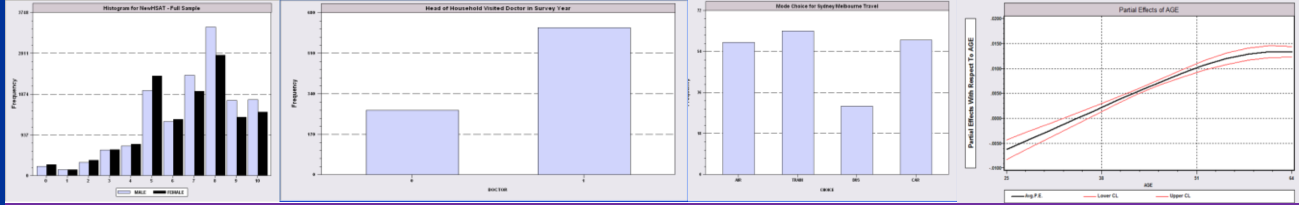
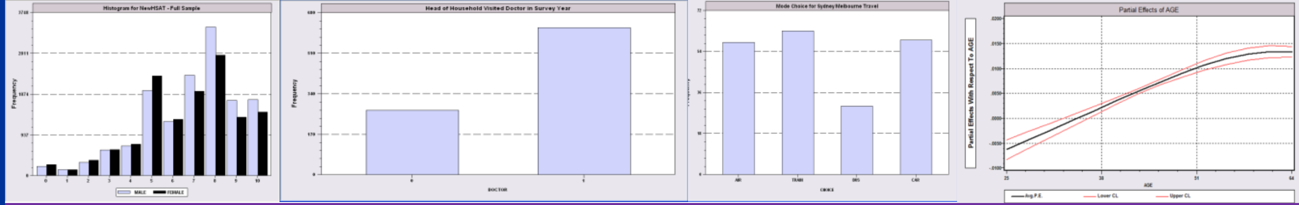


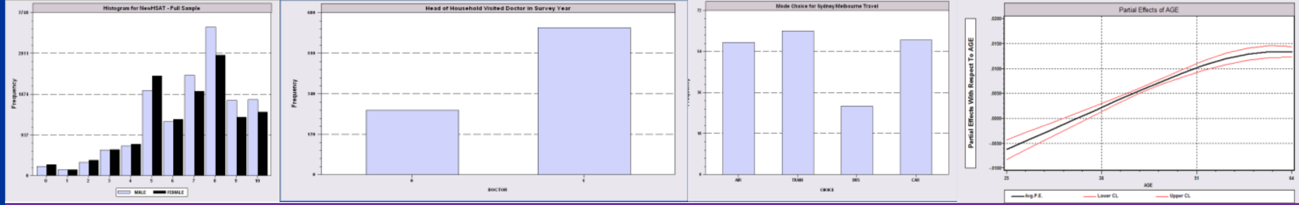
Discrete Choice Modeling

William Greene
Stern School of Business
New York University



Multivariate Binary Choice Models

- Bivariate Probit Models
  - Analysis of bivariate choices
  - Marginal effects
  - Prediction
- Simultaneous Equations and Recursive Models
- A Sample Selection Bivariate Probit Model
- The Multivariate Probit Model
  - Specification
  - Simulation based estimation
  - Inference
  - Partial effects and analysis
- The ‘panel probit model’

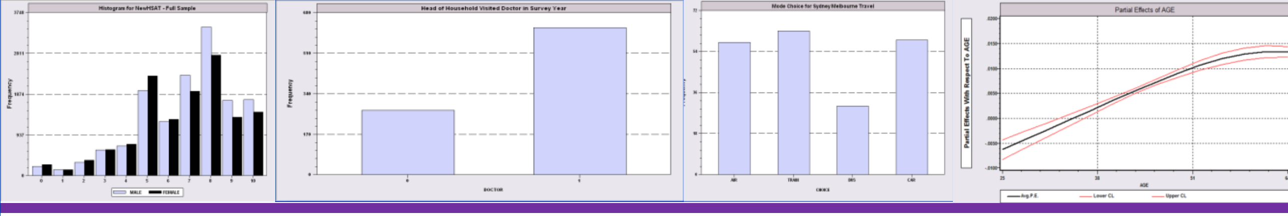


Application: Health Care Usage

German Health Care Usage Data, 7,293 Individuals, Varying Numbers of Periods
Variables in the file are

Data downloaded from Journal of Applied Econometrics Archive. This is an unbalanced panel with 7,293 individuals. They can be used for regression, count models, binary choice, ordered choice, and bivariate binary choice. This is a large data set. There are altogether 27,326 observations. The number of observations ranges from 1 to 7. (Frequencies are: 1=1525, 2=1079, 3=825, 4=926, 5=1051, 6=1000, 7=887). Note, the variable NUMOBS below tells how many observations there are for each person. This variable is repeated in each row of the data for the person.

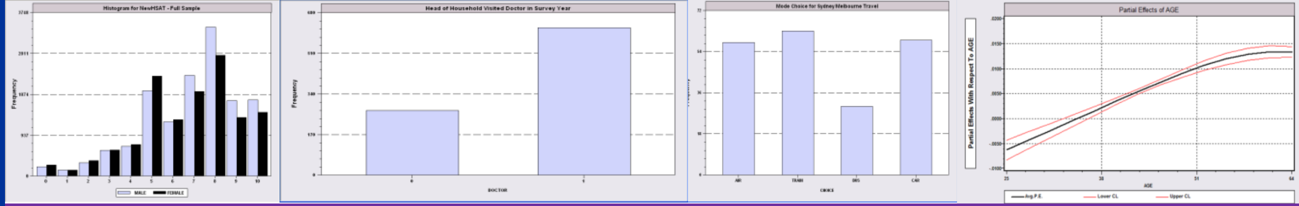
→ **DOCTOR = 1(Number of doctor visits > 0)**
→ **HOSPITAL = 1(Number of hospital visits > 0)**
HSAT = health satisfaction, coded 0 (low) - 10 (high)
DOCVIS = number of doctor visits in last three months
HOSPVIS = number of hospital visits in last calendar year
→ **PUBLIC = insured in public health insurance = 1; otherwise = 0**
ADDON = insured by add-on insurance = 1; otherswise = 0
HHNINC = household nominal monthly net income in German marks / 10000.
(4 observations with income=0 were dropped)
HHKIDS = children under age 16 in the household = 1; otherwise = 0
EDUC = years of schooling
AGE = age in years
MARRIED = marital status
EDUC = years of education



Gross Relation Between Two Binary Variables

Cross Tabulation Suggests Presence or Absence of a Bivariate Relationship

```
+------------------------------------------------------------------+
|Cross Tabulation                                                  |
|Row variable is DOCTOR    (Out of range 0-49:        0)           |
|Number of Rows =   2      (DOCTOR    =   0 to   1)                |
|Col variable is HOSPITAL (Out of range 0-49:        0)            |
|Number of Cols =   2      (HOSPITAL =   0 to   1)                 |
|Chi-squared independence tests:                                   |
|Chi-squared[   1] =  430.11235   Prob value =  .00000             |
|G-squared  [   1] =  477.27393   Prob value =  .00000             |
+------------------------------------------------------------------+
|              HOSPITAL                                            |
+--------+--------------+-------+                                  |
|  DOCTOR|      0      1| Total|                                   |
+--------+--------------+-------+                                  |
|       0|   9715    420| 10135|                                   |
|       1|  15216   1975| 17191|                                   |
+--------+--------------+-------+                                  |
|   Total|  24931   2395| 27326|                                   |
+------------------------------------------------------------------+
```



Tetrachoric Correlation

A correlation measure for two binary variables

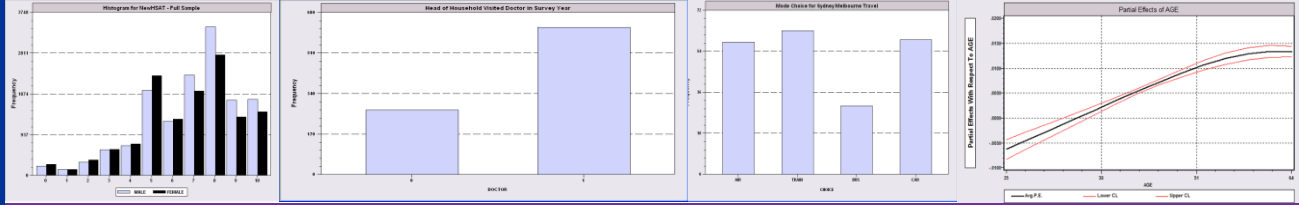
Can be defined implicitly

$$y_1^* = \mu_1 + \varepsilon_1,\ y_1 = 1(y_1^* > 0)$$

$$y_2^* = \mu_2 + \varepsilon_2,\ y_2 = 1(y_2^* > 0)$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$$

ρ is the **tetrachoric correlation** between y_1 and y_2



Log Likelihood Function

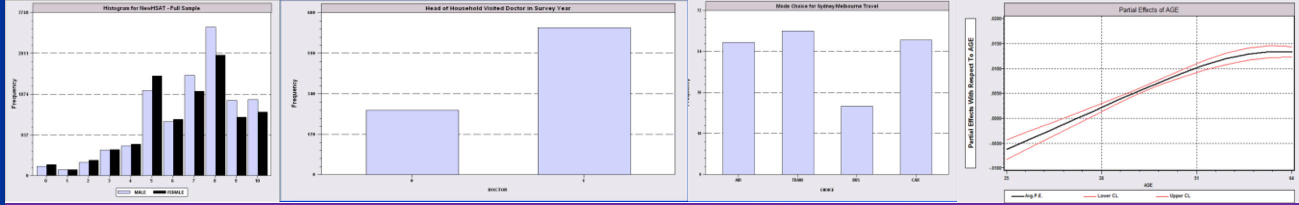
$$\log L = \sum_{i=1}^{n} \log \Phi_2 \left[(2y_{i1}-1)\mu_1, (2y_{i2}-1)\mu_2, (2y_{i1}-1)(2y_{i2}-1)\rho\right]$$

$$= \sum_{i=1}^{n} \log \Phi_2 \left[q_{i1}\mu_1, q_{i2}\mu_2, q_{i1}q_{i2}\rho\right]$$

Note: $q_{i1} = (2y_{i1}-1) = -1$ if $y_{i1} = 0$ and $+1$ if $y_{i1} = 1$.

Φ_2 = Bivariate normal CDF - must be computed using quadrature

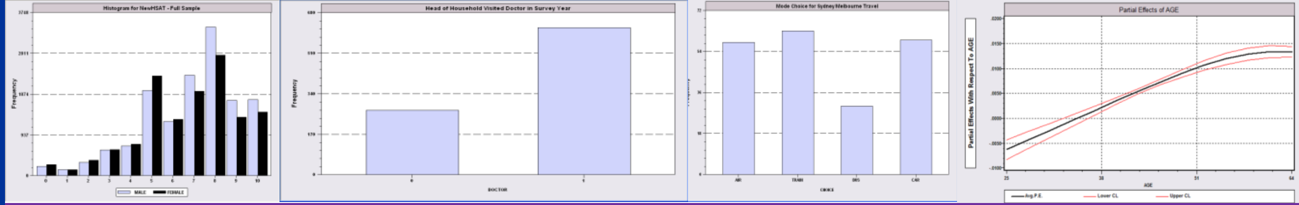
Maximized with respect to μ_1, μ_2 and ρ.



Estimation

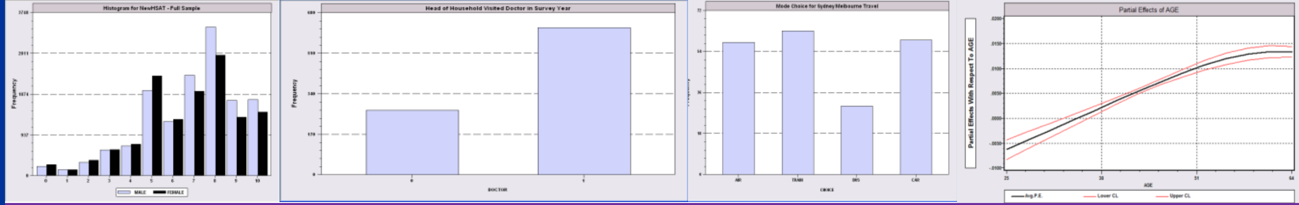
FIML Estimates of Bivariate Probit Model	
Maximum Likelihood Estimates	
Dependent variable	DOCHOS
Weighting variable	None
Number of observations	27326
Log likelihood function	-25898.27
Number of parameters	3

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]
Index equation for DOCTOR				
Constant	.32949128	.00773326	42.607	.0000
Index equation for HOSPITAL				
Constant	-1.35539755	.01074410	-126.153	.0000
Tetrachoric Correlation between DOCTOR and HOSPITAL				
RHO(1,2)	.31105965	.01357302	22.918	.0000



A Bivariate Probit Model

- Two Equation Probit Model
- (More than two equations comes later)
- No bivariate logit – there is no reasonable bivariate counterpart
- Why fit the two equation model?
  - Analogy to SUR model: Efficient
  - Make tetrachoric correlation conditional on covariates – i.e., residual correlation



Bivariate Probit Model

$$y_1^* = \boldsymbol{\beta}_1'\mathbf{x}_1 + \varepsilon_1,\ y_1 = 1(y_1^* > 0)$$

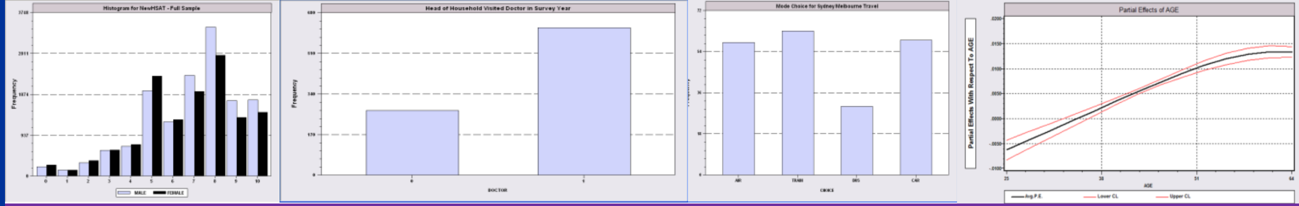
$$y_2^* = \boldsymbol{\beta}_2'\mathbf{x}_2 + \varepsilon_2, y_2 = 1(y_2^* > 0)$$

$$\begin{pmatrix}\varepsilon_1 \\ \varepsilon_2\end{pmatrix} \sim N\left[\begin{pmatrix}0 \\ 0\end{pmatrix}, \begin{pmatrix}1 & \rho \\ \rho & 1\end{pmatrix}\right]$$

The variables in $\mathbf{x}_2$ and $\mathbf{x}_2$ may be the same or different. There is no need for each equation to have its 'own variable.'

ρ is the conditional tetrachoric correlation between y_1 and y_2.

(The equations can be fit one at a time. Use FIML for (1) efficiency and (2) to get the estimate of ρ.)



ML Estimation of the Bivariate Probit Model

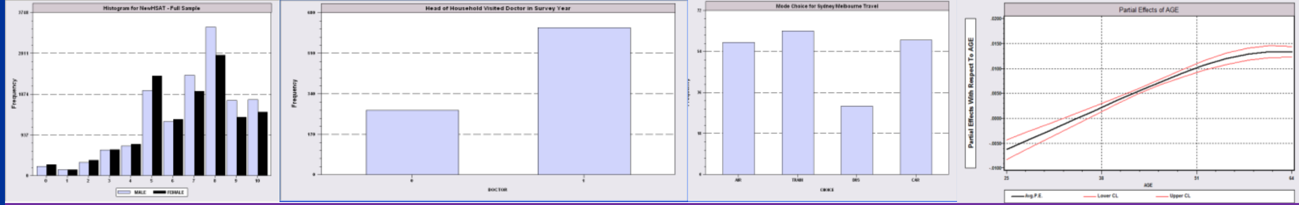
$$\log L = \sum_{i=1}^{n} \log \Phi_2 \begin{bmatrix} (2y_{i1}-1)\boldsymbol{\beta}_1'\mathbf{x}_{i1}, \\ (2y_{i2}-1)\boldsymbol{\beta}_2'\mathbf{x}_{i2}, \\ (2y_{i1}-1)(2y_{i2}-1)\rho \end{bmatrix}$$

$$= \sum_{i=1}^{n} \log \Phi_2 \left[q_{i1}\boldsymbol{\beta}_1'\mathbf{x}_{i1}, q_{i2}\boldsymbol{\beta}_2'\mathbf{x}_{i2}, q_{i1}q_{i2}\rho \right]$$

Note: $q_{i1} = (2y_{i1}-1) = -1$ if $y_{i1} = 0$ and $+1$ if $y_{i1} = 1$.

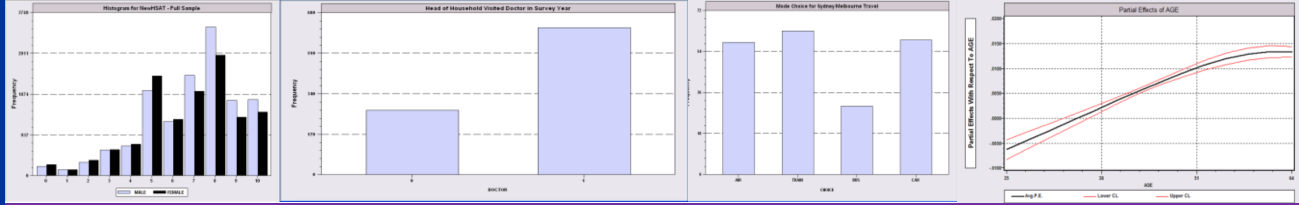
Φ_2 = Bivariate normal CDF - must be computed using quadrature

Maximized with respect to $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2$ and ρ.



Application to Health Care Data

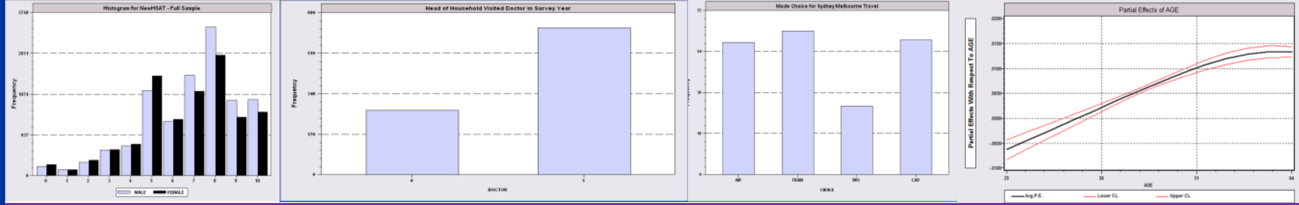
```
x1=one,age,female,educ,married,working
 x2=one,age,female,hhninc,hhkids
 BivariateProbit ;lhs=doctor,hospital
                 ;rh1=x1
                 ;rh2=x2;marginal effects $
```



Parameter Estimates

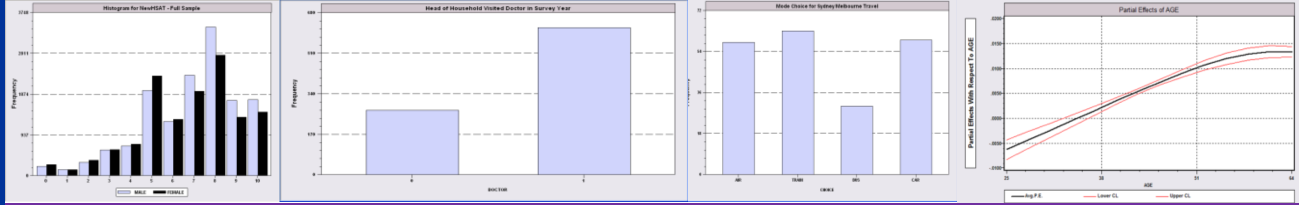
FIML Estimates of Bivariate Probit Model
Dependent variable DOCHOS
Log likelihood function -25323.63074
Estimation based on N = 27326, K = 12

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Index equation for DOCTOR				
Constant	-.20664***	.05832	-3.543	.0004	
AGE	.01402***	.00074	18.948	.0000	43.5257
FEMALE	.32453***	.01733	18.722	.0000	.47877
EDUC	-.01438***	.00342	-4.209	.0000	11.3206
MARRIED	.00224	.01856	.121	.9040	.75862
WORKING	-.08356***	.01891	-4.419	.0000	.67705
	Index equation for HOSPITAL				
Constant	-1.62738***	.05430	-29.972	.0000	
AGE	.00509***	.00100	5.075	.0000	43.5257
FEMALE	.12143***	.02153	5.641	.0000	.47877
HHNINC	-.03147	.05452	-.577	.5638	.35208
HHKIDS	-.00505	.02387	-.212	.8323	.40273
	Disturbance correlation				
RHO(1,2)	.29611***	.01393	21.253	.0000	



Marginal Effects

- What are the marginal effects
  - Effect of what on what?
  - Two equation model, what is the conditional mean?
- Possible margins?
  - Derivatives of joint probability = $\Phi_2(\boldsymbol{\beta}_1'\mathbf{x}_{i1}, \boldsymbol{\beta}_2'\mathbf{x}_{i2},\rho)$
  - Partials of $E[y_{ij}|\mathbf{x}_{ij}] = \Phi(\boldsymbol{\beta}_j'\mathbf{x}_{ij})$ (Univariate probability)
  - Partials of $E[y_{i1}|\mathbf{x}_{i1},\mathbf{x}_{i2},y_{i2}=1] = P(y_{i1},y_{i2}=1)/\text{Prob}[y_{i2}=1]$
- Note marginal effects involve both sets of regressors. If there are common variables, there are two effects in the derivative that are added.



Bivariate Probit Conditional Means

$$\text{Prob}[y_{i1}=1,y_{i2}=1]=\Phi_2(\boldsymbol{\beta}_1'\mathbf{x}_{i1},\boldsymbol{\beta}_2'\mathbf{x}_{i2},\rho)$$

This is not a conditional mean. For a generic $\mathbf{x}$ that might appear in either index function,

$$\frac{\partial \text{Prob}[y_{i1}=1,y_{i2}=1]}{\partial \mathbf{x}_i}=g_{i1}\boldsymbol{\beta}_1+g_{i2}\boldsymbol{\beta}_2$$

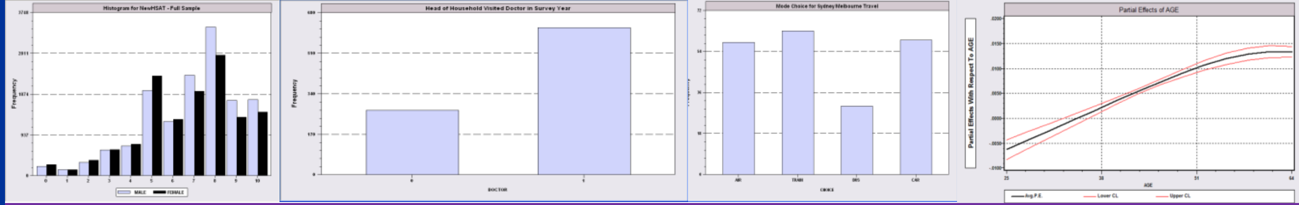
$$g_{i1}=\varphi(\boldsymbol{\beta}_1'\mathbf{x}_{i1})\Phi\left(\frac{\boldsymbol{\beta}_2'\mathbf{x}_{i2}-\rho\boldsymbol{\beta}_1'\mathbf{x}_{i1}}{\sqrt{1-\rho^2}}\right),g_{i2}=\varphi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})\Phi\left(\frac{\boldsymbol{\beta}_1'\mathbf{x}_{i1}-\rho\boldsymbol{\beta}_2'\mathbf{x}_{i2}}{\sqrt{1-\rho^2}}\right)$$

The term in $\boldsymbol{\beta}_1$ is 0 if $\mathbf{x}_i$ does not appear in $\mathbf{x}_{i1}$ and likewise for $\boldsymbol{\beta}_2$.

$$E[y_{i1}\,|\,\mathbf{x}_{i1},\mathbf{x}_{i2},y_{i2}=1]=\text{Prob}[y_{i1}=1\,|\,\mathbf{x}_{i1},\mathbf{x}_{i2},y_{i2}=1]=\frac{\Phi_2(\boldsymbol{\beta}_1'\mathbf{x}_{i1},\boldsymbol{\beta}_2'\mathbf{x}_{i2},\rho)}{\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})}$$

$$\frac{\partial E[y_{i1}\,|\,\mathbf{x}_{i1},\mathbf{x}_{i2},y_{i2}=1]}{\partial \mathbf{x}_i}=\frac{1}{\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})}(g_{i1}\boldsymbol{\beta}_1+g_{i2}\boldsymbol{\beta}_2)-\frac{\Phi_2(\boldsymbol{\beta}_1'\mathbf{x}_{i1},\boldsymbol{\beta}_2'\mathbf{x}_{i2},\rho)\varphi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})}{[\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})]^2}\boldsymbol{\beta}_2$$

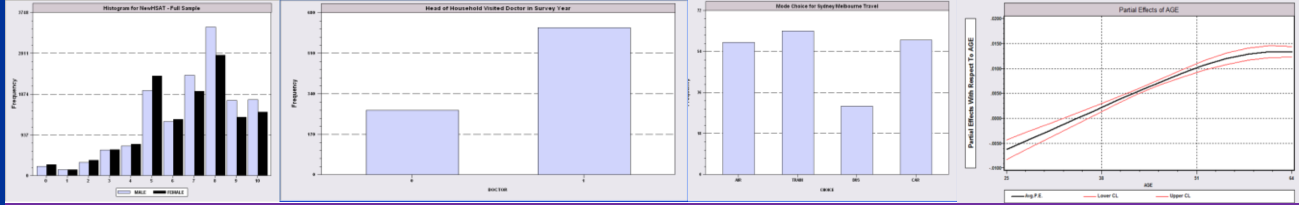
$$=\left[\frac{g_{i1}}{\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})}\right]\boldsymbol{\beta}_1+\left[\frac{g_{i2}}{\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})}-\frac{\Phi_2(\boldsymbol{\beta}_1'x_{i1},\boldsymbol{\beta}_2'x_{i2},\rho)\varphi(\boldsymbol{\beta}_2'x_{i2})}{[\Phi(\boldsymbol{\beta}_2'\mathbf{x}_{i2})]^2}\right]\boldsymbol{\beta}_2$$



Marginal Effects: Decomposition

Marginal Effects for Ey1|y2=1

Variable	Efct x1	Efct x2	Efct z1	Efct z2
AGE	.00383	-.00035	.00000	.00000
FEMALE	.08857	-.00835	.00000	.00000
EDUC	-.00392	.00000	.00000	.00000
MARRIED	.00061	.00000	.00000	.00000
WORKING	-.02281	.00000	.00000	.00000
HHNINC	.00000	.00217	.00000	.00000
HHKIDS	.00000	.00035	.00000	.00000

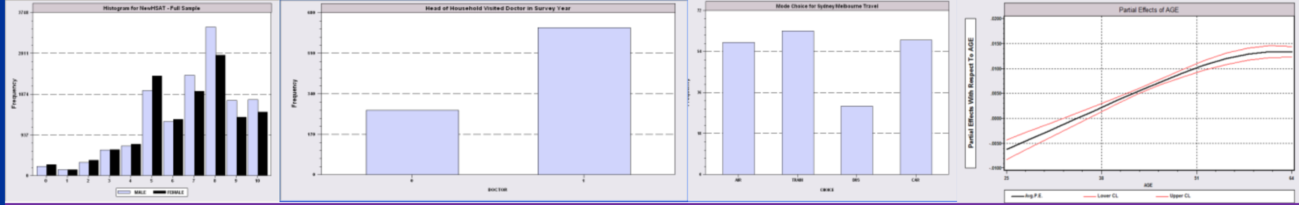


Direct Effects

Derivatives of $E[y_1|x_1,x_2,y_2=1]$ wrt x_1

```
+------------------------------------------+
| Partial derivatives of E[y1|y2=1] with   |
| respect to the vector of characteristics.|
| They are computed at the means of the Xs.|
| Effect shown is total of 4 parts above.  |
| Estimate of E[y1|y2=1] = .819898         |
| Observations used for means are All Obs. |
| These are the direct marginal effects.   |
+------------------------------------------+
```

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
AGE	.00382760	.00022088	17.329	.0000	43.5256898
FEMALE	.08857260	.00519658	17.044	.0000	.47877479
EDUC	-.00392413	.00093911	-4.179	.0000	11.3206310
MARRIED	.00061108	.00506488	.121	.9040	.75861817
WORKING	-.02280671	.00518908	-4.395	.0000	.67704750
HHNINC	.000000	......(Fixed Parameter).......			.35208362
HHKIDS	.000000	......(Fixed Parameter).......			.40273000

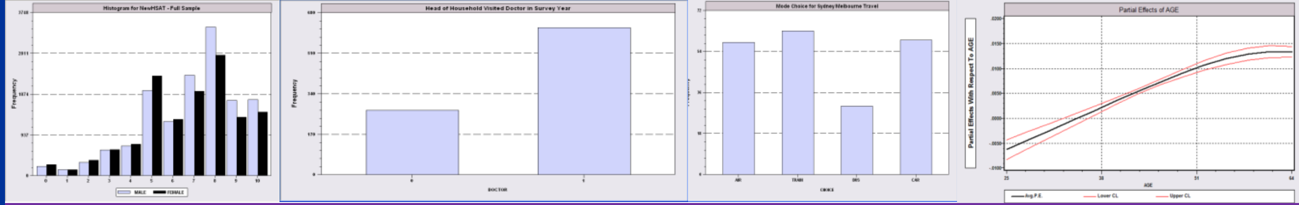


Indirect Effects

Derivatives of $E[y_1|x_1,x_2,y_2=1]$ wrt x_2

```
+------------------------------------------+
| Partial derivatives of E[y1|y2=1] with   |
| respect to the vector of characteristics.|
| They are computed at the means of the Xs.|
| Effect shown is total of 4 parts above.  |
| Estimate of E[y1|y2=1] = .819898         |
| Observations used for means are All Obs. |
| These are the indirect marginal effects. |
+------------------------------------------+
```

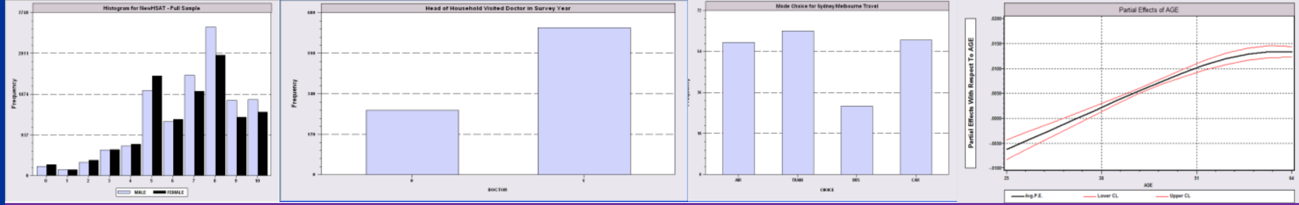
Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
AGE	-.00035034	.697563D-04	-5.022	.0000	43.5256898
FEMALE	-.00835397	.00150062	-5.567	.0000	.47877479
EDUC	.000000	......(Fixed Parameter).......			11.3206310
MARRIED	.000000	......(Fixed Parameter).......			.75861817
WORKING	.000000	......(Fixed Parameter).......			.67704750
HHNINC	.00216510	.00374879	.578	.5636	.35208362
HHKIDS	.00034768	.00164160	.212	.8323	.40273000



Marginal Effects: Total Effects
Sum of Two Derivative Vectors

```
+------------------------------------------+
| Partial derivatives of E[y1|y2=1] with    |
| respect to the vector of characteristics. |
| They are computed at the means of the Xs. |
| Effect shown is total of 4 parts above.   |
| Estimate of E[y1|y2=1] = .819898          |
| Observations used for means are All Obs.  |
| Total effects reported = direct+indirect. |
+------------------------------------------+
```

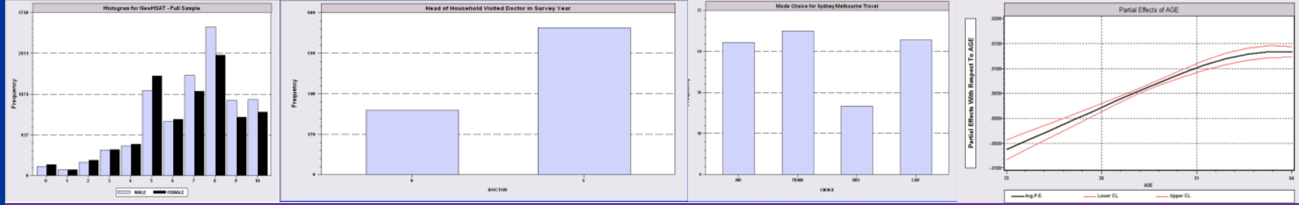
Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
AGE	.00347726	.00022941	15.157	.0000	43.5256898
FEMALE	.08021863	.00535648	14.976	.0000	.47877479
EDUC	-.00392413	.00093911	-4.179	.0000	11.3206310
MARRIED	.00061108	.00506488	.121	.9040	.75861817
WORKING	-.02280671	.00518908	-4.395	.0000	.67704750
HHNINC	.00216510	.00374879	.578	.5636	.35208362
HHKIDS	.00034768	.00164160	.212	.8323	.40273000



Marginal Effects: Dummy Variables Using Differences of Probabilities

Analysis of dummy variables in the model. The effects are computed using E[y1|y2=1,d=1] - E[y1|y2=1,d=0] where d is the variable. Variances use the delta method. The effect accounts for all appearances of the variable in the model.

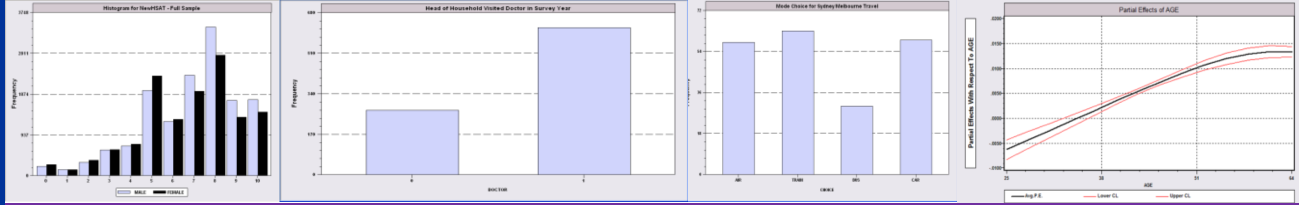
Variable	Effect	Standard error	t ratio	(deriv)
FEMALE	.079694	.005290	15.065	(.080219)
MARRIED	.000611	.005070	.121	(.000511)
WORKING	-.022485	.005044	-4.457	(-.022807)
HHKIDS	.000348	.001641	.212	(.000348)



Average Partial Effects

Partial Effects for Bivariate Probit E[y1|y2=1] function
Partial Effects Averaged Over Observations
* ==> Partial Effect for a Binary Variable

(Delta method)	Partial Effect	Standard Error	|t|	95% Confidence Interval	
AGE	.00346	.00023	14.84	.00300	.00392
* FEMALE	.07952	.00525	15.14	.06923	.08981

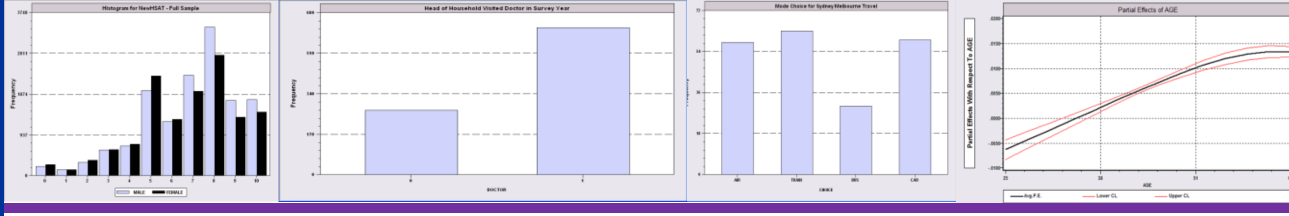


Model Simulation

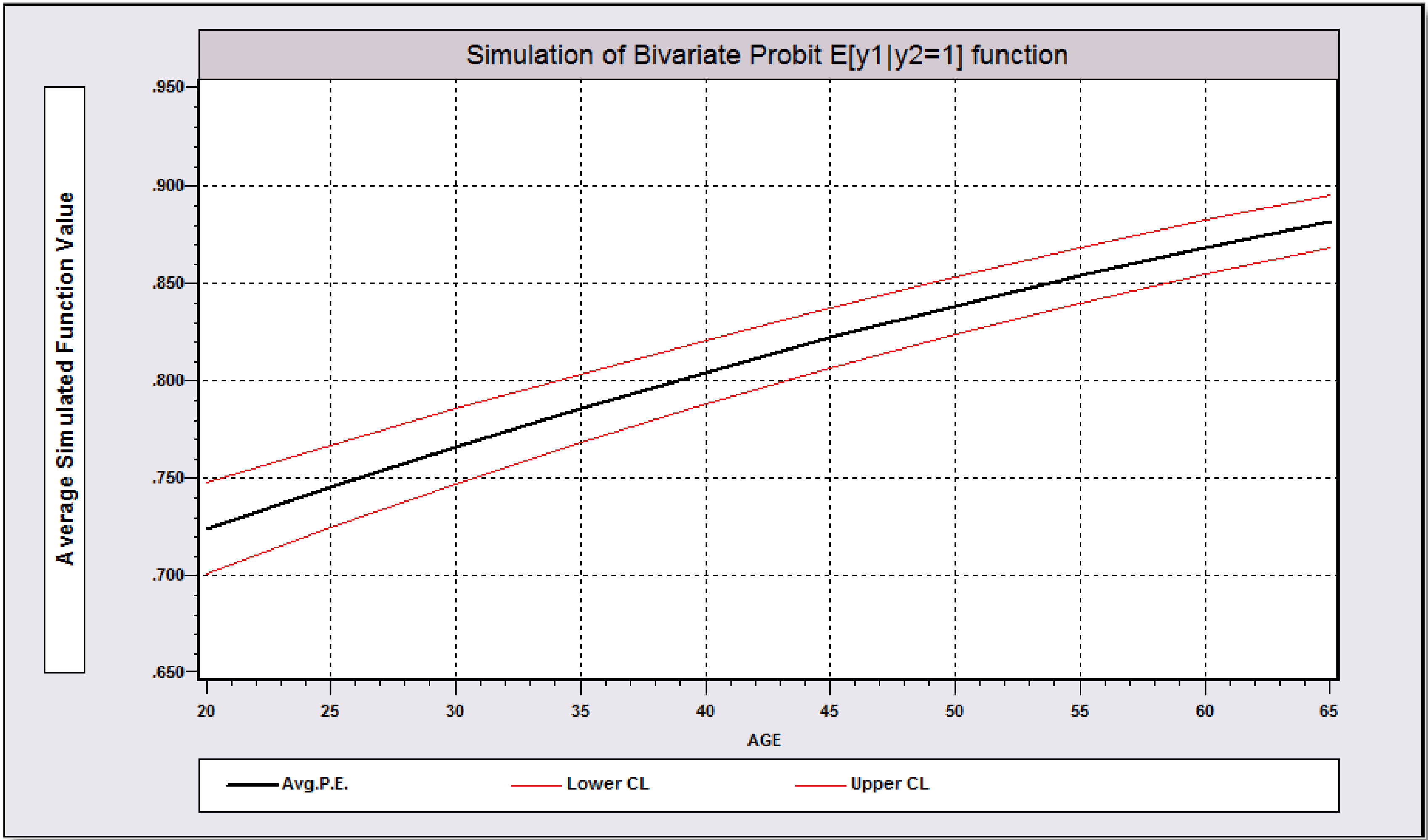
Model Simulation Analysis for Bivariate Probit E[y1|y2=1] function

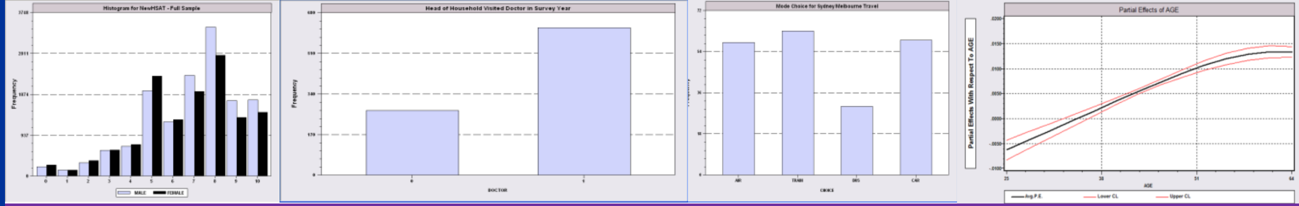
Simulations are computed by average over sample observations

User Function (Delta method)	Function Value	Standard Error	|t|	95% Confidence Interval	
Avrg. Function	.81236	.00801	101.36	.79665	.82807
AGE = 20.00	.72356	.01181	61.27	.70041	.74670
AGE = 25.00	.74491	.01074	69.35	.72386	.76597
AGE = 30.00	.76538	.00981	78.01	.74615	.78461
AGE = 35.00	.78490	.00903	86.94	.76721	.80260
AGE = 40.00	.80345	.00839	95.75	.78700	.81990
AGE = 45.00	.82100	.00789	104.02	.80553	.83647
AGE = 50.00	.83753	.00751	111.46	.82280	.85226
AGE = 55.00	.85303	.00723	117.93	.83886	.86721
AGE = 60.00	.86752	.00702	123.54	.85375	.88128
AGE = 65.00	.88099	.00685	128.52	.86755	.89442



Model Simulation





A Simultaneous Equations Model

Simultaneous Equations Model

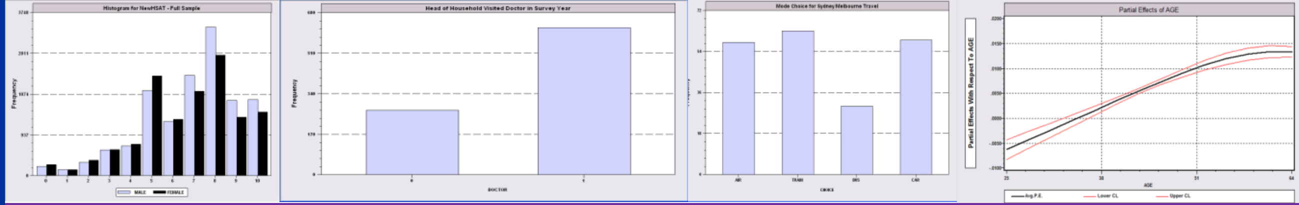
$$y_1^* = \boldsymbol{\beta}_1'\mathbf{x}_1 + \theta_1 y_2 + \varepsilon_1,\ y_1 = 1(y_1^* > 0)$$

$$y_2^* = \boldsymbol{\beta}_2'\mathbf{x}_2 + \theta_2 y_1 + \varepsilon_2,\ y_2 = 1(y_2^* > 0)$$

$$\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right]$$

This model is not identified. Incoherent.

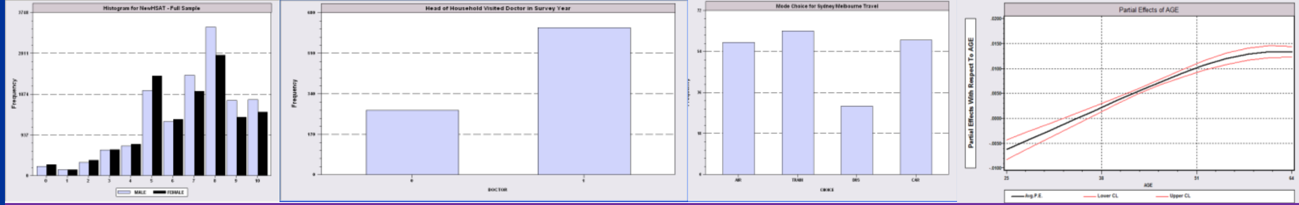
(Not estimable. The computer can compute 'estimates' but they have no meaning.)



Fully Simultaneous "Model"

FIML Estimates of Bivariate Probit Model
Dependent variable DOCHOS
Log likelihood function -20318.69455

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	Index equation for DOCTOR				
Constant	-.46741***	.06726	-6.949	.0000	
AGE	.01124***	.00084	13.353	.0000	43.5257
FEMALE	.27070***	.01961	13.807	.0000	.47877
EDUC	-.00025	.00376	-.067	.9463	11.3206
MARRIED	-.00212	.02114	-.100	.9201	.75862
WORKING	-.00362	.02212	-.164	.8701	.67705
HOSPITAL	2.04295***	.30031	6.803	.0000	.08765
	Index equation for HOSPITAL				
Constant	-1.58437***	.08367	-18.936	.0000	
AGE	-.01115***	.00165	-6.755	.0000	43.5257
FEMALE	-.26881***	.03966	-6.778	.0000	.47877
HHNINC	.00421	.08006	.053	.9581	.35208
HHKIDS	-.00050	.03559	-.014	.9888	.40273
DOCTOR	2.04479***	.09133	22.389	.0000	.62911
	Disturbance correlation				
RHO(1,2)	-.99996***	.00048	********	.0000	



A Recursive Simultaneous Equations Model

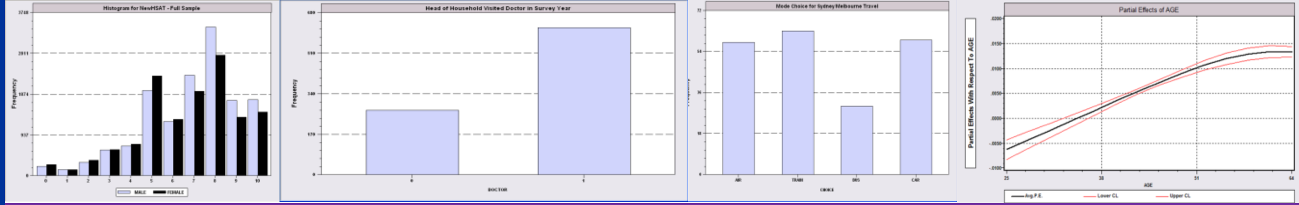
Recursive Simultaneous Equations Model

$$y_1^* = \boldsymbol{\beta}_1'\mathbf{x}_1 + \varepsilon_1, \; y_1 = 1(y_1^* > 0)$$
$$y_2^* = \boldsymbol{\beta}_2'\mathbf{x}_2 + \theta_2 y_1 + \varepsilon_2, \; y_2 = 1(y_2^* > 0)$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$$

This model is identified. It can be consistently and efficiently estimated by full information maximum likelihood. Treated as a bivariate probit model, ignoring the simultaneity.

Bivariate ; Lhs = y1,y2 ; Rh1=…,y2 ; Rh2 = … $



Application: Gender Economics at Liberal Arts Colleges

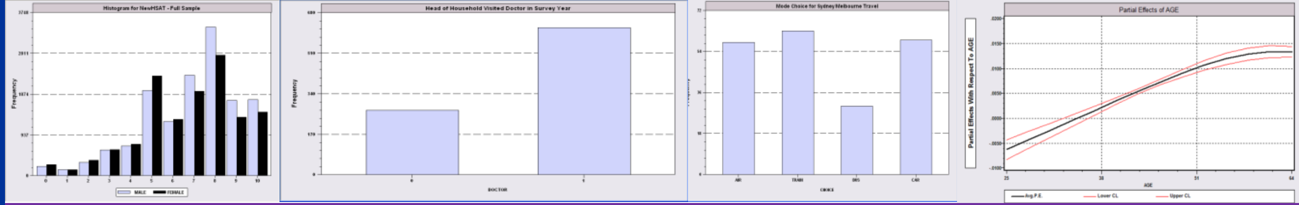
Equations:

$$\text{GndrEcon} = f_1(\text{AcRep}, \text{WomStud}, \text{EconFac}, \text{PctWEcn}, \text{Relig})$$
$$\text{WomStud} = f_2(\text{AcRep}, \text{PctWfac}, \text{Relig}, \text{Sou}, \text{Mid}, \text{Nor}, \text{West}).$$

Variable definitions:

GndrEcon	=	1 if gender economics class is offered, 0 otherwise;
WomStud	=	1 if the college includes a women's studies program, 0 otherwise;
AcRep	=	academic reputation, lower means better reputation, 1 is best;
EconFac	=	number of full-time economics faculty;
PctWEcn	=	percentage of female economics faculty;
Relig	=	1 if the college has a religious affiliation, 0 otherwise;
Sou	=	regional dummy variable, South;
Mid	=	regional dummy variable, Midwest;
Nor	=	regional dummy variable, North; and
West	=	regional dummy variable, West.

Journal of Economic Education, fall, 1998.

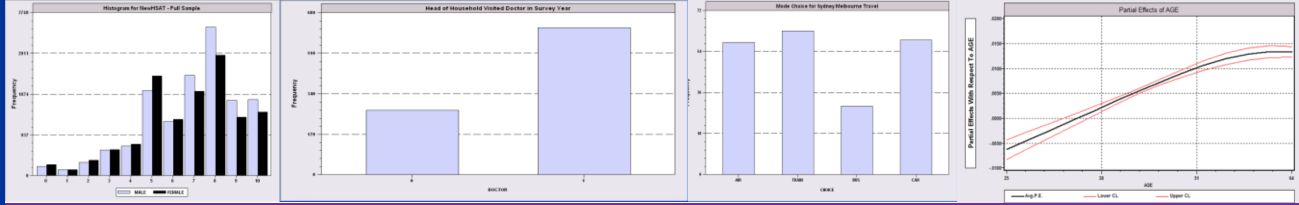


Estimated Recursive Model

Estimated Bivariate Probit Model

Variable	Coefficient	*SE*	*b/SE*	*P*[|Z| > *z*]	Mean of *x*
		Index equation for GNDRECON			
Constant	−1.4176	.8069	−1.757	.0789	
ACREP	−.01143E	0.04081E	−2.802	.0051	119.242
WOMSTUD	1.1095	.5674	1.955	.0505	.4394
ECONFAC	.06730E	.06874E	.979	.3275	6.7424
PCTWECN	2.5392	.9869	2.573	.0101	.2479
RELIG	−.3483	.4984	−.699	.4847	.5758
		Index equation for WOMSTUD			
ACREP	−.01957E	.005524E	−3.542	.0004	119.2424
PCTWFAC	1.9429	.8435	2.303	.0213	.3577
RELIG	−.4494	.3331	−1.349	.1774	.5758
SOU	1.3597	.6594	2.062	.0392	.2424
MID	2.3387	.8104	2.886	.0039	.2727
NOR	1.8867	.8204	2.300	.0215	.3333
WEST	1.8248	.8723	2.092	.0364	.1515
		Disturbance correlation			
RHO (1, 2)	.0000 (Fixed parameter)		Unrestricted value 0.1359		

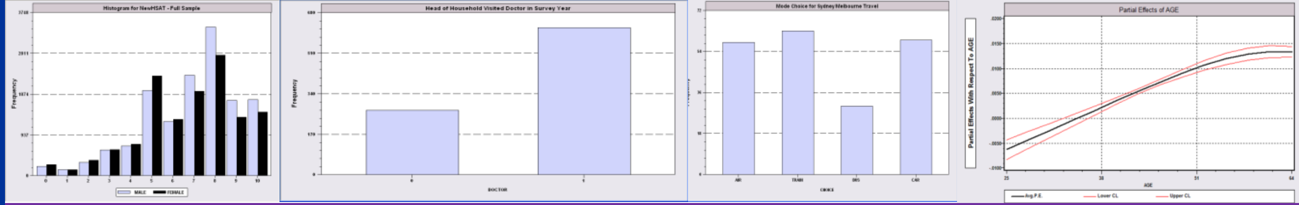
Note: N = 132. Log likelihood function = −85.6458. Log likelihood with ρ not zero = −85.6317.



Estimated Effects: Decomposition

Estimated Marginal Effects

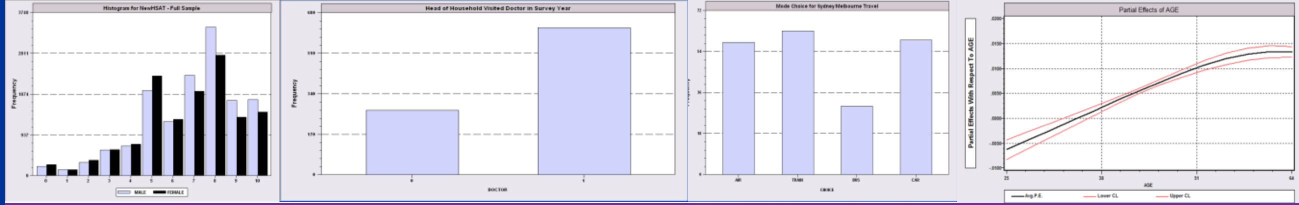
	Gender economic equation					
	Direct	Indirect	Total	*SE*	*t*	Type of variable
AcRep	–0.002022	–0.001453	–0.003476	.00126	–3.087	Continuous
PctWecon	+0.4491		+0.4491	.1568	2.864	Continuous
EconFac	+0.01190		+0.01190	.01292	0.922	Continuous
Relig	–0.07049	–0.03227	–0.1028	.1055	–0.974	Binary
WomStud	+0.1863		+0.1863	.0868	2.146	Endogenous
PctWfac		+0.013951	+0.13951	.08916	1.565	Continuous
Women's studies						
AcRep	–0.00754		–0.00754	.002187	–3.448	Continuous
PctWfac	+0.13789		+0.13789	.01002	13.76	Continuous
Relig	–0.13265		–0.13266	.18803	–0.706	Binary



HEALTH ECONOMICS
Health Econ. **18**: 1091–1108 (2009)
Published online 30 July 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hec.1536

THE EFFECTS OF AN INCENTIVE PROGRAM ON QUALITY OF CARE IN DIABETES MANAGEMENT

ANTHONY SCOTT, STEFANIE SCHURER*, PAUL H. JENSEN and PETER SIVEY
University of Melbourne, Melbourne Institute of Applied Economic and Social Research, Melbourne, Vic., Australia



We specify a bivariate probit model to estimate a set of probabilities depending on whether the GP j in the ith consultation works in a practice that joined the PIP program and provided quality of care. Two unobserved latent variables Y_{ij}^{*} and PIP_{ij}^{*} represent the utility of providing quality of care and the utility of joining the PIP program, respectively

$$Y_{ij}^{*} = \alpha_1 + \beta_1' X_{ij} + \beta_{\text{PIP}} \text{PIP}_{ij} + u_{1ij} \tag{4}$$

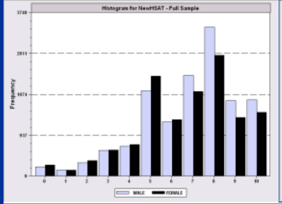
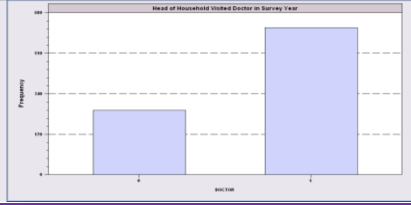
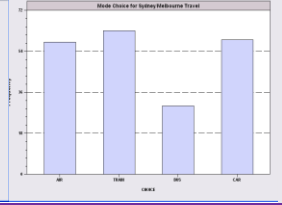
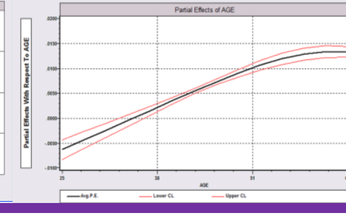
$$\text{PIP}_{ij}^{*} = \alpha_2 + \beta_2' X_{ij} + \pi' I_{ij} + u_{2ij} \tag{5}$$

The observable variables Y_{ij} and PIP_{ij} take the value 1 if the underlying latent variable is greater than 0, respectively, and 0 otherwise. In the main equation, the dependent variable $Y_{ij} = 1$ reflects the observation that GP j has ordered an HbA1c test during the recorded consultation i. In the reduced form equation, the dependent variable $\text{PIP}_{ij} = 1$ reflects the observation that the practice in which the GP j works has joined the PIP program.

We use the marginal probabilities from this bivariate probit, evaluated at the average of the sample, to calculate the marginal effects and the delta method to calculate their standard errors. We have chosen the analytical method to be able to calculate the standard errors of the marginal effect:[4]

$$\text{ME}_{\text{PIP}} = \beta_{\text{PIP}} \times \phi(\bar{X}\hat{\beta}) \tag{6}$$

where ϕ is the standard normal (marginal) density function. In our case, it would not make sense to assess the effect of PIP participation on the joint probability of PIP participation and HbA1c test.

Causal Inference?

$$Y^*_{ij} = \alpha_1 + \beta'_1 X_{ij} + \beta_{\text{PIP}}\text{PIP}_{ij} + u_{1ij}$$

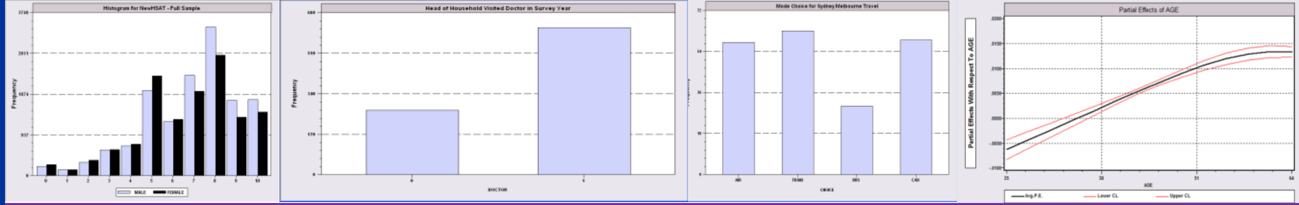
$$\text{PIP}^*_{ij} = \alpha_2 + \beta'_2 X_{ij} + \pi' I_{ij} + u_{2ij}$$

We use the marginal probabilities from this bivariate probit, evaluated at the average of the sample, to calculate the marginal effects and the delta method to calculate their standard errors. We have chosen the analytical method to be able to calculate the standard errors of the marginal effect:[4]

$$\text{ME}_{\text{PIP}} = \beta_{\text{PIP}} \times \phi(\bar{X}\hat{\beta}) \tag{6}$$

where ϕ is the standard normal (marginal) density function. In our case, it would not make sense to assess the effect of PIP participation on the joint probability of PIP participation and HbA1c test.

Causal Inference?
There is no partial (marginal) effect for PIP.
PIP cannot change partially (marginally). It changes because something else changes. (X or I or u_2.)
The calculation of ME_{PIP} does not make sense.



4.2. Exclusion restrictions for model identification

Although the model is formally identified by its non-linear functional form, as long as the full rank condition of the data matrix is ensured (Heckman, 1978; Wilde, 2000), we introduce exclusion restrictions to aid identification of the causal parameter β_{PIP} (Maddala, 1983; Monfardini and Radice, 2008). The row vector I_{ij} captures the variables included in the PIP participation Equation (5), but excluded from the outcome Equation (4).

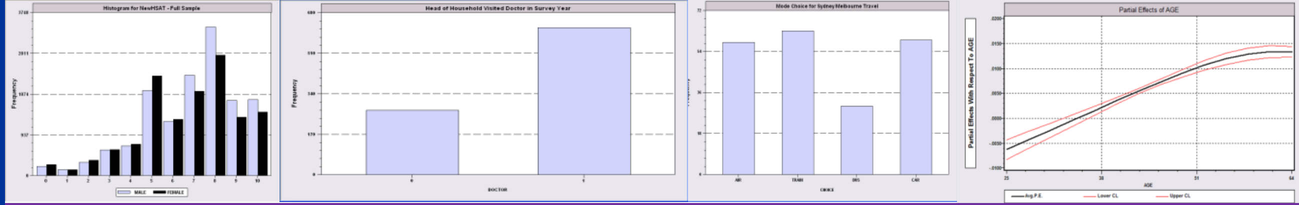
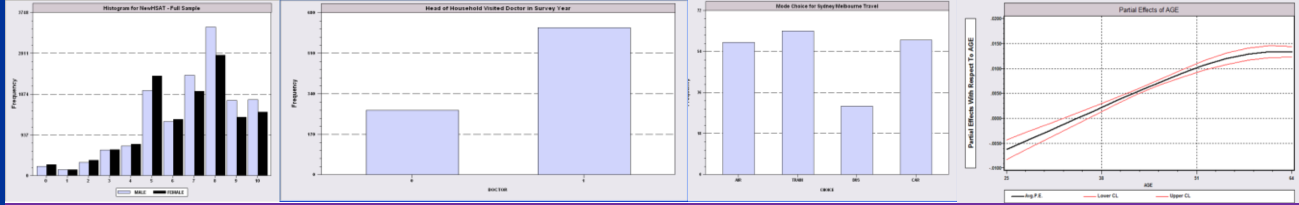


Table II. Marginal effects of PIP status on the probability of conducting an HbA1c test during the encounter for various model specifications; Model (1) is our preferred specification.

	Bivariate probits			Univariate probit
	Preferred model (1)	Sample restriction: type II diabetes (2)	Interaction of Treatment group 1 with ATSI status (3)	Assumption: Exogeneity of Treatment group 1 (4)
Treatment group (Reported: marginal effects)				
Treatment group 1[a]	0.202	0.198	0.197	0.028
	(0.072)***	(0.076)***	(0.071)***	(0.017)*
Treatment group 2[b]	0.028	0.030	0.026	0.028
	(0.017)*	(0.018)*	(0.017)	(0.017)*
Treatment group 1*ATSI status			0.168	
			(0.055)***	
Exclusion restrictions (Reported: coefficients)				
Prevalence of type 2 diabetes per 1000 population	0.013	0.013	0.013	
	(0.005)***	(0.005)***	(0.005)***	
Number of staff employed by Division divided by all practices	0.995	1.000	0.995	
	(0.407)**	(0.420)**	(0.407)**	
Test for strength of exclusion restrictions in participation decision				
$\chi^2(2)$	24.43***	23.881***	24.570***	
Correlation coefficient (ρ)	−0.352	−0.333	−0.352	
	(0.154)**	(0.157)**	(0.150)**	
Test for endogeneity of Treatment group 1 ($\rho = 0$)				
$\chi^2(1)$	7.867**	6.575**	8.075***	
Number of observations	12 187	11 247	12 187	12 187

[a] *Treatment group 1* refers to practices that are accredited and use IT for Internet, prescribing, and medical records.
[b] *Treatment group 2* refers to practices that are accredited and do not use IT for Internet, prescribing, and medical records. The control group includes all practices that are not accredited or that are accredited but do not use IT for Internet, prescribing, and medical records. *10%; **5%; ***1% significance level.

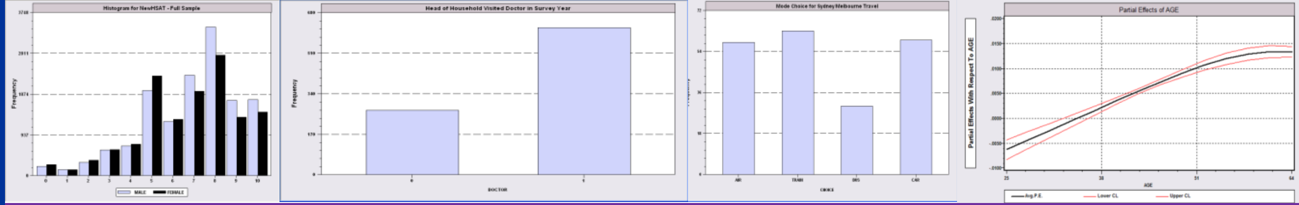


SPRING 2013 VOLUME 47, NUMBER 1 27

SHERRIE L.W. RHINE AND WILLIAM H. GREENE

Factors That Contribute to Becoming Unbanked

The proportion of US families that are unbanked (i.e., have no type of checking or savings account) has steadily declined for more than two decades. Nonetheless, more than nine million families still do not participate in the financial mainstream, and roughly half these unbanked families previously held a traditional bank account. This study uses the 2004 longitudinal Survey of Income Program Participation to examine the dynamic process within which changes in families' circumstances contribute to their becoming unbanked. Our findings suggest that families are significantly more likely to become unbanked when there is a decline in family income, loss of employment, or loss of health insurance coverage. Race and ethnicity, level of education or family income, and marital or housing status are also important determinants of whether families participate in the financial mainstream or not. To our knowledge, this is the first analysis of the dynamic process by which families change bank status.

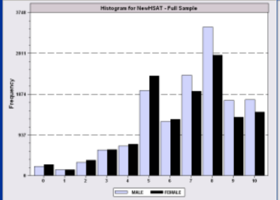
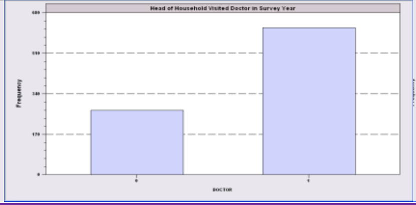
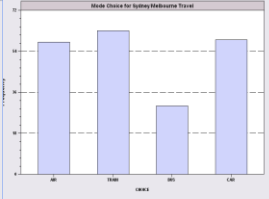
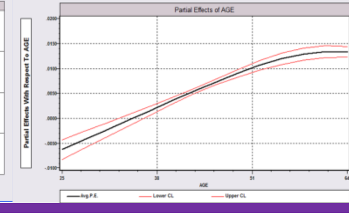


assumption that the coefficients are the same in the two periods. Though the model now accommodates changes that might motivate switching behavior, it does not account for habit persistence, or reluctance to switch that is not otherwise explained. A dynamic model that incorporates both of these ideas is:

$$y_{i0}^* = \beta_0' \mathbf{x}_{i0} + \varepsilon_{i0} + u_i \tag{2a}$$

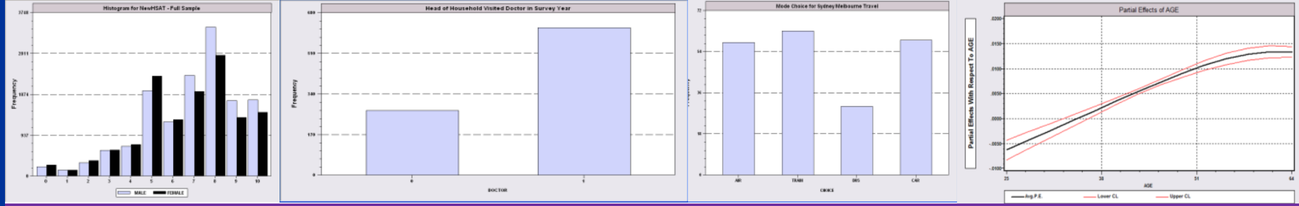
$$y_{i1}^* = \beta_1' \mathbf{x}_{i1} + \alpha'(\Delta \mathbf{x}_i) + \varepsilon_{i1} + \delta y_{i0} + u_i \tag{2b}$$

Although cast as a "panel data" (random effects) model, with two periods observed, this is mathematically, if not logically equivalent to a "recursive bivariate probit model." (The familiar recursive bivariate probit model would be atemporal.) We fit the model using full information maximum likelihood (Greene 2012, 745–756).

Covariates	Partial Effect
Age groups	
AGE45	0.0001
AGE64	0.016
AGE65	−0.028
Gender	
MALE	0.014
Marital status	
MARRIED	−0.079
Children in family	
NUMKIDS	0.008
Race/ethnicity	
BLACK	0.127
HISPANIC	0.115
ASIAN_OTHER	0.047
Education	
EDUC12	−0.060
SOME COLLEGE	−0.131
COLLEGE	−0.140
GRADSCHOOL	−0.149
Work status	
FULLTIME	−0.025
PARTTIME	−0.028
WORK VARIES	−0.014
Family income measure	
FAMINC	−0.002
Citizenship	
US CITIZEN	−0.007
Housing	
OWNER	−0.052
Geographic location	
WEST	−0.080
NORTHEAST	−0.060
MIDWEST	−0.041
Change factors	
MAR_NOMAR	−0.005
LOSTINC1	0.006
LOSTINC2	0.015
LOSTINC3	0.032
LOSTEMP	0.019
LOSTINS	0.061
OWN_TO_RENT	−0.018
NOPOV_POV	0.001

18. Table 3 contains the estimated partial effects for the specific variables discussed. The probability of interest in our study is $\text{Prob}(y_{i1} = 0 \mid y_{i0} = 1) = \text{Prob}(y_{i1} = 0, y_{i0} = 1)/\text{Prob}(y_{i0} = 1)$ $= \Phi_2(\beta_0' x_{i0}, \beta_1' x_{i1} + \alpha'(\Delta x_i) + \delta, -\rho)/\ \Phi_2(\beta_0' x_{i0})$. Average partial effects are obtained by averaging over the sample the derivatives of this expression with respect to the variables in question. The calculations are done using the PARTIALS procedure in NLOGIT Version 5.0.



A Sample Selection Model

Sample Selection Model

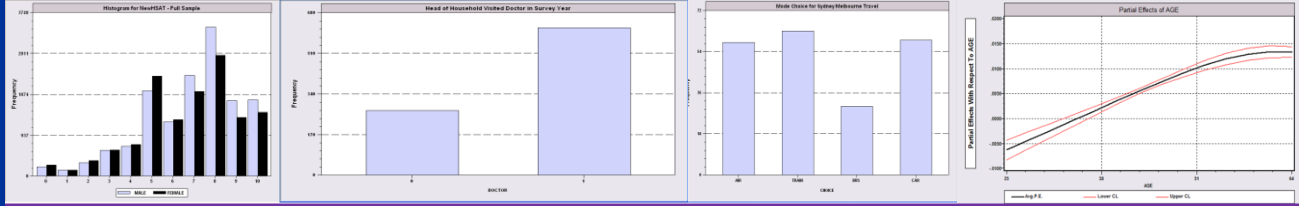
$$y_1^* = \boldsymbol{\beta}_1'\mathbf{x}_1 + \varepsilon_1,\; y_1 = 1(y_1^* > 0)$$

$$y_2^* = \boldsymbol{\beta}_2'\mathbf{x}_2 + \varepsilon_2,\; y_2 = 1(y_2^* > 0)$$

$$\begin{pmatrix}\varepsilon_1\\ \varepsilon_2\end{pmatrix} \sim N\left[\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right]$$

y_1 is only observed when $y_2 = 1$.

$$\begin{aligned} f(y_1,y_2) &= \text{Prob}[y_1 = 1 \mid y_2 = 1] * \text{Prob}[y_2 = 1] \quad (y_1 = 1, y_2 = 1)\\ &= \text{Prob}[y_1 = 0 \mid y_2 = 1] * \text{Prob}[y_2 = 1] \quad (y_1 = 0, y_2 = 1)\\ &= \text{Prob}[y_2 = 0] \quad (y_2 = 0)\end{aligned}$$



Sample Selection Model: Estimation

$$
\begin{aligned}
f(y_1,y_2) &= \text{Prob}[y_1 = 1 | y_2 = 1] * \text{Prob}[y_2 = 1] \quad (y_1 = 1, y_2 = 1) \\
&= \text{Prob}[y_1 = 0 | y_2 = 1] * \text{Prob}[y_2 = 1] \quad (y_1 = 0, y_2 = 1) \\
&= \text{Prob}[y_2 = 0] \quad (y_2 = 0)
\end{aligned}
$$

Terms in the log likelihood:

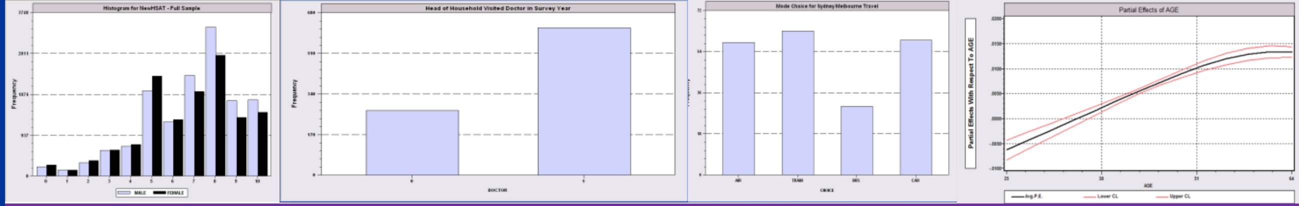
$(y_1 = 1, y_2 = 1)$ $\Phi_2(\boldsymbol{\beta}_1'\mathbf{x}_{i1}, \boldsymbol{\beta}_2'\mathbf{x}_{i2}, \rho)$ (Bivariate normal)

$(y_1 = 0, y_2 = 1)$ $\Phi_2(-\boldsymbol{\beta}_1'\mathbf{x}_{i1}, \boldsymbol{\beta}_2'\mathbf{x}_{i2}, -\rho)$ (Bivariate normal)

$(y_2 = 0)$ $\Phi(-\boldsymbol{\beta}_2'\mathbf{x}_{i2})$ (Univariate normal)

Estimation is by full information maximum likelihood.

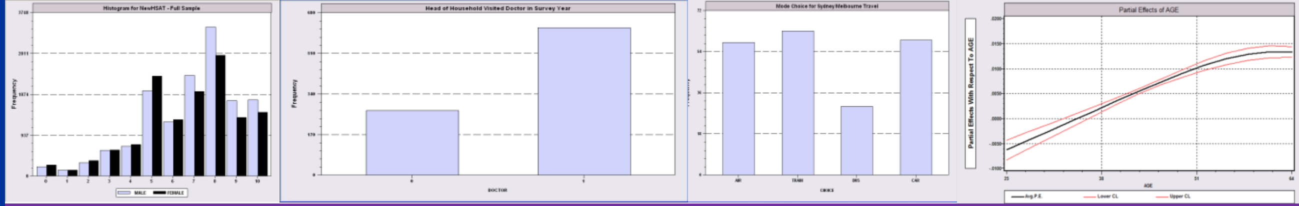
There is no "lambda" variable.



Application: Credit Scoring

- American Express: 1992
- **N = 13,444 Applications**
  - Observed application data
  - Observed acceptance/rejection of application
- **N_1 = 10,499 Cardholders**
  - Observed demographics and economic data
  - Observed default or not in first 12 months

Full Sample is in AmEx.lpj; description shows when imported.



The Multivariate Probit Model

Multiple Equations Analog to SUR Model for M Binary Variables

$$y_1^* = \boldsymbol{\beta}_1'\mathbf{x}_1 + \varepsilon_1, \quad y_1 = 1(y_1^* > 0)$$

$$y_2^* = \boldsymbol{\beta}_2'\mathbf{x}_2 + \varepsilon_2, \quad y_2 = 1(y_2^* > 0)$$

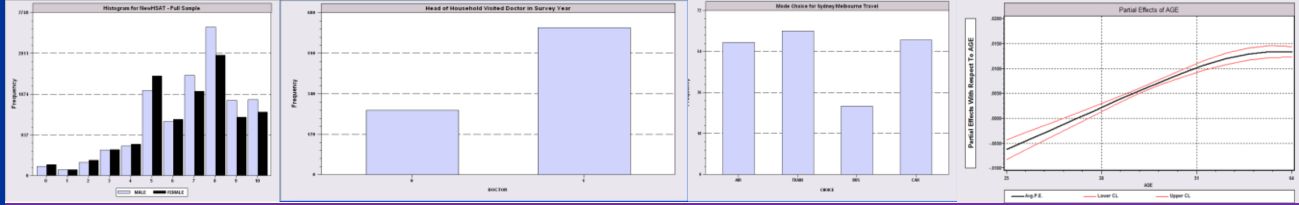
...

$$y_M^* = \boldsymbol{\beta}_M'\mathbf{x}_M + \varepsilon_M, \; y_M = 1(y_M^* > 0)$$

$$\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_M \end{pmatrix} \sim N_M\left[\begin{pmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1M} \\ \rho_{12} & 1 & \cdots & \rho_{2M} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{1M} & \rho_{2M} & \cdots & 1 \end{pmatrix}\right]$$

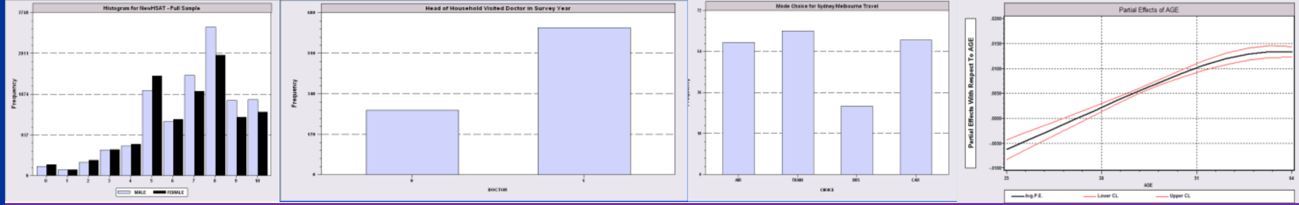
$$\log L = \sum_{i=1}^{N} \log \Phi_M[q_{i1}\boldsymbol{\beta}_1'\mathbf{x}_{i1}, q_{i2}\boldsymbol{\beta}_2'\mathbf{x}_{i2}, \ldots, q_{iM}\boldsymbol{\beta}_M'\mathbf{x}_{iM} \mid \boldsymbol{\Sigma}^*]$$

$$\boldsymbol{\Sigma}_{\mathbf{mn}}^* = 1 \text{ if } m = n \text{ or } q_{im}q_{in}\rho_{mn} \text{ if not.}$$



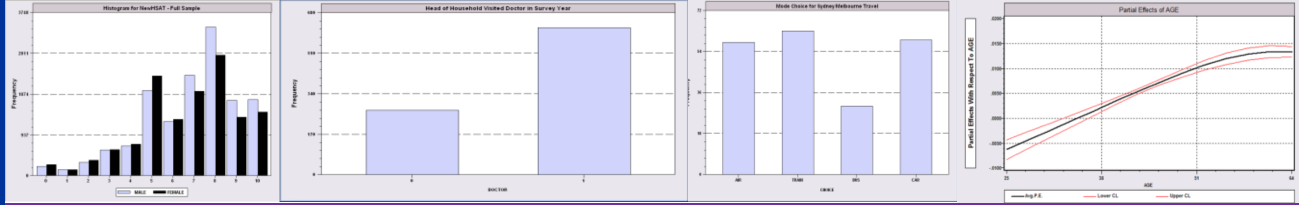
MLE: Simulation

- Estimation of the multivariate probit model requires evaluation of M-order Integrals
- The general case is usually handled with the GHK simulator. Much current research focuses on efficiency (speed) gains in this computation.
- The "Panel Probit Model" is a special case.
  - (Bertschek-Lechner, JE, 1999) – Construct a GMM estimator using only first order integrals of the univariate normal CDF
  - (Greene, Emp.Econ, 2003) – Estimate the integrals with simulation (GHK) anyway.



Multivariate Probit Model: 3 equations.
Dependent variable MVProbit
Log likelihood function -4751.09039

Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X	
	Index function for DOCTOR					
Constant	-.35527**	.16715	-2.125	.0335		[-0.29987 .16195]
AGE	.01664***	.00194	8.565	.0000	43.9959	[0.01644 .00193]
FEMALE	.30931***	.04812	6.427	.0000	.47935	[0.30643 .04767]
EDUC	-.01566	.01024	-1.530	.1261	11.0909	[-0.01936 .00962]
MARRIED	-.04487	.05112	-.878	.3801	.78911	[-0.04423 .05139]
WORKING	-.14712***	.05075	-2.899	.0037	.63345	[-0.15390 .05054]
	Index function for HOSPITAL					
Constant	-1.61787***	.15729	-10.286	.0000		[-1.58276 .16119]
AGE	.00717**	.00283	2.536	.0112	43.9959	[0.00662 .00288]
FEMALE	-.00039	.05995	-.007	.9948	.47935	[-0.00407 .05991]
HHNINC	-.41050	.25147	-1.632	.1026	.29688	[-0.41080 .22891]
HHKIDS	-.01547	.06551	-.236	.8134	.44915	[-0.03688 .06615]
	Index function for PUBLIC					
Constant	1.51314***	.18608	8.132	.0000		[1.53542 .17060]
AGE	.00661**	.00289	2.287	.0222	43.9959	[0.00646 .00268]
HSAT	-.06844***	.01385	-4.941	.0000	6.90062	[-0.07069 .01266]
MARRIED	-.00859	.06892	-.125	.9008	.78911	[-.00813 .06908]
	Correlation coefficients					
R(01,02)	.28381***	.03833	7.404	.0000		[was 0.29611]
R(01,03)	.03509	.03768	.931	.3517		
R(02,03)	-.04100	.04831	-.849	.3960		



Marginal Effects

- There are M equations: “Effect of what on what?”
- NLOGIT computes E[y1|all other ys, all xs]
- Marginal effects are derivatives of this with respect to all xs. (EXTREMELY MESSY)
- Standard errors are estimated with bootstrapping.