Kernel matching with automatic bandwidth selection

Ben Jann

University of Bern, ben.jann@soz.unibe.ch

2017 London Stata Users Group meeting
London, September 7–8, 2017

Contents

What is Matching?

- Matching is an approach to "condition on X" between a treatment group and a control group.
- Basic idea:
  1. For each observation in the treatment group, find "statistical twins" in the control group with the same (or at least very similar) X values.
  2. The Y values of these matching observations are then used to compute the counterfactual outcome without treatment for the observation at hand.
  3. An estimate for the average treatment effect can be obtained as the mean of the differences between the observed values and the "imputed" counterfactual values over all observations.

What is Matching?

- Formally:

$$\widehat{ATT} = \frac{1}{N^{T=1}} \sum_{i|T=1} \left[Y_i - \hat{Y}_i^0 \right] \quad \text{with} \quad \hat{Y}_i^0 = \sum_{j|T=0} w_{ij} Y_j$$

$$\widehat{ATC} = \frac{1}{N^{T=0}} \sum_{i|T=0} \left[\hat{Y}_i^1 - Y_i \right] \quad \text{with} \quad \hat{Y}_i^1 = \sum_{j|T=1} w_{ij} Y_j$$

$$\widehat{ATE} = \frac{N^{T=1}}{N} \cdot \widehat{ATT} + \frac{N^{T=0}}{N} \cdot \widehat{ATC}$$

- Different matching algorithms use different definitions of w_{ij}.

ATE: average treatment effect; ATT: a.t.e. on the treated; ATC: a.t.e. on the untreated
T: treatment indicator (0/1)
Y: observed outcome; Y^1; potential outcome with treatment; Y^0: p.o. without treatment

Exact Matching

- Exact matching:
$$w_{ij} = \begin{cases} 1/k_i & \text{if } X_i = X_j \\ 0 & \text{else} \end{cases}$$
with k_i as the number of observations for which $X_i = X_j$ applies.
- The result equivalent to "perfect stratification" or "subclassification" (see, e.g., Cochran 1968).
- Problem: If X contains several variables there is a large probability that no exact matches can be found for many observations (the "curse of dimensionality").

Multivariate Distance Matching (MDM)

- An alternative is to match based on a distance metric that measures the proximity between observations in the multivariate space of X.
- The idea then is to use observations that are “close”, but not necessarily equal, as matches.
- A common approach is to use

$$MD(X_i, X_j) = \sqrt{(X_i - X_j)' \Sigma^{-1} (X_i - X_j)}$$

  as distance metric, where Σ is an appropriate scaling matrix.
  - Mahalanobis matching: Σ is the covariance matrix of X.
  - Euclidean matching: Σ is the identity matrix.
  - Mahalanobis matching is equivalent to Euclidean matching based on standardized and orthogonalized X.

Propensity Score Matching (PSM)

- $(Y^0, Y^1) \perp\!\!\!\perp T \,|\, X$ implies $(Y^0, Y^1) \perp\!\!\!\perp T \,|\, \pi(X)$, where $\pi(X)$ is the treatment probability conditional on X (the "propensity score") (Rosenbaum and Rubin 1983).
- This simplifies the matching task as we can match on one-dimensional $\pi(X)$ instead of multi-dimensional X.
- Procedure
  - Step 1: Estimate the propensity score, e.g. using a Logit model.
  - Step 2: Apply a matching algorithm using differences in the propensity score, $|\hat{\pi}(X_i) - \hat{\pi}(X_j)|$, instead of multivariate distances.
- PSM is very popular
  - https://scholar.google.ch/scholar?q="propensity+score"+AND+(matching+OR+matched+OR+match)

Matching Algorithms

- Various matching algorithms can be used to find potential matches based on MD or $\hat{\pi}(X)$ and determine the matching weights w_{ij}.
- Pair matching (one-to-one matching without replacement)
  - For each observation in the treatment group find the closest observation in the control group. Each control is only used once.
- Nearest-neighbor matching (with replacement)
  - For each observation in the treatment group find the k closest observations in the control group. A single control can be used multiple times. In case of ties, use all ties as matches. k is set by the researcher.
- Caliper matching
  - Like nearest-neighbor matching, but only use controls with a distance smaller than some threshold c.

Matching Algorithms

- Radius matching
  - Use *all* controls with a distance smaller than some threshold c.
- Kernel matching
  - Like radius matching, but give larger weight to controls with smaller distances (using some kernel function such as, e.g., the Epanechnikov kernel).
- Optional: remove remaining imbalance after matching using regression adjustment (a.k.a. "bias correction" in the context of nearest-neighbor matching).

“Why PSM Should Not Be Used for Matching”

- The message of a recent paper by Gary King and Richard Nielsen is: Do not use PSM, it is really, really bad.
  - The paper: http://j.mp/1sexgVw
  - Slides: https://gking.harvard.edu/presentations/why-propensity-scores-should-not-be-used-matching-6
  - Watch it: https://www.youtube.com/watch?v=rBv39pK1iEs
- Their argument goes about as follows:
  - In experimental language, PSM approximates *complete randomization*.
  - Other methods such as MDM approximate *fully blocked randomization*.
  - A fully blocked design is more efficient. It leads to less data imbalance and less “model dependence” (dependence of results on modeling decisions by the researcher).
  - Hence, procedures such as MDM dominate PSM.
  - King and Nielsen provide evidence suggesting that PSM performs shockingly bad.

Types of Experiments

Balance Covariates:	*Complete Randomization*	*Fully Blocked*
Observed	On average	Exact
Unobserved	On average	On average

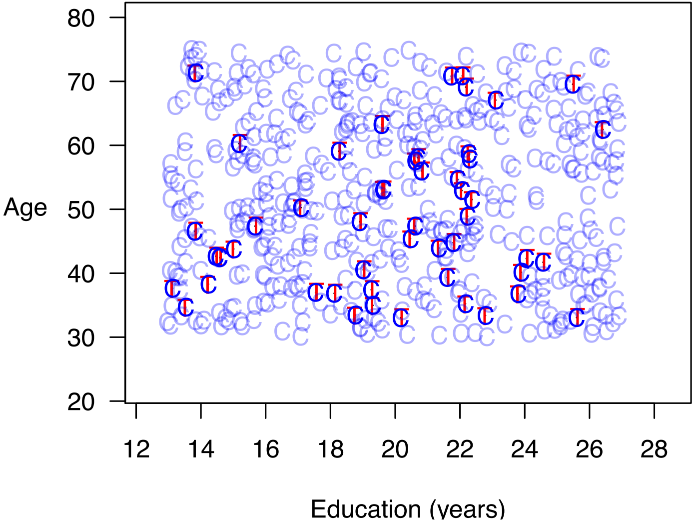
⇝ *Fully blocked* dominates *complete randomization* for: imbalance, model dependence, power, efficiency, bias, research costs, robustness. E.g., Imai, King, Nall 2009: SEs 600% smaller!

Goal of Each Matching Method (in Observational Data)

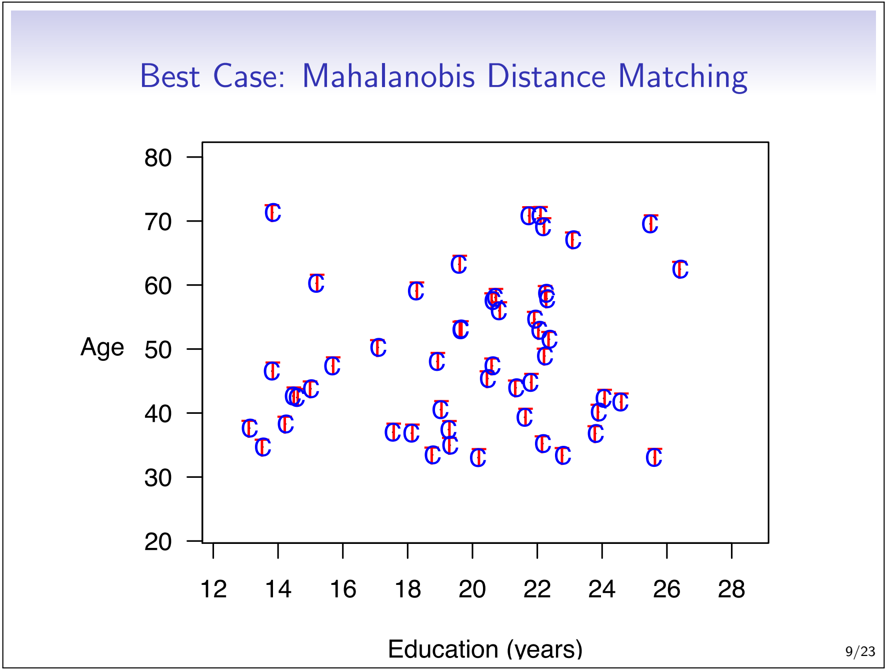
- PSM: *complete randomization*
- Other methods: *fully blocked*
- Other matching methods dominate PSM (wait, it gets worse)

(slides by King and Nielsen)

Best Case: Mahalanobis Distance Matching



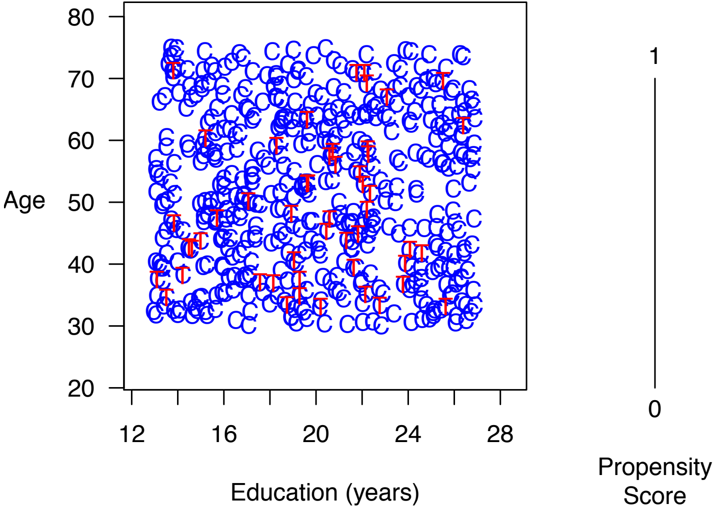
(slides by King and Nielsen)



Best Case: Mahalanobis Distance Matching

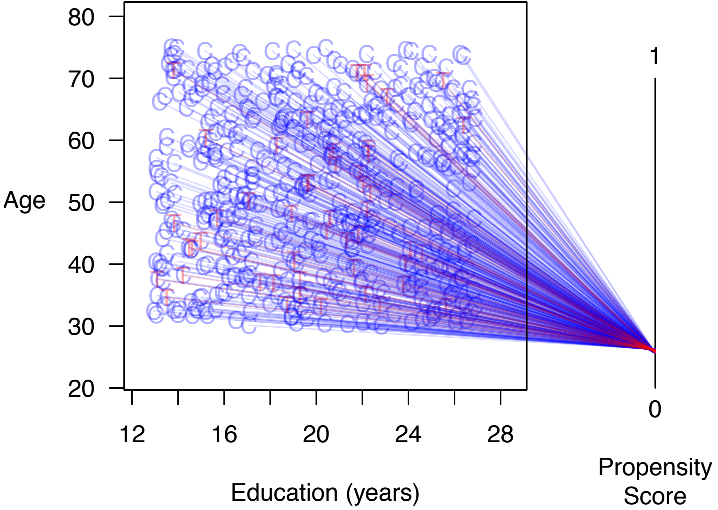
(slides by King and Nielsen)

Best Case: Propensity Score Matching



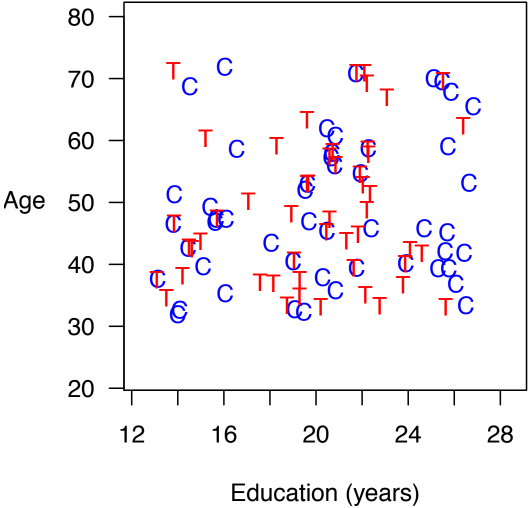
(slides by King and Nielsen)

Best Case: Propensity Score Matching



(slides by King and Nielsen)

Best Case: Propensity Score Matching is Suboptimal



(slides by King and Nielsen)

“Why PSM Should Not Be Used for Matching”

- Are King and Nielsen right?
  - For a given sample size (as in an experiment with fixed budget), fully blocked randomization is more efficient than complete randomization. Things are less clear if blocking reduces the sample size, as in matching.
  - The complete randomization analogy only works for *observations with the same propensity score*. If X has a strong effect on T, there is a lot of blocking also in PSM.
  - King and Nielson’s examples illustrating the bad performance of PSM seem to be based on *pair matching without replacement*. Pair matching throws away a lot of data. For PSM, pair matching is particularly bad because a lot of good data (i.e. observations with the same PS) is thrown away (“random pruning”).
  - The performance of PSM should be alright for matching algorithms that do not engage in random pruning, such as radius or kernel matching.

The kmatch command

- New matching software for Stata.
- Partly written in response to the paper by King and Nielsen.
- Available from SSC (`ssc install kmatch`).

Key Features

- Type of matching
  - Multivariate Distance Matching (MDM)
  - Propensity Score Matching (PSM)
  - MDM combined with PSM
  - MDM and PSM combined with exact matching
- Matching algorithms
  - Kernel matching, including ridge and local-linear matching
  - Nearest-neighbor matching, optionally with caliper
  - Optional regression adjustment
- Several automatic bandwidth selectors for kernel matching
- Joint analysis of multiple subgroups and multiple outcome variables
- Various post-estimation commands for balancing and common-support diagnostics
- Computationally efficient

Examples: Mahalanobis-Distance Kernel Matching

- Estimation of the “effect” of union membership on wages using the NLSW 1988 data.

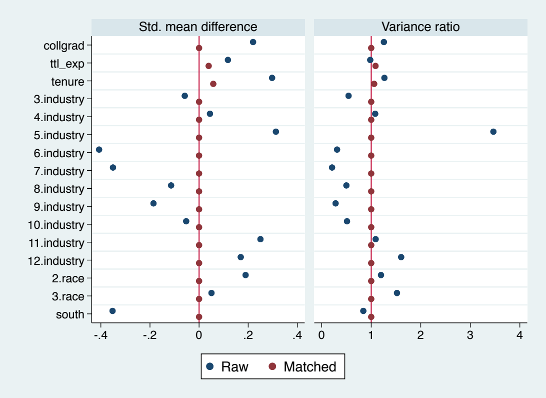
```
. sysuse nlsw88, clear
(NLSW, 1988 extract)

. drop if industry==2
(4 observations deleted)

. kmatch md union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage), nate att
(computing bandwidth ... done)

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
            |           Matched         |           Controls        |   Band-
            |     Yes       No    Total |    Used   Unused    Total |   width
------------+---------------------------+---------------------------+---------
    Treated |     432       25      457 |    1105      291     1396 |  1.3394
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
       wage |      Coef.
------------+-----------------------------------------------------------------
        ATT |   .6059013
       NATE |   1.432913
------------------------------------------------------------------------------
```

Examples: Balancing Statistics

```
. kmatch summarize
(refitting the model using the generate() option)

─────────────┬────────────────────────────────┬────────────────────────────────
             │              Raw               │          Matched(ATT)
       Means │   Treated  Untrea~d    StdDif  │   Treated  Untrea~d    StdDif
─────────────┼────────────────────────────────┼────────────────────────────────
    collgrad │   .321663   .224212   .219912  │   .319444   .319444         0
     ttl_exp │   13.2685   12.7323   .117584  │   13.3205   13.1425   .039036
      tenure │   7.89205   6.17658    .29735  │   7.91744   7.58347   .057888
  3.industry │   .006565   .012178  -.058246  │    .00463    .00463         0
  4.industry │   .183807   .166905   .044425  │   .185185   .185185         0
  5.industry │   .105033   .027937   .312944  │   .085648   .085648         0
  6.industry │   .045952   .169771  -.407129  │   .048611   .048611         0
  7.industry │   .019694   .102436  -.350657  │   .020833   .020833         0
  8.industry │   .017505   .035817  -.113785  │   .009259   .009259         0
  9.industry │   .010941   .040115  -.185669  │   .011574   .011574         0
 10.industry │   .004376   .008596  -.052551  │   .002315   .002315         0
 11.industry │   .479212   .356734   .250073  │   .506944   .506944         0
 12.industry │   .122538    .07235   .169707  │    .12037    .12037         0
      2.race │   .330416   .244986   .189418  │     .3125     .3125         0
      3.race │   .017505   .011461   .050566  │   .006944   .006944         0
       south │   .297593   .466332  -.352408  │   .291667   .291667         0
─────────────┴────────────────────────────────┴────────────────────────────────

─────────────┬────────────────────────────────┬────────────────────────────────
             │              Raw               │          Matched(ATT)
   Variances │   Treated  Untrea~d     Ratio  │   Treated  Untrea~d     Ratio
─────────────┼────────────────────────────────┼────────────────────────────────
    collgrad │   .218674   .174066   1.25628  │   .217904   .217904         1
     ttl_exp │   20.5898   21.0001   .980459  │   19.8177   18.2323   1.08696
      tenure │   37.2044   29.3629   1.26706  │   37.0399   34.9543   1.05966
  3.industry │   .006536   .012038   .542928  │   .004619   .004619         1
  4.industry │   .150351   .139148   1.08052  │   .151242   .151242         1
  5.industry │   .094207   .027176   3.46656  │   .078494   .078494         1
  6.industry │   .043936    .14105   .311496  │   .046355   .046355         1
  7.industry │   .019348   .092008   .210287  │   .020447   .020447         1
  8.industry │   .017237   .034559   .498769  │   .009195   .009195         1
  9.industry │   .010845   .038533   .281445  │   .011467   .011467         1
```

Examples: Make a Graph of the Balancing Statistics

```
. mat M = r(M)

. mat V = r(V)

. coefplot matrix(M[,3]) matrix(M[,6]) || matrix(V[,3]) matrix(V[,6]) || , ///
>     bylabels("Std. mean difference" "Variance ratio") ///
>     noci nolabels byopts(xrescale)

. addplot 1: , xline(0) norescaling legend(order(1 "Raw" 2 "Matched"))

. addplot 2: , xline(1) norescaling
```

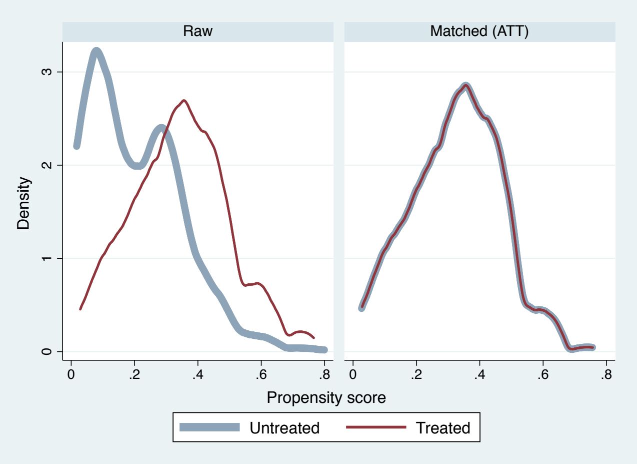


Examples: Propensity-Score Kernel Matching

```
. kmatch ps union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage), nate att
(computing bandwidth ... done)

Propensity-score kernel matching                 Number of obs     =      1,853
                                                 Kernel            =       epan

Treatment : union = 1
Covariates: collgrad ttl_exp tenure i.industry i.race south
PS model  : logit (pr)

Matching statistics
------------------------------------------------------------------------------
             |          Matched           |          Controls          |  Band-
             |     Yes         No   Total |    Used    Unused    Total |  width
-------------+----------------------------+----------------------------+--------
     Treated |     431         26     457 |    1214       182     1396 | .00188
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .3887224
        NATE |   1.432913
------------------------------------------------------------------------------
```

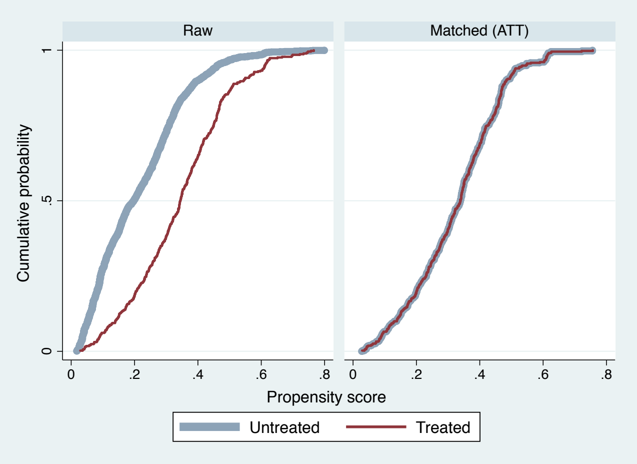
Examples: Density Balancing Plot

```
. kmatch density, lw(*6 *2) lc(*.5 *1)
(refitting the model using the generate() option)
(applying 0-1 boundary correction to density estimation of propensity score)
(bandwidth for propensity score = .06803989)
```



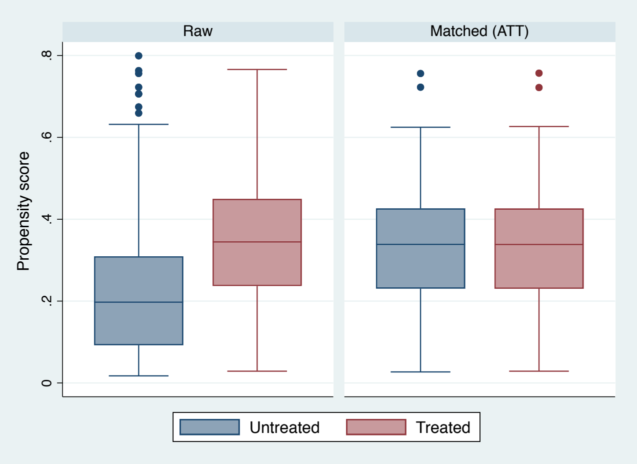
Examples: Cumulative Distribution Balancing Plot

```
. kmatch cumul, lw(*6 *2) lc(*.5 *1)
(refitting the model using the generate() option)
```



Examples: Balancing Box Plot

```
. kmatch box
(refitting the model using the generate() option)
```



Examples: Standard Errors

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage), nate ate att atc vce(bootstrap)
(computing bandwidth for treated ... done)
(computing bandwidth for untreated ... done)

(running kmatch on estimation sample)

Bootstrap replications (50)
----+--- 1 ---+--- 2 ---+--- 3 ---+--- 4 ---+--- 5
..................................................    50

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Replications      =         50
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
             |           Matched           |          Controls           | Band-
             |      Yes        No    Total |     Used   Unused     Total | width
-------------+-----------------------------+-----------------------------+---------
     Treated |      432        25      457 |     1105      291      1396 | 1.3394
   Untreated |     1386        10     1396 |      455        2       457 | 3.3975
    Combined |     1818        35     1853 |     1560      293      1853 |      .
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
             |   Observed   Bootstrap                         Normal-based
        wage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         ATE |   .4095729    .1920853     2.13   0.033      .0330928    .7860531
         ATT |   .6059013    .2472069     2.45   0.014      .1213846    1.090418
         ATC |   .3483797    .1893653     1.84   0.066     -.0227695    .7195289
        NATE |   1.432913    .2333282     6.14   0.000      .9755981    1.890228
------------------------------------------------------------------------------
```

Examples: Postestimation Tests

```
. lincom ATT-NATE
 ( 1)  ATT - NATE = 0
------------------------------------------------------------------------------
        wage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -.8270117   .1810415    -4.57   0.000    -1.181847   -.4721768
------------------------------------------------------------------------------
. test ATT = ATC
 ( 1)  ATT - ATC = 0
           chi2(  1) =    2.42
         Prob > chi2 =    0.1200
```

Examples: Nearest-Neighbor Matching (1 Neighbor)

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), att nn

Multivariate-distance nearest-neighbor matching

                                                Number of obs     =      1,853
                                                Neighbors:    min =          1
Treatment : union = 1                                         max =          1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
-------------------------------------------------------------------------------
            |          Matched          |          Controls          |  Band-
            |     Yes       No    Total |    Used  Unused     Total  |  width
------------+---------------------------+----------------------------+----------
    Treated |     457        0      457 |     328     1068     1396  |       .
------------+---------------------------+----------------------------+----------

Treatment-effects estimation
-------------------------------------------------------------------------------
        wage |      Coef.
-------------+-----------------------------------------------------------------
         ATT |   .7246969
-------------+-----------------------------------------------------------------

. teffects nnmatch (wage collgrad ttl_exp tenure i.industry i.race south) (union), atet

Treatment-effects estimation                   Number of obs      =      1,853
Estimator      : nearest-neighbor matching     Matches: requested =          1
Outcome model  : matching                                     min =          1
Distance metric: Mahalanobis                                  max =          1
--------------------------------------------------------------------------------------
                     |              AI Robust
                wage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
ATET                 |
               union |
(union vs nonunion)  |   .7246969   .2942952     2.46   0.014      .147889    1.301505
---------------------+----------------------------------------------------------------
```

Examples: Nearest-Neighbor Matching (5 Neighbors)

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), att nn(5)
Multivariate-distance nearest-neighbor matching
                                                Number of obs     =      1,853
                                                Neighbors:    min =          5
Treatment : union = 1                                         max =          5
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south
Matching statistics
─────────────┬───────────────────────────┬───────────────────────────┬─────────
             │          Matched          │          Controls         │   Band-
             │     Yes        No    Total│     Used   Unused    Total│   width
─────────────┼───────────────────────────┼───────────────────────────┼─────────
     Treated │     457         0      457│      870      526     1396│        .
─────────────┴───────────────────────────┴───────────────────────────┴─────────

Treatment-effects estimation
──────────────┬────────────────────────────────────────────────────────────────
         wage │      Coef.
──────────────┼────────────────────────────────────────────────────────────────
          ATT │   .5590823
──────────────┴────────────────────────────────────────────────────────────────

. teffects nnmatch (wage collgrad ttl_exp tenure i.industry i.race south) (union), atet nn(5)
Treatment-effects estimation                   Number of obs      =      1,853
Estimator      : nearest-neighbor matching     Matches: requested =          5
Outcome model  : matching                                     min =          5
Distance metric: Mahalanobis                                  max =          6
──────────────────────┬────────────────────────────────────────────────────────────────
                      │              AI Robust
                 wage │      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
──────────────────────┼────────────────────────────────────────────────────────────────
ATET                  │
                union │
  (union vs nonunion) │   .5590823   .2381752     2.35   0.019     .0922675    1.025897
──────────────────────┴────────────────────────────────────────────────────────────────
```

Examples: Regression Adjustment

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage = collgrad ttl_exp tenure i.industry i.race south), att nn(5)

Multivariate-distance nearest-neighbor matching
                                                Number of obs     =      1,853
                                                Neighbors:    min =          5
Treatment : union = 1                                         max =          5
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
             |           Matched         |           Controls        |   Band-
             |     Yes       No    Total |    Used   Unused    Total |   width
-------------+---------------------------+---------------------------+--------
     Treated |     457        0      457 |     870      526     1396 |       .
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .5288023
------------------------------------------------------------------------------
adjusted for collgrad ttl_exp tenure i.industry i.race south

. teffects nnmatch (wage collgrad ttl_exp tenure i.industry i.race south) ///
>     (union), atet nn(5) biasadj(collgrad ttl_exp tenure i.industry i.race south)

Treatment-effects estimation                   Number of obs      =      1,853
Estimator      : nearest-neighbor matching     Matches: requested =          5
Outcome model  : matching                                     min =          5
Distance metric: Mahalanobis                                  max =          6

--------------------------------------------------------------------------------------
                     |              AI Robust
                wage |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------------+----------------------------------------------------------------
ATET                 |
               union |
(union vs nonunion)  |   .5288023   .2420635     2.18   0.029     .0543666    1.003238
--------------------------------------------------------------------------------------
```

Examples: MDM and PSM combined

```
. kmatch md union collgrad ttl_exp tenure (wage), att ///
>     psvars(i.industry i.race south) psweight(3)
(computing bandwidth ... done)

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis (modified)
Covariates: collgrad ttl_exp tenure
PS model  : logit (pr)
PS covars : i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
            |            Matched       |            Controls       |   Band-
            |      Yes        No  Total|     Used   Unused    Total|   width
------------+--------------------------+---------------------------+---------
    Treated |      439        18    457|     1258      138     1396|  .83886
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .6408443
------------------------------------------------------------------------------
```

Examples: MDM with Exact Matching

```
. kmatch md union collgrad ttl_exp tenure (wage), att ematch(industry race south)
(computing bandwidth ... done)

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure
Exact     : industry race south

Matching statistics
-------------------------------------------------------------------------------
            |            Matched        |           Controls        |    Band-
            |       Yes        No  Total|     Used  Unused     Total|    width
------------+---------------------------+---------------------------+----------
    Treated |       432        25    457|     1103     293      1396|   1.3013
-------------------------------------------------------------------------------

Treatment-effects estimation
-------------------------------------------------------------------------------
        wage |      Coef.
-------------+-----------------------------------------------------------------
         ATT |   .6047374
-------------------------------------------------------------------------------
```

Examples: Bandwidth Selection

- Default: 1.5 times the 90% quantile of the (non-zero) distances in pair matching with replacement (Huber et al. 2013, 2015).

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), ///
>     att bwidth(pm)
(computing bandwidth ... done)

Multivariate-distance kernel matching           Number of obs     =      1,853
                                                Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
            |               Matched |                Controls |       Band-
            |      Yes        No      Total |      Used    Unused      Total |       width
------------+---------------------------+-------------------------------+-----------
    Treated |      432        25        457 |      1105       291       1396 |      1.3394
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .6059013
------------------------------------------------------------------------------
```

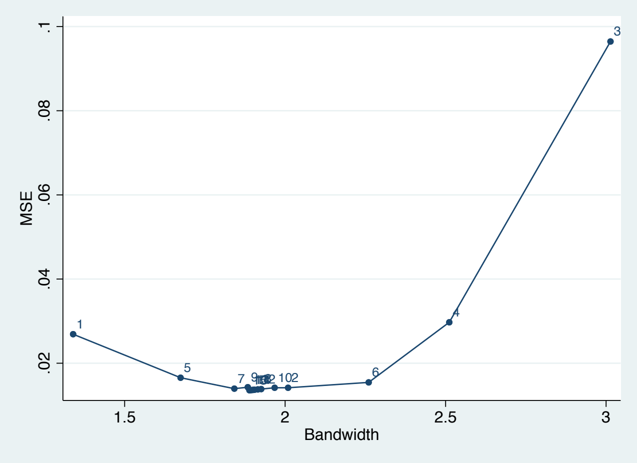
Examples: Bandwidth Selection

- Cross validation with respect to the means of X.

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), ///
>     att bwidth(cv)
(computing bandwidth ................ done)

Multivariate-distance kernel matching           Number of obs     =      1,853
                                                Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
             |         Matched          |          Controls          |  Band-
             |    Yes       No    Total |    Used  Unused     Total |  width
-------------+--------------------------+---------------------------+---------
     Treated |    448        9      457 |    1184     212      1396 |  1.8888
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .6651578
------------------------------------------------------------------------------
```

Examples: Bandwidth Selection

```
. kmatch cvplot, ms(o) index mlabposition(1) sort
```



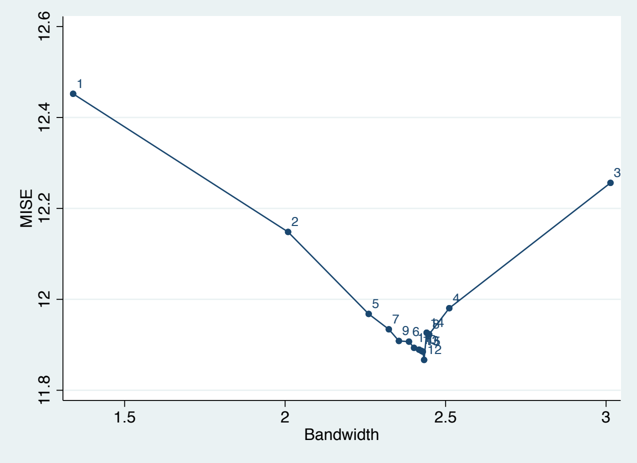
Examples: Bandwidth Selection

- Cross validation with respect to Y (Frölich 2004, 2005).

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), ///
>     att bwidth(cv wage)
(computing bandwidth ................ done)

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
             |           Matched           |           Controls          | Band-
             |      Yes         No    Total|    Used   Unused      Total | width
-------------+-----------------------------+-----------------------------+------
     Treated |      453          4      457|    1289      107       1396 | 2.433
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .6928956
------------------------------------------------------------------------------
```

Examples: Bandwidth Selection

```
. kmatch cvplot, ms(o) index mlabposition(1) sort
```



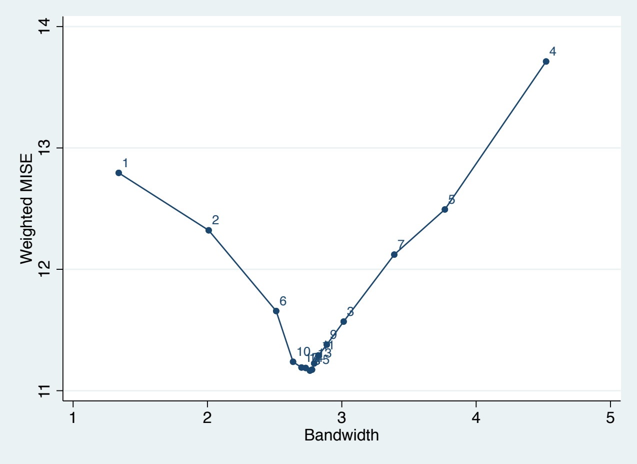
Examples: Bandwidth Selection

- Weighted cross validation with respect to Y (Galdo et al. 2008, Section 4.2).

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), ///
>     att bwidth(cv wage, weighted)
(computing bandwidth ................ done)

Multivariate-distance kernel matching           Number of obs     =      1,853
                                                Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
            |             Matched       |            Controls       |  Band-
            |      Yes        No   Total|     Used   Unused    Total|  width
------------+---------------------------+---------------------------+---------
    Treated |      455         2     457|     1356       40     1396|  2.7626
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
        wage |      Coef.
-------------+----------------------------------------------------------------
         ATT |   .7308166
------------------------------------------------------------------------------
```

Examples: Bandwidth Selection

```
. kmatch cvplot, ms(o) index mlabposition(1) sort
```



Examples: Common Support Diagnostics

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south (wage), ///
>     att bwidth(0.5)

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
──────────────────────────────────────────────────────────────────────────
            │           Matched          │          Controls         │ Band-
            │      Yes        No    Total│    Used   Unused    Total │ width
────────────┼────────────────────────────┼───────────────────────────┼───────
     Treated│      366        91      457│     701      695     1396 │    .5
──────────────────────────────────────────────────────────────────────────

Treatment-effects estimation
──────────────────────────────────────────────────────────────────────────
        wage │      Coef.
─────────────┼────────────────────────────────────────────────────────────
         ATT │   .3303161
──────────────────────────────────────────────────────────────────────────

. kmatch csummarize
(refitting the model using the generate() option)

───────────────────────────────────────────────────────────────────────
             │   Common support (treated)   │  Standardized difference
       Means │   Matched  Unmatc~d    Total │  (1)-(3)   (2)-(3)   (1)-(2)
─────────────┼──────────────────────────────┼──────────────────────────
    collgrad │   .322404   .318681  .321663 │  .001585  -.006376   .007962
     ttl_exp │   13.3929   12.7682  13.2685 │  .027413  -.110253   .137666
      tenure │   8.12614   6.95055  7.89205 │  .038378  -.154356   .192734
  3.industry │   .002732   .021978  .006565 │ -.047404   .190657  -.238061
  4.industry │   .191257   .153846  .183807 │  .019212  -.077269   .096481
  5.industry │   .062842   .274725  .105033 │ -.137462   .552867  -.690329
  6.industry │   .057377         0  .045952 │  .054507  -.219225   .273732
  7.industry │   .019126   .021978  .019694 │ -.004083   .016423  -.020506
  8.industry │   .005464   .065934  .017505 │ -.091714   .368871  -.460585
  9.industry │   .010929   .010989  .010941 │ -.000115   .000462  -.000577
 10.industry │         0   .021978  .004376 │ -.066227   .266363  -.332589
 11.industry │   .554645   .175824  .479212 │   .15083  -.606636   .757467
 12.industry │   .092896   .241758  .122538 │ -.090299   .363181   -.45348
      2.race │   .243169   .681319  .330416 │ -.185284   .745209  -.930494
      3.race │   .002732   .076923  .017505 │ -.112525   .452572  -.565097
       south │    .29235   .318681  .297593 │ -.011456   .046074   -.05753
```

Examples: Make a Graph of Common Support Statistics

```
. mat M = r(M)

. coefplot matrix(M[,4]), noci nolabels xline(0) ///
>     title("Std. difference between matched and original")
```

Examples: Multiple Outcome Variables

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage hours), nate att
(computing bandwidth ... done)

Multivariate-distance kernel matching          Number of obs     =      1,852
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
            |           Matched        |            Controls        |   Band-
            |     Yes       No    Total|     Used    Unused    Total|   width
------------+--------------------------+----------------------------+---------
    Treated |     432       25      457|     1104       291     1395|  1.3392
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
             |      Coef.
-------------+----------------------------------------------------------------
wage         |
         ATT |   .6021049
        NATE |   1.430823
-------------+----------------------------------------------------------------
hours        |
         ATT |   1.263759
        NATE |   1.450303
------------------------------------------------------------------------------
```

Examples: Varying Regression-Adjustment Equations

```
. kmatch md union collgrad ttl_exp tenure i.industry i.race south ///
>     (wage = collgrad ttl_exp tenure) ///
>     (hours = i.industry i.race), nate att
(computing bandwidth ... done)

Multivariate-distance kernel matching           Number of obs     =      1,852
                                                Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race south

Matching statistics
------------------------------------------------------------------------------
             |            Matched         |           Controls         | Band-
             |      Yes        No   Total |   Used   Unused      Total | width
-------------+----------------------------+----------------------------+-------
     Treated |      432        25     457 |   1104      291       1395 | 1.3392
------------------------------------------------------------------------------

Treatment-effects estimation
------------------------------------------------------------------------------
             |      Coef.
-------------+----------------------------------------------------------------
wage         |
         ATT |   .5152752
        NATE |   1.430823
-------------+----------------------------------------------------------------
hours        |
         ATT |   1.263759
        NATE |   1.450303
------------------------------------------------------------------------------
wage: adjusted for collgrad ttl_exp tenure
hours: adjusted for i.industry i.race
```

Examples: Treatment Effects by Subpopulation

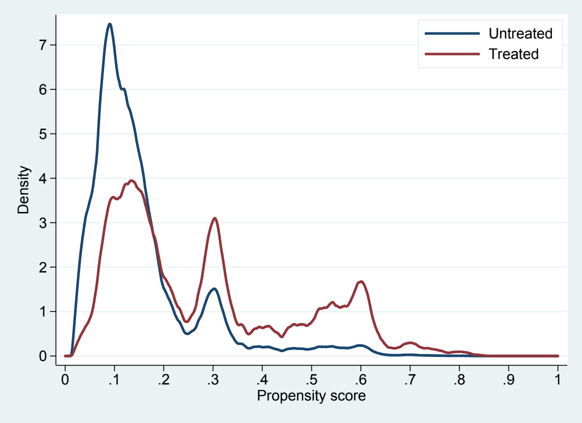
```
. kmatch md union collgrad ttl_exp tenure i.industry i.race (wage), ///
>     att vce(boot) over(south)
(south=0: computing bandwidth ... done)
(south=1: computing bandwidth ... done)

(running kmatch on estimation sample)

Bootstrap replications (50)
────┼─── 1 ───┼─── 2 ───┼─── 3 ───┼─── 4 ───┼─── 5
..................................................    50

Multivariate-distance kernel matching          Number of obs     =      1,853
                                               Replications      =         50
                                               Kernel            =       epan

Treatment : union = 1
Metric    : mahalanobis
Covariates: collgrad ttl_exp tenure i.industry i.race

         0: south = 0
         1: south = 1

Matching statistics
─────────────┬───────────────────────────┬───────────────────────────┬──────────
             │          Matched          │          Controls         │  Band-
             │     Yes       No    Total │    Used   Unused    Total │   width
─────────────┼───────────────────────────┼───────────────────────────┼──────────
0            │                           │                           │
     Treated │     306       15      321 │     625      120      745 │  1.3199
─────────────┼───────────────────────────┼───────────────────────────┼──────────
1            │                           │                           │
     Treated │     126       10      136 │     473      178      651 │  1.3398
─────────────┴───────────────────────────┴───────────────────────────┴──────────

Treatment-effects estimation
─────────────┬──────────────────────────────────────────────────────────────────
             │   Observed   Bootstrap                         Normal-based
        wage │      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
─────────────┼──────────────────────────────────────────────────────────────────
0            │
         ATT │   .4586332    .2808206    1.63   0.102    -.0917652    1.009032
─────────────┼──────────────────────────────────────────────────────────────────
1            │
         ATT │   .9518705     .334356    2.85   0.004     .2965449    1.607196
─────────────┴──────────────────────────────────────────────────────────────────

. test [0]ATT = [1]ATT

 ( 1)  [0]ATT - [1]ATT = 0

           chi2(  1) =    1.36
         Prob > chi2 =    0.2433

. lincom [1]ATT - [0]ATT

 ( 1)  - [0]ATT + [1]ATT = 0

─────────────┬──────────────────────────────────────────────────────────────────
        wage │      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
─────────────┼──────────────────────────────────────────────────────────────────
         (1) │   .4932373    .4227171    1.17   0.243     -.335273    1.321748
─────────────┴──────────────────────────────────────────────────────────────────
```

Simulation

- Population data from Swiss census of 2000.
- Outcome: Treiman occupational prestige (recoded from ISCO codes of the current job using command `iskotrei` by Hendrickx 2002) (values from 6 to 78; mean 44).
- Estimand: ATT of nationality on occupational prestige, with resident aliens as the treatment group and Swiss nationals as the control group.
- Control variables: gender, age, and highest educational degree.
- Population restricted to people between 24 to 60 years old who are working.
- 2'308'006 individuals, of which 17.5% belong to the treatment group.
- Draw random samples ($N = 500$ or 5000) from population and compute various matching estimators.

Simulation

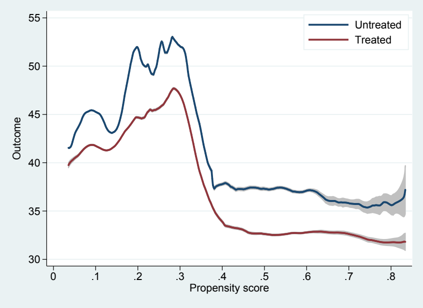
- Substantial differences between resident aliens and Swiss nationals on all three covariates.
- Propensity score in population (computed from fully stratified data)

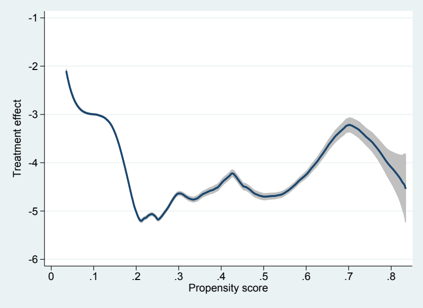


- McFadden $R^2 = 0.121$

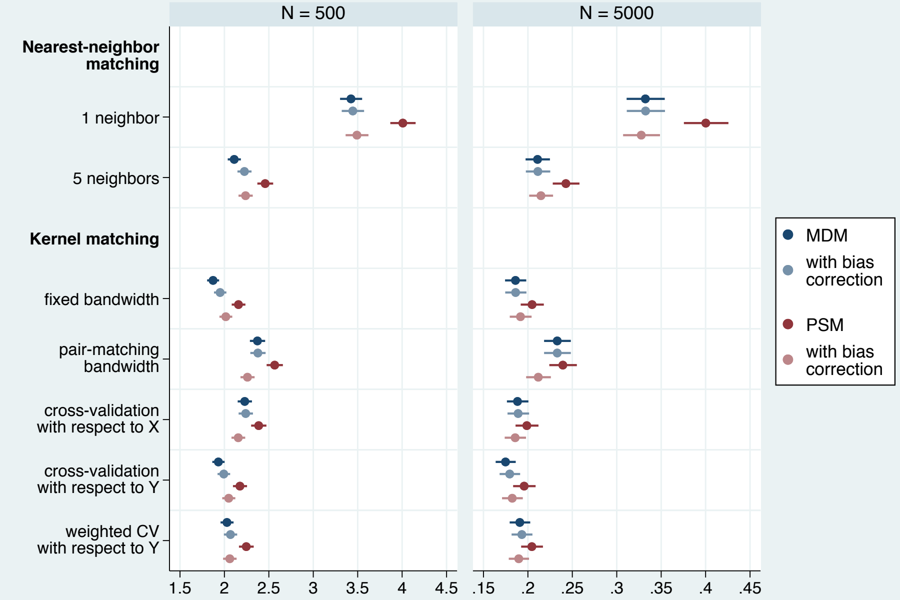
Simulation

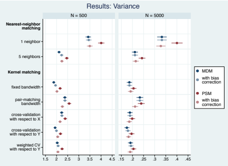
- Raw mean difference in occupational prestige (NATE): −4.79
- Population ATT (computed from fully stratified data): −3.96
- There is some treatment effect heterogeneity (ATE = −3.51, ATC = −3.41)





Results: Variance



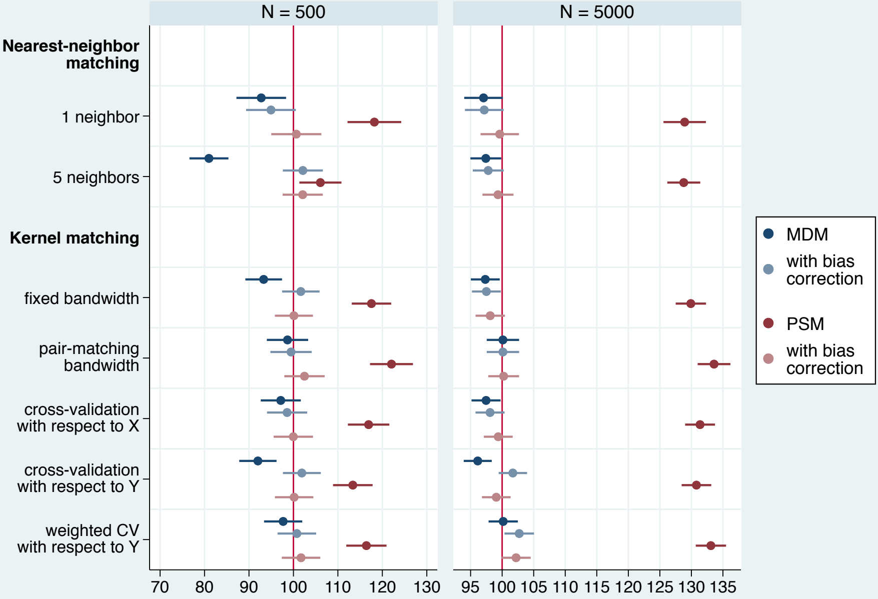


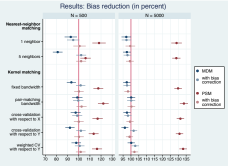
Results: Variance

In this slide we can see that for the same algorithm PSM typically is somewhat less efficient than MDM, but that across algorithms PSM can also be much more efficient than MDM. For example, kernel matching PSM has a much smaller variance than 1-nearest-neighbor MDM. That is, the choice of algorithm matters much more than the choice between PSM and MDM.

For kernel matching the efficiency differences between PSM and MDM are only small; additional post-matching regression adjustment further reduces the differences.

Results: Bias reduction (in percent)

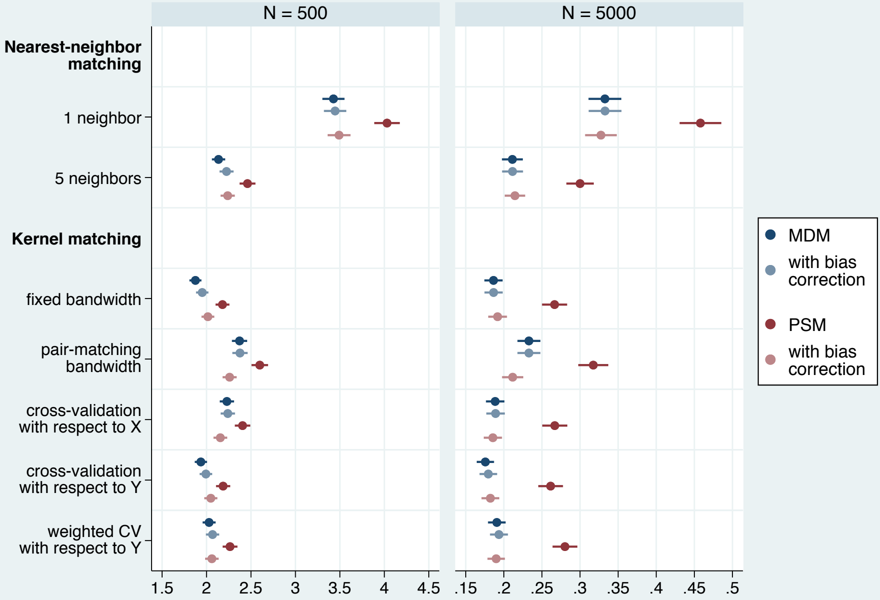




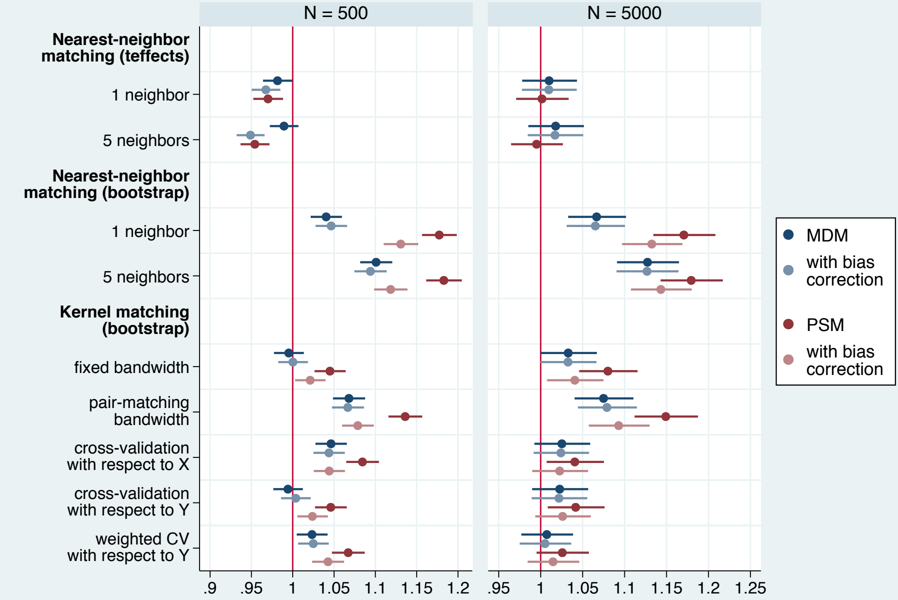
Here we see that PSM has a bias that does not vanish as the sample size increases. The reason is that the same propensity-score model specification is used for both sample sizes. The model is rather simple (linear effect of age, no interactions) and due to the specific pattern of the data (in particular, the sharp drop in the outcome variable after propensity score 0.3) small imprecisions can have substantial effects on the results. In practice, one would probably use a more refined specification in the large-sample situation, which would reduce bias.

The bias also vanishes once post-matching regression adjustment is applied.

Results: Mean squared error



Results: Relative standard error

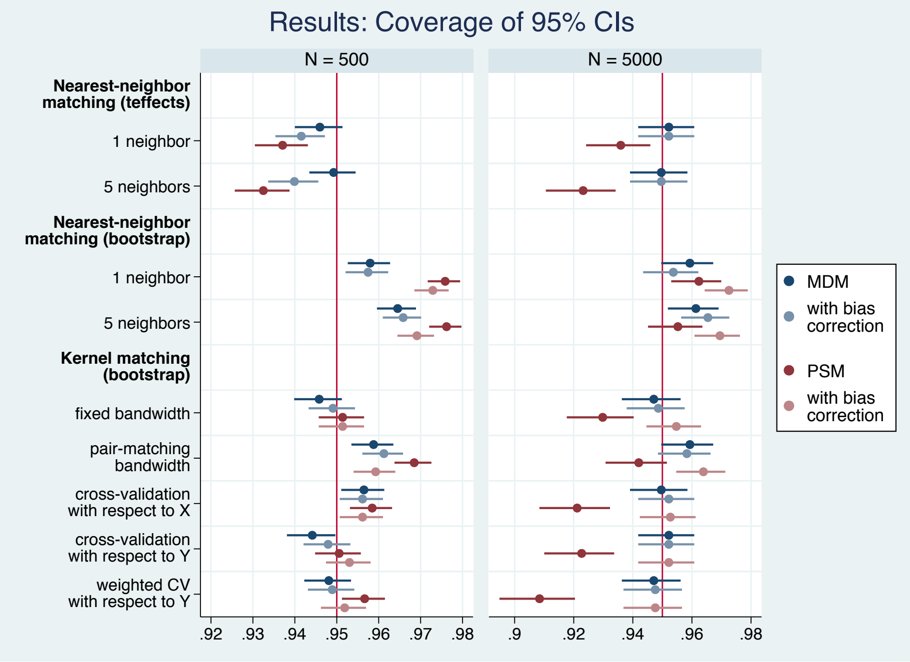


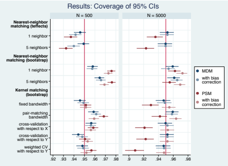
Results: Relative standard error

Here we can observe the well-known result that bootstrap standard errors are biased (too large) for nearest-neighbor matching.

In small samples, also the `teffects` standard errors seem to be slightly off (too low) for PSM and for MDM with bias-correction.

For kernel matching, bootstrap standard standard errors are often somewhat too large, especially in the small sample. The bias is most pronounced for the estimates using the pair-matching bandwidth selector. Results are better if the bandwidth is selected by cross-validation.


Results: Coverage of 95% CIs
N = 500
N = 5000
Nearest-neighbor matching (teffects)
1 neighbor
5 neighbors
Nearest-neighbor matching (bootstrap)
1 neighbor
5 neighbors
Kernel matching (bootstrap)
fixed bandwidth
pair-matching bandwidth
cross-validation with respect to X
cross-validation with respect to Y
weighted CV with respect to Y
.92
.93
.94
.95
.96
.97
.98
.9
.92
.94
.96
.98
MDM
with bias correction
PSM
with bias correction



Results: Coverage of 95% CIs

Coverage of `teffects` CIs is a bit too low for PSM (and for MDM with bias-correction in the small sample).

Bootstrap CIs are too conservative for nearest-neighbor matching.

For kernel matching, coverage is mostly okay, being a bit too conservative in case of the pair-matching bandwidth selector and considerably off (anti-conservative) for the PSM estimates without bias-correction (due to the pronounced bias in these estimates).

Conclusions

- Overall, I agree with King and Nielsen that MDM has some advantages over PSM, but it also has some disadvantages. In applied research the choice may not be that clear.
  - 👍 MDM leaves less scope for bias due to post-matching modeling decisions.
  - 👍 Theoretical results (see, e.g., Frölich 2007) suggest that MDM will generally tend to outperform PSM in terms of efficiency (but differences are likely to be small).
  - 👍 Less restrictions in terms of possible post-matching analyses.
  - 👎 Choice of scaling matrix largely arbitrary.
  - 👎 Computational complexity.
- One clear conclusion we can draw, however, is:

Do not use propensity scores for pair matching!

(But don't use pair matching anyhow.)

Conclusions

- Some additional conclusions from the simulation
  - For PSM, application of regression-adjustment seems like a great idea (reduction of bias and variance); for MDM the advantages of regression-adjustment are less clear.
  - Bootstrap standard error/confidence interval estimation seems to be mostly ok for kernel/ridge matching; this is in contrast to nearest-neighbor matching, where bootstrap standard errors are clearly biased.
- To do
  - Run some more simulations.
  - Variance estimation based on influence functions?
  - Better (and faster) bandwidth selection algorithms?
  - Explore potential of adaptive bandwidths?

References I

- Cochran, W.G. 1968. The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies. *Biometrics* 24(2):295–313.
- Frölich, M. 2004. Finite-sample properties of propensity-score matching and weighting estimators. *The Review of Economics and Statistics* 86(1):77–90.
- Frölich, M. 2005. Matching estimators and optimal bandwidth choice. *Statistics and Computing* 15:197-215.
- Frölich, M. 2007. On the inefficiency of propensity score matching *AStA* 91:279–290.
- Galdo, J.C., J. Smith, D. Black. 2008. Bandwidth selection and the estimation of treatment effects with unbalanced data. *Annales d'Économie et de Statistique* 91/92:89-216.

References II

- Hendrickx, J. 2002. ISKO: Stata module to recode 4 digit ISCO-88 occupational codes. Statistical Software Components S425802, Boston College Department of Economics.

- Huber, M., M. Lechner, A. Steinmayr. 2015. Radius matching on the propensity score with bias adjustment: tuning parameters and finite sample behaviour. *Empirical Economics* 49:1-31.

- Huber, M., M. Lechner, C. Wunsch. 2013. The performance of estimators based on the propensity score. *Journal of Econometrics* 175:1-21.

- King, G., R. Nielsen. 2016. Why Propensity Scores Should Not Be Used for Matching. Working Paper. Available from http://j.mp/1sexgVw.

- Rosenbaum, P.R., D.B. Rubin. 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* 70:41–55.