PEP Technical note

By: Araar, A.: April 2015

The treatment effect: Comparing the ETR, ESR and PSM methods with an artificial example

The main aim of this note is help the user to understand and to assimilate the similarities and dissimilarities of some popular econometric approaches, used to assess the causality or treatment effect. Our approach here is based on the progressive manner of assimilation through the use of a simple illustrative data, and for which we can define beforehand the exact level of the impact. The three popular models for which we focus in this technical note are the Endogenous Treatment Regression (ETR: Stata 13), the Endogenous Switching Regression (ESR) and the popular Propensity Score Matching method. The primitive reaction of the applicant is: Starting from the fact that all methods tries to assess the same aspect, and based on the complexity of real data and noise that contaminates the net effect, how can one decide about the relevance of each model? In which case the different models are equivalent?

PART A: *The artificial example*

We assume that the number of observations is 1000:

```
set more off
clear all
set seed 1234
set obs 1000
```

Also, we assume that we have three regions:

```
gen      region    = 1 in  1/300
replace region   = 2 in 301/600
replace region   = 3 in 601/1000
```

It is assumed that the first region has more working age population:

```
gen      age  = min(int(runiform()*65+15), 65)
replace age  = age+5 if region==1

gen      educ = min(int(runiform()*5+1), 6)
```

It is assumed that the program is not randomly attributed and the population in region_1 have more probability to be selected. Also, it is assumed that the selection depends partially on age:

```
set seed 7421
gen treatment=3*runiform()*(region==1)+0.5*runiform()*(region==2)+0.5*runiform()*(region==3)+(0.2+0.8*runiform())*(age> 30)
replace treatment    =  treatment > 1

local a =       0.60
local b =       0.01
gen   e= `a'*runiform()

sum e if treatment ==1
qui replace e = e - r(mean) if  treatment ==1
sum e if treatment ==0
qui replace e = e - r(mean) if  treatment ==0
```

The outcome (income) depends on education, age, and the treatment. The parameter **a** enables to control for the predictive power of the two outcome models with the ESR method. The higher is **a,** the lower is the predictive power of the model. The parameter **b** enables to control for the contribution of the variable endogeneity (age). The higher is **b**, the higher is the endogeneity. In this artificial example, we know the exact value of the effect of the program, which is equal to **2**:

```
local at = 2
gen income = 60+0.5*educ+`b'*age+`at'*treatment + e
```

It is assumed that the variable age is not observed, but it affects jointly the program selection and the outcome. This raises the endogeneity problem, and we will need to use or to construct an instrumental variable (inst). The latter is assumed to be not explained by the outcome:

```
gen income0 =income -`at'*treatment
gen ins = (0.5+uniform())*age
regress ins income0
predict  inst, res
```

PART B: ***Estimating the effect with PSM and ESR methods:***

```
gen pw=1
xi: psmatch2 treatment i.region inst , outcome(income) cal(0.01) pw(pw) ate
local att_psm = r(att)
local atu_psm = r(atu)
gen lincome = log(income)
set seed 5241
xi: movestay lincome educ   , select(treatment  i.region inst )
msat treatment, expand(yes)

gen region2= region==2
gen region3= region==3
etregress income  i.treatment#c.( educ ) , treat(treatment =  region2 region3 inst) vce(robust)
margins r.treatment , contrast(nowald) subpop(treatment)   vce(unconditional)
local att_etr= el(r(b),1,1)
gen untreatment = treatment!=1
margins r.treatment , contrast(nowald) subpop(untreatment) vce(unconditional)
local atu_etr= el(r(b),1,1)
```

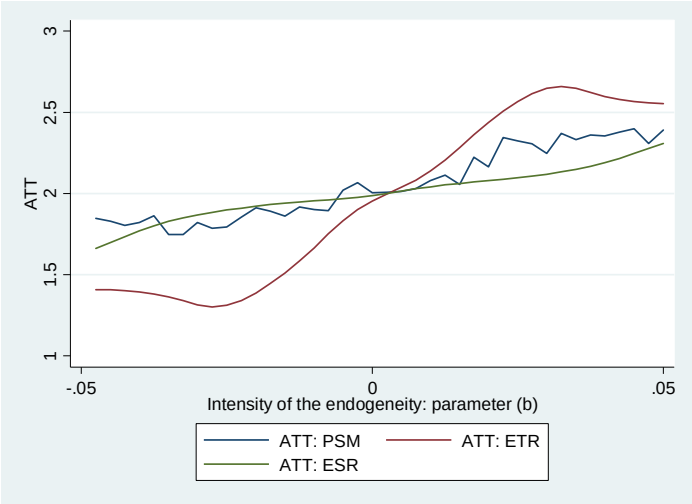
	PSM	ETR	ESR
ATT	2.01357	2.213564	2.05258
ATU	2.23400	2.212639	2.07934

Based on the results above, the first conclusion is that the three models succeed to well capture the effect when the predictive power of outcome models is high (lower level of parameter a, that control the residual part of the model).

Estimating the ATT, models and endogeneity

Now, we show how the results are affected by the level of endogeneity (parameter b) where the latter varies between -0.05 and 0.05. For this end, we select a moderate level of the parameter a (a=0.6).

Figure 1: Estimating the ATT, models and endogeneity

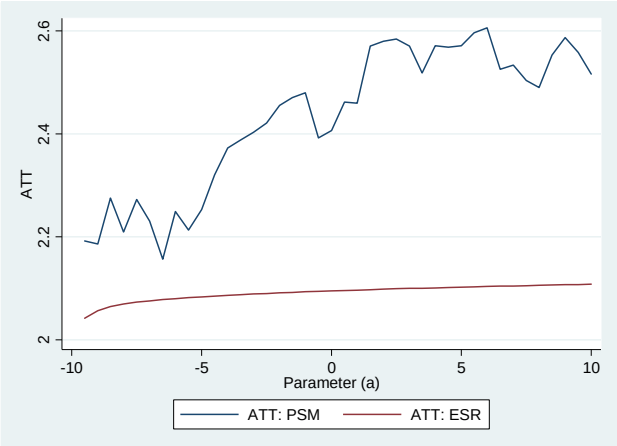


Compared to the ETR, the ESR has a refined form for modeling the covariance between residuals. As shown in figure 1, this helps to gives better estimates of the ATT. Indeed, the ETR imposes restriction, because it is assumed that the correlations between the unobservable determinants in the outcome model and those in the treatment model are equal for the treated and the untreated groups. Obviously, this restriction limits the correction of the endogeneity bias. Thus, better than the ETR, the ESR model is more appropriate in the case of presence of a high level of endogeneity.

The ATT, models and the predictive power of the outcome model

To check the sensitivity of results with the predictive power of outcome models (which is inversely linked with the parameter a), we estimate the ATT according to a, and this, when b is fixed to 0.01. Again here the ESR shows a better performance.

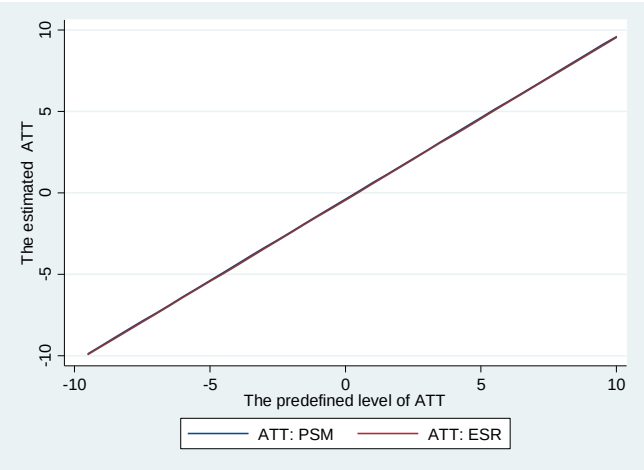
Figure 2: Estimating the ATT, models and the predictive power of the outcome model



The ATT, models and the importance of the fixed impact

Now, we control the parameters **a** and **b**, (a=0.6 and b=0.01) and we vary the predefined ATT (see the command to generate the income and how to initialise the value of the parameter **at**)

Figure 3: Estimating the ATT, models and the importance of the homogeneous impact

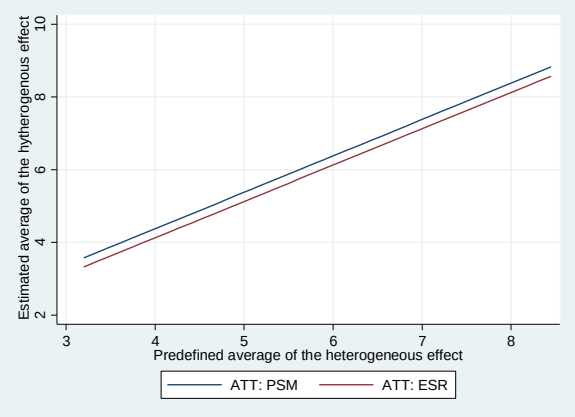


As we can observe, with the high predictive power of models and with a low endogeneity, the two models are practically equivalent and estimate accurately the treatment effect.

The ATT, models and the importance of impact heterogeneity

In the previous investigations, it was assumed a homogenous treatment effect. Obviously, in this case, even a simple econometric model is enough to estimate the impact. Now assume that the treatment effect depends on the observed and on the unobserved individual characteristics (at=3***v****educt+0.01*age). We show in Figure 4 the estimates when the parameter **v** varies between 0.1 and 0.3, a=0.6 and b=0.03. Note that, for each value of **v**, we can compute beforehand the predefined average treatment of that have the heterogeneous form (the value in x-axis). Further, we can use the ESR and PSM models to estimate the average treatment effect.

Figure 4: Estimating the ATT, models and the importance of the heterogeneous impact



The PSM show a low performance to treat the endogeneity problem based in the non-randomisation of the heterogeneous impact.