

## Lecture 5: Propensity Score, Part II

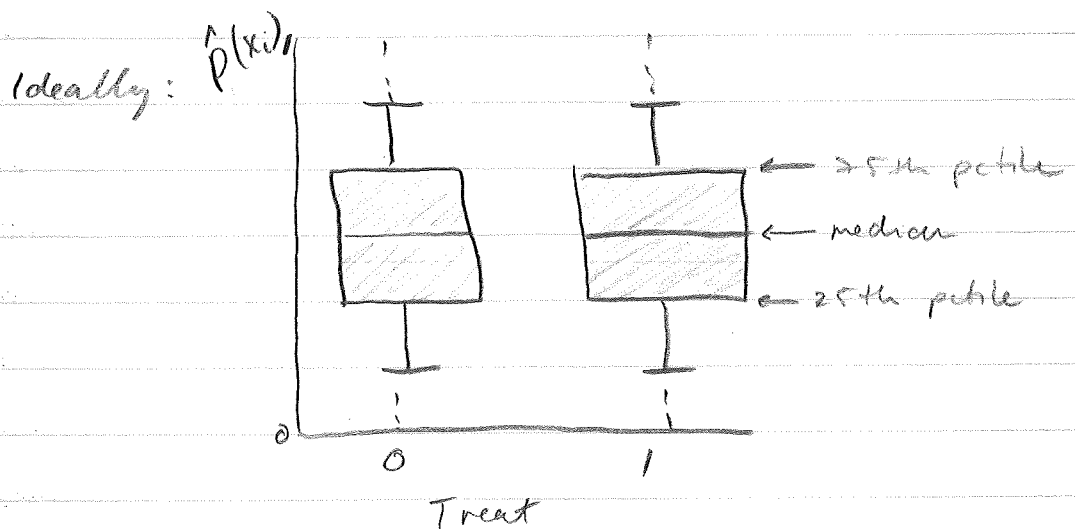
### "Using the Propensity Score"

Previous lecture: how to calculate the propensity score; justifications for using it.  
→ Now we want to see how to use it.

Once you get the propensity score:

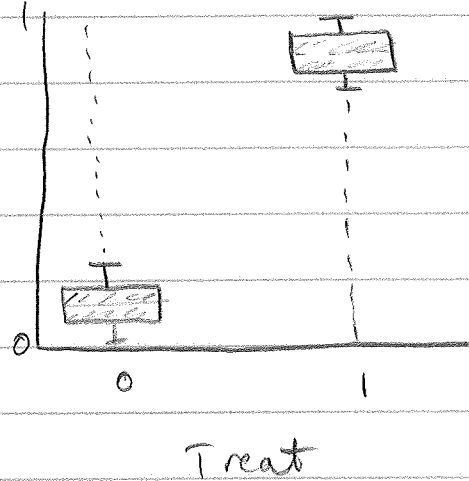
1a) graph box  $\hat{p}(x_i)$ , over (treat)

Why? We want to see if there's a lot of overlap between estimated propensity scores for treatment and control group



lots of overlap

Not good:  $\hat{p}(x_i)$



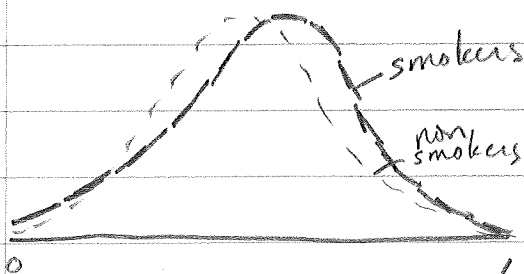
Very little overlap: comparing "apples to oranges".  $\Rightarrow$  Pure selection on X's. Extrapolating across non-comparable populations.

If you have good overlap, continue.

Otherwise, STOP! you really can't do propensity-score methods without good overlap.

What's "good" overlap? We need to have some commonality in p-scores between treat and control:

1b) `kdensity  $\hat{p}(x_i)$  if treat==1,`  
`addplot(kdensity  $\hat{p}(x_i)$  if treat==0).`



Are the distributions the same?

Kolmogorov-Smirnov Test:

$H_0$ : Distributions are the same

$H_a$ : They are not.

`ksmirnov  $\hat{p}(x_i)$ , by (treat).`

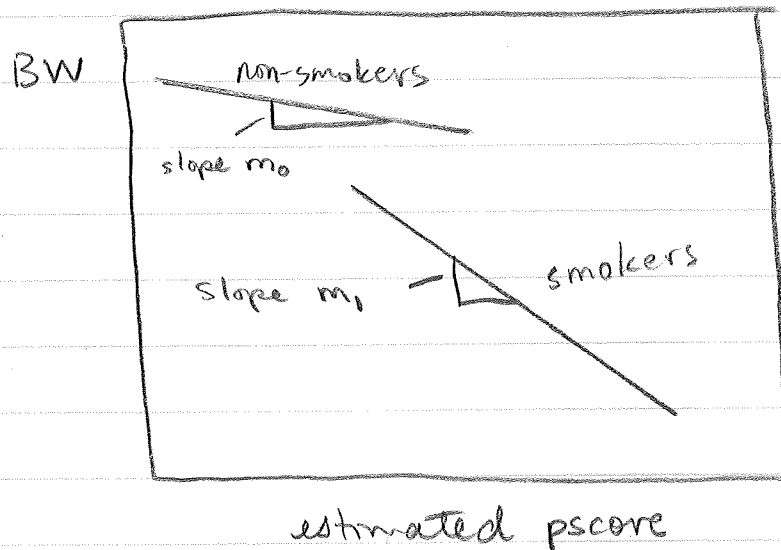
If corrected  $p\text{-val} < .05$  on corrected Combined K-S, reject  $H_0$  in favor of  $H_a$ .  
 ↳ Trim/match data until can no longer reject  $H_0$ !

2) Use the estimated propensity score as a control.

$$y_i = \alpha + \theta T + \delta_1 \hat{p}_i + \delta_2 T_i (\hat{p}_i - \hat{\mu}_p) + \varepsilon_i$$

control  
for  
pscore

deviation  
from mean  
p-score  $\times$  Treat



$\delta_1 = m_0$  } This is just the effect on BW of being likely to select into treatment.  
 $\delta_2 = m_1 - m_0$  } This is the marginal effect of selection into treatment by smokers.

If  $\delta_2 \neq 0$ , allows for a differential slope.  $\Rightarrow$  Differential selection into treatment.

- 3) Use the blocks generated from estimating the propensity score to match treat and control individuals.

$$y_i = \sum_b \alpha_b + \theta T + \varepsilon_i$$

~~~~~  
different  
intercept  
for each  
block

In stata `areg y treat, absorb(block)`.

Really easy, pure matching on propensity score and then averaging the differences in  $(\bar{y}_1 - \bar{y}_0)$ .

Essentially: 
$$\frac{\sum_b (\bar{y}_1 - \bar{y}_0)_b}{b}$$

Note, though, that we're still assuming  $\theta_b = \theta$  for all  $b$ .

4) What if  $\theta_b \neq \theta$  for all  $b$ ? why don't we let  $\theta_b$  vary by propensity score block?

$$y_i = \sum_b \alpha_b D_b + \sum_b \gamma_b T \cdot D_b + \varepsilon_i$$

↓  
a dummy  
for each block

↘  
This is how much  
BW differs between  
smokers and nonsmokers  
in each block.

5) Since  $\theta_b \neq \theta$  for all  $b$ , let's calculate  $E(y_{1i} - y_{0i} | T=1) \equiv \text{Treatment on Treated}$ .

↳ what would be the effect of treatment specifically for those who are really likely to select into treatment?

weight controls by:  $\frac{p(x_i)}{1-p(x_i)}$

higher  $p_i \Rightarrow$  more weight.

Example:  $p_i = 0.6 \Rightarrow w_i = \frac{0.6}{1-0.6} = 1.5$

$p_i = 0.5 \Rightarrow w_i = \frac{0.5}{1-0.5} = 1$

$p_i = 0.4 \Rightarrow w_i = \frac{0.4}{0.6} = 0.66$

So more weight on control obs. that are more likely to select into treatment (but didn't), and less weight on control obs. less likely to select into treatment.

weight treatment by 1.

Then run:

reg  $y = \alpha + \theta T$  [pw=TOT weight]

[Plot kdensities! Notice the overlap between distributions. Where is the mean?]

6) How does this differ from the Average Treatment Effect (ATE)?

$ATE = E(y_{1i} - y_{0i})$

weight treatment by  $\frac{1}{p_i} \Rightarrow$  more weight on lower  $p_i$  treatment obs.

weight control by  $\frac{1}{1-p_i} \Rightarrow$  more weight on higher  $p_i$  control obs.

Conceptually, we want to add more weight to "average" individuals.

Note: ksmirnov test not possible when using weights.

reg  $y = \alpha + \theta T$  [pw = ATE weight]  
[Plot & densities! Notice the overlap between distributions. Where is the mean? How does it compare to when weighting with TOT weights]  
Why might ATE and TOT effects be different?

- 7) Find the mean pscore for treat and control obs by centile. Graph against mean BW for treat and control obs by centile.  
 $\hookrightarrow$  "Non-parametric" approach.

```

1 cd "/Users/lfeler/Documents/teaching/Applied Econometrics Course/Notes/
  Weeks 4-5"
2 clear
3 use smoking2
4 set seed 1000
5 sample 20
6
7
8 gen dmeduc2=dmeduc^2
9 gen dimageXdmeduc=dimage*dmeduc
10 gen dmeducXdfeduc=dmeduc*dfeduc
11 gen dmeducXmblack=dmeduc*mblack
12 gen dfeducXfblack=dfeduc*fblack
13
14 pscore tobacco dimage dmeduc dmar ddivord nprevist dfage dfeduc anemia
  diabetes phyper alcohol ///
15 drink foreignb plural deadkids mblack motherr mhispan fblack fotherr
  fhispan first ///
16 dmeduc2 dimageXdmeduc dmeducXdfeduc dmeducXmblack dfeducXfblack, ///
17 logit pscore(phat) blockid(block) numblo(20)
18
19 graph box phat, over(tobacco)
20
21 kdensity phat if tobacco==1, legend(label(1 "Smokers"))
  addplot(kdensity phat if tobacco==0, legend(label(2 "Non Smokers")))
22 ksmirnov phat, by(tobacco)
23 /*We can reject that smokers and non-smokers have the same
  distribution of propensity scores. What we do now is either trim the
  data or re-weight the data
24 until we can no longer reject the null that distributions are the
  same. More on this below*/
25
26 egen meanphat=mean(phat)
27 gen devmeanphat=phat-meanphat
28 gen treatXdevmeanphat=tobacco*devmeanphat
29
30
31 reg dbirwt tobacco, robust
32
33
34 reg dbirwt tobacco dimage dmeduc dmar ddivord nprevist dfage dfeduc
  anemia diabetes ///

```



```

35 phyper alcohol drink foreignb plural deadkids mblack motherr mhispan
   fblack fotherr fhispan first, robust
36
37 reg dbirwt tobacco phat, robust
38
39 reg dbirwt tobacco phat treatXdevmeanphat, robust
40
41
42 areg dbirwt tobacco, absorb(block) robust
43
44
45 xi i.block, prefix(_bl) noomit
46
47 reg dbirwt tobacco _bl*, robust
48
49
50 foreach var of varlist _bl* {
51   gen TreatX`var'=tobacco*`var'
52 }
53
54 reg dbirwt _bl* TreatX_bl*, noconstant robust
55
56
57 gen TOTweight=phat/(1-phat)
58 replace TOTweight=1 if tobacco==1
59
60 reg dbirwt tobacco [pw=TOTweight], robust
61
62 gen ATEweight=(1/phat)
63 replace ATEweight=(1/(1-phat)) if tobacco==0
64
65 reg dbirwt tobacco [pw=ATEweight], robust
66
67 **Trim by top and bottom 1% of pscore and re-estimate ATE and TOT to
   show sensitivity to pscore extreme values**
68 sum phat, d
69 gen trim=(phat<`r(p1)'\|phat>`r(p99)')
70 reg dbirwt tobacco if trim==0 [pw=TOTweight], robust
71 reg dbirwt tobacco if trim==0 [pw=ATEweight], robust
72
73 *Once we trim, are distributions the same*
74 kdensity phat if tobacco==1&trim==0, legend(label(1 "Smokers"))

```

```

addplot(kdensity phat if tobacco==0&trim==0, legend(label(2 "Non
Smokers"))))
75 ksmirnov phat if trim==0, by(tobacco)
76
77 *Notice what happens to the densities when we use the ATE weight*
78 kdensity phat if tobacco==1&trim==0 [aw=ATEweight], legend(label(1
"Smokers")) addplot(kdensity phat if tobacco==0&trim==0 [aw=ATEweight],
legend(label(2 "Non Smokers"))))
79
80 *Notice what happens to the densities when we use the TOT weight*
81 kdensity phat if tobacco==1&trim==0 [aw=TOTweight], legend(label(1
"Smokers")) addplot(kdensity phat if tobacco==0&trim==0 [aw=TOTweight],
legend(label(2 "Non Smokers"))))
82
83 xtile c200=phat, nq(200)
84 sort c200
85 by c200: egen meanbwsmo=mean(dbirwt) if tobacco==1
86 by c200: egen meanpscoresmo=mean(phat) if tobacco==1
87 by c200: egen meanbwnsmo=mean(dbirwt) if tobacco==0
88 by c200: egen meanpscorensmo=mean(phat) if tobacco==0
89
90 preserve
91 collapse (mean) meanbwsmo meanpscoresmo meanbwnsmo meanpscorensmo,
by(c200)
92
93 label var meanbwsmo smokers
94 label var meanbwnsmo nonsmokers
95
96 graph twoway scatter meanbwsmo meanpscoresmo ///
97 || scatter meanbwnsmo meanpscorensmo, msymbol(+) ///
98 ytitle(birth weight) xtitle(pscore)
99
100 restore
101
102
103
104

```