Stata Syntax for Section 11.2, Chapter 11

Models 1.1 and 1.1.1 (Selection on the observables, OLS with and without z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Models 1.1 and 1.1.1 OLS

//u & z correlated, u & v NOT correlated

cd “D:\psa\_e2\Chapter11\data”

capture log close

set more off

log using Model\_11\_&111, replace

clear

program model1\_1

 postfile sim rzu ruv rwu rzv b1 t1 b2 t2 b3 t3 using sim1\_1, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: regress y x1 x2 x3 z ww

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,5)

 mat t1=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b1

 svmat t1

 quietly:regress y x1 x2 x3 ww

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 ww

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,3)

 mat t3=el(e(b),1,3)/sqrt(el(v,1,3))

 svmat b3

 svmat t3

 post sim (rzu) (ruv) (rwu) (rzv) (b1) (t1) (b2) (t2) (b3) (t3)

 }

 postclose sim

end

set seed 1000

model1\_1

use sim1\_1, replace

sum rzu ruv rwu b1 b2

quietly: sum b1

display r(sd)^2

quietly: sum b2

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

gen v2= (b2-.5)^2

quietly: sum v2

gen mse2=r(sum)/10000

sum mse1 mse2

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 1.2 (Selection on the observables, PSM)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 1.2

//u & z correlated, u & v NOT correlated - PSM

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_12, replace

clear

program model1\_2

 postfile sim1 rzu ruv rwu rzv ps25 n b1 t1 b2 t2 b3 t3 using psm12, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u ,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

 quietly: sum ps, detail

 gen ps25=r(sd)\*.25

 gen id=\_n

 save original, replace

 quietly: psmatch2 ww, caliper(.086) pscore(ps) logit quietly noreplacement descending

 sort \_id

 g match=id[\_n1]

 g treat=id if \_nn==1

 drop if treat==.

 save t1, replace

 sort match

 keep match

 rename match id

 save id1, replace

 use original, replace

 sort id

 merge id using id1

 drop if \_merge!=3

 append using t1

 save tm1, replace

 quietly: regress y ww x1 x2 x3 z

 gen n=e(N)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,1)

 mat t1=el(e(b),1,1)/sqrt(el(v,1,1))

 svmat b1

 svmat t1

 quietly: regress y x1 x2 x3 ww

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 ww

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,3)

 mat t3=el(e(b),1,3)/sqrt(el(v,1,3))

 svmat b3

 svmat t3

 post sim1 (rzu) (ruv) (rwu) (rzv) (ps25) (n) (b1) (t1) (b2) (t2) (b3) (t3)

 }

 postclose sim1

erase t1.dta

erase id1.dta

erase original.dta

erase tm1.dta

end

set seed 1000

model1\_2

use psm12, replace

sum rzu ruv ps25 b1

quietly: sum b1

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum mse1

program drop model1\_2

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 1.2.1 (Selection on the observables, PSM omitted z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 1.2.1

//u & z correlated, u & v NOT correlated - PSM

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_121, replace

clear

program model121

 postfile sim1 ruv rzu rzv ps25 n b1 t1 b2 t2 b3 t3 using psm121, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rzu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3

 predict ps

 quietly: sum ps, detail

 gen ps25=r(sd)\*.25

 gen id=\_n

 save original, replace

 quietly: psmatch2 ww, caliper(.06) pscore(ps) logit quietly noreplacement descending

 sort \_id

 g match=id[\_n1]

 g treat=id if \_nn==1

 drop if treat==.

 save t1, replace

 sort match

 keep match

 rename match id

 save id1, replace

 use original, replace

 sort id

 merge id using id1

 drop if \_merge!=3

 append using t1

 save tm1, replace

 quietly: regress y ww x1 x2 x3 z

 gen n=e(N)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,1)

 mat t1=el(e(b),1,1)/sqrt(el(v,1,1))

 svmat b1

 svmat t1

 quietly: regress y x1 x2 x3 ww

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 ww

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,3)

 mat t3=el(e(b),1,3)/sqrt(el(v,1,3))

 svmat b3

 svmat t3

 post sim1 (ruv) (rzu) (rzv) (ps25) (n) (b1) (t1) (b2) (t2) (b3) (t3)

 }

 postclose sim1

erase t1.dta

erase id1.dta

erase original.dta

erase tm1.dta

end

set seed 1000

model121

use psm121, replace

sum rzu ruv ps25 b2

quietly: sum b2

display r(sd)^2

gen v2= (b2-.5)^2

quietly: sum v2

gen mse=r(sum)/10000

sum mse

program drop model121

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 1.3 (Selection on the observables, Heckit Treatment Effect Model)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 1.3

//u & z correlated, u & v NOT correlated, treatreg (maximum likelihood)

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_13, replace

clear

program model1\_3

 postfile sim rzu ruv rwu rzv b1 t1 athrho using tx13, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: treatreg y x1 x2 x3,treat(ww=z)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,4)

 mat t1=el(e(b),1,4)/sqrt(el(v,1,4))

 mat athrho=el(e(b),1,8)

 svmat b1

 svmat t1

 svmat athrho

 post sim (rzu) (ruv) (rwu) (rzv) (b1) (t1) (athrho)

 }

 postclose sim

end

set seed 1000

model1\_3

use tx13, replace

gen rho=(exp(2\*athrho)-1)/(1+exp(2\*athrho))

sum rzu ruv b1

quietly: sum b1

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum mse1

program drop model1\_3

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 1.3.1 (Selection on the observables, Heckit Treatment Effect Model omitted z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 1.3.1

//u & z correlated, u & v NOT correlated, treatreg (maximum likelihood)

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_131, replace

clear

program model\_131

 postfile sim rzu ruv rwu rzv b1 t1 athrho using tx131, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: treatreg y x1 x2 x3,treat(ww=x1 x2 x3)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,4)

 mat t1=el(e(b),1,4)/sqrt(el(v,1,4))

 mat athrho=el(e(b),1,8)

 svmat b1

 svmat t1

 svmat athrho

 post sim (rzu) (ruv) (rwu) (rzv) (b1) (t1) (athrho)

 }

 postclose sim

end

set seed 1000

model\_131

use tx131, replace

gen rho=(exp(2\*athrho)-1)/(1+exp(2\*athrho))

sum rzu ruv b1

quietly: sum b1

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum mse1

program drop model\_131

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Models 1.4 & 1.4.1 (Selection on the observables, Matching Estimators with & without z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Models 1.4 & 1.4.1 Matching

// u & z correlated, u & v NOT correlated, nnmatch

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Models\_14\_&\_141, replace

clear

program model14\_141

 postfile sim2 ruv rzu rzv b1 z1 bb4 zz4 using sim14\_141, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rzu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: nnmatch y ww x1 x2 x3 z, m(4) tc(ate) bias(bias) robust(4)

 mat b1=el(e(b),1,1)

 mat z1=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat b1

 svmat z1

 quietly: nnmatch y ww x1 x2 x3, m(4) tc(ate) bias(bias) robust(4)

 mat bb4=el(e(b),1,1)

 mat zz4=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat bb4

 svmat zz4

 post sim2 (ruv) (rzu) (rzv) (b1) (z1) (bb4) (zz4)

 }

 postclose sim2

end

set seed 1000

model14\_141

use sim14\_141, replace

gen v1=(b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum ruv rzu b1 mse1

use sim14\_141, replace

gen v4=(bb4-.5)^2

quietly: sum v4

gen mse4=r(sum)/10000

sum ruv rzu bb4 mse4

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Models 1.5 & 1.5.1 (Selection on the observables, PSW with & without z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Models 1.5 & 1.5.1 PSW

// u & z correlated, u & v NOT correlated

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Models\_15\_&\_151, replace

clear

program model1\_5

 postfile sim rzu ruv rwu rzv b1 t1 b2 t2 b3 t3 b4 t4 using sim1\_5, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

 //Create various types of weights

 gen ate\_w=1/(1-ps) if ww==0

 replace ate\_w=1/ps if ww==1

 gen att\_w=ps/(1-ps) if ww==0

 replace att\_w=1 if ww==1

 //PSW

 quietly: regress y x1 x2 x3 z ww [pweight=ate\_w]

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,5)

 mat t1=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b1

 svmat t1

 quietly: regress y x1 x2 x3 ww [pweight=ate\_w]

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 x3 z ww [pweight=att\_w]

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,5)

 mat t3=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b3

 svmat t3

 quietly: regress y x1 x2 x3 ww [pweight=att\_w]

 mat v=vecdiag(e(V))

 mat b4=el(e(b),1,4)

 mat t4=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b4

 svmat t4

 post sim (rzu) (ruv) (rwu) (rzv) (b1) (t1) (b2) (t2) (b3) (t3) (b4) (t4)

 }

 postclose sim

end

set seed 1000

model1\_5

use sim1\_5, replace

sum rzu ruv rwu b1 b2 b3 b4

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

gen v2= (b2-.5)^2

quietly: sum v2

gen mse2=r(sum)/10000

gen v3= (b3-.5)^2

quietly: sum v3

gen mse3=r(sum)/10000

gen v4= (b4-.5)^2

quietly: sum v4

gen mse4=r(sum)/10000

sum mse1 mse2 mse3 mse4

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Models 1.6 & 1.6.1 (Selection on the observables, PSS with & without z)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Models 1.6 & 1.6.1 PSS

//u & z correlated, u & v NOT correlated

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_16\_&161, replace

clear

program model1\_6

 postfile sim b1 t1 b2 t2 using sim1\_6, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u, /\*

 \*/ corr(1,.2,.3,0,0\.2,1,0,0,0\.3,0,1,0,0\0,0,0,1,.4\0,0,0,.4,1) /\*

 \*/ means (3 2 10 5 0) sds (.5 .6 9.5 2 1)

 gen v=invnormal(uniform())

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

save "C:\tmp\t", replace

//PSS

set more off

sort ps

quietly: sum ps

gen x\_=r(N)/5

g pg=0

replace pg=1 if \_n <= x\_

replace pg=2 if \_n > x\_ & \_n <= 2\*x\_

replace pg=3 if \_n > 2\*x\_ & \_n <= 3\*x\_

replace pg=4 if \_n > 3\*x\_ & \_n <= 4\*x\_

replace pg=5 if \_n > 4\*x\_

save "c:\tmp\t", replace

//pg=1

use "c:\tmp\t", clear

drop if pg !=1

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg1", replace

use "c:\tmp\t", clear

drop if pg !=1

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg1"

drop \_merge

save "c:\tmp\pg1", replace

//pg=2

use "c:\tmp\t", clear

drop if pg !=2

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg2", replace

use "c:\tmp\t", clear

drop if pg !=2

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg2"

drop \_merge

save "c:\tmp\pg2", replace

//pg=3

use "c:\tmp\t", clear

drop if pg !=3

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg3", replace

use "c:\tmp\t", clear

drop if pg !=3

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg3"

drop \_merge

save "c:\tmp\pg3", replace

//pg=4

use "c:\tmp\t", clear

drop if pg !=4

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg4", replace

use "c:\tmp\t", clear

drop if pg !=4

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg4"

drop \_merge

save "c:\tmp\pg4", replace

//pg=5

use "c:\tmp\t", clear

drop if pg !=5

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg5", replace

use "c:\tmp\t", clear

drop if pg !=5

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg5"

drop \_merge

save "c:\tmp\pg5", replace

use "c:\tmp\pg1", clear

append using "c:\tmp\pg2"

append using "c:\tmp\pg3"

append using "c:\tmp\pg4"

append using "c:\tmp\pg5"

gen b1=b1\_B[1]\*(n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[2]\*(n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[3]\*(n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[4]\*(n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[5]\*(n[5]/(n[1]+n[2]+n[3]+n[4]+n[5]))

gen b1\_se\_total=sqrt( ///

 b1\_var[1]\*(((n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[2]\*(((n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[3]\*(((n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[4]\*(((n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[5]\*(((n[5]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2))

gen b2=b2\_B[1]\*(n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[2]\*(n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[3]\*(n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[4]\*(n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[5]\*(n[5]/(n[1]+n[2]+n[3]+n[4]+n[5]))

gen b2\_se\_total=sqrt( ///

 b2\_var[1]\*(((n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[2]\*(((n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[3]\*(((n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[4]\*(((n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[5]\*(((n[5]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2))

gen t1=b1/b1\_se\_total

gen t2=b2/b2\_se\_total

 post sim (b1) (t1) (b2) (t2)

 }

 postclose sim

end

set seed 1000

model1\_6

use sim1\_6, replace

sum

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

gen v2= (b2-.5)^2

quietly: sum v2

gen mse2=r(sum)/10000

sum mse1 mse2

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.1 (Selection on the unobservables, OLS)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 2.1 OLS

//u & z are NOT correlated, BUT u & v correlated, OLS

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_21, replace

clear

program model2\_1

 postfile sim1 rzu ruv rwu rzv b1 t1 b2 t2 b3 t3 using sim2\_1, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: regress y x1 x2 x3 z ww

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,5)

 mat t1=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b1

 svmat t1

 quietly:regress y x1 x2 x3 ww

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 ww

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,3)

 mat t3=el(e(b),1,3)/sqrt(el(v,1,3))

 svmat b3

 svmat t3

 post sim1 (rzu) (ruv) (rwu) (rzv) (b1) (t1) (b2) (t2) (b3) (t3)

 }

 postclose sim1

end

set seed 1000

model2\_1

use sim2\_1, replace

sum rzu ruv b1

quietly: sum b1

display r(sd)^2

quietly: sum b2

display r(sd)^2

quietly: sum b3

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

gen v2= (b2-.5)^2

quietly: sum v2

gen mse2=r(sum)/10000

gen v3= (b3-.5)^2

quietly: sum v3

gen mse3=r(sum)/10000

sum mse1

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.2 (Selection on the unobservables, PSM)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 2.2

//u & z NOT correlated, u & v correlated - PSM

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_22, replace

clear

program model2\_2

 postfile sim1 rzu ruv rwu rzv ps25 n b1 t1 b2 t2 b3 t3 using psm22, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u ,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

 quietly: sum ps, detail

 gen ps25=r(sd)\*.25

 gen id=\_n

 save original, replace

 quietly: psmatch2 ww, caliper(.086) pscore(ps) logit quietly noreplacement descending

 sort \_id

 g match=id[\_n1]

 g treat=id if \_nn==1

 drop if treat==.

 save t1, replace

 sort match

 keep match

 rename match id

 save id1, replace

 use original, replace

 sort id

 merge id using id1

 drop if \_merge!=3

 append using t1

 save tm1, replace

 quietly: regress y ww x1 x2 x3 z

 gen n=e(N)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,1)

 mat t1=el(e(b),1,1)/sqrt(el(v,1,1))

 svmat b1

 svmat t1

 quietly: regress y x1 x2 x3 ww

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 ww

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,3)

 mat t3=el(e(b),1,3)/sqrt(el(v,1,3))

 svmat b3

 svmat t3

 post sim1 (rzu) (ruv) (rwu) (rzv) (ps25) (n) (b1) (t1) (b2) (t2) (b3) (t3)

 }

 postclose sim1

erase t1.dta

erase id1.dta

erase original.dta

erase tm1.dta

end

set seed 1000

model2\_2

use psm22, replace

sum rzu ruv ps25 b1

quietly: sum b1

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum mse1

program drop model2\_2

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.3 (Selection on the unobservables, Heckit Treatment Effect Model)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 2.3

//u & z are NOT correlated, BUT u & v correlated, treatreg (maximum likelihood)

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_23, replace

clear

program model2\_3

 postfile sim1 rzu ruv rwu rzv b1 t1 athrho using tx23, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: treatreg y x1 x2 x3,treat(ww=z)

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,4)

 mat t1=el(e(b),1,4)/sqrt(el(v,1,4))

 mat athrho=el(e(b),1,8)

 svmat b1

 svmat t1

 svmat athrho

 post sim1 (rzu) (ruv) (rwu) (rzv) (b1) (t1) (athrho)

 }

 postclose sim1

end

set seed 1000

model2\_3

use tx23, replace

gen rho=(exp(2\*athrho)-1)/(1+exp(2\*athrho))

sum rzu ruv b1

quietly: sum b1

display r(sd)^2

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

sum mse1

program drop model2\_3

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.4 (Selection on the unobservables, Matching Estimators)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Section 11.2 Model 2.4

//u & z are NOT correlated, BUT u & v correlated, - nnmatch

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model2\_4, replace

clear

program model2\_4

 postfile sim2 ruv rzu rzv b1 z1 bb4 zz4 bt1 zt1 bbt4 zzt4 using sim2\_4, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: nnmatch y ww x1 x2 x3 z, m(4) tc(ate) bias(bias) robust(4)

 mat b1=el(e(b),1,1)

 mat z1=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat b1

 svmat z1

 quietly: nnmatch y ww x1 x2 x3, m(4) tc(ate) bias(bias) robust(4)

 mat bb4=el(e(b),1,1)

 mat zz4=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat bb4

 svmat zz4

 quietly: nnmatch y ww x1 x2 x3 z, m(4) tc(att) bias(bias) robust(4)

 mat bt1=el(e(b),1,1)

 mat zt1=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat bt1

 svmat zt1

 quietly: nnmatch y ww x1 x2 x3, m(4) tc(att) bias(bias) robust(4)

 mat bbt4=el(e(b),1,1)

 mat zzt4=el(e(b),1,1)/sqrt(el(e(V),1,1))

 svmat bbt4

 svmat zzt4

 post sim2 (ruv) (rzu) (rzv) (b1) (z1) (bb4) (zz4) (bt1) (zt1) (bbt4) (zzt4)

 }

 postclose sim2

end

set seed 1000

model2\_4

use sim2\_4, replace

gen v1=(b1-.5)^2

quietly: sum v1

gen mse=r(sum)/10000

sum ruv rzu b1 mse

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.5 (Selection on the unobservables, PSW with OLS)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Models 2.5 and 2.5.1 PSW with OLS

//u & z are NOT correlated, BUT u & v correlated, OLS

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_25\_&251, replace

clear

program model2\_5

 postfile sim25 rzu ruv rwu rzv b1 t1 b2 t2 b3 t3 b4 t4 using sim2\_5, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

 //Create various types of weights

 gen ate\_w=1/(1-ps) if ww==0

 replace ate\_w=1/ps if ww==1

 gen att\_w=ps/(1-ps) if ww==0

 replace att\_w=1 if ww==1

 //PSW

 quietly: regress y x1 x2 x3 z ww [pweight=ate\_w]

 mat v=vecdiag(e(V))

 mat b1=el(e(b),1,5)

 mat t1=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b1

 svmat t1

 quietly: regress y x1 x2 x3 ww [pweight=ate\_w]

 mat v=vecdiag(e(V))

 mat b2=el(e(b),1,4)

 mat t2=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b2

 svmat t2

 quietly: regress y x1 x2 x3 z ww [pweight=att\_w]

 mat v=vecdiag(e(V))

 mat b3=el(e(b),1,5)

 mat t3=el(e(b),1,5)/sqrt(el(v,1,5))

 svmat b3

 svmat t3

 quietly: regress y x1 x2 x3 ww [pweight=att\_w]

 mat v=vecdiag(e(V))

 mat b4=el(e(b),1,4)

 mat t4=el(e(b),1,4)/sqrt(el(v,1,4))

 svmat b4

 svmat t4

 post sim25 (rzu) (ruv) (rwu) (rzv) (b1) (t1) (b2) (t2) (b3) (t3) (b4) (t4)

 }

 postclose sim25

end

set seed 10000

model2\_5

use sim2\_5, replace

sum rzu ruv rwu b1 b2 b3 b4

gen v1= (b1-.5)^2

quietly: sum v1

gen mse1=r(sum)/10000

gen v2= (b2-.5)^2

quietly: sum v2

gen mse2=r(sum)/10000

gen v3= (b3-.5)^2

quietly: sum v3

gen mse3=r(sum)/10000

gen v4= (b4-.5)^2

quietly: sum v4

gen mse4=r(sum)/10000

sum mse1 mse2 mse3 mse4

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Model 2.6 (Selection on the unobservables, PSS with OLS)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

//Models 2.6 and 2.6.1 PSS with OLS

//u & z are NOT correlated, BUT u & v correlated

cd "D:\psa\_e2\Chapter11\data"

capture log close

set more off

log using Model\_26\_&261, replace

clear

program model2\_6

 postfile sim26 b3 t3 b4 t4 using sim2\_6, replace

 forvalues i = 1/10000 {

 drop \_all

 set obs 500

 corr2data x1 x2 x3 z u vv, /\*

 \*/ corr(1,.2,.3,0,0,0\.2,1,0,0,0,0\.3,0,1,0,0,0\0,0,0,1,0,0\0,0,0,0,1,.7\0,0,0,0,.7,1) /\*

 \*/ means (3 2 10 5 0 0) sds (.5 .6 9.5 2 1 1)

 gen v=invnormal(uniform())+.15\*vv

 gen w=.5\*z+.1\*x3+v

 quietly: sum w,detail

 gen ww=1

 replace ww=0 if w<= r(p50)

 gen y=100+.5\*x1+.2\*x2-.05\*x3+.5\*ww+u

 quietly: correlate z u,means

 gen rzu=r(rho)

 quietly: correlate u v,means

 gen ruv=r(rho)

 quietly: correlate ww u,means

 gen rwu=r(rho)

 quietly: correlate z v,means

 gen rzv=r(rho)

 quietly: logistic ww x1 x2 x3 z

 predict ps

 save "C:\tmp\t", replace

//PSS

set more off

sort ps

quietly: sum ps

gen x\_=r(N)/5

g pg=0

replace pg=1 if \_n <= x\_

replace pg=2 if \_n > x\_ & \_n <= 2\*x\_

replace pg=3 if \_n > 2\*x\_ & \_n <= 3\*x\_

replace pg=4 if \_n > 3\*x\_ & \_n <= 4\*x\_

replace pg=5 if \_n > 4\*x\_

save "c:\tmp\t", replace

//pg=1

use "c:\tmp\t", clear

drop if pg !=1

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg1", replace

use "c:\tmp\t", clear

drop if pg !=1

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg1"

drop \_merge

save "c:\tmp\pg1", replace

//pg=2

use "c:\tmp\t", clear

drop if pg !=2

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg2", replace

use "c:\tmp\t", clear

drop if pg !=2

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg2"

drop \_merge

save "c:\tmp\pg2", replace

//pg=3

use "c:\tmp\t", clear

drop if pg !=3

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg3", replace

use "c:\tmp\t", clear

drop if pg !=3

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg3"

drop \_merge

save "c:\tmp\pg3", replace

//pg=4

use "c:\tmp\t", clear

drop if pg !=4

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg4", replace

use "c:\tmp\t", clear

drop if pg !=4

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg4"

drop \_merge

save "c:\tmp\pg4", replace

//pg=5

use "c:\tmp\t", clear

drop if pg !=5

quietly: regress y x1 x2 x3 z ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,5)

 mat ww\_var=el(v,1,5)

 svmat ww

 svmat ww\_var

rename ww1 b1\_B

rename ww\_var1 b1\_var

g n=\_N

keep if \_n==1

keep pg b1\_B b1\_var n

sort pg

save "c:\tmp\b1\_pg5", replace

use "c:\tmp\t", clear

drop if pg !=5

quietly: regress y x1 x2 x3 ww

mat v=vecdiag(e(V))

 mat ww=el(e(b),1,4)

 mat ww\_var=el(v,1,4)

 svmat ww

 svmat ww\_var

rename ww1 b2\_B

rename ww\_var1 b2\_var

g n=\_N

keep if \_n==1

keep pg b2\_B b2\_var

sort pg

merge pg using "c:\tmp\b1\_pg5"

drop \_merge

save "c:\tmp\pg5", replace

use "c:\tmp\pg1", clear

append using "c:\tmp\pg2"

append using "c:\tmp\pg3"

append using "c:\tmp\pg4"

append using "c:\tmp\pg5"

gen b3=b1\_B[1]\*(n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[2]\*(n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[3]\*(n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[4]\*(n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b1\_B[5]\*(n[5]/(n[1]+n[2]+n[3]+n[4]+n[5]))

gen b3\_se\_total=sqrt( ///

 b1\_var[1]\*(((n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[2]\*(((n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[3]\*(((n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[4]\*(((n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b1\_var[5]\*(((n[5]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2))

gen b4=b2\_B[1]\*(n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[2]\*(n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[3]\*(n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[4]\*(n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])) ///

 +b2\_B[5]\*(n[5]/(n[1]+n[2]+n[3]+n[4]+n[5]))

gen b4\_se\_total=sqrt( ///

 b2\_var[1]\*(((n[1]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[2]\*(((n[2]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[3]\*(((n[3]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[4]\*(((n[4]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2) ///

+b2\_var[5]\*(((n[5]/(n[1]+n[2]+n[3]+n[4]+n[5])))^2))

gen t3=b3/b3\_se\_total

gen t4=b4/b4\_se\_total

 post sim26 (b3) (t3) (b4) (t4)

 }

 postclose sim26

end

set seed 10000

model2\_6

use sim2\_6, replace

sum

gen v3= (b3-.5)^2

quietly: sum v3

gen mse3=r(sum)/10000

gen v4= (b4-.5)^2

quietly: sum v4

gen mse4=r(sum)/10000

sum mse3 mse4

log close

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_