age2 + tenure + slack + abol + seasonal + head + married + dkids + dykids + smsa + nwhite + yrdispl + school12 + male + statemb + stateur, family = binomial(link = "probit"),   
 data = database)  
summary (probit)

# LPM -> MQO Normal Linear Regression

lpm <- lm(y ~ rr + rr2 + age + age2 + tenure + slack + abol + seasonal + head + married + dkids + dykids + smsa + nwhite + yrdispl + school12 + male + statemb + stateur, data = database)  
summary(lpm)