This exercise examines the following research question: What is the effect of maternal smoking during pregnancy on infant birth weight and death? The required reading for this problem set is Almond, Chay, and Lee (2005). The data extract, “lbw.dta”, is from the 1989 Linked National Natality-Mortality Details Files, which are an annual census of births in the U.S., derived from Certificates of Live Birth. Information on subsequent infant death within a year of birth is derived from Death Certificates. This extract consists of all births in Pennsylvania in 1989. The observational unit of the data is the mother-infant outcome match. There are 139,149 observations and 32 variables. For this problem set, observations with missing values for any of the variables below were dropped from the original sample (about 17%).

The key variables are:
- dbirwt: birth weight of the infant (in grams)
- death: indicator equal to one if the infant died within one year of birth and zero, otherwise
- tobacco: indicator equal to one if the mother smoked during pregnancy and zero, otherwise

The relevant control variables are:

- Mother’s attributes:
  - dmage: (mother’s age), dmeduc: (mother’s educational attainment), mblack: (indicator=1 if mother is Black), motherr: (=1 if neither Black nor White), mhispan: (=1 if Hispanic), dmar: (=1 if mother is unmarried), foreignb: (=1 if mother is foreign born)

- Father’s attributes:
  - dfage: (father’s age), dfeduc: (father’s educational attainment), fblack: (indicator=1 if father is Black), fotherr: (=1 if neither Black nor White), fhispan: (=1 if Hispanic)

- Other risky behavior:
  - alcohol: (indicator=1 if mother drank alcohol during pregnancy), drink: (number of drinks per week)

- Medical care:
  - tripre1, tripre2, tripre3: (indicators=1 if 1st prenatal care visit in 1st, 2nd, or 3rd trimester, respectively), tripre0: (=1 if no prenatal care visits), nprevist: (total number of prenatal care visits)

- Pregnancy history and maternal health:
  - first: (=1 if first-born), dlivord: (birth order), deadkids: (number previous births where newborn died), disllb: (months since last birth), preterm: (=1 if previous birth premature or small for gestational age), pre4000: (=1 if previously had > 4000 gram newborn), plural: (=1 if twins or greater birth), phyper: (=1 if mother had pregnancy associated hypertension), diabete: (=1 if mother diabetic), anemia: (=1 if mother anemic)

a. Under what conditions can one identify the average treatment effect of maternal smoking by comparing the unadjusted difference in mean birth weight of infants of smoking and non-smoking mothers? Under the assumption that maternal smoking is randomly assigned, estimate its impact on birth weight. Provide some evidence for or against the hypothesis that maternal smoking is randomly assigned.

---

1 This problem set requires that you download a user-written command “pscore” and its associated files. First, type in “net search pscore” in Stata, and then click the link “st0026_2 from http://www.stata-journal.com/software/sj5-3” shown on the Stata result window.
b. Suppose that maternal smoking is randomly assigned conditional on the other observable determinants of infant birth weight.

(1) What does this imply about the relationship between maternal smoking and unobservable determinants of birth weight conditional on the observables?

(2) Use a basic linear regression model to estimate the impact of smoking and report your estimates.

- *Stata code:* `reg dbirwt tobacco dmage dmeduc mblack motherr mhispan dmar foreignb dfage dfeduc fblack fotherr fhispan alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist first dlvord deadkids disllb preterm pre4000 plural phyper diabete anemia, vce(robust)`

(3) Under what conditions is the average treatment effect identified?

c. Under the assumption of random assignment conditional on the observables:

(1) What are the sources of misspecification bias in the estimates generated by the linear model estimated in part (b)?

(2) Use an approach in the spirit of multivariate matching, that is, estimate the smoking effects using a flexible functional form for the control variables (e.g., higher order terms and interactions).

- *Stata code:* `reg dbirwt tobacco dmage dmeduc mblack motherr mhispan dmar foreignb dfage dfeduc fblack fotherr fhispan alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist first dlvord deadkids disllb preterm pre4000 plural phyper diabete anemia dmage2 dmage3 dmeduc2 dfage2 dfage3 dfeduc2 dmage dmeduc dmage_mblack dmage_motherr dmage_mhispan dmage_dmage_dfage_dfeduc dfage_fblack dfage_fotherr dfage_fhispan dmage_alcohol dmage_drink, vce(robust)`

(3) What are the benefits and drawbacks to this approach?

d. Describe the propensity score approach to the problem of estimating the average treatment effect of smoking when the treatment is randomly assigned conditional on the observables. How does it reduce the dimensionality problem of multivariate matching?

e. Implement the propensity score approach (hints below) to the evaluation problem using two methods:

(1) Method 1: control directly for the estimated propensity scores in a regression model.

(2) Method 2: use the estimated propensity score in a classification scheme to stratify the sample.

(3) Provide empirical evidence that your implementation is reasonable and evidence on the overlap of the observables of smokers and non-smokers. Present your findings and interpret the results. (This is an open-ended question, so show me what you know and be creative and thoughtful.)

- Below is my Stata code for estimating the propensity score without specifying the number of blocks at the beginning. Stata generated 33 blocks in the end. For balancing results, please see the attached log file.

  pscore tobacco dmage dmeduc mblack motherr mhispan dmar foreignb dfage dfeduc fblack fotherr fhispan alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist first dlvord deadkids disllb preterm pre4000 plural phyper diabete anemia dmage2 dmage3 dmeduc2 dfage2 dfage3 dfeduc2 dmage dmage_mblack dmage_motherr dmage_mhispan dmage_dmage_dfage_dfeduc dfage_fblack dfage_fotherr dfage_fhispan dmage_alcohol dmage_drink, pscore(pscore) blockid(block) logit level(0.01)
• Below is my Stata code for estimating the propensity score with 100 blocks.

```
pscore tobacco dimage dmeduc mblack motherr mhispan dmar foreignb dfage dfeduc fblack fotherr fhispan alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist first dlivord deadkids disllb preterm pre4000 plural phyper diabete anemia dimage2 dmeduc2 dfage2 dfeduc2 dimage_mblack dimage_motherr dimage_mhispan dimage_dmar dfage_dfeduc dfage_fblack dfage_fotherr dfage_fhispan dimage_alcohol dimage_drink, pscore(pscore100) blockid(block100) logit level(0.001) numblo(101)
```

• Below is my Stata code for estimating the propensity score with 200 blocks.

```
pscore tobacco dimage dmeduc mblack motherr mhispan dmar foreignb dfage dfeduc fblack fotherr fhispan alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist first dlivord deadkids disllb preterm pre4000 plural phyper diabete anemia dimage2 dmeduc2 dfage2 dfeduc2 dimage_mblack dimage_motherr dimage_mhispan dimage_dmar dfage_dfeduc dfage_fblack dfage_fotherr dfage_fhispan dimage_alcohol dimage_drink, pscore(pscore200) blockid(block200) logit level(0.005) numblo(201)
```

f. Use the estimated propensity scores to reweigh the outcomes of non-smokers and estimate the average treatment effect.

(1) Compare the estimates to those in part (e) and interpret your findings.

(2) What are the benefits and drawbacks of approaches that use the estimated propensity scores as weights?

g. A more informative way to describe the birth weight effects of smoking is to estimate the nonparametric conditional mean of birth weight as a function of the estimated propensity score, separately for smokers and non-smokers.

(1) To do this, simply stratify the smokers into 100 equal-sized cells based on their propensity scores and calculate the mean birth weight and propensity score in each cell. Do the same for the non-smokers. Plot these two conditional mean functions, with the mean scores on the x-axis and mean birth weight on the y-axis.

(2) Interpret your findings and relate them to the results in part (e) and part (f).

(3) Redo the above using 200 equal-sized cells for smokers and non-smokers.

h. Low birth weights (less than 2500 grams) are considered particularly undesirable since they comprise a large share of infant deaths.

(1) Redo part (g) using an indicator for low birth weight as the outcome of interest.

(2) Interpret your findings.

i. Estimate the impact of maternal smoking on infant death using the methods in parts (a), (b), and (g), using the 100 equal-sized cells for smokers and non-smokers.

(1) Interpret your findings.

(2) From your results, what might you conclude about the relationship between birth weight and infant death?

j. Concisely and coherently summarize your results above. Describe the estimated effects of maternal smoking on infant-birth weight and infant mortality, and whether you think your “best” estimate of the effects of smoking is credibly identified. State why or why not.
References

```plaintext
/* Trial versions *******************************************/
pscore tobacco
dmage dmeduc mblack motherr mhispan dmar foreignb
dfage dfeduc fblack fotherr fhispan
alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist
first divord deadkids disllb preterm pre4000 plural phyper diabete anemia
,pscore(p_score) blockid(block_id) detail logit comsup;

pscore tobacco
dmage dmeduc mblack motherr mhispan dmar foreignb
dfage dfeduc fblack fotherr fhispan
alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist
first divord deadkids disllb preterm pre4000 plural phyper diabete anemia
,pscore(p_score) blockid(block_id) logit;

pscore tobacco
dmage dmeduc mblack motherr mhispan dmar foreignb
dfage dfeduc fblack fotherr fhispan
alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist
first divord deadkids disllb preterm pre4000 plural phyper diabete anemia
,pscore(p_score) blockid(block_id) logit comsup;

pscore tobacco
dmage dmeduc mblack motherr mhispan dmar foreignb
dfage dfeduc fblack fotherr fhispan
alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist
first divord deadkids disllb preterm pre4000 plural phyper diabete anemia
,pscore(p_score) blockid(block_id) logit level(0.001);

************************************************************************************ */

. gen dmage2 = dmage^2;
. gen dmeduc2 = dmeduc^2;
. gen dfage2 = dfage^2;
. gen dfeduc2 = dfeduc^2;
. gen dmage_dmeduc = dmage*dmeduc;
. gen dmage_mblack = dmage*mblack;
. gen dmage_motherr = dmage*motherr;
. gen dmage_mhispan = dmage*mhispan;
```
. gen dmage_dmar = dmage*dmar;
. gen dfage_dfeduc = dfage*dfeduc;
. gen dfage_fblack = dfage*fblack;
. gen dfage_fotherr = dfage*fotherr;
. gen dfage_fhispan = dfage*fhispan;
. gen dmage_alcohol = dmage*alcohol;
. gen dmage_drink = dmage*drink;
. pscore tobacco
> dmage dmeduc mblack motherr mhispan dmar foreignb
> dfage dfeduc fblack fotherr fhispan
> alcohol drink tripre1 tripre2 tripre3 tripre0 nprevist
> first divord deadkids disllb preterm pre4000 plural phyper diabete anemia
> dmage2 dmeduc2 dfage2 dfeduc2
> dmage dmeduc dmage_mblack dmage_motherr dmage_mhispan dmage_dmar
> dfage dfeduc dfage_fblack dfage_fotherr dfage_fhispan
> dmage_alcohol dmage_drink
> ,pscore(pscore) blockid(block) logit level(0.01);

*******************************************************************************
Algorithm to estimate the propensity score
*******************************************************************************

The treatment is tobacco

<table>
<thead>
<tr>
<th>indicator=1</th>
<th>Freq.</th>
<th>Percent</th>
<th>Cum.</th>
</tr>
</thead>
<tbody>
<tr>
<td>if the mother smoked during pregnancy and zero, otherwise</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>112,782</td>
<td>81.29</td>
<td>81.29</td>
</tr>
<tr>
<td>1</td>
<td>25,957</td>
<td>18.71</td>
<td>100.00</td>
</tr>
<tr>
<td>Total</td>
<td>138,739</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

Estimation of the propensity score

note: tripre3 dropped due to collinearity
Iteration 0:   log likelihood = -66869.313
Iteration 1:   log likelihood = -58462.913
Iteration 2:   log likelihood = -57422.224
Iteration 3:   log likelihood = -57356.981
Iteration 4:   log likelihood = -57356.052
Iteration 5:   log likelihood = -57356.051

Logit estimates

| Number of obs = 138739 |
| LR chi2(43)      = 19026.52 |
| Prob > chi2      = 0.0000 |
| Pseudo R2        = 0.1423 |

Log likelihood = -57356.051
| tobacco          | Coef.   | Std. Err. | z     | P>|z|  | [95% Conf. Interval] |
|------------------|---------|-----------|-------|------|---------------------|
| dmage            | -.1068329 | .0168385 | -6.34 | 0.000 | -.1398358            | -.07383 |
| dmeduc           | .3566266 | .0372407 | 9.58  | 0.000 | .2836362             | .429617 |
| mblack           | -1.634605 | .1438595 | -11.36 | 0.000 | -1.916565            | -1.352646 |
| mother            | -1.16326 | .555444 | 2.09  | 0.036 | 2.252107              | -.0744127 |
| mhispanic        | -1.872902 | .2752192 | -6.81 | 0.000 | -2.1462321            | -.1332468 |
| dmar             | -.450019 | .0907707 | -4.96 | 0.000 | -.6279263             | -.2721118 |
| foreigngb        | -0.6493657 | .0593651 | -10.94 | 0.000 | -.7657192             | -.5330122 |
| dfage            | .0493512 | .0121155 | 4.36  | 0.000 | .0271732              | .0715293 |
| dfeduc           | .2903256 | .0305901 | 9.52  | 0.000 | .2305288              | .3501224 |
| fblack           | -1.610809 | .120269 | -5.13 | 0.000 | -.8528038             | -.381358 |
| fother           | -0.3703641 | .5542399 | -0.67 | 0.504 | -1.456654             | .7159261 |
| fhispan          | -0.6086838 | .2169956 | -2.81 | 0.005 | .0133937              | -.1833803 |
| alcohol          | 3.189513 | .3753923 | 8.50  | 0.000 | 2.453757              | .392526 |
| drink            | -1.376371 | .0893637 | -1.54 | 0.124 | -3.127867             | .0375125 |
| tripre1          | -1.029287 | .0424955 | -2.42 | 0.015 | -.1862183             | .0196391 |
| tripre2          | .0777792 | .0421629 | 1.84  | 0.065 | -.0048584             | .1604169 |
| tripre0          | .3218592 | .0694966 | 4.63  | 0.000 | .1856484              | .45807 |
| npvisit          | -0.0174786 | .0251919 | -6.94 | 0.000 | -.0224174             | -.0125399 |
| first            | -0.0132039 | .0278816 | -0.47 | 0.636 | -.0678509             | .0414431 |
| dlidvor          | .048795 | .0100513 | 4.85  | 0.000 | .0290949              | .0684952 |
| deadkids         | .1839926 | .0098299 | 18.72 | 0.000 | .1647264              | .2032589 |
| disllb           | .0040391 | .0003119 | 12.95 | 0.000 | .0034278              | .0046503 |
| preterm          | .5363455 | .051284 | 10.46 | 0.000 | .4358308              | .6396602 |
| pre4000          | .7662002 | .0872055 | -8.79 | 0.000 | -.3971198             | -.592806 |
| plural           | -1.351962 | .064253 | -2.10 | 0.035 | -.2611297             | .0992627 |
| phyper           | -0.4634038 | .0510504 | -9.08 | 0.000 | -.5634608             | -.3633468 |
| diabetes         | .0742374 | .0604983 | 1.23  | 0.220 | -.044337 | .1928118 |
| anemia           | .3486489 | .0611455 | 5.64  | 0.000 | .2250222              | .4647079 |
| dimage2          | -.0002672 | .0002931 | -8.96 | 0.000 | -.0032016             | -.0002052 |
| dmeduc2          | -.0386059 | .0017458 | -22.11 | 0.000 | -.0420276             | -.0351843 |
| dfage2           | -.0007918 | .000132 | -6.00 | 0.000 | -.0010504             | -.0005331 |
| dfeduc2          | -.0188084 | .0011074 | -16.98 | 0.000 | -.0209789             | -.016638 |
| dimage_dmeduc    | .0155629 | .0010838 | 14.36 | 0.000 | .0134383              | .0176686 |
| dimage_mblack    | .0408797 | .0054963 | 7.44  | 0.000 | .030107 | .0516523 |
| dimage_moth-r    | .0335603 | .0204456 | 1.64  | 0.101 | -.0065124             | .07363 |
| dimage_mhis-n    | .0472308 | .0113564 | 4.16  | 0.000 | .0249727              | .0694888 |
| dimage_dmar      | .0574987 | .003609 | 15.93 | 0.000 | .0504252              | .0645722 |
| dfage_dfeduc     | .0019303 | .0006574 | 2.94  | 0.003 | .0006418              | .0032188 |
| dfage_fblack     | .0182842 | .0038912 | 4.70  | 0.000 | .0106576              | .0259109 |
| dfage_foth-r     | -.0131404 | .0184187 | -0.71 | 0.476 | -.0492405             | .0229596 |
| dfage_fhis-n     | .010875 | .0078056 | 1.39  | 0.164 | -.0044237             | .0261736 |
| dimage_alco-l    | -.0552314 | .013497 | -4.09 | 0.000 | -.0816849             | -.0287778 |
| dimage_drink     | .0074777 | .0033175 | 2.25  | 0.024 | .0009756              | .0139798 |
| _cons            | -.2444141 | .3329251 | -7.34 | 0.000 | -.3096662             | -.179162 |

Description of the estimated propensity score

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Smallest</th>
<th>Estimated propensity score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1%</td>
<td>0.0111771</td>
<td>.0000215</td>
</tr>
<tr>
<td>5%</td>
<td>0.0219342</td>
<td>.0000304</td>
</tr>
<tr>
<td>10%</td>
<td>0.0310177</td>
<td>.0000586</td>
</tr>
<tr>
<td>25%</td>
<td>0.0725522</td>
<td>.0000612</td>
</tr>
<tr>
<td>50%</td>
<td>0.1643964</td>
<td>Mean 1.1870923</td>
</tr>
<tr>
<td>75%</td>
<td>0.2479226</td>
<td>.983068</td>
</tr>
<tr>
<td>90%</td>
<td>0.3873616</td>
<td>.983151</td>
</tr>
<tr>
<td>95%</td>
<td>0.4789602</td>
<td>.986755</td>
</tr>
<tr>
<td>99%</td>
<td>0.6690482</td>
<td>.989007</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kurtosis</td>
</tr>
</tbody>
</table>
Step 1: Identification of the optimal number of blocks
Use option detail if you want more detailed output

The final number of blocks is 33

This number of blocks ensures that the mean propensity score is not different for treated and controls in each block.

Step 2: Test of balancing property of the propensity score
Use option detail if you want more detailed output

Variable dmeduc is not balanced in block 1
Variable mblack is not balanced in block 1
Variable dmar is not balanced in block 1
Variable dfeduc is not balanced in block 1
Variable fblack is not balanced in block 1
Variable fotherr is not balanced in block 1
Variable tripre0 is not balanced in block 1
Variable dmeduc2 is not balanced in block 1
Variable dfeduc2 is not balanced in block 1
Variable dage_dmeduc is not balanced in block 1
Variable dage_mblack is not balanced in block 1
Variable dage_dmar is not balanced in block 1
Variable dfage_dfeduc is not balanced in block 1
Variable dfage_fblack is not balanced in block 1
Variable dfage_fotherr is not balanced in block 1
Variable mblack is not balanced in block 2
Variable fblack is not balanced in block 2
Variable alcohol is not balanced in block 2
Variable drink is not balanced in block 2
Variable tripre0 is not balanced in block 2
Variable first is not balanced in block 2
Variable dage2 is not balanced in block 2
Variable dfage2 is not balanced in block 2
Variable dage_mblack is not balanced in block 2
Variable dimage_alcohol is not balanced in block 2
Variable dimage_drink is not balanced in block 2
Variable mblack is not balanced in block 3
Variable fblack is not balanced in block 3
Variable first is not balanced in block 3
Variable divord is not balanced in block 3
Variable disllb is not balanced in block 3
Variable dimage_mblack is not balanced in block 3
Variable dfage_fblack is not balanced in block 3
Variable dfeeduc is not balanced in block 4
Variable fblack is not balanced in block 4
Variable dfeeduc2 is not balanced in block 4
Variable dimage_mblack is not balanced in block 4
Variable dfage_fblack is not balanced in block 4
Variable dimage_mblack is not balanced in block 5
Variable dfage_fblack is not balanced in block 5
Variable dfeeduc is not balanced in block 7
Variable dfeeduc2 is not balanced in block 7
Variable dimage is not balanced in block 8
Variable deadkids is not balanced in block 8
Variable dimage2 is not balanced in block 8
Variable dimage_dmeduc is not balanced in block 8
Variable dimage is not balanced in block 9
Variable dfage is not balanced in block 9
Variable disllb is not balanced in block 9
Variable dimage2 is not balanced in block 9
Variable dfage2 is not balanced in block 9
Variable dimage_dmeduc is not balanced in block 9
Variable dimage is not balanced in block 10
Variable dfage is not balanced in block 10
Variable dimage2 is not balanced in block 10
Variable dfage2 is not balanced in block 10
Variable dimage_dmeduc is not balanced in block 10
Variable dfage_dfeeduc is not balanced in block 10
Variable fotherr is not balanced in block 12
Variable fhispan is not balanced in block 12
Variable drink is not balanced in block 12
Variable dfage_fotherr is not balanced in block 12
Variable dfage_fhispan is not balanced in block 12
Variable dimage_drink is not balanced in block 12
Variable anemia is not balanced in block 13
Variable drink is not balanced in block 14
Variable first is not balanced in block 14
Variable dimage_drink is not balanced in block 14
Variable dfage_fblack is not balanced in block 16
Variable dfeduc is not balanced in block 17
Variable dmeduc is not balanced in block 18
Variable dfeduc is not balanced in block 18
Variable first is not balanced in block 18
Variable dmeduc2 is not balanced in block 18
Variable dfeduc2 is not balanced in block 18
Variable dfage_dfeduc is not balanced in block 18
Variable motherr is not balanced in block 19
Variable dfeduc is not balanced in block 19
Variable nprevist is not balanced in block 19
Variable dfeduc2 is not balanced in block 19
Variable dimage_motherr is not balanced in block 19
Variable dfage_dfeduc is not balanced in block 19
Variable dfeduc is not balanced in block 20
Variable tripred is not balanced in block 20
Variable dfeduc2 is not balanced in block 20
Variable dmeduc is not balanced in block 21
Variable deadkids is not balanced in block 21
Variable pre4000 is not balanced in block 21
Variable dmeduc2 is not balanced in block 21
Variable tripred is not balanced in block 23
Variable dfeduc is not balanced in block 24
Variable deadkids is not balanced in block 24
Variable fblack is not balanced in block 27
Variable dmage_mblack is not balanced in block 27
Variable dfage_fblack is not balanced in block 27
Variable dmage is not balanced in block 28
Variable dmeduc is not balanced in block 28
Variable mblack is not balanced in block 28
Variable dfage is not balanced in block 28
Variable disllb is not balanced in block 28
Variable dmage2 is not balanced in block 28
Variable dmeduc2 is not balanced in block 28
Variable dfage2 is not balanced in block 28
Variable dmage_dmeduc is not balanced in block 28
Variable dmage_mblack is not balanced in block 28
Variable dfage_dfeduc is not balanced in block 28
Variable dfage_fblack is not balanced in block 28
Variable dmeduc is not balanced in block 29
Variable dfage is not balanced in block 29
Variable tripre1 is not balanced in block 29
Variable disllb is not balanced in block 29
Variable dmeduc2 is not balanced in block 29
Variable dfage2 is not balanced in block 29
Variable dmage_dmeduc is not balanced in block 29
Variable dfage_dfeduc is not balanced in block 29
Variable dmage is not balanced in block 30
Variable dmeduc is not balanced in block 30
Variable deadkids is not balanced in block 30
Variable disllb is not balanced in block 30
Variable dmage2 is not balanced in block 30
Variable dmeduc2 is not balanced in block 30
Variable dmage_dmeduc is not balanced in block 30
Variable deadkids is not balanced in block 31
Variable anemia is not balanced in block 32
The balancing property is not satisfied
Try a different specification of the propensity score
Inferior block of pscore

<table>
<thead>
<tr>
<th></th>
<th>indicator=1 if the mother smoked during pregnancy and zero, otherwise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>8,916</td>
</tr>
<tr>
<td>.025</td>
<td>9,168</td>
</tr>
<tr>
<td>.0375</td>
<td>6,373</td>
</tr>
<tr>
<td>.05</td>
<td>5,119</td>
</tr>
<tr>
<td>.0625</td>
<td>4,611</td>
</tr>
<tr>
<td>.075</td>
<td>4,260</td>
</tr>
<tr>
<td>.0875</td>
<td>4,118</td>
</tr>
<tr>
<td>.1</td>
<td>8,406</td>
</tr>
<tr>
<td>.125</td>
<td>8,209</td>
</tr>
<tr>
<td>.15</td>
<td>4,132</td>
</tr>
<tr>
<td>.1625</td>
<td>4,545</td>
</tr>
<tr>
<td>.175</td>
<td>5,028</td>
</tr>
<tr>
<td>.1875</td>
<td>4,752</td>
</tr>
<tr>
<td>.2</td>
<td>8,512</td>
</tr>
<tr>
<td>.225</td>
<td>6,116</td>
</tr>
<tr>
<td>.25</td>
<td>2,260</td>
</tr>
<tr>
<td>.2625</td>
<td>1,774</td>
</tr>
<tr>
<td>.275</td>
<td>2,961</td>
</tr>
<tr>
<td>.3</td>
<td>4,468</td>
</tr>
<tr>
<td>.35</td>
<td>1,699</td>
</tr>
<tr>
<td>.375</td>
<td>1,413</td>
</tr>
<tr>
<td>.4</td>
<td>1,199</td>
</tr>
<tr>
<td>.425</td>
<td>536</td>
</tr>
<tr>
<td>.4375</td>
<td>144</td>
</tr>
<tr>
<td>.440625</td>
<td>119</td>
</tr>
<tr>
<td>.44375</td>
<td>211</td>
</tr>
<tr>
<td>.45</td>
<td>1,507</td>
</tr>
<tr>
<td>.5</td>
<td>997</td>
</tr>
<tr>
<td>.55</td>
<td>540</td>
</tr>
<tr>
<td>.6</td>
<td>442</td>
</tr>
<tr>
<td>.7</td>
<td>145</td>
</tr>
<tr>
<td>.8</td>
<td>84</td>
</tr>
<tr>
<td>.9</td>
<td>18</td>
</tr>
</tbody>
</table>

Total 112,782 25,957 138,739

************************************************
End of the algorithm to estimate the pscore
************************************************

.sort block pscore;
.drop dimage2 dmeduc2 dfage2 dfeduc2
> dimage_dmeduc dimage_mblack dimage_mother dimage_mhispan dimage_dmar
> dfage_deduc dfage_fblack dfage_fother dfage_fhispan
> dimage_alcohol dimage_drink;
.save "$dr\files\data_pscore.dta",replace;