A COMPARISON OF LINEAR PROBABILITY, LOGIT AND THE PROBIT MODELS WITH THE ESTIMATION OF LABOR FORCE PARTICIPATION MODEL .

Emine Bilgili
Erciyes University, Faculty of Economics and Administrative Sciences,
Department of Economics, Kayseri 38039

Abstract

Emprical studies of labor force participation based on individuals as the unit of observation normally utilize a binary dependent variable, coded "one" if the person participates in the labor force, and "zero" if the person does not participate. Researchers use either "linear probability models (LPM with OLS)" or "nonlinear probability models (such as logit, probit and tobit models) in estimation of probabilities of these "0 and 1." Due to some undesirable statistical properties of LPM's, the use of probit and logit models has become quite common whenever dependent variable in a regression is qualitative. These models have been used intensively in mikroeconometric studies especialy in labor economics in 1990s. The purpose of this paper is to estimate a typical labor force participation model and compare OLS, probit and logit estimates of that model. In conclusion, I will give an explanation of the statistical problems of OLS, logit and probit procedures.

I. Introduction

Especially last two decades have experienced many researches on economic analyses of employment patterns, such as Bowen and Finegan (1969), Cohen, Rea, and Lerman (1970,1971), Boskin (1973), Gunderson (1974 and 1980), Schmidt and Strauss (1975), Gronau (1976), Heckmann and Wills (1977), Kahn and (1979), Smith (1979), Leuthold (1978), Schultz (1980), Cogan (1980), Layard et al.(1980), Hausman (1981), Mroz(1987). More detailed list of these studies can be obtained from Mroz (1987), Gunderson (1980) and Amemiya (1981).

Empirical studies of labor force participation based on individuals as the unit of observation normally utilize a binary dependent variable, coded "one" if the person participates in the labor force, and "zero" if the person does not participate. Researchers use either “linear probability models (LPM with OLS)” or “nonlinear Probability models (such as logit, probit and tobit models) in estimation of probabilities of these “0 and 1.” Due to some undesirable statistical properties of LPM’s, the use of probit and logit models has become quite common whenever dependent variable in a regression is qualitative. This progress was one of most important developments in econometrics; in the area of qualitative response models or categorical or discrete models. These models have been used intensively in microeconometric studies especially in labor economics in1990s as well as 1980s and 1970s.

The purpose of this paper is to estimate a typical labor force participation model and compare OLS, probit and logit estimates of that model. Section II gives the source of the data and definition of the variables. Sections III, IV and V estimate the parameters by using the OLS, logit and probit techniques respectively. Section VI compares the maximized log-likelihoods of logit and probit models and compares two approximation methods to transform logit estimates to comparable probit estimates. In conclusion, I will give an explanation of the statistical problems of OLS, logit and probit procedures.

II. Model and Data

Model deals with the female labor force participation in 1975. The variables are given in Table 1.

Table 1: The Variables of Labor Force Participation Model

LFP	: Labor force participation dummy variable that equals 1 if the woman's hours of work in 1975 were positive; otherwise, it equals zero.
WHRS	: Wife's hours of work in 1975.
KL6	: Number of children in the household under age six.
K618	: Number of children in the household between age six and eighteen.
WA	: Wife's age in years.
WE	: Wife's educational attainment in years of schooling.
WW	: Wife's 1975 average hourly earnings in 1975 dollars.
FAMINC	: Family Income in 1975 dollars.
UN	: Unemployment rate in the country of residence, in percentage points.
CIT	: Dummy variable that equals 1 if the family lives in a large city(a Standard Metropolitan Statistical Area, SMSA), otherwise it equals zero.
AX	: Wife's previous labor market experience.

In the Labor Force Participation Model (LFPM), the dependent variable LFP, is a labor force participation dummy variable that equals 1 if the woman's hours of work in

1975 were positive; otherwise it equals 0. Data was obtained by the University of Michigan Panel Study of Income Dynamics for the year 1975. The interview year is 1976. Data file contains 753 observations on married white women aged 30-60 in 1975 for sixteen variables. The first 428 observations are those for women whose hours of work in 1975 were positive, while the final 325 observations are those for women who did not work for pay in 1975 (Mroz, 1987, p.769). Data is available in the ASCII data disk that comes with the Berndt's book, The Practice of Econometrics (1991). Mroz (1987) studies on a systematic analysis of female labor supply that has been analyzed many times in 1970s and 1980s. He finds out that economic and statistical assumptions can have a substantial impact on the estimates of the behavioral labor supply parameters and that the tobit models exaggerate both the income and wage effects in the labor force model. His conclusion is that economic factors such as wage rates, taxes, and non labor incomes have a small impact on the labor supply behavior of working married woman. He suggest that some other factors such as "search costs", "imperfectly elastic labor demand schedules", "labor force participation and dynamic behavior" and "nonpecuniary benefits" should be considered to be able to estimate the behavior of married woman. In this study, I will fit the LFPM by estimating the behavioral labor supply parameters, and make statistical inferences by using linear probability, logit and probit models.

III. Linear Probability Model

The linear probability model; $LFP_i = B_1 + B_2X_i + u_i$ (1)

LFP=1 if women participates in labor force

LFP=0 if women does not participates in labor force

X_i = explanatory variables

The dichotomous variable, LFP, is a linear function of the explanatory variables X_i. In this section, I will estimate the parameters of a linear probability model in which LFP is related linearly to an intercept term, the WW1, KL6, K618, WA, WE, UN, CIT and PRIN by using OLS estimation procedure.

First, I took the natural logarithm of the WW variable (for the first 428 observations) and called this log-transformed variable LWW. Then, for the entire sample of 753 observations, I constructed the squares of the wife's experience, AX, and called this AX2. I regressed LWW on a constant term, WA, WE, CIT, AX and AX2. Then I used the parameter estimates from this equation and the values of the WA, WE, CIT, AX and AX2 variables for the 325 women in the nonworking sample to generate the predicted or fitted log-wage for the nonworkers. I called this fitted log-wage variable for the nonworkers FLWW. Finally, for the entire sample of 753 observations, I generated a variable called LWW1 for which the first 428 observations (the working sample) LWW1= LWW and for which the last 325 observations (the nonworking sample) LWW1=FLWW. In this way, LWW1 variable includes either the actual or a predicted wage for each individual in the sample. It is assumed that in making woman's labor supply decision, she takes as given the household's entire nonlabor income plus her husbands labor income. Call this the wife's property income. Therefore, I used another variable, property income variable called PRIN in the equation. It was computed as PRIN=FAMINC-(WHRS × WW).

I regressed LFP on variables given in Table 2 by OLS procedure. I used The Econometrics Toolkit (ET) software , version 3.0 and SAS-Syslin, SAS-Proc Logistic and SAS-Proc Probit procedures from SAS statistical package, version 6.03

Table 2:Linear Probability Model of LFPM with OLS Estimation Procedure
Dependent variable is LFP

Independent Variable	Parameter Estimate	Standard Error	t for H_0	Prob > |t|
Constant	0.692297	0.1626	4.255	0.0001
LWW1	0.093301	0.0319	2.917	0.0036
KL6	-0.290638	0.0357	-8.136	0.0001
K618	-0.008387	0.0139	-0.600	0.5484
WA	-0.011609	0.0025	-4.554	0.0001
WE	0.042096	0.0086	4.866	0.0001
UN	-0.003487	0.0054	-0.636	0.5247
CIT	-0.004479	0.0367	-0.122	0.9029
PRIN	-0.000067	0.0000	-4.402	0.0001
F Value(8, 744)	17.281			
Prob>F	0.0001			
R-square	0.1567			
Akaike Information	-1.5517			

The signs of the estimated parameters make sense. As KL6, K618, WA, UN, PRIN increase one unit and CIT=1, the probability that woman work will (P=1) decrease by the numerical values of parameters. Wife has to take care of children at home, therefore she has to allocate some of her time to her children rather than working outside, provided other things being equal. The same logic can be applied to the other variables. As wife gets older and as unemployment rate in the country increases, the probability that she can find job will decrease. As wife's property increases, she decreases the her positive working hours. As for CIT, it's negative value is questionable but depends on the some other assumptions. For instance if CIT and UN are highly correlated each other and the value of UN is high, then I expect less positive working hours. On the other hand, opportunities of working and life style in a big city might lead to some more positive working hours rather than declining in working hours. Intercept term, LWW1 and WE variables have positive impact on the LFP. With zero values of all other variables, wife must have at least some positive working hours. As economic theory of labor market indicates, labor supply would increase if the wage rate increased. And as the years of schooling increases, the participation in labor force will increase.

The intercept of 0.692297 gives the probability that woman participates in the labor force when all explanatory variables in the model are zero. The coefficient of 0.093301 attached to the variable LWW1 means, holding all other factors constant, for a unit change, $1, in average hourly earnings of wife (either the actual or a predicted wage for each individual in the sample), the probability of participation in the labor force by wife increases by 0.093301 or about 1 percent. The coefficient of -0.008387 attached to the variable K618 means, holding all other factors constant, as the number of children in the household between age six and eighteen increases, the probability of participation in the labor force by wife decreases by -0.008387. If we want to find the probability for woman earning $7 in one hour, with 2 children under age six, with 1 child between age six and eighteen, aged 32, with 13 years of schooling, with an unemployment rate of 5 percent, living in large city and with property income of $ 25,000, we obtain

$$0.692297 + 7(0.093301) - 2(0.290638) - 0.008387 - 32(0.011609) \\ + 13(0.042096) -5(0.003487) - 0.004479 - 25(0.000067) \\ = 0.907912$$

The probability of labor force participation by woman with the preceding characteristics is estimated to be about 90 percent. Suppose that the hourly wage is $10 instead of $7 and that woman has 15 years of schooling instead of 13. Then the resulting probability is 1.27. Since this value is greater than 1 and since probability can not be greater than 1, we treat this value as zero.

Can one be confident with the results of LPM? Among some statistical problems of LPM, one can mention two positive statements of OLS. The first one is that OLS estimates will tend to indicate the correct sign of the effect of independent variable on dependent variable (Aldrich and Nelson, 1984: 27). The second positive statement about LPM is that expectation of disturbance term is zero; $E(u_i)=0$ and as a result, OLS estimates of b_K will be unbiased. However, variance of u_i is not constant anymore and it varies systematically with the values of independent variables (Aldrich and Nelson, 1984: 13).

$$\begin{aligned} v(u_i) = E(u_i)^2 &= P(Y_i = 0)\left[-\sum b_k X_{iK}\right]^2 + P(Y_i = 1)\left[1 - \sum b_K X_{iK}\right]^2 \\ &= \left[1 - P(Y_i = 1)\right]\left[\mathrm{P}(\mathrm{Y_i} = 1)\right]^2 + P(Y_i = 1)\left[1 - P(Y_i = 1)\right]^2 \\ &= P(Y_i = 1)\left[1 - P(Y_i = 1)\right] \\ &= \left[\sum b_K X_{iK}\right]\left[1 - \sum b_K X_{iK}\right] \end{aligned} \qquad (2)$$

Hence, estimate of b_K will be unbiased but will not have the smallest possible sampling variance. In other words, it is not the best among the others and any hypothesis test, e.g., the t and F tests or confidence intervals based on these sampling variances will be invalid, even for very large samples. Therefore, OLS estimates for LPM models are unbiased, not very desirable (Aldrich and Nelson, 1984: 13).

The null hypothesis that K618, UN, CIT are not statistically different from zero may not be rejected at the 95 percent confidence level. Other coefficients are statistically significant at 95 percent level. And at 0.05 significance level, the F test indicates that all coefficients are jointly statistically different from zero. However as indicated above, due to heteroscedasticity, the t and F tests might be biased (In heteroscedasticity case, one can run two-step procedure to correct for heteroscedasticity).

Table 3: Fitted values < 0 and >1

	i^{th} observation	less than zero	greater than one
1	007	-	1.0379
2	040	-	1.1696
3	056	-	1.0210
4	058	-	1.0362
5	203	-	1.1169
6	300	-	1.0870
7	304	-	1.0103
8	381	-	1.0752
9	398	-	1.0103
10	484	-0.4476	-
11	531	-0.2549	-
12	562	-0.0006	-
13	596	-0.0483	-
14	605	-0.1026	-
15	617	-0.0412	-
16	715	-0.1640	-

As seen from the Table 3, the fitted the values of first nine observations are greater than one and the fitted the values of other seven observations are negative. This complicates the interpretation of the model, since the probabilities can not exceed one or can not have the negative values. There are 16 observations out of 753 observations that violate this basic property of probabilities. Hence one can treat them as zero values.

The R^2 in this model is 0.1567. In other words the %15.67 variation of LFP can be explained by explanatory variables given in the model. This R^2 is very low. In qualitative and limited dependent variables case, R^2 is likely to be very low for this kind of regression, suggesting that R^2 should not be used to judge the model (Kennedy, 1992: 22).

From Figure 1, for a given X, Y takes on two values; 1 or 0. Therefore all the Y values lie along X axis or along the line corresponding to 1. Hence one can not expect Linear probability Model (LPM) to fit this scatter well. As a result, the conventionally computed R^2 is likely to be much lower than 1 for such models. In most practical applications the R^2 ranges between 0.2 to 0.6. The R^2 is high if it is higher than ,say, 0.8 (Gujarati, 1995: 545-546).

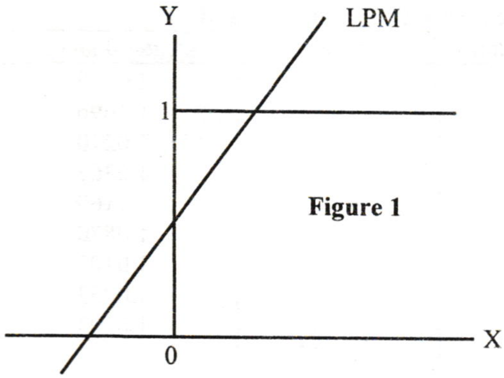


Figure 1

Source: P. Kennedy (1992, p.228)
Aldrich and Nelson (1984, p.27)

IV. Logit Model

n this section, I estimated parameters by logit maximum likelihood procedure. LFP is dependent variable and LWW1, KL6, K618, WA, WE, UN, CIT and PRIN are explanatory variables. I used the entire sample of 753 observations in the MROZ data file.

Table 4: MLE Results for Logit Model on Mroz Data.

Independent Variable	Parameter Estimate *	Standard Error *	t for H_0 *	Wald Chi-square **	Pr> Chi-square **
Constant	0.95089	0.80423	1.1824	1.3983	0.2370
LWW1	0.46449	0.19851	2.9541	8.7286	0.0031
KL6	-1.46897	0.19851	-7.4001	54.771	0.0001
K618	-0.05119	0.06842	-0.7481	0.5601	0.4542
WA	-0.05825	0.01291	-4.5131	20.372	0.0001
WE	0.21193	0.04411	4.8045	23.088	0.0001
UN	-0.01857	0.02646	-0.7017	0.4925	0.4828
CIT	0.01127	0.17816	0.0714	0.0051	0.9430
PRIN	-0.00004	0.00001	- 4.3548	18.968	0.0001

Chi-squared[8] *	: 130.800 prob. : 0.00000
Log L with only a constant term *	:-514.8732
Iterations *	: 4
AIC Intercept Only **	: 1031.746
SC Intercept Only **	: 1036.370
-2 LOG L intercept only **	: 1029.746
Chi-square for Covariates **	: 130.800 with 8 df (p=0.0001)
Iterations **	: 5

* ET output ** SAS output

From Table 4, we see that signs of the estimated parameters are the same as those I obtained in LPM with OLS procedure. Again, the number of children, (KL6, K618), wife's age in years (WA), percentage change in unemployment (UN) and dummy variable (CIT) have negative effects on the probability of having positive working hours for women. The constant term, wage rate (LWW1) and wife's schooling years in education (WE) have positive effects on the probability of having positive working hours for women.

From Table 4, by observing the Chi-square and P>chi-square, I can indicate that the constant term and variables K618, UN, CIT are statistically not different from zero at 95 percent confidence level and that all other variables are statistically significant at 0.05 significance level.

Table 4 states that nonlinear logit computational algorithm in ET converged after 4 iterations and in SAS converged in 5 iterations.

ET provides the chi-square, log-likelihood, Log L with only a constant term and the number of iterations as "goodness of fit" information.

Maddala (1983: 37-41) gives the summary of alternative measuring goodness of fit. A reasonable R^2 measure for goodness-of fit is "McFadden's R^2" which is same as "likelihood ratio index" defined in Green (1990, 682).

$$LRI = 1 - [L_{UR} / L_R] \quad (3)$$

L_{UR} = the log likelihood value with independent variables plus constant term.

L_R = the log likelihood value with only a constant term.

L_{UR} = -449. 4732

L_R = -514.8732

LRI = 1- 0.8279 = 0.1270

Although the values between 0 and 1 do not have natural interpretations, LRI increases as the fit of the model improves and one can state that LRI = % 12.71 is not reasonable for such models. In most practical applications the R^2 ranges between 0.2 to 0.6. The R^2 is high if it is higher than ,say, 0.8 (Gujarati, 1995: 545-546).

SAS provides the pseudo F and t^2 statistics in the SAS-Proc Cluster Procedure (SAS-Stat Users Guide, 1988: 289). SAS also provides the Akaike Information Criterion (AIC) information for 'Goodness-of-fit" measurement with the SAS Proc Logistic procedure. According to Amemiya (1985, 147) the idea is to choose the model for which AIC is the smallest.

For goodness of fit, in regression analysis , an F statistics can be used to test the joint hypothesis that all coefficients except the intercept are zero. A corresponding test for logit and probit for the same purpose is likelihood ratio test (Aldrich and Nelson, 1984, p.55). For instance, using either ET or SAS output; the likelihood ratio test (LR) = $-2(L_R-L_{UR}) = 130.799$. The 95% critical value from the chi-squared distribution with 8 degree of freedom is 15.50 which is less than computed LR value of 130.79. Then the hypothesis that all coefficients are jointly equal to zero would be rejected.

Another measurement is the Psuedo R^2 for the goodness of fit. It can be calculated by using the ET and SAS Outputs.

$$\text{Psuedo } R^2 = [LR / (N + LR)] \quad (4)$$
$$= [130.739 / (753 + 130.739)]$$
$$= 0.1480$$

Psuedo R^2 ranges between 0 and 1. A Psuedo R^2 approaching 0 means that the quality of the fit diminishes (i.e. as LR approaches 0). A Psuedo R^2 approaching 1 means that the quality of the fit improves. It is not, however, universally accepted and does not incorporate a penalty for increasing the number exogenous variable (Aldrich and Nelson, 1984: 57).

In logit model, in order to find the probability of labor force participation by woman, first, the "odds ratio" is calculated.

$$L_i = \ln\left(\frac{P_i}{1-P_i}\right) = B_1 + B_2X_i \quad (5)$$

In eq. (5), L is the log of the odds ratio and is called the logit. Pi is the probability that woman participates in labor force and 1- Pi is the probability that woman does not participates in labor force. The intercept of 0.95089 gives the probability that woman participates in the labor force when all explanatory variables in the model are zero. The coefficient of LWW1 means, holding all other factors constant, for a unit change, $1, in average hourly earnings of wife, the probability of participation in the labor force by wife increases by 0.46449. If we want to find the probability for women earning $7 in one hour, with 2 children under age six, with 1 child between age six and eighteen, aged 32, with 13 years of schooling, with an unemployment rate of 5 percent, living in large city and with property income of $ 25,000, we obtain $L_i = 2.021$. Taking the antilog of 2.021, we obtain 7.5458. Therefore $P_i = 0.8829$, that is the probability that a woman with preceding characteristics will participate in labor force is about 0.88.

The estimated slope coefficients of LWW1, KL6, WA, WE and PRIN are significant at the 1% level. The estimated coefficients have the same signs as those obtained by LPM in Table 2 except the variable CIT. The coefficients of the logit model are not the same as those given in Table 2. Takeshi Amemiya (1981) states that each of the OLS slope parameter estimates should approximately equal to 0.25 times the corresponding logit slope parameters.

Table 5:

	LPM			LOGIT			$0.25B_L$
Independent Variable	Parameter Estimate	Standard Error	t for H_0	Parameter Estimate	Standard Error	t for H_0	*
Constant	0.692297	0.1626	4.255	0.95089	0.80423	1.182	0.6377
LWW1	0.093301	0.0319	2.917	0.46449	0.19851	2.954	0.1161
KL6	-0.290638	0.0357	-8.136	-1.46897	0.19851	-7.400	-0.3672
K618	-0.008387	0.0139	-0.600	-0.05119	0.06842	-0.748	-0.0127
WA	-0.011609	0.0025	-4.554	-0.05825	0.01291	-4.513	-0.0145
WE	0.042096	0.0086	4.866	0.21193	0.04411	4.804	0.0529
UN	-0.003487	0.0054	-0.636	-0.01857	0.02646	-0.701	-0.0046
CIT	-0.004479	0.0367	-0.122	0.01127	0.17816	0.071	0.0028
PRIN	-0.000067	0.0000	-4.402	-0.00004	0.00001	-4.354	-0.00001

* shows the Amemiya's approximation (Amemiya, 1981: 1489);

$b_{LPM} \cong \mathbf{0.25} b_L$ except for the constant term,

$b_{LPM} \cong \mathbf{0.25} b_L + \mathbf{0.5}$ for the constant term.

From the last column, one can state that slope coefficients of LPM are roughly equal to those of logit in this sample.

V. Probit Model

In Section IV, we saw the cumulative logistic probability function. Another cumulative probability function is the cumulative normal probability function. This estimation procedure that is used in the estimation of dichotomous dependent variable models, that is based on the assumption that the cumulative distribution of the disturbances is normal, is known as the probit model or the normit model. In this section, I estimated the same parameters using probit maximum likelihood procedure. Results are given in Table 6.

In probit model, we first obtain an unobservable utility index Z_i which is determined by explanatory variables X_i

$$Z_i = B_1 + B_2X_i \tag{6}$$

The larger the value of the index Z_i, the greater the probability of woman will participate in labor force. Let Z_i^* represent critical level of the index and eq.(7) below

$$P_i = Pr(LFP=1) = Pr(Z_i^* \leq Z_i) = F(I_i) \tag{7}$$

can be computed from the cumulative normal probability function.

Table 6: MLE Results for Probit Model on Mroz Data.

Independent Variable	Parameter Estimate *	Standard Error *	t for H_0 *	Chi-square **	Pr> Chi-squre **
Constant	0.56790	0.48209	1.1780	1.3877	0.2388
LWW1	0.28211	0.09277	3.0409	9.2475	0.0024
KL6	-0.88080	0.11466	-7.6822	59.017	0.0001
K618	-0.02972	0.04077	-0.7290	0.5314	0.4660
WA	-0.03497	0.00768	-4.5539	20.738	0.0001
WE	0.12770	0.02597	4.9182	24/188	0.0001
UN	-0.01106	0.01598	-0.6919	0.4787	0.4890
CIT	0.01000	0.10763	0.0929	0.0086	0.9259
PRIN	-0.0002	0.00000	-4.5062	20.306	0.0001

Chi-squared[8] *	: 130.957 prob. : 0.00000
Log likelihood *	: -449.3946
Log L with only a constant term*	:-514.8732
Iterations *	: 4
LR. chi-square **	: 898.7892 with df = 744 (p=0.4747)
Pearson Chi-square **	: 745.7794 with df = 744 (p=0.0001)
Iterations **	: 4

* ET output, ** SAS output.

As, for instance, LWW1 increases by a unit, on average, Zi increases by 28 percent. Again if we want to find the probability for women earning $7 in one hour, with 2 children under age six, with 1 child between age six and eighteen, aged 32, with 13 years of schooling, with an unemployment rate of 5 percent, living in large city and with property income of $ 25,000, we obtain $Z_i = 1.2421$. Looking at the standardized normal distribution table, we obtain the P_i value of 0.3929 + 0.5 = 0.8929. In other words the probability (LFP=1) = 0.8929.

Signs of the estimated parameters are the same as those I obtained from the logit model. From Table 6, by observing the Chi-square and P>chi-square, one can state that constant term and variables K618, UN and CIT are statistically not different from zero at 95 percent confidence level and all other variables are statistically significant at 95 percent confidence level. It is seen from Table 6 that the nonlinear logit computational algorithm in ET and SAS converged after 4 iterations. Convergence in ET is same for both logit and probit. In SAS, however, convergence is more rapid in probit model than the logit model.

For goodness-of fit again "McFadden's R^2" (Maddala, 1983) or "likelihood ratio index" Green (1990).

$$LRI = 1 - [L_{UR} / L_R]$$
$$L_{UR} = -449.4732$$
$$L_R = -514.8732$$

LRI = 0.1271 indicating that the model is not a good fit to data of LFP.

For discrete choice models, among others stated by Amemiya (1981, 1503-1507) and Maddala (1983, 37-41), Cramer (1991, 158) states that the traditional goodness of fit is Pearson's chi-square test for the agreement of a frequency distribution with a probability distribution. SAS Proc-Probit gives both Pearson's chi-square and LR chi-square. Table 6 shows both criteria. Prob> Pearson's chi-sq of indicates that the joint hypothesis that all coefficients except the intercept are zero would be accepted whereas prob> LR chi-square states to accept the alternative hypothesis at 95% confidence level.

VI. Comparasion of Logit and Probit Models

In this section, I basically compare logit model based on cumulative logistic distribution and probit model based on cumulative normal distribution. Since these distributions are close to each other, in most cases the logit and probit estimated model models will be quite similar (Berndt, 1991: 657). Sample maximized log-likelihoods in my estimated logit and probit models are similar.

The similarities in the shapes of the logistic and normal distributions suggest that results of probit and logit analysis will differ by very little. Indeed, the inferences drawn from the two methods applied to the same data are invariably similar, and even parameter estimates from the two models will agree, approximately, up to a factor of proportionality. Logit coefficients tend to exceed probit coefficients by a scale factor in the range 1.6 to 1.8 (Nelson, 1990: 137).

The estimated effect of a change in a regressor on the probability of participating in the labor force (in other words their marginal effects) are:

$\partial P/\partial X = P(1-P)\beta_L$, in the logit model and

$\partial / \partial X = f(P)\beta_P$, in the probit model.

P= 0.568 (working sample/ Total sample)

and f(P)= cumulative normal density = 0.393

Table 7 : Marginal Effects when P= 0.568 and f(P)=0.393

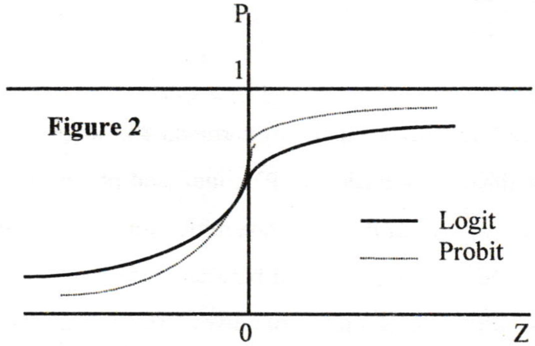
	Probit		Logit	
Independent Variable	Parameter Estimate	Marginal effects	Parameter Estimate	Marginal effects
Constant	0.56790	0.2231	0.95089	0.2333
LWW1	0.28211	0.1108	0.46449	0.1139
KL6	-0.88080	-0.3460	-1.46897	-0.3604
K618	-0.02972	-0.0116	-0.05119	-0.0125
WA	-0.03497	-0.0137	-0.05825	-0.0142
WE	0.12770	0.0501	0.21193	0.0.052
UN	-0.01106	-0.0039	-0.01857	0.0045
CIT	0.01000	0.0039	0.01127	0.0027
PRIN	-0.0002	0.0000	-0.00004	0.0000

Table 8: Marginal Effects when P=0.9 and f(P)=0.175

	Probit		Logit	
Independent Variable	Parameter Estimate	Marginal effects	Parameter Estimate	Marginal effects
Constant	0.56790	0.0993	0.95089	0.8558
LWW1	0.28211	0.0493	0.46449	0.4179
KL6	-0.88080	-0.1541	-1.46897	-1.3212
K618	-0.02972	-0.0052	-0.05119	-0.0460
WA	-0.03497	-0.0061	-0.05825	-0.0524
WE	0.12770	0.0223	0.21193	0.1930
UN	-0.01106	-0.0019	-0.01857	-0.0167
CIT	0.01000	0.0017	0.01127	0.0101
PRIN	-0.0002	0.0000	-0.00004	0.0000

As we see from the Table 7, marginal effects are almost same of logit and probit models. If P=0.9 and f(P)=0.175, as results are shown in Table 8, the marginal effects differ substantially.

Figure 2



Source: Pindyck and Rubinfeld (1991, 259)
Gujarati (1995, 568)

The slope coefficients of logit and the probit models are not directly comparable. The main difference between them is that cumulative logistic has slightly flatter tails and that cumulative normal curve approaches the axes more qoickly than the logistic curve (Figure 2).There are two approximations to transform logit estimates to comparable probit estimates.

1- $\sqrt{\dfrac{3}{\pi}} \cong \dfrac{1.73205}{3.14159} \cong 0.5513 B_L \cong B_P$ (8)

2- $0.625 B_L \cong B_P$ (Amemiya's approximation) (9)

Table 9: Two Approximations

Independent Variable	Approximation (1); eq.(8) Logit	Approximation (1); eq.(8) Probit	Amemiya's approximation; eq.(9) Logit	Amemiya's approximation; eq.(9) Probit	Actual probit
Constant	0.95089	0.52422	0.95089	0.59430	0.56790
LWW1	0.46449	0.25602	0.46449	0.29030	0.28211
KL6	-1.46897	-0.80980	-1.46897	-0.91810	-0.88080
K618	-0.05119	-0.02822	-0.05119	-0.03199	-0.02972
WA	-0.05825	-0.03211	-0.05825	-0.03640	-0.03497
WE	0.21193	0.11683	0.21193	0.13245	0.12770
UN	-0.01857	-0.01023	-0.01857	-0.01160	-0.01106
CIT	0.01127	0.00621	0.01127	0.00704	0.01000
PRIN	-0.00004	0.00000	-0.00004	0.00000	-0.0002

As we see from the Table 9, Amemiya's approximation is better than the first approximation. Because the first approximation gives the logistic distribution with zero mean and unit variance, it may be considered as a best, however by trail and error one finds that "0.625" does better since logistic distribution has slightly heavier tails (Amemiya, 1981: 1487).

VI. Conclusion

I found the probabilities that woman will participate in labor force given several characteristics using the estimation procedures LPM, logit and probit, that are used in the estimation of dichotomous variable LFP. The probabilities for woman earning $7 in one hour, with 2 children under age six, with 1 child between age six and eighteen, aged 32, with 13 years of schooling, with an unemployment rate of 5 percent, living in large city and with property income of $ 25,000 in LPM, logit and probit models are; 0.90, 0.88 and 0.89, respectively.

There are, however, some statistical problems with LPM, logit and probit models. I can summarize these problems as follows. The problems of LPM are

1-$E(Y_i) = P(Y_i=1)=\sum b_K X_{iK}$ can take values greater than one or less than zero. We treat them as zero since probabilities can not take negative and grater than one values.

2-$E(e^2_i) = \sigma^2$ is no longer tenable, and OLS estimate will not have smallest variance. Any hypothesis test based on these sampling variance will be invalid. Since the variance of e_i depends on i, the e_i are heteroscedastic hence OLS will result inefficient estimates and imprecise predictions.

3-Usual tests of significance for the estimated coefficients do not apply, estimated standard errors are not consistent, and the R^2 is no longer meaningful.

4-The slopes of the parameters are constant. And these have same effect on probability whether the initial value is near zero or near one. But dichotomous variable can take only on two variables 0 or 1.

The problems of Logit and Probit models can be counted as;

1- Logit and probit models are sensitive to misspecifications. In particular, in contrast to the OLS, estimators will be inconsistent if an explanatory variable is omitted or if there is heteroscedaticity.

2-In the logit model, the personal critical values are distributed as a hyperbolic-secant-square($sech^2$) distribution, the cumulative distribution of which is logistic function .The logistic function is easy to calculate, whereas the cumulative normal distribution is difficult to compute, involving the calculation of an integral.

3-An estimated slope parameter in a logit or probit does not estimate the change in the probability of y=1 due to a unit change in the relevant explanatory variable. This is given by marginal effect of variable on probability. However this effect (derivative of the expression for P(y=1) with respect to X) can give misleading estimates of probability changes in the contexts in which an explanatory variable is postulated to change by an amount that is not infinitesimal.

4- There is no universally accepted goodness-of-fit measure for probit and logit models. Maddala(1983) and Amemiya(1981) give several criteria for measurement for goodness-of-fit.

5- In many applications of the logit model, probit, or LPM, it happens that the number of observations in one of the groups is much smaller than the number in other group. For instance, in a study of unemployment, the number of non-working sample is smaller than the that of working sample. Therefore a larger sample size is needed. In this case, we either utilize data from census tapes or we have to sample the two groups at different sampling rates. In such cases a question arises as to how one should analyze the data. It has been commonly suggested that one should use Weighted LPM, weighted logit or weighted probit and by doing so we correct procedure if there is heteroscedasticity problem.

All three techniques yield estimates that have quite similar properties (e.g. asymptotically unbiased, efficient and normal), so that there is little gain to choose among three of them. LPM differs from probit and logit in being in linear, so that the linearity assumption may be key to that choice. As for preference between logit and probit model, only question is which distribution is to be chosen; cumulative logistic or cumulative normal, they are very similar each other. On the margin, then, the computer programs might effect the decision on which one to chose Today many computer programs such as ET, LIMDEP, SAS, RATS, TSP, SHAZAM and NUPROLD. run LPM, logit and probit models.

References

ALDRICH, J. H. AND NELSON, F. D.(1984), **Linear Probability, Logit, and Probit Models**, Series: Quantitative Applications in the Social Sciences, Sage University Publications, Newbury Park.

AMEMIYA, T. (1981), "Qualitative Response Models: A Survey," **Journal of Economic Literature,** Vol. XIX, 1483-1536.

______. (1985), **Advanced Econometrics**, Harvard University Press, Cambridge, Massachusetts.

BECKER, W.E. AND KENNEDY, P.E. (1992), "A Graphical Exposition of The Ordered Probit," **Econometric Theory**, Vol. 8, 127-134.

BERNDT, ERNST R. (1991), **The Practice of Econometrics: Classic and Contemporary**, Addison Wesley Publishing Company, New York.

COHEN, M. S., REA, S. A. AND LERMAN, R. L. (1970), **A Micro Model of Labor Supply**, BLS Staff Paper 4, U.S. Department of Labor.

CRAMER, J. S. (1991), **Econometric Applications of Maximum Likelihood Methods**, Cambridge University Press, Massachusetts.

GUJARATI, D. N. (1995), **Basic Econometrics**, McGraw-Hill, New York.

GUNDERSON, M. (1980), "Probit and Logit Estimates of Labor Force Participation," **Industrial Relations**, Vol. 19, No.2, The Regents of The University of California.

JUDGE, G., E.T. ALL. (1985), **The Theory and Practice of Econometrics,** John Wiley & Sons., New York.

KENNEDY, P. (1994), **A Guide to Econometrics**, The MIT Press, Cambridge, Massachusetts.

MADDALA, G.S. (1983), **Limited-Dependent and Qualitative Variables in Econometrics** Cambridge, England: Cambridge University Press, Econometric Society Monographs No.3. Massachusetts.

MROZ, THOMAS A. (1987), "The Sensitivity of an Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions", **Econometrica,** Vol.55, No.4, 765-799.

NELSON, FORREST D. (1990), "Logit, Probit and Tobit", **The New Palgrave: Econometrics**, W.W. Norton & Company, New York, 136-144,

SAS/STAT User's Guide, Release 6.03 Edition, SAS Institute, 1988.

SCHMIDT, P. AND STRAUSS, R.P. (1975), "The Prediction of Occupation Using Multiple Logit Models", **International Economic Review**, Vol. 16, No.2, 471-486.