

Estimating effects from extended regression models

David M. Drukker

Executive Director of Econometrics
Stata

Canadian Stata User Group Meeting
Ottawa
June 9, 2017

Extended regression models

Extended regression model (ERM) is a Stata term for a class of regression models.

- The outcome can be continuous (linear), probit, ordered probit, or censored (tobit).
- Some of the covariates may be endogenous.
 - The endogenous covariates may be continuous, probit, or ordered probit
- Endogenous sample-selection may be modeled.
- Exogenous or endogenous treatment assignment may be modeled.
- The new-in-Stata-15 commands `eregress`, `eprobit`, `eoprobit`, and `eintreg` fit ERMs

Extended regression models

- Some of the covariates may be endogenous.
 - The endogenous covariates may be continuous, binary, or ordinal.
 - Polynomials of the endogenous covariates are allowed.
 - Interactions among the endogenous covariates and interactions between the endogenous covariates and the exogenous covariates are allowed.

Outline

- I cannot do justice to ERMs in this short talk.
- I discuss examples in which I
 - define some of the terms that I have already used (endogenous covariate and endogenous sample selection)
 - illustrate some command syntax
 - illustrate how to estimate some effects using postestimation commands

- Fictional data on wellness program from large company

```
. use wprogram
```

```
. describe
```

Contains data from wprogram.dta

```
obs:      3,000
```

```
vars:      6
```

```
11 Jun 2017 20:05
```

```
size:     72,000
```

variable name	storage type	display format	value 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

Sorted by:

- Weight change levels and program participation

```
. tabulate wchange prog
```

Weight change level	Participate in wellness program		Total
	No	Yes	
Loss	239	909	1,148
No change	468	605	1,073
Gain	593	186	779
Total	1,300	1,700	3,000

- Program appears to help
- But this data is observational
 - Table does not account for how observed covariates and/or unobserved errors that affect program participation also affect the outcome variable

- If only observed covariates age over and phealth affect program participation and wchange (with or without program), we could use an ordinal probit model

```
. eoprobit wchange prog age over phealth, vsquish nolog
```

```
Extended ordered probit regression          Number of obs      =       3,000
                                           Wald chi2(4)       =       736.09
Log likelihood = -2866.5688                Prob > chi2        =       0.0000
```

wchange	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
prog	-.8668405	.0460018	-18.84	0.000	-.9570023	-.7766787
age	.097322	.0677733	1.44	0.151	-.0355113	.2301552
over	.3433724	.0360858	9.52	0.000	.2726456	.4140992
phealth	-.3983531	.0385081	-10.34	0.000	-.4738276	-.3228786
cut1	-.8871706	.0539205			-.9928528	-.7814885
cut2	.2358913	.0522019			.1335775	.3382051

$$\mathbf{x}\beta = \beta_2\text{age} + \beta_3\text{over} + \beta_4\text{phealth}$$

$$\mathbf{w}\beta = \beta_1\text{prog} + \mathbf{x}\beta$$

$$wchange = \begin{cases} \text{"Loss"} & \text{if } \mathbf{w}\beta + \epsilon \leq cut1 \\ \text{"No change"} & \text{if } cut1 < \mathbf{w}\beta + \epsilon \leq cut2 \\ \text{"Gain"} & \text{if } cut2 < \mathbf{w}\beta + \epsilon \end{cases}$$

$\epsilon \sim N(0, 1)$ yields

$$\Pr(wchange = \text{"Loss"}) = \Phi(cut1 - \mathbf{w}\beta)$$

$$\Pr(wchange = \text{"No change"}) = \Phi(cut2 - \mathbf{w}\beta) - \Phi(cut1 - \mathbf{w}\beta)$$

$$\Pr(wchange = \text{"Gain"}) = 1 - \Phi(cut2 - \mathbf{w}\beta)$$

- I want to estimate the how changing prog=0 to prog=1 changes each of the probabilities

$\Pr(\text{wchange} = \text{"Loss"})$

$\Pr(\text{wchange} = \text{"No change"})$

$\Pr(\text{wchange} = \text{"Gain"})$

When I type

`eoprobit wchange prog age over phealth, vsquish nolog`

I am assuming that `prog` is independent of ϵ in

$$wchange = \begin{cases} \text{"Loss"} & \text{if } \beta_1 \text{prog} + \mathbf{x}\boldsymbol{\beta} + \epsilon \leq cut1 \\ \text{"No change"} & \text{if } cut1 < \beta_1 \text{prog} + \mathbf{x}\boldsymbol{\beta} + \epsilon \leq cut2 \\ \text{"Gain"} & \text{if } cut2 < \beta_1 \text{prog} + \mathbf{x}\boldsymbol{\beta} + \epsilon \end{cases}$$

In other words, I am assuming that `prog` is exogenous

If `prog` is not independent of ϵ , `prog` is endogenous

If prog is endogenous, I must model the dependence. Consider

$$wchange = \begin{cases} \text{"Loss"} & \text{if } \beta_1 prog + \mathbf{x}\beta + \epsilon \leq cut1 \\ \text{"No change"} & \text{if } cut1 < \beta_1 prog + \mathbf{x}\beta + \epsilon \leq cut2 \\ \text{"Gain"} & \text{if } cut2 < \beta_1 prog + \mathbf{x}\beta + \epsilon \end{cases}$$

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

ϵ and η are joint normal

$$\mathbf{x}\gamma = \gamma_2 age + \gamma_3 over + \gamma_4 phealth$$

Fit by: eoprobit wchange age over phealth ,
endog(prog = age over phealth wtime, probit)

```

. eoprobit wchange age over phealth , ///
> endog(prog = age over phealth wtprog, probit) ///
> vsquish nolog
Extended ordered probit regression      Number of obs      =      3,000
                                         Wald chi2(4)       =      409.97
Log likelihood = -4401.0952             Prob > chi2        =      0.0000

```

	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
age	.2155906	.0705048	3.06	0.002	.0774037	.3537776
over	.4349946	.0387185	11.23	0.000	.3591078	.5108814
phealth	-.4933361	.0411866	-11.98	0.000	-.5740603	-.412612
prog						
Yes	-.3624996	.1031408	-3.51	0.000	-.5646519	-.1603473
prog						
age	-.9341234	.0840002	-11.12	0.000	-1.098761	-.7694861
over	-1.058621	.0514252	-20.59	0.000	-1.159412	-.9578294
phealth	.9001108	.0504804	17.83	0.000	.801171	.9990507
wtprog	1.631615	.0780834	20.90	0.000	1.478574	1.784656
_cons	.0090842	.0535434	0.17	0.865	-.095859	.1140274
/wchange						
cut1	-.5897304	.0781626			-.7429264	-.4365345
cut2	.5029323	.068292			.3690825	.6367821
corr(e.prog, e.wchange)	-.3478179	.0604422	-5.75	0.000	-.4603282	-.2243109

cut1	-.5897304	.0781626			-.7429264	-.4365345
cut2	.5029323	.068292			.3690825	.6367821
corr(e.prog, e.wchange)	-.3478179	.0604422	-5.75	0.000	-.4603282	-.2243109

- The nonzero correlation between e.prog and e.wchange indicates that prog is endogenous

```

. margins r.prog, ///
>   predict(fix(prog) outlevel(#1)) ///
>   predict(fix(prog) outlevel(#2)) ///
>   predict(fix(prog) outlevel(#3)) ///
>   contrast(nowald)

```

Contrasts of predictive margins

Model VCE : OIM

```

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

```

	Delta-method		
	Contrast	Std. Err.	[95% Conf. Interval]
prog@_predict			
(Yes vs No) 1	.1259899	.0356631	.0560914 .1958883
(Yes vs No) 2	-.0185024	.0055583	-.0293965 -.0076084
(Yes vs No) 3	-.1074874	.0306512	-.1675628 -.0474121

- When everyone joins the program instead of when no one participants in the program,
 - On average, the probability of “Loss” goes up
 - On average, the probability of “No change” goes down by a small amount
 - On average, the probability of “Gain” goes down

- `predict(fix(prog))` tells margins to specify `fix(prog)` to predict when computing each predicted probability
- Holding `prog` fixed at the specific values of zero or one, margins out the part of ϵ that is correlated with `prog` fixed.
 - 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 (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))

Standard errors for population versus sample

- The delta-method standard errors reported by `margins` hold the covariates fixed at their sample values
 - The delta-method standard errors are for a sample-average treatment effect instead of a population-averaged treatment effect
 - The sample-averaged treatment effect is for those individuals that showed up in that run of the treatment
 - The population-averaged treatment effect is for a random draw of individuals from the population
- To get standard errors for the population-average treatment effect, specify `vce(robust)` to the estimation command and specify `vce(unconditional)` to `margins`


```

. quietly eoprobit wchange age over phealth , ///
> endog(prog = age over phealth wtprog, probit) ///
> vce(robust)

. margins r.prog, ///
> predict(fix(prog) outlevel(#1)) ///
> predict(fix(prog) outlevel(#2)) ///
> predict(fix(prog) outlevel(#3)) ///
> contrast(nowald) vce(unconditional)

```

Contrasts of predictive margins

```

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

```

	Unconditional		
	Contrast	Std. Err.	[95% Conf. Interval]
prog@_predict			
(Yes vs No) 1	.1259899	.0349061	.0575753 .1944045
(Yes vs No) 2	-.0185024	.0054389	-.0291624 -.0078424
(Yes vs No) 3	-.1074874	.0300866	-.1664561 -.0485188

```

. matrix b = r(b)

```

Effects sum to zero

- The categories are mutually exclusive, so the effects sum to zero

```
. matrix list b
b[1,3]
      1.prog#      1.prog#      1.prog#
      1._predict 2._predict 3._predict
r1      .12598987  -.01850244  -.10748744
. display b[1,1] + b[1,2] + b[1,3]
1.110e-16
```

Interacting an endogenous variable with other covariates

```
. eoprobit wchange i.prog i.prog#c.(age over phealth) , ///
>      endog(prog = age over phealth wtprog, nomain probit)  ///
>      vce(robust) vsquish nolog
```

Extended ordered probit regression

Log pseudolikelihood = -4394.2174

Number of obs = 3,000
Wald chi2(7) = 433.24
Prob > chi2 = 0.0000

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
prog						
Yes	-.1941662	.1278507	-1.52	0.129	-.444749	.0564165
prog#c.age						
No	.3356884	.1039017	3.23	0.001	.1320449	.539332
Yes	.0966954	.0921944	1.05	0.294	-.0840023	.2773931
prog#c.over						
No	.4306879	.0537107	8.02	0.000	.325417	.5359589
Yes	.4201763	.0496839	8.46	0.000	.3227976	.5175549
prog#						
c.phealth						
No	-.3358392	.0682247	-4.92	0.000	-.4695571	-.2021212
Yes	-.582711	.0476667	-12.22	0.000	-.6761361	-.4892859

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
prog						
Yes	-.1941662	.1278507	-1.52	0.129	-.444749	.0564165
prog#c.age						
No	.3356884	.1039017	3.23	0.001	.1320449	.539332
Yes	.0966954	.0921944	1.05	0.294	-.0840023	.2773931
prog#c.over						
No	.4306879	.0537107	8.02	0.000	.325417	.5359589
Yes	.4201763	.0496839	8.46	0.000	.3227976	.5175549
prog#						
c.phealth						
No	-.3358392	.0682247	-4.92	0.000	-.4695571	-.2021212
Yes	-.582711	.0476667	-12.22	0.000	-.6761361	-.4892859
prog						
age	-.936958	.0828918	-11.30	0.000	-1.099423	-.7744929
over	-1.058691	.0499303	-21.20	0.000	-1.156552	-.9608291
phealth	.8942599	.0498163	17.95	0.000	.7966218	.9918979
wtprog	1.633958	.0759464	21.51	0.000	1.485106	1.782811
_cons	.0119034	.05316	0.22	0.823	-.0922882	.116095
/wchange						
cut1	-.4986848	.0900524			-.6751843	-.3221854
cut2	.5998864	.0815039			.4401418	.759631
corr(e.prog, e.wchange)	-.3355574	.0599945	-5.59	0.000	-.447512	-.2132471

```

. margins r.prog,          ///
>   predict(fix(prog) outlevel(#1)) ///
>   predict(fix(prog) outlevel(#2)) ///
>   predict(fix(prog) outlevel(#3)) ///
>   contrast(nowald) vce(unconditional)

```

Contrasts of predictive margins

```

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

```

	Unconditional			
	Contrast	Std. Err.	[95% Conf. Interval]	
prog@_predict				
(Yes vs No) 1	.1405727	.0362828	.0694597	.2116856
(Yes vs No) 2	-.0276059	.0072535	-.0418225	-.0133893
(Yes vs No) 3	-.1129667	.0310018	-.173729	-.0522044

- When everyone joins the program instead of when no one participants in the program,
 - On average, the probability of “Loss” goes up
 - On average, the probability of “No change” goes down by a small amount
 - On average, the probability of “Gain” goes down

Endogenous treatment model

```
. eoprobit wchange (age over phealth) , ///
>      entreat(prog = age over phealth wtprog ) ///
>      vce(robust) vsquish nolog
Extended ordered probit regression      Number of obs      =      3,000
                                         Wald chi2(6)        =      236.09
Log pseudolikelihood = -4389.0839      Prob > chi2         =      0.0000
```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
prog#c.age						
No	.3318583	.1010243	3.28	0.001	.1338543	.5298624
Yes	.0991993	.0944861	1.05	0.294	-.08599	.2843886
prog#c.over						
No	.4241221	.0527011	8.05	0.000	.3208298	.5274144
Yes	.4310345	.051217	8.42	0.000	.3306509	.5314181
prog#						
c.phealth						
No	-.3323793	.0665871	-4.99	0.000	-.4628875	-.201871
Yes	-.5973977	.0486921	-12.27	0.000	-.6928325	-.501963
prog						
age	-.9365184	.0828734	-11.30	0.000	-1.098947	-.7740895
over	-1.058259	.0499251	-21.20	0.000	-1.15611	-.9604071
phealth	.8938228	.0498157	17.94	0.000	.7961859	.9914598
wtprog	1.632994	.0760053	21.49	0.000	1.484026	1.781961

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
prog#c.age						
No	.3318583	.1010243	3.28	0.001	.1338543	.5298624
Yes	.0991993	.0944861	1.05	0.294	-.08599	.2843886
prog#c.over						
No	.4241221	.0527011	8.05	0.000	.3208298	.5274144
Yes	.4310345	.051217	8.42	0.000	.3306509	.5314181
prog# c.phealth						
No	-.3323793	.0665871	-4.99	0.000	-.4628875	-.201871
Yes	-.5973977	.0486921	-12.27	0.000	-.6928325	-.501963
prog						
age	-.9365184	.0828734	-11.30	0.000	-1.098947	-.7740895
over	-1.058259	.0499251	-21.20	0.000	-1.15611	-.9604071
phealth	.8938228	.0498157	17.94	0.000	.7961859	.9914598
wtprog	1.632994	.0760053	21.49	0.000	1.484026	1.781961
_cons	.0130243	.0531604	0.24	0.806	-.0911683	.1172168
/wchange						
prog#c.cut1						
No	-.438465	.0907758			-.6163823	-.2605477
Yes	-.3395413	.0697649			-.4762781	-.2028046
prog#c.cut2						
No	.5647753	.0811816			.4056623	.7238883
Yes	.849572	.0817844			.6892775	1.009867
corr(e.prog, wchange)	-.3419207	.0610844	-5.60	0.000	-.4556747	-.2171783

prog#c.age	No	.3318583	.1010243	3.28	0.001	.1338543	.5298624
	Yes	.0991993	.0944861	1.05	0.294	-.08599	.2843886
prog#c.over	No	.4241221	.0527011	8.05	0.000	.3208298	.5274144
	Yes	.4310345	.051217	8.42	0.000	.3306509	.5314181
prog# c.phealth	No	-.3323793	.0665871	-4.99	0.000	-.4628875	-.201871
	Yes	-.5973977	.0486921	-12.27	0.000	-.6928325	-.501963
prog	age	-.9365184	.0828734	-11.30	0.000	-1.098947	-.7740895
	over	-1.058259	.0499251	-21.20	0.000	-1.15611	-.9604071
	phealth	.8938228	.0498157	17.94	0.000	.7961859	.9914598
	wtprog	1.632994	.0760053	21.49	0.000	1.484026	1.781961
	_cons	.0130243	.0531604	0.24	0.806	-.0911683	.1172168
/wchange	prog#c.cut1						
	No	-.438465	.0907758			-.6163823	-.2605477
	Yes	-.3395413	.0697649			-.4762781	-.2028046
	prog#c.cut2						
	No	.5647753	.0811816			.4056623	.7238883
	Yes	.849572	.0817844			.6892775	1.009867
corr(e.prog, e.wchange)		-.3419207	.0610844	-5.60	0.000	-.4556747	-.2171783


```

. margins r.prog,          ///
>   predict(fix(prog) outlevel(#1)) ///
>   predict(fix(prog) outlevel(#2)) ///
>   predict(fix(prog) outlevel(#3)) ///
>   contrast(nowald) vce(unconditional)

```

Contrasts of predictive margins

```

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

```

	Unconditional			
	Contrast	Std. Err.	[95% Conf. Interval]	
prog@_predict				
(Yes vs No) 1	.1087061	.038293	.0336531	.1837591
(Yes vs No) 2	.0288781	.0190952	-.0085478	.0663039
(Yes vs No) 3	-.1375842	.0322663	-.200825	-.0743433

- When everyone joins the program instead of when no one participants in the program,
 - On average, the probability of “Loss” goes up
 - On average, the probability of “No change” does not change
 - On average, the probability of “Gain” goes down

```
. 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)
```

	Unconditional Margin	Std. Err.	z	P> z	[95% Conf. Interval]	
ATE_Pr0 prog (Yes vs No)	.1087061	.038293	2.84	0.005	.0336531	.1837591
ATE_Pr1 prog (Yes vs No)	.0288781	.0190952	1.51	0.130	-.0085478	.0663039
ATE_Pr2 prog (Yes vs No)	-.1375842	.0322663	-4.26	0.000	-.200825	-.0743433

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 participation
 - 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

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

$$Pr(wchange == \text{"No change"})$$

$$= \begin{cases} Pr(cut1_0 < \mathbf{x}\beta_0 + \epsilon \leq cut2_0) & \text{if } prog == 0 \\ Pr(cut1_1 < \mathbf{x}\beta_1 + \epsilon \leq cut2_1) & \text{if } prog == 1 \end{cases}$$

(Analogously define probabilities Loss, and Gain)

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

ξ, ϵ and η are joint normal

```
Fit by: eoprobit wchange (age over phealth) ,
        entreat(prog = age over phealth wtprog )
        select(samp = age over phealth wtsamp )
        vce(robust)
```

Data

```
. use wprogram2
```

```
. describe
```

```
Contains data from wprogram2.dta
```

```
obs:      3,000
```

```
vars:      8
```

```
size:     96,000
```

```
11 Jun 2017 20:05
```

variable name	storage type	display format	value 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:
```

```

. eoprobit wchange (age over phealth) ,      ///
>       entreat(prog = age over phealth wtprog ) ///
>       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)   =       180.18
                                             Prob > chi2    =       0.0000

Log pseudolikelihood = -4484.2347

```

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
wchange						
prog#c.age						
No	.3833806	.1306121	2.94	0.003	.1273856	.6393756
Yes	-.0630257	.1084828	-0.58	0.561	-.2756482	.1495967
prog#c.over						
No	.4734046	.0775788	6.10	0.000	.321353	.6254561
Yes	.2110918	.0774768	2.72	0.006	.0592401	.3629436
prog#						
c.phealth						
No	-.3910086	.0839247	-4.66	0.000	-.5554981	-.2265192
Yes	-.8192438	.0675175	-12.13	0.000	-.9515757	-.6869119
insamp						
age	-.0239454	.0805554	-0.30	0.766	-.181831	.1339403
over	-.7639015	.045083	-16.94	0.000	-.8522626	-.6755404
phealth	.7762507	.0467149	16.62	0.000	.6846911	.8678104
wtsamp	2.614852	.2666563	9.81	0.000	2.092215	3.137489
cons	.2834801	.0516434	5.49	0.000	.1822608	.3846994

c.phealth							
No		-.3910086	.0839247	-4.66	0.000	-.5554981	-.2265192
Yes		-.8192438	.0675175	-12.13	0.000	-.9515757	-.6869119
insamp							
age		-.0239454	.0805554	-0.30	0.766	-.181831	.1339403
over		-.7639015	.045083	-16.94	0.000	-.8522626	-.6755404
phealth		.7762507	.0467149	16.62	0.000	.6846911	.8678104
wtsamp		2.614852	.2666563	9.81	0.000	2.092215	3.137489
_cons		.2834801	.0516434	5.49	0.000	.1822608	.3846994
prog							
age		-.9366997	.0818766	-11.44	0.000	-1.097175	-.7762245
over		-1.062399	.0491499	-21.62	0.000	-1.158731	-.9660671
phealth		.8913551	.0494733	18.02	0.000	.7943892	.988321
wtprog		1.645957	.0729847	22.55	0.000	1.502909	1.789004
_cons		.016411	.0527191	0.31	0.756	-.0869166	.1197386
/wchange							
prog#c.cut1							
No		-.3561706	.1278314			-.6067155	-.1056258
Yes		-.4668053	.0986098			-.660077	-.2735335
prog#c.cut2							
No		.6304983	.1309562			.3738289	.8871677
Yes		.721275	.1430156			.4409696	1.00158
corr(e.insamp,							
e.wchange)		-.5691458	.0901455	-6.31	0.000	-.7199754	-.366975
corr(e.prog,							
e.wchange)		-.5310186	.0707529	-7.51	0.000	-.6553935	-.3786047
corr(e.prog,							
e.insamp)		.4757458	.0296034	16.07	0.000	.4156943	.5316674

	1.748337	1.0723317	22.88	0.000	1.802309	1.748337
_cons	.016411	.0527191	0.31	0.756	-.0869166	.1197386
/wchange						
prog#c.cut1						
No	-.3561706	.1278314			-.6067155	-.1056258
Yes	-.4668053	.0986098			-.660077	-.2735335
prog#c.cut2						
No	.6304983	.1309562			.3738289	.8871677
Yes	.721275	.1430156			.4409696	1.00158
corr(e.insamp,						
e.wchange)	-.5691458	.0901455	-6.31	0.000	-.7199754	-.366975
corr(e.prog,						
e.wchange)	-.5310186	.0707529	-7.51	0.000	-.6553935	-.3786047
corr(e.prog,						
e.insamp)	.4757458	.0296034	16.07	0.000	.4156943	.5316674

- Nonzero correlation between e.insamp and e.wchange implies endogenous sample selection
- Nonzero correlation between e.prog and e.wchange implies 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)
```

	Unconditional Margin	Std. Err.	z	P> z	[95% Conf. Interval]	
ATE_Pr0 prog (Yes vs No)	.1552606	.051808	3.00	0.003	.0537189	.2568024
ATE_Pr1 prog (Yes vs No)	.006893	.0300435	0.23	0.819	-.0519913	.0657772
ATE_Pr2 prog (Yes vs No)	-.1621536	.038066	-4.26	0.000	-.2367616	-.0875457

- When everyone joins the program instead of when no one participants in the program,
 - On average, the probability of “Loss” goes up
 - On average, the probability of “No change” does not change
 - On average, the probability of “Gain” goes down

- Blundell, R. W., and J. L. Powell. 2003. Endogeneity 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. Violating ignorability of treatment by controlling for too many factors. *Econometric Theory* 21(5): 1026.
- . 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.