Estimating the ATE of an endogenously assigned treatment from a sample with endogenous selection by regression adjustment using an extended regression models

David M. Drukker

Executive Director of Econometrics
Stata

2018 Italian Stata Users Group meeting
15 November 2018

- Fictional data on wellness program from large company

```
. use wprogram2
. describe
Contains data from wprogram2.dta
  obs:         3,000
 vars:             8                          28 Jul 2017 07:13
 size:        96,000
--------------------------------------------------------------------------------
              storage   display    value
variable name   type    format     label      variable label
--------------------------------------------------------------------------------
wchange         float   %9.0g      changel    Weight change level
age             float   %9.0g                 Years over 50
over            float   %9.0g                 Overweight (tens of pounds)
phealth         float   %9.0g                 Prior health score
prog            float   %9.0g      yesno      Participate in wellness program
wtprog          float   %9.0g      yesno      Offered work time to participate
                                                in program
wtsamp          float   %9.0g                 Offered work time to participate
                                                in sample
insamp          float   %9.0g                 In sample: attended initial and
                                                final weigh in
--------------------------------------------------------------------------------
Sorted by:
```

- Three levels of `wchange`

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program No	Yes	Total
Loss	154	960	1,114
No change	251	299	550
Gain	184	36	220
Total	589	1,295	1,884

- Three levels of `wchange`

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program No	Yes	Total
Loss	154	960	1,114
No change	251	299	550
Gain	184	36	220
Total	589	1,295	1,884

- Data are observational

Dealing with observational data

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program No	Yes	Total
Loss	154	960	1,114
No change	251	299	550
Gain	184	36	220
Total	589	1,295	1,884

- Table does not account for
  - how observed covariates that affect program participation also affect the potential outcome variables
    - Assume the treatment is as good as random after conditioning on covariates
    - Conditional mean independence
    - Exogenous treatment assignment
    - `teffects`

Dealing with observational data

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program No	Yes	Total
Loss	154	960	1,114
No change	251	299	550
Gain	184	36	220
Total	589	1,295	1,884

- Table does not account for
  - how observed unobserved error that affect program participation also affect the potential outcome variables
    - Endogenous treatment assignment
    - `eteffects` and `etregress` for continuous outcomes
    - `etpoisson` for count outcomes
    - Need Stata command for ordinal outcome

Dealing with observational data

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program No	Yes	Total
Loss	154	960	1,114
No change	251	299	550
Gain	184	36	220
Total	589	1,295	1,884

- Table does not account for
  - the possibility that unobserved errors in the process that caused some of 3,000 individuals not to show for the final weigh in may also affect the potential outcome variables
    - Endogenous loss to follow up
    - Endogenous sample selection

Ordinal Potential outcomes

- Because the outcome `wchange` is ordinal, there are really three binary outcomes
  - `wchange`==“Loss”,
  - `wchange`==“No Change”, and
  - `wchange`==“Gain”

Ordinal Potential outcomes

- In the potential outcome framework, there is an outcome for each person when they participate and when the do not participate

Ordinal Potential outcomes

- In the potential outcome framework, there is an outcome for each person when they participate and when the do not participate
- Thus, there are really three binary outcomes for each potential outcome

Participate			Not participate		
$\texttt{wchange}_p$	$==$	"Loss"	$\texttt{wchange}_{np}$	$==$	"Loss"
$\texttt{wchange}_p$	$==$	"No change"	$\texttt{wchange}_{np}$	$==$	"No change"
$\texttt{wchange}_p$	$==$	"Gain"	$\texttt{wchange}_{np}$	$==$	"Gain"

Potential outcome framework

- For each outcome (Loss, No change, and Gain), we only observe one of these two potential outcomes for each individual

Potential outcome framework

- For each outcome (Loss, No change, and Gain), we only observe one of these two potential outcomes for each individual
- We estimate the parameters of a model and use the estimated parameters to predict what each person does in the unobserved potential outcome
    - Regression adjustment

Average treatment effects

- In the case of one outcome, the average treatment effect (ATE) is

$$\mathbf{E}[y_p - y_{np}]$$

Average treatment effects

- In the case of one outcome, the average treatment effect (ATE) is
$$\mathbf{E}[y_p - y_{np}]$$
- As there are three outcomes, there are three ATEs
    - one for "Loss", one for "No Change", and one for "Gain"

Average treatment effects

- In the case of one outcome, the average treatment effect (ATE) is

$$\mathbf{E}[y_p - y_{np}]$$

- As there are three outcomes, there are three ATEs
  - one for "Loss", one for "No Change", and one for "Gain"

$$\begin{aligned}
ATE_{Loss} &= \mathbf{E}[(\texttt{wchange}_p == \text{``}\mathit{Loss}\text{''}) - (\texttt{wchange}_{np} == \text{``}\mathit{Loss}\text{''})] \\
ATE_{Nochange} &= \mathbf{E}[(\texttt{wchange}_p == \text{``}\mathit{No\ change}\text{''}) - \\
&\qquad (\texttt{wchange}_{np} == \text{``}\mathit{No\ change}\text{''})] \\
ATE_{Gain} &= \mathbf{E}[(\texttt{wchange}_p == \text{``}\mathit{Gain}\text{''}) - (\texttt{wchange}_{np} == \text{``}\mathit{Gain}\text{''})]
\end{aligned}$$

Average treatment effects

- I will provide some details about the average treatment effect for "Loss"
- The details for the outcomes of "No change" and "Gain" are analogous

- the average treatment effect (ATE) of the program on the `Loss` outcome ATE_{Loss}

$$ATE_{Loss} = \mathbf{E}[(\texttt{wchange}_p == \text{``}Loss\text{''}) - (\texttt{wchange}_{np} == \text{``}Loss\text{''})]$$

- The first line says that ATE_{Loss} is the mean diffence in the outcomes when everyone participates instead of no one participates

- the average treatment effect (ATE) of the program on the `Loss` outcome ATE_{Loss}

$$
\begin{aligned}
ATE_{Loss} &= \mathbf{E}[(\texttt{wchange}_p == \text{``}Loss\text{''}) - (\texttt{wchange}_{np} == \text{``}Loss\text{''})] \\
&= \mathbf{E}[\texttt{wchange}_p == \text{``}Loss\text{''}] - \mathbf{E}[\texttt{wchange}_{np} == \text{``}Loss\text{''}]
\end{aligned}
$$

- The second line says that the mean of the differences is the difference in the means

- the average treatment effect (ATE) of the program on the `Loss` outcome ATE_{Loss}

$$
\begin{aligned}
ATE_{Loss} &= \mathbf{E}[(\texttt{wchange}_p == \text{"}Loss\text{"}) - (\texttt{wchange}_{np} == \text{"}Loss\text{"})] \\
&= \mathbf{E}[\texttt{wchange}_p == \text{"}Loss\text{"}] - \mathbf{E}[\texttt{wchange}_{np} == \text{"}Loss\text{"}] \\
&= \Pr[\texttt{wchange}_p == \text{"}Loss\text{"}] - \Pr[\texttt{wchange}_{np} == \text{"}Loss\text{"}]
\end{aligned}
$$

- The third line says that because the mean of binary outcome is the probability that the event is true, the ATE_{Loss} is the difference in the probability an individual is in the state of "Loss" when everyone participates instead of no one participates

- I am going to use the ERM comand `eoprobit` to estimate the parameters of $\Pr[\texttt{wchange}_p == \text{“}\textit{Loss}\text{”}|\mathbf{x}]$ and $\Pr[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{”}|\mathbf{x}]$ and

- I am going to use the ERM comand `eoprobit` to estimate the parameters of $\Pr[\texttt{wchange}_p == \text{“}\textit{Loss}\text{”}|\mathbf{x}]$ and $\Pr[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{”}|\mathbf{x}]$ and
- Then I use `margins` or `estat teffects` to estimate

$$
\begin{aligned}
&\mathbf{E}[\Pr[\texttt{wchange}_p == \text{“}\textit{Loss}\text{”}|\mathbf{x}]] - \mathbf{E}[\Pr[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{”}|\mathbf{x}]] \\
&\quad = \Pr[\texttt{wchange}_p == \text{“}\textit{Loss}\text{”}] - \Pr[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{”}] \\
&\quad = \textit{ATE}_{\textit{Loss}}
\end{aligned}
$$

- I am going to use the ERM comand `eoprobit` to estimate the parameters of $\mathrm{Pr}[\texttt{wchange}_p == \text{“}\textit{Loss}\text{''}|\mathbf{x}]$ and $\mathrm{Pr}[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{''}|\mathbf{x}]$ and
- Then I use `margins` or `estat teffects` to estimate

$$
\begin{aligned}
&\mathbf{E}[\mathrm{Pr}[\texttt{wchange}_p == \text{“}\textit{Loss}\text{''}|\mathbf{x}]] - \mathbf{E}[\mathrm{Pr}[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{''}|\mathbf{x}]] \\
&\quad = \mathrm{Pr}[\texttt{wchange}_p == \text{“}\textit{Loss}\text{''}] - \mathrm{Pr}[\texttt{wchange}_{np} == \text{“}\textit{Loss}\text{''}] \\
&\quad = ATE_{Loss}
\end{aligned}
$$

- The ATE_{Loss} is the mean difference in the probability an individual is in the state of “Loss” when everyone participates instead of no one participates

Models for the ordinal outcome

- For exogenous treatment, we do a one-step equivalent to fitting two separate ordinal probit models
  - One fit to partipants
  - Another fit to non partipants

Model for partipants

$$wchange = \begin{cases} \text{“}\mathit{Loss}\text{”} & \text{if} \quad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le cut1_0 \\ \text{“}\mathit{No\ change}\text{”} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le cut2_0 \\ \text{“}\mathit{Gain}\text{”} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which `prog`=0, and

ϵ_0, is standard normal

Model for nonparticipants

$$
wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if} \qquad \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut1_1 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut2_1 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases}
$$

$$
\mathbf{x}\boldsymbol{\beta}_1 = \beta_{1,1}\texttt{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth}
$$

for the observations at which $\texttt{prog}=1$

ϵ_1 is standard normal

$$wchange = \begin{cases} \text{"}\mathit{Loss}\text{"} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut1_0 \\ \text{"}\mathit{No\ change}\text{"} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut2_0 \\ \text{"}\mathit{Gain}\text{"} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which `prog`=0, and

$$wchange = \begin{cases} \text{"}\mathit{Loss}\text{"} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut1_1 \\ \text{"}\mathit{No\ change}\text{"} & \text{if } cut1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut2_1 \\ \text{"}\mathit{Gain}\text{"} & \text{if } cut2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_1 = \beta_{1,1}\texttt{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth}$$

for the observations at which `prog`=1

ϵ_0, and ϵ_1 are normal

$corr(\epsilon_0, \epsilon_1)$ is not identified or estimated

```
. eoprobit wchange age over phealth, extreat(prog) vsquish nolog
Extended ordered probit regression              Number of obs     =      1,884
                                                Wald chi2(6)      =      99.08
Log likelihood = -1434.5465                     Prob > chi2       =     0.0000

------------------------------------------------------------------------------
     wchange |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
  prog#c.age |
         No  |   .2180787   .1464522     1.49   0.136    -.0689623    .5051196
        Yes  |  -.2356064   .1196215    -1.97   0.049    -.4700603   -.0011526
 prog#c.over |
         No  |   .2156394   .0784599     2.75   0.006     .0618609    .3694179
        Yes  |  -.0352986   .0781835    -0.45   0.652    -.1885355    .1179383
       prog# |
   c.phealth |
         No  |  -.0746153   .0844652    -0.88   0.377    -.2401641    .0909334
        Yes  |  -.6229527   .0669733    -9.30   0.000    -.7542181   -.4916874
-------------+----------------------------------------------------------------
/wchange     |
 prog#c.cut1 |
         No  |  -.4960282   .0978731                     -.6878559   -.3042005
        Yes  |   .0712884   .0810525                     -.0875716    .2301484
 prog#c.cut2 |
         No  |    .642945   .0988945                      .4491153    .8367747
        Yes  |   1.421407   .0984319                      1.228484     1.61433
------------------------------------------------------------------------------

. estimates store oprobit
```

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0        : Pr(wchange=0=Loss)
ATE_Pr1        : Pr(wchange=1=No change)
ATE_Pr2        : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.4374574	.0238647	18.33	0.000	.3906834	.4842314
ATE_Pr1						
prog						
(Yes vs No)	-.1688022	.0244607	-6.90	0.000	-.2167443	-.1208601
ATE_Pr2						
prog						
(Yes vs No)	-.2686552	.0198483	-13.54	0.000	-.3075572	-.2297532

- When everyone joins the program instead of when no one participants in the program,

```
. estat teffects
Predictive margins                          Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.4374574	.0238647	18.33	0.000	.3906834	.4842314
ATE_Pr1						
prog						
(Yes vs No)	-.1688022	.0244607	-6.90	0.000	-.2167443	-.1208601
ATE_Pr2						
prog						
(Yes vs No)	-.2686552	.0198483	-13.54	0.000	-.3075572	-.2297532

- When everyone joins the program instead of when no one participants in the program,
    - On average, the probablity of “Loss” goes up by .44

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0      : Pr(wchange=0=Loss)
ATE_Pr1      : Pr(wchange=1=No change)
ATE_Pr2      : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.4374574	.0238647	18.33	0.000	.3906834	.4842314
ATE_Pr1						
prog						
(Yes vs No)	-.1688022	.0244607	-6.90	0.000	-.2167443	-.1208601
ATE_Pr2						
prog						
(Yes vs No)	-.2686552	.0198483	-13.54	0.000	-.3075572	-.2297532

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .44
  - On average, the probablity of "No change" goes down by .17

```
. estat teffects
Predictive margins                                  Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.4374574	.0238647	18.33	0.000	.3906834	.4842314
ATE_Pr1						
prog						
(Yes vs No)	-.1688022	.0244607	-6.90	0.000	-.2167443	-.1208601
ATE_Pr2						
prog						
(Yes vs No)	-.2686552	.0198483	-13.54	0.000	-.3075572	-.2297532

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .44
  - On average, the probablity of "No change" goes down by .17
  - On average, the probablity of "Gain" goes down .27

None lost to follow up

- Some observations on `wchange` are missing
- No observations on covariates are missing
- Can do predictions for all cases

ATE: How (1)

```
. generate prog_original = prog
. replace  prog = 0
(1,700 real changes made)
. predict double pr_loss_0 , outlevel("Loss")
(option pr assumed; predicted probabilities)
. replace  prog = 1
(3,000 real changes made)
. predict double pr_loss_1 , outlevel("Loss")
(option pr assumed; predicted probabilities)
. replace  prog = prog_original
(1,300 real changes made)
. drop prog_original
. mean pr_loss_0 pr_loss_1
Mean estimation                   Number of obs    =      3,000

-------------------------------------------------------------
             |       Mean   Std. Err.     [95% Conf. Interval]
-------------+-----------------------------------------------
   pr_loss_0 |   .2721432   .0009077      .2703634     .273923
   pr_loss_1 |   .7096007   .0020206      .7056388    .7135625
-------------------------------------------------------------
```

ATE: How (2)

```
. estimates restore oprobit
(results oprobit are active now)

. margins prog,                              ///
>         predict(outlevel("Loss"))          ///
>         predict(outlevel("No change"))     ///
>         predict(outlevel("Gain")) noesample

Predictive margins                              Number of obs     =      3,000
Model VCE    : OIM

1._predict   : Pr(wchange==Loss), predict(outlevel("Loss"))
2._predict   : Pr(wchange==No change), predict(outlevel("No change"))
3._predict   : Pr(wchange==Gain), predict(outlevel("Gain"))

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
_predict#prog|
       1#No  |   .2721432   .0191116    14.24   0.000     .2346853    .3096012
      1#Yes  |   .7096007   .0142655    49.74   0.000     .6816407    .7375606
       2#No  |   .4260522   .0203869    20.90   0.000     .3860947    .4660097
      2#Yes  |     .25725   .0133175    19.32   0.000     .2311483    .2833518
       3#No  |   .3018046   .0191367    15.77   0.000     .2642973    .3393118
      3#Yes  |   .0331493   .0055184     6.01   0.000     .0223334    .0439652
------------------------------------------------------------------------------
```

ATE: How (3)

```
. margins r.prog,                           ///
>         predict(outlevel("Loss"))         ///
>         predict(outlevel("No change"))    ///
>         predict(outlevel("Gain"))         ///
>         contrast(nowald)                  ///
>         noesample
Contrasts of predictive margins
Model VCE    : OIM
1._predict   : Pr(wchange==Loss), predict(outlevel("Loss"))
2._predict   : Pr(wchange==No change), predict(outlevel("No change"))
3._predict   : Pr(wchange==Gain), predict(outlevel("Gain"))
```

	Contrast	Delta-method Std. Err.	[95% Conf.	Interval]
prog@_predict				
(Yes vs No) 1	.4374574	.0238486	.390715	.4841999
(Yes vs No) 2	-.1688022	.0243512	-.2165296	-.1210748
(Yes vs No) 3	-.2686552	.0199165	-.3076908	-.2296196

Endogenous Treatment model

The potential-outcome model for an endogenous treatment

- Allows the coefficients to differ for the treated and not-treated state
- Allows the cut offs to differ for the treated and not-treated state
- Allows for distinct (nonzero) correlations between the errors driving treatment assignment and the errors driving the ordinal outcomes for the treated and not-treated states

$$prog = (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)$$

$$prog = (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)$$

$$wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if} \qquad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut1_0 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut2_0 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which `prog`=0, and

$$
\begin{aligned}
\mathit{prog} &= (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0) \\
\mathit{wchange} &= \begin{cases} \text{“}\mathit{Loss}\text{”} & \text{if} \quad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le \mathit{cut}1_0 \\ \text{“}\mathit{No\ change}\text{”} & \text{if } \mathit{cut}1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le \mathit{cut}2_0 \\ \text{“}\mathit{Gain}\text{”} & \text{if } \mathit{cut}2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases} \\
\mathbf{x}\boldsymbol{\beta}_0 &= \beta_{1,0}\text{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth} \\
&\quad \text{for the observations at which } \texttt{prog}{=}0\text{, and} \\
\mathit{wchange} &= \begin{cases} \text{“}\mathit{Loss}\text{”} & \text{if} \quad \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \le \mathit{cut}1_1 \\ \text{“}\mathit{No\ change}\text{”} & \text{if } \mathit{cut}1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \le \mathit{cut}2_1 \\ \text{“}\mathit{Gain}\text{”} & \text{if } \mathit{cut}2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases} \\
\mathbf{x}\boldsymbol{\beta}_1 &= \beta_{1,1}\text{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth} \\
&\quad \text{for the observations at which } \texttt{prog}{=}1
\end{aligned}
$$

$$prog = (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)$$

$$wchange = \begin{cases} \text{``}\textit{Loss}\text{''} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut1_0 \\ \text{``}\textit{No change}\text{''} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut2_0 \\ \text{``}\textit{Gain}\text{''} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\mathsf{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which $\texttt{prog}=0$, and

$$wchange = \begin{cases} \text{``}\textit{Loss}\text{''} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut1_1 \\ \text{``}\textit{No change}\text{''} & \text{if } cut1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut2_1 \\ \text{``}\textit{Gain}\text{''} & \text{if } cut2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_1 = \beta_{1,1}\mathsf{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth}$$

for the observations at which $\texttt{prog}=1$

$\epsilon_0,\ \epsilon_1,$ and η are correlated and joint normal

ρ_0 correlation between ϵ_0 and η

ρ_1 correlation between ϵ_1 and η

Endogenous treatment model

```
. eoprobit wchange age over phealth ,                          ///
>         entreat(prog = age over phealth wtprog, pocorr ) ///
>         vce(robust) vsquish nolog
Extended ordered probit regression                Number of obs     =      1,884
                                                  Wald chi2(6)      =     137.27
Log pseudolikelihood = -2335.2213                 Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
wchange      |
  prog#c.age |
          No |   .4919782   .1357859     3.62   0.000     .2258427    .7581137
         Yes |  -.1111304   .1183412    -0.94   0.348    -.3430749    .1208142
 prog#c.over |
          No |   .4659558   .0789709     5.90   0.000     .3111757    .6207359
         Yes |   .0458895   .0794788     0.58   0.564     -.109886    .2016651
         prog#
   c.phealth |
          No |  -.3162974   .0872579    -3.62   0.000    -.4873198    -.145275
         Yes |  -.6880971   .0713535    -9.64   0.000    -.8279474   -.5482467
-------------+----------------------------------------------------------------
prog         |
         age |  -.9224146   .1057226    -8.72   0.000    -1.129627   -.7152021
        over |  -.9957274   .0675412   -14.74   0.000    -1.128106    -.863349
     phealth |   .7483889   .0604543    12.38   0.000     .6299007    .8668771
      wtprog |   1.718043   .1160706    14.80   0.000     1.490549    1.945537
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
wchange						
prog#c.age						
No	.4919782	.1357859	3.62	0.000	.2258427	.7581137
Yes	-.1111304	.1183412	-0.94	0.348	-.3430749	.1208142
prog#c.over						
No	.4659558	.0789709	5.90	0.000	.3111757	.6207359
Yes	.0458895	.0794788	0.58	0.564	-.109886	.2016651
prog# c.phealth						
No	-.3162974	.0872579	-3.62	0.000	-.4873198	-.145275
Yes	-.6880971	.0713535	-9.64	0.000	-.8279474	-.5482467
prog						
age	-.9224146	.1057226	-8.72	0.000	-1.129627	-.7152021
over	-.9957274	.0675412	-14.74	0.000	-1.128106	-.863349
phealth	.7483889	.0604543	12.38	0.000	.6299007	.8668771
wtprog	1.718043	.1160706	14.80	0.000	1.490549	1.945537
_cons	.3398047	.0690413	4.92	0.000	.2044863	.475123
/wchange						
prog#c.cut1						
No	.1953761	.1544741			-.1073875	.4981397
Yes	-.133868	.0985578			-.3270377	.0593017
prog#c.cut2						
No	1.193014	.111908			.9736779	1.412349
Yes	1.170747	.1289195			.9180695	1.423425
corr(e.prog, e.wchange)						
prog						

Yes	-.1111304	.1183412	-0.94	0.348	-.3430749	.1208142
prog#c.over						
No	.4659558	.0789709	5.90	0.000	.3111757	.6207359
Yes	.0458895	.0794788	0.58	0.564	-.109886	.2016651
prog# c.phealth						
No	-.3162974	.0872579	-3.62	0.000	-.4873198	-.145275
Yes	-.6880971	.0713535	-9.64	0.000	-.8279474	-.5482467
prog						
age	-.9224146	.1057226	-8.72	0.000	-1.129627	-.7152021
over	-.9957274	.0675412	-14.74	0.000	-1.128106	-.863349
phealth	.7483889	.0604543	12.38	0.000	.6299007	.8668771
wtprog	1.718043	.1160706	14.80	0.000	1.490549	1.945537
_cons	.3398047	.0690413	4.92	0.000	.2044863	.475123
/wchange						
prog#c.cut1						
No	.1953761	.1544741			-.1073875	.4981397
Yes	-.133868	.0985578			-.3270377	.0593017
prog#c.cut2						
No	1.193014	.111908			.9736779	1.412349
Yes	1.170747	.1289195			.9180695	1.423425
corr(e.prog, e.wchange)						
prog						
No	-.6325687	.1073524	-5.89	0.000	-.7992197	-.3755982
Yes	-.4199058	.1042067	-4.03	0.000	-.6015292	-.1970056

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0      : Pr(wchange=0=Loss)
ATE_Pr1      : Pr(wchange=1=No change)
ATE_Pr2      : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1082033	.0606482	1.78	0.074	-.0106649	.2270715
ATE_Pr1						
prog (Yes vs No)	-.0066579	.0439074	-0.15	0.879	-.0927147	.079399
ATE_Pr2						
prog (Yes vs No)	-.1015455	.0233349	-4.35	0.000	-.147281	-.0558099

- When everyone joins the program instead of when no one participants in the program,

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0      : Pr(wchange=0=Loss)
ATE_Pr1      : Pr(wchange=1=No change)
ATE_Pr2      : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1082033	.0606482	1.78	0.074	-.0106649	.2270715
ATE_Pr1						
prog (Yes vs No)	-.0066579	.0439074	-0.15	0.879	-.0927147	.079399
ATE_Pr2						
prog (Yes vs No)	-.1015455	.0233349	-4.35	0.000	-.147281	-.0558099

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .1

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1082033	.0606482	1.78	0.074	-.0106649	.2270715
ATE_Pr1						
prog (Yes vs No)	-.0066579	.0439074	-0.15	0.879	-.0927147	.079399
ATE_Pr2						
prog (Yes vs No)	-.1015455	.0233349	-4.35	0.000	-.147281	-.0558099

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .1
  - On average, the probablity of "No change" does not change by much

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.1082033	.0606482	1.78	0.074	-.0106649	.2270715
ATE_Pr1						
prog						
(Yes vs No)	-.0066579	.0439074	-0.15	0.879	-.0927147	.079399
ATE_Pr2						
prog						
(Yes vs No)	-.1015455	.0233349	-4.35	0.000	-.147281	-.0558099

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .1
  - On average, the probablity of "No change" does not change by much
  - On average, the probablity of "Gain" goes down .09

```
. margins r.prog,                                      ///
>         predict(fix(prog) outlevel("Loss"))         ///
>         predict(fix(prog) outlevel("No change")) ///
>         predict(fix(prog) outlevel("Gain"))         ///
>         contrast(nowald) vce(unconditional) noesample

Contrasts of predictive margins

1._predict    : Pr(wchange==Loss), predict(fix(prog) outlevel("Loss"))
2._predict    : Pr(wchange==No change), predict(fix(prog) outlevel("No
                change"))
3._predict    : Pr(wchange==Gain), predict(fix(prog) outlevel("Gain"))

-------------------------------------------------------------------
                |            Unconditional
                |   Contrast   Std. Err.     [95% Conf. Interval]
----------------+--------------------------------------------------
 prog@_predict  |
(Yes vs No) 1   |   .1082033   .0606482     -.0106649    .2270715
(Yes vs No) 2   |  -.0066579   .0439074     -.0927147     .079399
(Yes vs No) 3   |  -.1015455   .0233349      -.147281   -.0558099
-------------------------------------------------------------------
```

- `fix(prog)` gets us the effect of the program that is not contaminated by the selection effect/correlation between ϵ and η that increases the participation among people more likely to lose weight

- `fix(prog)` gets us the effect of the program that is not contaminated by the selection effect/correlation between ϵ and η that increases the participation among people more likely to lose weight
- `predict(fix(prog))` tells `margins` to specify `fix(prog)` to `predict` when computing each predicted probability

- `fix(prog)` gets us the effect of the program that is not contaminated by the selection effect/correlation between ϵ and η that increases the participation among people more likely to lose weight
- `predict(fix(prog))` tells `margins` to specify `fix(prog)` to `predict` when computing each predicted probability

- `fix(prog)` causes the value of `prog` not to affect ϵ, even though they are correlated

- `fix(prog)` gets us the effect of the program that is not contaminated by the selection effect/correlation between ϵ and η that increases the participation among people more likely to lose weight
- `predict(fix(prog))` tells `margins` to specify `fix(prog)` to `predict` when computing each predicted probability

- `fix(prog)` causes the value of `prog` not to affect ϵ, even though they are correlated
    - `fix(prog)` specifies that the part of ϵ that is correlated with `prog` be integrated out

- This type of prediction is sometimes called the structural prediction or an average structural function; see Blundell and Powell (2003), Blundell and Powell (2004), Wooldridge (2005), Wooldridge (2010), and Wooldridge (2014),
- The difference between the mean of the average of the structural predictions when `prog=1` and the mean of the average of the structural predictions when `prog=0` is an average treatment effect (Blundell and Powell (2003) and Wooldridge (2014))

Endogenous sample selection

- Reconsider our fictional weight-loss program

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up
  - If unobservables that affect whether someone is lost to follow up

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up
  - If unobservables that affect whether someone is lost to follow up
    - are independent of the unobservables that affect program participantion

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up
  - If unobservables that affect whether someone is lost to follow up
    - are independent of the unobservables that affect program participantion
    - and they are independent of the unobservables that affect the outcomes with and without the program,

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up
  - If unobservables that affect whether someone is lost to follow up
    - are independent of the unobservables that affect program participantion
    - and they are independent of the unobservables that affect the outcomes with and without the program,
  - the previously discussed estimator consistently estimates the effects

Endogenous sample selection

- Reconsider our fictional weight-loss program
  - Some program participants and some nonparticipants will not show up for the final weigh in
    This is commonly known as lost to follow up
  - If unobservables that affect whether someone is lost to follow up
    - are independent of the unobservables that affect program participantion
    - and they are independent of the unobservables that affect the outcomes with and without the program,
  - the previously discussed estimator consistently estimates the effects
- Any dependence among the unobservables must be modeled

$$insamp = (\mathbf{x}\boldsymbol{\alpha} + \alpha_1 \texttt{wtsamp} + \xi > 0)$$

$$
\begin{aligned}
\textit{insamp} &= (\mathbf{x}\boldsymbol{\alpha} + \alpha_1 \texttt{wtsamp} + \xi > 0) \\
\textit{prog} &= (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)
\end{aligned}
$$

$$\begin{aligned}
\textit{insamp} &= (\mathbf{x}\boldsymbol{\alpha} + \alpha_1\texttt{wtsamp} + \xi > 0) \\
\textit{prog} &= (\mathbf{x}\boldsymbol{\gamma} + \gamma_1\texttt{wtprog} + \eta > 0) \\
\textit{wchange} &= \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if } \qquad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le \textit{cut}1_0 \\ \text{“}\textit{No change}\text{”} & \text{if } \textit{cut}1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \le \textit{cut}2_0 \\ \text{“}\textit{Gain}\text{”} & \text{if } \textit{cut}2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases} \\
\mathbf{x}\boldsymbol{\beta}_0 &= \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}
\end{aligned}$$

for the observations at which `prog`=0, and

$$insamp = (\mathbf{x}\boldsymbol{\alpha} + \alpha_1 \texttt{wtsamp} + \xi > 0)$$

$$prog = (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)$$

$$wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut1_0 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut2_0 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which `prog`=0, and

$$wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if } \quad \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut1_1 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut2_1 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_1 = \beta_{1,1}\texttt{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth}$$

for the observations at which `prog`=1

$$insamp = (\mathbf{x}\boldsymbol{\alpha} + \alpha_1 \texttt{wtsamp} + \xi > 0)$$

$$prog = (\mathbf{x}\boldsymbol{\gamma} + \gamma_1 \texttt{wtprog} + \eta > 0)$$

$$wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if } \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut1_0 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \leq cut2_0 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_0 < \mathbf{x}\boldsymbol{\beta}_0 + \epsilon_0 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_0 = \beta_{1,0}\texttt{age} + \beta_{2,0}\texttt{over} + \beta_{3,0}\texttt{phealth}$$

for the observations at which `prog`=0, and

$$wchange = \begin{cases} \text{“}\textit{Loss}\text{”} & \text{if } \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut1_1 \\ \text{“}\textit{No change}\text{”} & \text{if } cut1_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \leq cut2_1 \\ \text{“}\textit{Gain}\text{”} & \text{if } cut2_1 < \mathbf{x}\boldsymbol{\beta}_1 + \epsilon_1 \end{cases}$$

$$\mathbf{x}\boldsymbol{\beta}_1 = \beta_{1,1}\texttt{age} + \beta_{2,1}\texttt{over} + \beta_{3,1}\texttt{phealth}$$

for the observations at which `prog`=1

$\xi, \epsilon_0,\ \epsilon_1,$ and η are correlated and joint normal

distinct correlations between each treatment error and others

```
. eoprobit wchange age over phealth ,                           ///
>         entreat(prog = age over phealth wtprog, pocorr ) ///
>         select(insamp = age over phealth wtsamp )  ///
>         vce(robust) vsquish nolog
Extended ordered probit regression              Number of obs     =      3,000
                                                      Selected    =      1,884
                                                      Nonselected =      1,116
                                                Wald chi2(6)      =     163.70
Log pseudolikelihood = -4483.9683               Prob > chi2       =     0.0000
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
wchange      |
  prog#c.age |
         No  |   .4174575   .1335097     3.13   0.002     .1557832    .6791318
        Yes  |  -.0779536   .1120819    -0.70   0.487    -.2976301     .141723
 prog#c.over |
         No  |   .5046857   .0836683     6.03   0.000     .3406989    .6686725
        Yes  |   .1930521   .0973183     1.98   0.047     .0023118    .3837924
       prog# |
   c.phealth |
         No  |  -.4250361    .091857    -4.63   0.000    -.6050726   -.2449996
        Yes  |  -.8098627   .0753678   -10.75   0.000    -.9575809   -.6621444
-------------+----------------------------------------------------------------
insamp       |
         age |  -.0231005   .0805424    -0.29   0.774    -.1809607    .1347597
        over |  -.7639994   .0450909   -16.94   0.000     -.852376   -.6756229
     phealth |   .7765721   .0467569    16.61   0.000     .6849303    .8682139
      wtsamp |   2.611108   .2660121     9.82   0.000     2.089734    3.132483
       _cons |   .2832551   .0516926     5.48   0.000     .1819395    .3845707
```

c.phealth						
No	-.4250361	.091857	-4.63	0.000	-.6050726	-.2449996
Yes	-.8098627	.0753678	-10.75	0.000	-.9575809	-.6621444
insamp						
age	-.0231005	.0805424	-0.29	0.774	-.1809607	.1347597
over	-.7639994	.0450909	-16.94	0.000	-.852376	-.6756229
phealth	.7765721	.0467569	16.61	0.000	.6849303	.8682139
wtsamp	2.611108	.2660121	9.82	0.000	2.089734	3.132483
_cons	.2832551	.0516926	5.48	0.000	.1819395	.3845707
prog						
age	-.9371024	.0818803	-11.44	0.000	-1.097585	-.7766199
over	-1.060975	.0492229	-21.55	0.000	-1.15745	-.9645
phealth	.890558	.0494954	17.99	0.000	.7935487	.9875673
wtprog	1.644504	.0731516	22.48	0.000	1.501129	1.787878
_cons	.0153225	.0527572	0.29	0.771	-.0880796	.1187247
/wchange						
prog#c.cut1						
No	-.2754667	.1708586			-.6103433	.05941
Yes	-.4323606	.1401249			-.7070003	-.1577208
prog#c.cut2						
No	.6797857	.1534354			.3790578	.9805137
Yes	.7803365	.2260056			.3373737	1.223299
corr(e.ins~p, e.wchange)						
prog						
No	-.5779184	.1004465	-5.75	0.000	-.7420068	-.3484981
Yes	-.5355424	.1948537	-2.75	0.006	-.81217	-.0623165
corr(e.prog, e.wchange)						
prog						

prog#c.cut1						
No	-.2754667	.1708586			-.6103433	.05941
Yes	-.4323606	.1401249			-.7070003	-.1577208
prog#c.cut2						
No	.6797857	.1534354			.3790578	.9805137
Yes	.7803365	.2260056			.3373737	1.223299
corr(e.ins~p, e.wchange)						
prog						
No	-.5779184	.1004465	-5.75	0.000	-.7420068	-.3484981
Yes	-.5355424	.1948537	-2.75	0.006	-.81217	-.0623165
corr(e.prog, e.wchange)						
prog						
No	-.6031412	.1119322	-5.39	0.000	-.7790275	-.3392526
Yes	-.4940044	.0934446	-5.29	0.000	-.6547774	-.2904625
corr(e.prog, e.insamp)	.4745668	.0298397	15.90	0.000	.4140283	.5309257

- Nonzero correlations between `e.insamp` and `e.wchange` imply endogenous sample selection for outcomes
- Nonzero correlations between `e.prog` and `e.wchange` imply endogenous treatment assignment

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0      : Pr(wchange=0=Loss)
ATE_Pr1      : Pr(wchange=1=No change)
ATE_Pr2      : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1406344	.0785061	1.79	0.073	-.0132346	.2945035
ATE_Pr1						
prog (Yes vs No)	.0210902	.0369635	0.57	0.568	-.0513569	.0935372
ATE_Pr2						
prog (Yes vs No)	-.1617246	.0642328	-2.52	0.012	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog						
(Yes vs No)	.1406344	.0785061	1.79	0.073	-.0132346	.2945035
ATE_Pr1						
prog						
(Yes vs No)	.0210902	.0369635	0.57	0.568	-.0513569	.0935372
ATE_Pr2						
prog						
(Yes vs No)	-.1617246	.0642328	-2.52	0.012	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
    - On average, the probablity of “Loss” goes up by .14

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1406344	.0785061	1.79	0.073	-.0132346	.2945035
ATE_Pr1						
prog (Yes vs No)	.0210902	.0369635	0.57	0.568	-.0513569	.0935372
ATE_Pr2						
prog (Yes vs No)	-.1617246	.0642328	-2.52	0.012	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .14
  - On average, the probablity of "No change" does not change

```
. estat teffects
Predictive margins                              Number of obs     =      3,000
ATE_Pr0       : Pr(wchange=0=Loss)
ATE_Pr1       : Pr(wchange=1=No change)
ATE_Pr2       : Pr(wchange=2=Gain)
```

	Margin	Unconditional Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE_Pr0						
prog (Yes vs No)	.1406344	.0785061	1.79	0.073	-.0132346	.2945035
ATE_Pr1						
prog (Yes vs No)	.0210902	.0369635	0.57	0.568	-.0513569	.0935372
ATE_Pr2						
prog (Yes vs No)	-.1617246	.0642328	-2.52	0.012	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .14
  - On average, the probablity of "No change" does not change
  - On average, the probablity of "Gain" goes down .16

```
. margins r.prog,                                              ///
>         predict(fix(prog) outlevel("Loss"))      ///
>         predict(fix(prog) outlevel("No change")) ///
>         predict(fix(prog) outlevel("Gain"))      ///
>         contrast(nowald) vce(unconditional) noesample

Contrasts of predictive margins

1._predict    : Pr(wchange==Loss), predict(fix(prog) outlevel("Loss"))
2._predict    : Pr(wchange==No change), predict(fix(prog) outlevel("No
                change"))
3._predict    : Pr(wchange==Gain), predict(fix(prog) outlevel("Gain"))
```

	Contrast	Unconditional Std. Err.	[95% Conf.	Interval]
prog@_predict				
(Yes vs No) 1	.1406344	.0785061	-.0132346	.2945035
(Yes vs No) 2	.0210902	.0369635	-.0513569	.0935372
(Yes vs No) 3	-.1617246	.0642328	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,

```
. margins r.prog,                                             ///
>         predict(fix(prog) outlevel("Loss"))          ///
>         predict(fix(prog) outlevel("No change")) ///
>         predict(fix(prog) outlevel("Gain"))          ///
>         contrast(nowald) vce(unconditional) noesample
Contrasts of predictive margins

1._predict   : Pr(wchange==Loss), predict(fix(prog) outlevel("Loss"))
2._predict   : Pr(wchange==No change), predict(fix(prog) outlevel("No
               change"))
3._predict   : Pr(wchange==Gain), predict(fix(prog) outlevel("Gain"))
```

	Contrast	Unconditional Std. Err.	[95% Conf.	Interval]
prog@_predict				
(Yes vs No) 1	.1406344	.0785061	-.0132346	.2945035
(Yes vs No) 2	.0210902	.0369635	-.0513569	.0935372
(Yes vs No) 3	-.1617246	.0642328	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of “Loss” goes up by .14

```
. margins r.prog,                                       ///
>         predict(fix(prog) outlevel("Loss"))           ///
>         predict(fix(prog) outlevel("No change"))      ///
>         predict(fix(prog) outlevel("Gain"))           ///
>         contrast(nowald) vce(unconditional) noesample

Contrasts of predictive margins

1._predict   : Pr(wchange==Loss), predict(fix(prog) outlevel("Loss"))
2._predict   : Pr(wchange==No change), predict(fix(prog) outlevel("No
               change"))
3._predict   : Pr(wchange==Gain), predict(fix(prog) outlevel("Gain"))
```

	Contrast	Unconditional Std. Err.	[95% Conf.	Interval]
prog@_predict				
(Yes vs No) 1	.1406344	.0785061	-.0132346	.2945035
(Yes vs No) 2	.0210902	.0369635	-.0513569	.0935372
(Yes vs No) 3	-.1617246	.0642328	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .14
  - On average, the probablity of "No change" does not change

```
. margins r.prog,                                    ///
>         predict(fix(prog) outlevel("Loss"))        ///
>         predict(fix(prog) outlevel("No change")) ///
>         predict(fix(prog) outlevel("Gain"))        ///
>         contrast(nowald) vce(unconditional) noesample

Contrasts of predictive margins

1._predict   : Pr(wchange==Loss), predict(fix(prog) outlevel("Loss"))
2._predict   : Pr(wchange==No change), predict(fix(prog) outlevel("No
               change"))
3._predict   : Pr(wchange==Gain), predict(fix(prog) outlevel("Gain"))
```

	Contrast	Unconditional Std. Err.	[95% Conf.	Interval]
prog@_predict				
(Yes vs No) 1	.1406344	.0785061	-.0132346	.2945035
(Yes vs No) 2	.0210902	.0369635	-.0513569	.0935372
(Yes vs No) 3	-.1617246	.0642328	-.2876187	-.0358305

- When everyone joins the program instead of when no one participants in the program,
  - On average, the probablity of "Loss" goes up by .14
  - On average, the probablity of "No change" does not change
  - On average, the probablity of "Gain" goes down .16

More about ERM commands

- The commands `eregress`, `eprobit`, and `eintreg` fit ERMs handle continuous-and-unbounded, binary, and censored/corner outcomes
- Look at

  http://www.stata.com/manuals/erm.pdf

  for more examples and a wealth of details

Blundell, R. W., and J. L. Powell. 2003. Endogeity in nonparametric and semiparametric regression models. In *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, ed. M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, vol. 2, 312–357. Cambridge: Cambridge University Press.

———. 2004. Endogeneity in semiparametric binary response models. *Review of Economic Studies* 71: 655–679.

Wooldridge, J. M. 2005. Unobserved heterogeneity and estimation of average partial effects. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. K. Andrews and J. H. Stock, 27–55. Cambridge, UK: Cambridge: Cambridge University Press.

———. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, Massachusetts: MIT Press.

———. 2014. Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics* 182: 226–234.