Modeling Female Labor Force Participation with the Bivariate Probit Model

Okochab Igohn Chike & Isaac Didi Essi

Department of Mathematics/Computer Science Rivers State University, Port Harcourt. Rivers State, Nigeria

Abstract

This research focuses on binary choice models and its applications with emphasis, on variables that influence the participation of females in the work force. We consider factors, for example, spouse's training, wife's age, family income as well as the number of children in the family. So as to substantiate our case, we adopt a three-step approach in accomplishing our results. Firstly, we developed a probit and bivariate probit model of labor force participation (lfp) that involves kl6 (number of children under 6years old), k618 (number of children between 6 to 18 years), Wa (wife's age), We (wife's education), Faminc (family income excluding wife's wage) and W (wife's wage). Our findings with the real life data suggests that components, for example, a lady's age, the number of children she has and her family pay lessens her chances of participating in the labor force.

Key Word: Modeling, Female, Labor, Force, Participation

1.1 Introduction

Individuals and corporate bodies make decisions that affect their activities on the long run. One way or the other, most decisions made by individuals and corporate organizations are based on choice. These choices are made to estimate the value attributed to items by revealing their preferences and utilities. One particular choice is the binary choice model. Binary choice models are models that are used to relate and select an option among various alternatives (Greene, 2000).

In economic theory, binary choice models comply with Lancaster's consumer theory which states that an individual purchases an item based on the features that are attached to the items. Similarly, McFadden (1974) argues that a consumer's behavior is best described by the attributes of the product he wishes to purchase.

Consider, for instance, the choice regardless of whether to make a noteworthy buy or the choice of which course is best for transportation, or modeling labor participation.

A marketing company may want to predict consumer's loyalty to different brand of products. For instance, customers may need to choose where to shop, when to buy a product and which brand to purchase. Marketing practitioners had over the years, studied the factors that drive consumer's choices.

In healthcare, a patient might need to decide whether or not to purchase health insurance. One of such studies was carried out by Cardon and Hendell (2001), when they structurally estimated a model for health insurance and healthcare choices, using data on single and married patients. They tested for observable links between insurance status and healthcare consumption.

Brouwer *et al.*, (2013) models surge hazard protection in beach front zones in Vietnam. They made use of a binary choice model with data from the European Commission Research

Programme in Vietnam. Their point was to test the conditions under which family units support small scale surge protection. They found that location was significant factor.

Similarly, Ryan and Gerard (2003) provide a good review about binary choice models as it relates to healthcare economics. For the labor force participation, Mroz (1987) suggests that factors such as education, marital status or location could determine an individual's choice of being active in the work force. Furthermore, binary choice models play an important role in our everyday life. In modeling binary choice, we consider a Bernoulli model, where the chances of an event occurring is given as p and the chances of it not occurring is given as (1-p). In most cases, further estimation is usually done by maximizing the likelihood of an event occurring. For example, an individual's employment status could be full time, part-time or none while, transportation might involve a choice between travelling by air, land or sea. We refer to such a model as a multinomial choice model.

In addition, there are other models that provide a qualitative base for decision-making such as operations research models and game theory. Both approaches deal with the selection of an optimal course of action given the possible payoffs and their associated probabilities. However, there are certain real –life problems in which the fractional value of the decision variable has no significance. In such a situation, a binary choice model becomes applicable.

There are two commonly used discrete choice models namely: Logit and Probit models and their variants and extensions. Both models elucidate the attributes of the alternatives available to an individual. Consider, for example, modeling labor force participation. We might wish to studythe effect of age, educational qualification of an individual and that of their parents, marital status etc. A simple kind of one of these models is the ordinary Probit model where there exists one binary dependent variable with only one latent variable.

However, situations might arise where there are two binary dependent variables as well as two latent variables and there is need to model them jointly as a function of some explanatory variable i.e. whether to work or not. In such a situation, we refer to such a model as a bivariate probit model (Green (2003)). It goes further to provide a specification for modeling cases where there is an endogenous binary variable in one of the equations.

Chun and Mun (2012) assert that the bivariate probit model is a very useful model for estimation since it considers taste variation.

1.3 Aim and objectives of the study

The aim of this research is modeling binary choice models with applications. In other to achieve this, we consider the following objectives:

- **1.** To specify and estimate binary choice models.
- **2.** To simulate binary choice models with real life data.
- **3.** To investigate factors influencing the labor force participation of women.

2. Methodology

We explain the methods used in carrying out this work and the rationale behind using the specified models.

Our data set was obtained from a study on the working hours of married women. (Source: (Mroz (1987)).

2.1 Probit and Logit Models Specification

According to Wang (2006), the logit model is based on the odds that an event taking place. Logit (p) = $\ln(\frac{P}{1-P})$ (2.2)

If $P=Pr(Y=1|X\beta)$ is the probability of an event happening, Thenln $\left(\frac{P}{(1-P)}\right)$ is the corresponding log odds. The logit model states the log odds of an event happening are a linear function of a given set of explanatory variables, i.e. $\ln(\frac{P}{1-P}) = X\beta$ (2.3)The probability P=Pr(Y=1|X\beta) is solved by using P=P(Y=1|X\beta)= $\frac{\exp(X\beta)}{1+\exp(X\beta)}$ (2.4)The probability of Y being 0 is given as $P(Y=0|X\beta) = 1 - P(Y=1|X\beta) = 1 - \frac{\exp(X\beta)}{1 + (X\beta)}$ Therefore, the likelihood of the logit model is $L(B) = \prod_{i}^{n} \left[\frac{\exp(X\beta)}{1 + \exp(X\beta)} \right]^{Y_{i}} \left[1 - \frac{\exp(X\beta)}{1 + \exp(X\beta)} \right]^{1 - Y_{i}}$ (2.5)The log likelihood LL(B) is LL(B) = $\sum_{i=1}^{n} Y_i \ln \frac{\exp(X\beta)}{1 + \exp(X\beta)} + (1 - Y_i) \ln \left[1 - \frac{\exp(X\beta)}{1 + \exp(X\beta)}\right]$ (2.6)Coefficient estimates are obtained by maximizing LL(B) The Probit model is given as $P(Y=1|XB)=\Phi(XB) = \int_{-\infty}^{XB} \phi(z)dz$ (2.7)Where Y is a discrete variable that takes the value 1 or 0; X is a vector of explanatory variables, B is a vector of coefficients. $\Phi(z)$ is the cumulative normal distribution. The likelihood function L(B) is given as $L(B) = \prod_{i}^{n} [\Phi(XB)]^{Y_{i}} [1 - \Phi(XB)]^{1 - Y_{i}}$ (2.8)

And the log likelihood LL(B) for the probit model is given as $LL(B) = \sum_{i=1}^{n} Y_i \ln[\Phi(XB) + (1 - Y_i)\ln[1 - \Phi(XB)]$ (2.9) By maximizing LL(B), we obtain estimates of the coefficients.

In this study, we shall be using the probit model.

2.10 Bivariate Probit Model

The bivariate probit model is an extension of the probit regression model, where there are two binary dependent variables as well as two latent variables and there is need to model them jointly as a function of some explanatory variable where the distribution of the regression equations is assumed to be correlated in the same spirit as the seemingly uncorrelated regression. (Maddala, 1983).

Given two unobserved latent variables

 $y_1^* = x_1'\beta_1 + y_2\gamma + \varepsilon_1, y_1 = 1$ if $y_1^* > 0, 0$ otherwise $y_2^* = x_2'\beta_2 + \varepsilon_2, y_2 = 1$ if $y_2^* > 0, 0$ otherwise

Where y_1 is, a binary dependent variable of interest in equation 1, y_2 is a discrete dependent variable of equation 2 that is included in the first equation as an endogenous variable. x'_1 and x'_2 are the vectors of the explanatory variables of the two regression equations. The error terms of the two equations are assumed to be independent and independently distributed and follow the bivariate standard normal distribution with correlation coefficient ρ . The model collapses to two separate probit models y_1 and y_2 when the value of $\rho=0$.

We estimate a bivariate probit model using real life data. The analysis is given some interpretation in the light of socio-economic realities.

3. Data Analysis

3.1 Probit Models With Real Life Data

In this section, we conduct a probit regression using real life data. The first regression is in Table 3.2, where labour force participation (lfp) is regressed against kl6 (no of kids less than 6 years), k618 (no of kids between 6 and 18 years), Wa (wife's age), We (wifes education), Faminc (family income excluding wife's wage and W(wife's wage).W is a binary variable with value equal1 if the wife's receives wage and 0 otherwise.

Variable	Coefficient	P-value
Constant	-0.918	0.1364
K16	-0.6771	0.0000
K618	-0.4102	0.4229
Wa	-0.2488	0.0109
We	0.1258	0.0001
Faminc	0.0000	0.9739
W	2.5402 0.0000	
AIC = 0.701	Prob [ChiSqd>value]	= 0.0000

Table 3.2 Probit Regression 1 (Real Data)Dependent variable: lfp

The second probit regression has W2 as a dependent variable and the result with its independent variable are there in the table.

Table 4.7: Probit Regression 2 (Real Data)
--

Dependent variable: W			
Variable	Coefficient	P-value	
Constant	0.7417	0.1017	
K16	-0.7474	0.0000	
K618	-0.6077	0.1239	
Wa	-0.3552	0.0000	
We	0.6021	0.0088	
Faminc	0.0000	0.1497	
AIC =1.3018	Prob[ChiSqd> value]	= 0.0000	

Table 3.4 takes care of bivariate probit regression result. In the first equation of the bivariate regression, lfp is the dependent variable. Whereas in the second equation, W is the dependent variable.

IIARD – International Institute of Academic Research and Development

Dependent variable for First equation: lfp			
Variable	Coefficient	P-value	
Constant	-1.527	0.1587	
Kl6	-0.4005	0.4464	
K618	-0.1814	0.7654	
Wa	-0.115	0.6665	
We	-0.0969	0.1404	
Faminc	0.0000	0.7379	
W	3.2127	0.0000	

Table 3.4: Bivariate Probit Regression (Real Data)Dependent variable for First equation: lfp

Dependent variable for Second equation: W	
---	--

Variable		Coefficient	P-value
Constant	0.7532		0.0986
K16	-0.7524		0.0000
K618	-0.0603		0.1193
Wa	-0.0355		0.000
We	0.0589		0.0118
Faminc	0.0000		0.1463
Rho(1,2)	-0.5613		0.5397

In these tables, we have

Lfp= paid labour force: 1=Yes, 0=No kl6= kids less than 6 years k618= kids between 6-18 years Wa= wife's age We= wife's educational attainment faminc= income of the family, excluding wife's W= reported wife's wage: 1=yes, 0= no rpwg= reported wife's wage

3.5: Effect of Increasing Number of Kids Less than 6 Years (kl6)

We study the effect of increasing number of kids below 6 years old in each family by 1, 2 and 3 kids. The scenarios and the results of these increments are presented in Tables 3.6.

Outcome	Base case	Under Scenario	Change
0	53.70%	55.50%	1.8
1	46.22%	44.49%	-1.8
Total	100.00%	100.00%	0

Table 3.6: Scenario 1: Increasing kl6 by 1

Outcome	Base case	Under Scenario	Change
0	53.70%	57.90%	-4.2
1	46.22%	42.10%	-4.2
Total	100.00%	100.00%	0

Scenario 2: Increasing kl6 by 2

Scenario 3: Increasing kl6 by 3

Outcome	Base case	Under Scenario	Change
0	53.78%	79.81%	26.03
1	46.22%	20.19%	-26.03
Total	100.00%	100.00%	0

4. Discussion

1st Probit Regression (Real Data))

From Table 3.2, the negative sign of the estimate of coefficient associated with the labor force participation (kl6= -0.6771) suggests that having kids that are less than 6 years impacts negatively on a woman's participation. This could mean that women, who have kids that are less than 6 years, have fewer opportunities to seek employment. Similarly, a negative sign on the coefficient (k618 = -0.4102) signifies that when a woman have kids that are aged between 6 to 18 years, they impact negatively on her participation in the labor force thereby reducing her chances of seeking employment. The negative value on the coefficient, wife's age (Wa = -0.2488) is an indication that as a woman's age increases, her chances of seeking employment reduces. Meanwhile, the positive sign of the coefficient associated with the wife's income (We = 0.1258) suggests that the more income a woman earns the higher her chances of seeking employment. Furthermore, the positive sign of the coefficient associated with wife's education (We =0.1258) suggests that the coefficient has a positive effect on the dependent variable. This could mean that the more educated a woman becomes the more chances she has to be employed. The value of the coefficient (Faminc=0) suggests that family income does not affect the employment opportunities of a woman. However, it is unclear why it is so. Although, factors such as, cultural and religious norms, etc, might as well be considered. Among the various selection criterions, we choose the AIC (0.701) because it has the least error among other information criterions. In addition, the p-values of coefficients (we and w) are all significant.

2ndProbitRegression (Real Data)

From the Table, we observe a negative sign on the coefficient associated with a woman having kids less than 6 years (kl6 = -0.7417). This suggests that the more a woman have kids that are less than 6 years, the lesser her chances in applying for a job. Similarly, a negative coefficient (k618 = -0.6077) suggests that having kids between 6 to18 years reduces the chances of mothers to participate in the labor force. The negative value for a woman's age (wa= -0.3552) implies that as a woman gets old, her chances of seeking employment reduces while a positive value for wife's education (we =0.6021) implies that the more educated a woman is, the higher her chances of being employed in an organization. Moreover, a p-value (p<0.05) also suggests that a the variable is significant. The zero value for family income (faminc=0.000) suggests that a

woman's family income does not affect her employment opportunities. Regarding the selection criteria, we choose the AIC (1.3018) because it has the least error value.

Bivariate Regression (Real data)

From the 1st equation in the table, we observe a negative sign on the coefficient associated with a woman having kids less than 6 years (kl6 = -0.7417). This suggests that the more a woman have kids that are less than 6 years, the lesser the chances she has in seeking for a job. Similarly, a negative coefficient (k618 = -0.6077) suggests that having kids between 6 to 18 years has a negative impact on the employment opportunities of a woman. A negative coefficient for a woman's age (Wa= -0.1150) implies that as a woman's age increases, her chances of securing a job decreases while a positive value for wife's education (We = 0.6021) implies that the more educated a woman is, the higher her chances of being employed. Furthermore, a p-value (p<0.05) also suggests that the variable is significant. The zero value for family income (Faminc=0.000) suggests that family income does not play a role in a woman's decision to be employed. The positive value of the coefficient for wife's wage (W=3.2127) suggests that wage has a positive effect on the employment opportunities of a woman. Regarding the selection criteria, we choose the AIC (1.3018) because it has the least error value while for the 2nd equation, a negative sign on the coefficient associated with women having kids less than 6 years (k16 = -0.7532). This suggests that the more a woman have kids that are less than 6 years, the lower her chances of seeking employment. Similarly, a negative coefficient (k618 = -0.0603) suggests that women who have kids between 6 to18 years stand the risk of not being focused on their job. The negative coefficient for a woman's age (wa= -0.0355) implies that as a woman gets old, her chances of seeking employment reduces. Meanwhile, the positive value for wife's education (We =0.0589) implies that the more educated a woman becomes, the higher her chances of getting a job. Moreover, a p-value (p<0.05) also suggests that the variable is significant. The zero value for family income (Faminc=0.000) suggests that family income does not affect a woman's chances of seeking for employment. We choose the AIC (1.3018) because it has the least error value.

Effect of increasing the number of kids below 6 years (kl6) by 1, 3 and 5

From Table 3.5, Scenario 1, we observe that when kl6 is increased by 1 unit, the probability of occurrence reduces from 0.46 to 0.44. This signifies a reduction in her chances of seeking employment, when increased by one unit. Similarly for scenario 2, when kl6 is increased by 3 units, the probability of a woman applying for a job reduces from 0.46 to 0.42. Finally, for scenario 3, when kl6 is increased by 5units, the probabilities of occurrence shifts from 0.46 to 0.20.We observe here, that as kl6 increases, a woman's chances of seeking for employment reduces.

5. Conclusion and Recommendation

From our analysis, we have achieved the following:

- 1. Specifying and estimating binary choice models.
- 2. Simulated binary choice models with real life data.
- 3. Investigated factors/variables influencing labor force participation of women.

5.1 Recommendation

The model in our methodology can be extended to other cases such as:

- **1.** Labor force participation of men.
- 2. The impact of religion and other cultural norms on the labor force participation of women.

References

- Aitchison, J., & Brown, J. A. C. (1957). *The Lognormal Distribution*. Number 5 in University of Cambridge, Department of Applied Economics Monographs. Cambridge: Cambridge University Press.
- Alba, A. Avare, Z. G., & Carrasco, R. (2009). On the Estimating of the Effect of Labor Participation of Fertility. *Spanish Economic Review*. Vol. pp 1.22.
- Aromolaran, A. (2004). "Female Schooling, Non-Market Productivity, and Labor Market Participation". A comparative analysis using the generations and gender survey', *Demographic research*, 27, 53-84.
- Arpino B, Pronzato C.D, Iavares H.P. (2014). The Effect of Grand Parental Support on Mothers Labor Market Participation: An Instrumental Variable Approach: *European Journal of Population*, Vol 30, pp 369-390
- Assad R (2002) The timing of marriage, fertility, and female labor force participation in Morroco. *The Economics of women and works in the Middle East and North Africa*. Vol 4, pp74-87
- Assev: A., Arpino B, Goisis A. (2012) Grand parenting and Mother's Labor Force Participations: A Comparative Analysis Using the Generations and Gender Survey Demographic ch, Vol. 27, pp 53-84.
- Azam, A. Rafiq, M. (2014). Female labour force participation in third world countries: An empirical Analysis. *International Journal of Economics and Management Engineering*, 8, 14-20.
- Bell, Richard & Dudgeon, Paul &Mcgorry, Patrick & Jackson, Henry. (1998). The dimensionality of schizophrenia concepts in first-episode psychosis. *Actapsychiatrica Scandinavica*. Vol 97.pp 334-42
- Bergrath, E., Pavlova, M., Groot, W(2014) Attracting Health Insurance Buyers through Selective . *Risks* ,vol 2,pp 146-170
- Bicakova, A. (2016). Gender unemployment gaps in the EU; Blame the family. 12A *Journal* of European Labour Studies, 5, 22-29
- Bliss, C. I. (1934). The method of probits. Science, 79,38-39
- Bliss, C.I. (1935) The Calculation of the dosage-mortality curve. *Annals of Applied Biology*, vol 22, pp 134-167
- Borhequez, A., Becarra, L., &Giraldo, M. (2016) Health Status and Labor Force Participation: evidence for urban low and middle income individuals in Columbia. *Portuguese Economics Journal*, 15,33-55
- Brazier, J. Rowen, D. Yang, Y., &Tsuchiya, A. (2012). Comparison of health state utility values derived using time trade off, rank ask discrete choice data anchored on the full health dead scale. Scale. *European Journal of health Economics*, 13, 57-5587
- Brouwer, R., Tinh, B.D., Tuan.T (2013). Modelling demand for catastrophic flood risk insurance in coastal zones in Vietnam using choice experiment. *Environmental and Development Economics*, vol 19,228-249
- Cardon, J. H., &Hendel, I. (2001). Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *RAND Journal of Economics*, *32*(3), 408-427. DOI: 10.2307/2696362
- Castellano R, Punzo. G., Rocca, A. (2013) Women's Job Search Propensity and Selection Effect in European Labour Markets. *MetoloskiZvezki*, Vol. 10, pp 83-98.
- Chun, H., &Mun, S. (2012) Determinations of Cooperation in Small and Medium Enterprises, Small Business Economics, 39, 419-436.
- Clementine, S. (2016). Working and Women Empowerment in the Egyptian housedhold: The type of work and location matter: <halsh-01525220>.

IIARD – International Institute of Academic Research and Development

- Coast, J., Flynn, T. N., &Natarafan, L. (2008). Valuing the ICECAP capability index for older people. Social science and medincine, 67,874-882
- Colon, M. S. Moles M.C. (2013) Housing and Labor Decisions of Households. *Review of Economics Household* Vol. pp 55-82.
- Compton D, Polak K.A. (2013) Family Proximity, Childcare, and Women's Labor Force Attachment .Journal of Urban Economics, http://dx.doi.org/10.1016j.june.2013.03.007
- Compton, J,Polak.R. (2014) Family proximity childcare and women labor force attachment. *Journal of Urban Economics*, vol 79,pp 72-90
- Connelly, R. Kimmel J. (2001) Participation and Welfare Recipients of Single Mothers: Implications for Welfare Reform: Upjohn Institute Working Paper 07-69.
- Corman H. Noonan, K. Reichman N. (2005) Mothers Labor Supply in Tragile Families: The Role of Child Health. *Eastern Economic Journal*, Vol 31, pp 1-616.
- Crespo, L. (2006). "Caring for parents and employment status of European mid-wife women." Paper presented at the Second Workshop on Economics of the Family, held at the University of Zaragoza.
- Dobis, E.(2011) A Socio Demographic Analysis of Health Care Utilization in the United States of America.
- Ekert O. Stier. H. (2009) Normative or Economic Behaviour? Fertility Women's Employment in Isreal. *Social Science Research* Vol. 38, pp 644-655.
- Ettner, S. L. (1996). "The opportunity costs of elder care." *Journal of Human Resources, 31* (1), 189-205
- Fechner, G. T. (1860). Elemente der Psychophusik. Leipzig: Breitkpof and Hartel.
- Finney, D. (1947). Probit Analysis (1st edition) Cambridge University Press, Cambridge, UK.
- Finney, D. (1971). *Probit Analysis* (Thirded). Cambridge: Cambridge University Press. First edition in 1947.
- Frontiers in Econometrics. New York: Academic Press.
- Garganta S. Gaspararini L. Marchison M. Tappata M. (2017) The Effect of Cash Transfers on Fertility: Evidence from Argentina. *Population Residents Policy Review* Vol 36, pp 1-24.
- Gensler, H. Walls D. (2006) Labor Force and Welfare Program Participation: The Effects of Welfare. *American Journal of Economics and Sociology*. Vol. 56, pp 229-241.
- Gorlich, D. Mahmond T.O., Trebesh C. (2007) Explaining Labour Market Inactivity in Migrant – Sending Families: Household, Hammock, or Higher Education? *Kiel Working Paper No. 1391*.
- Green, W., (2003) Econometric Analysis, Prentice Hall, Upper Saddle River. New Jersey.
- Gujurati, D.(2003) Basic Econometrics 4th edition, Toronto: McGrawHill
- Habib, K. (2015). An investigation on mode choice and travel distance demand of older people in the National Capital Region (NCR) of Canada: application of a utility theoretic joint econometric model. *Transportation*, 42 143-61
- Hander, A. Jahangri A. (2017) How Trust Towards Family Decreases Female Labor Force Participation. *Journal of Labor Research* Vol. 38, pp 122-144.
- Hardon, D. C., Hay, R. D. Uebersax, J., &Hauber, T, (1992). Improving Task comprehension in the measurement of health state preferences. A trial of informational cartoon figures and paried-comparison task. *Journal of clinical epidemiology*, 45,233-243.

- Hayo, B., Caris, T (2013) Female Labor Force Participation in Mena Region: The Role of Identify: *Review of Middle East Economic and Finance* Vol. 9, pp 271-292.
- Heintz, J., Kabar, N., & Mahmud, S. (2017). Cultural norms, economic incentives and women's labour market beaviour: Empirical insights from Bangladesh. Oxford Development studies. ISSN 1360-0818. DOI: 10. 1080/13600818.20017.1382464.
- Heitmueller, A., & Michaud, P.C. (2006). Informed care and employment in England: Evidence from the British household Panel Survey, IZA Discussion Paper No.2010, IZA, Bonn.
- Herrate, A. Carcedo, Saez .F., (2012) The Impact of Child Birth on Spanish Women's Decisions to Leave the Labor Market. *Review of Economic Household*. Vol. 10,