PoliSci 4782 Political Analysis II

Matching

Seamus Wagner

The Ohio State University

Idea of Matching

A preprocessing step to find hidden randomized/fully blocked experiments in observational data

References:

- King, Gary, "Matching." Gov2001 Quantitative Social Science Method I (2020). Harvard University: `https://projects.iq.harvard.edu/gov2001/lecture-videos`
- King, Gary and Richard Nielsen. "Why Propensity Scores Should Not Be Used for Matching." Political Analysis 27, 4 (May 2019): 435–454

Outline

How to find matched observations or remove unmatched controls?

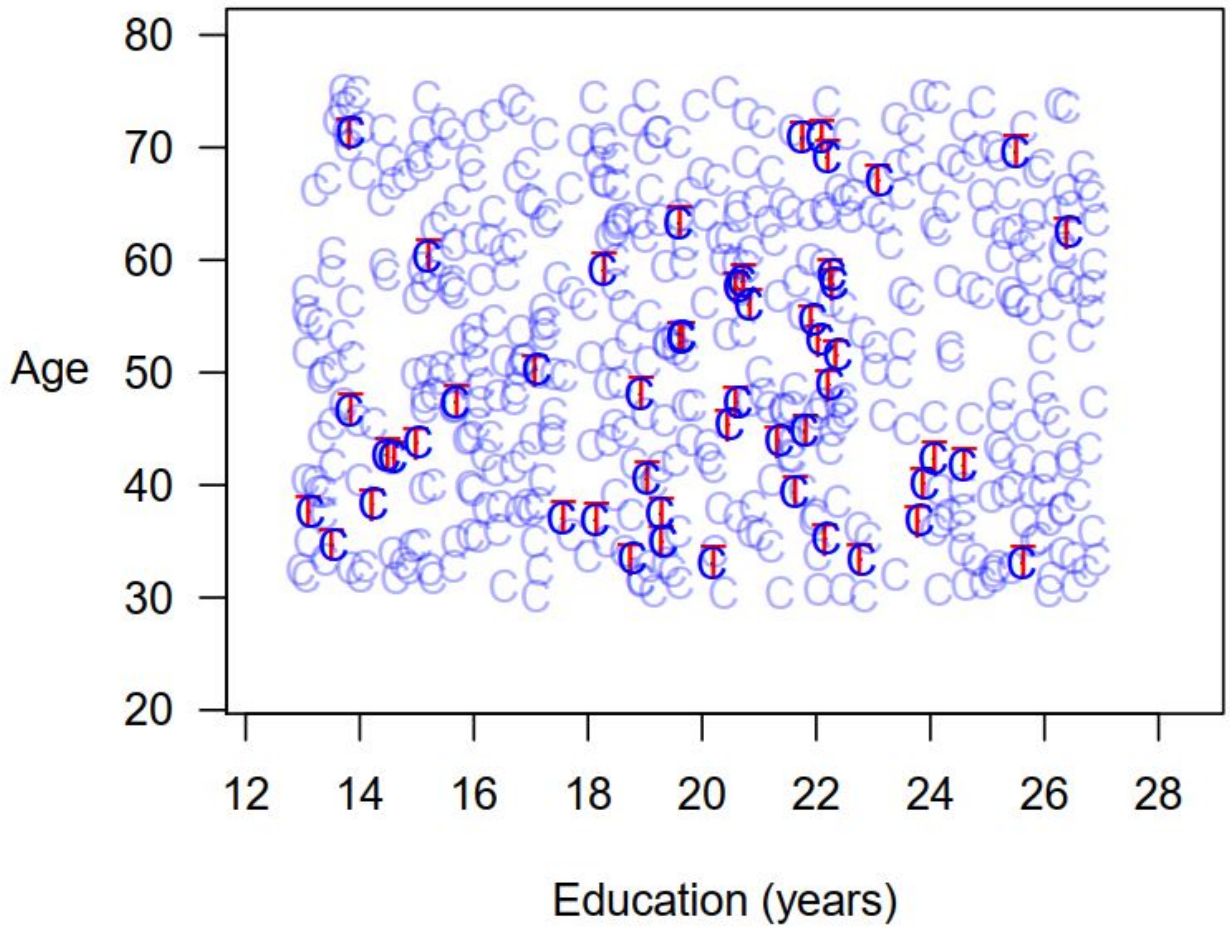
- Exact matching
- Mahalanobis distance matching
- Coarsened exact matching
- Propensity score matching

Exact Matching

Compare and preprocess observations on the values of covariates X

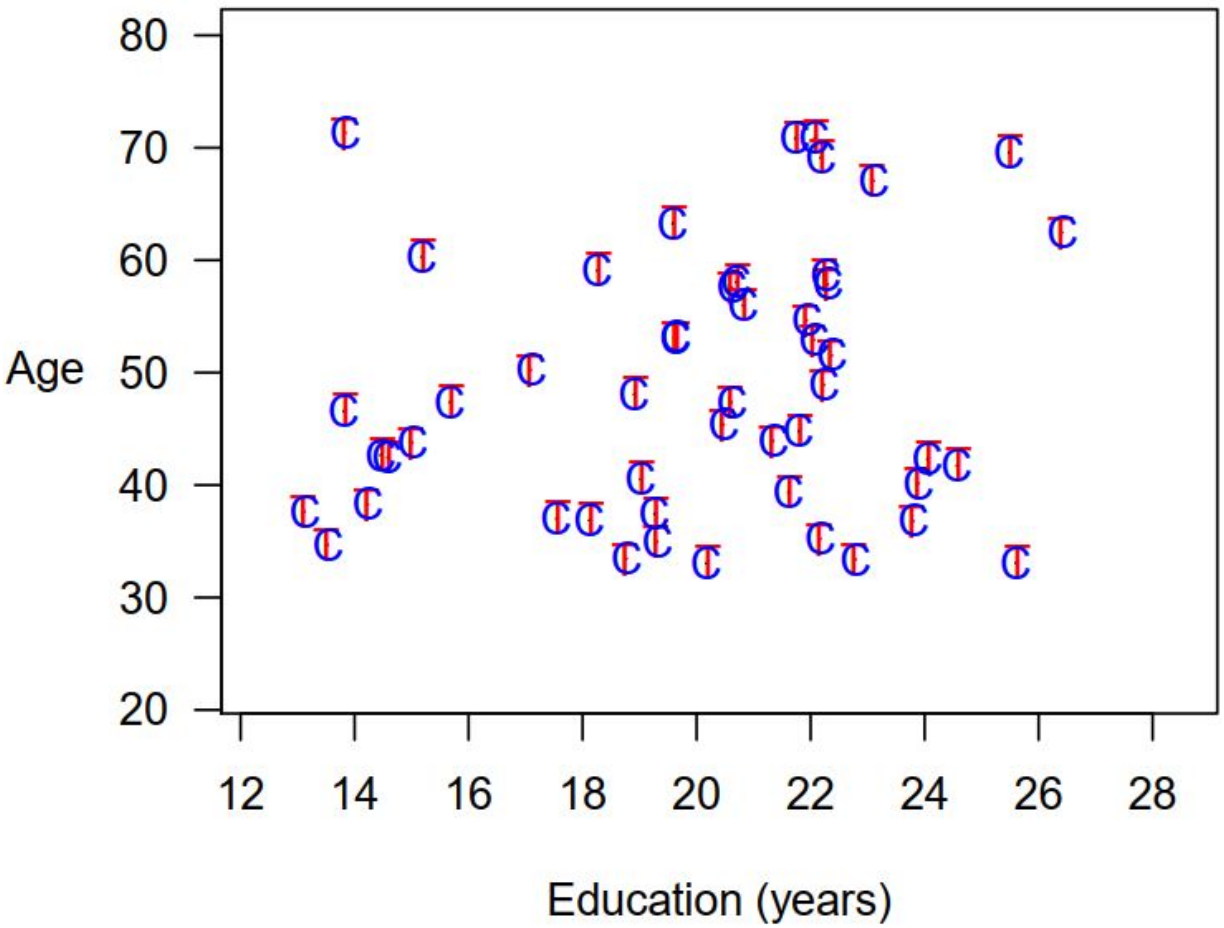
- Match each treated and control units at exact values in X
- Remove unmatched units
- Might not be feasible

Exact Matching



Source: King (2020)

Exact Matching



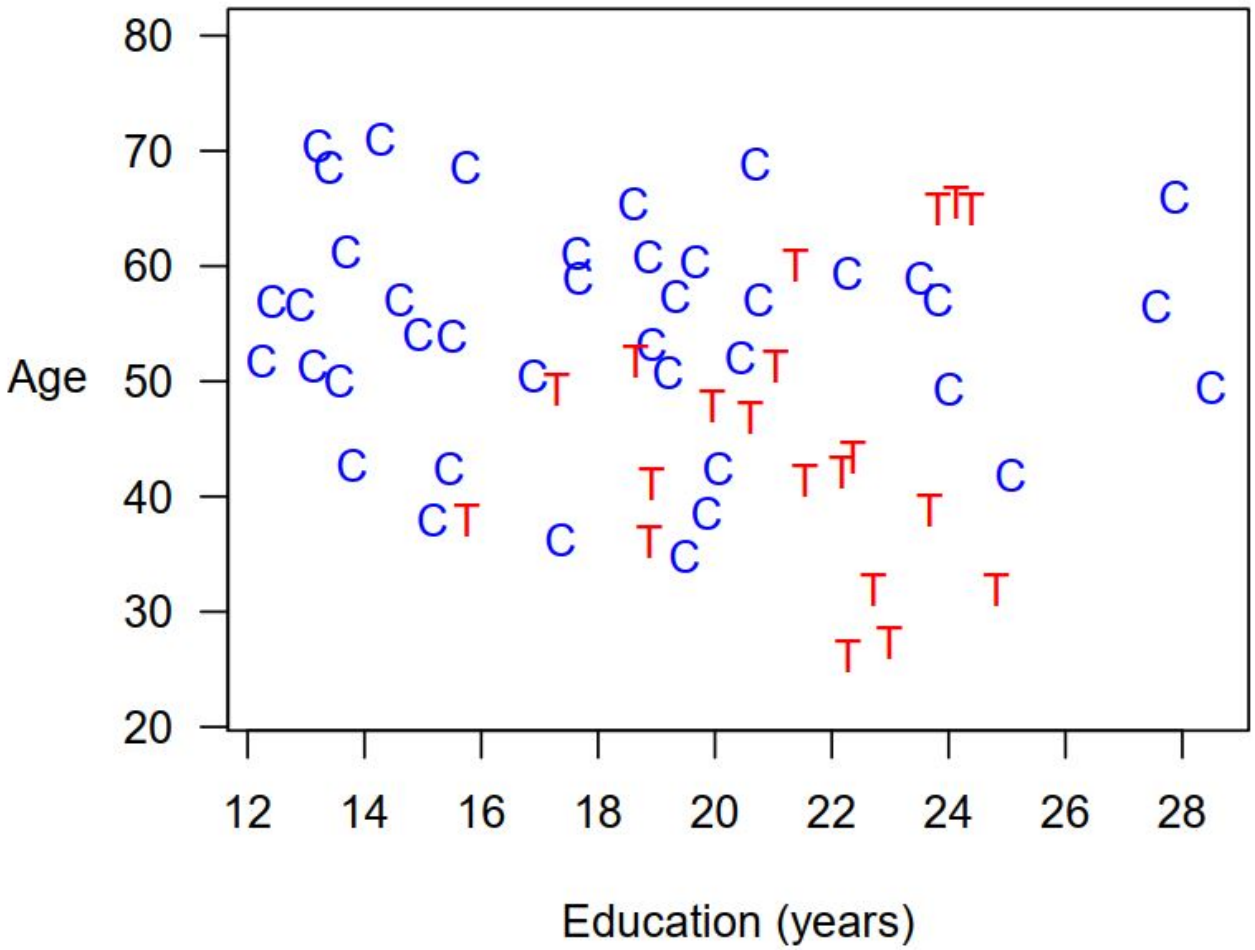
Source: King (2020)

Mahalanobis Distance Matching

Compare and preprocess observations on the distance measure of covariates X

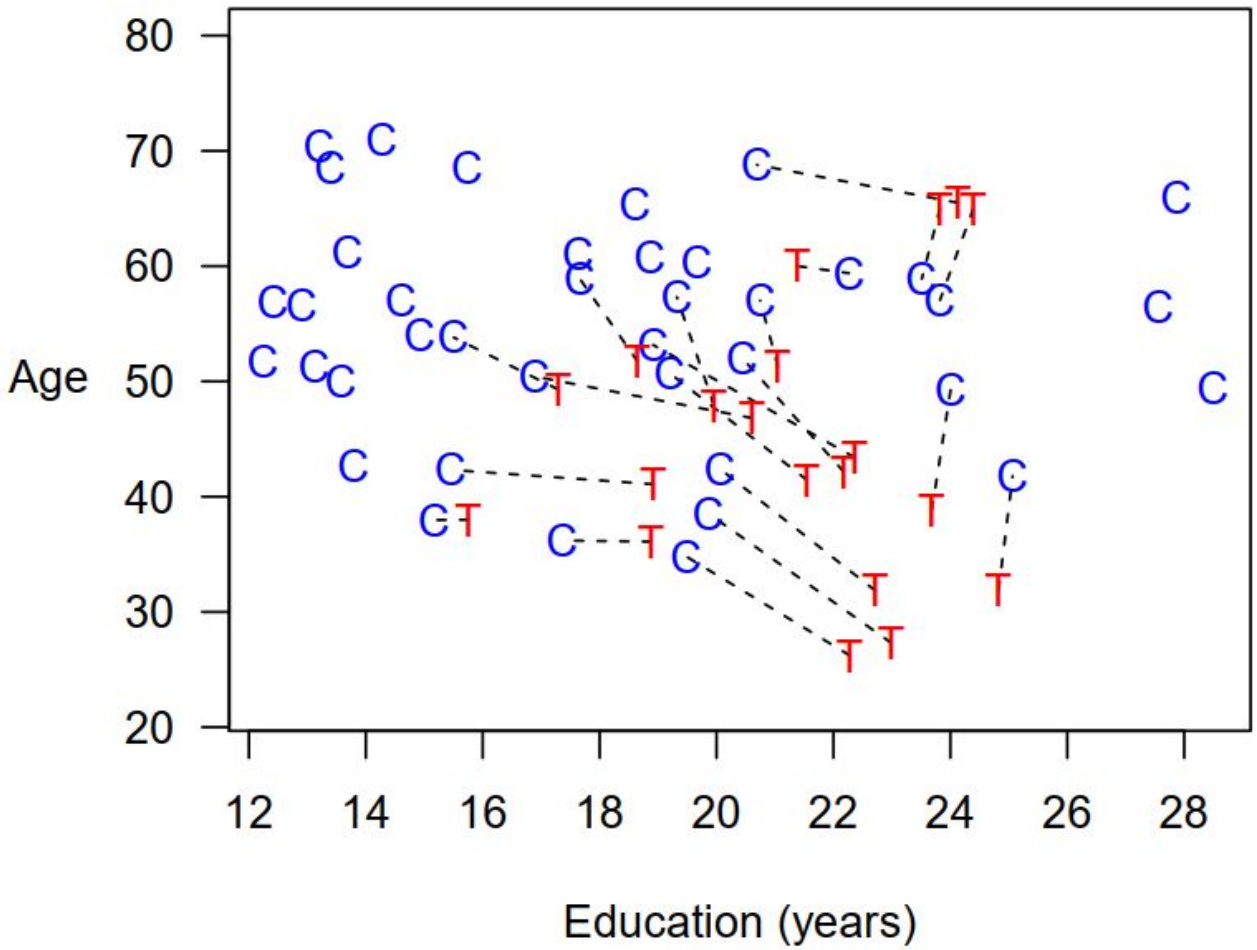
- Mahalanobis distance: $\sqrt{(X_c - X_t)' S^{-1} (X_c - X_t)}$
  - it is Euclidean distance with standardization
- Match each treated unit to the nearest control unit
- Set a caliper and remove matched pairs when distance $>$ caliper
- Remove unused control units

Mahalanobis Distance Matching



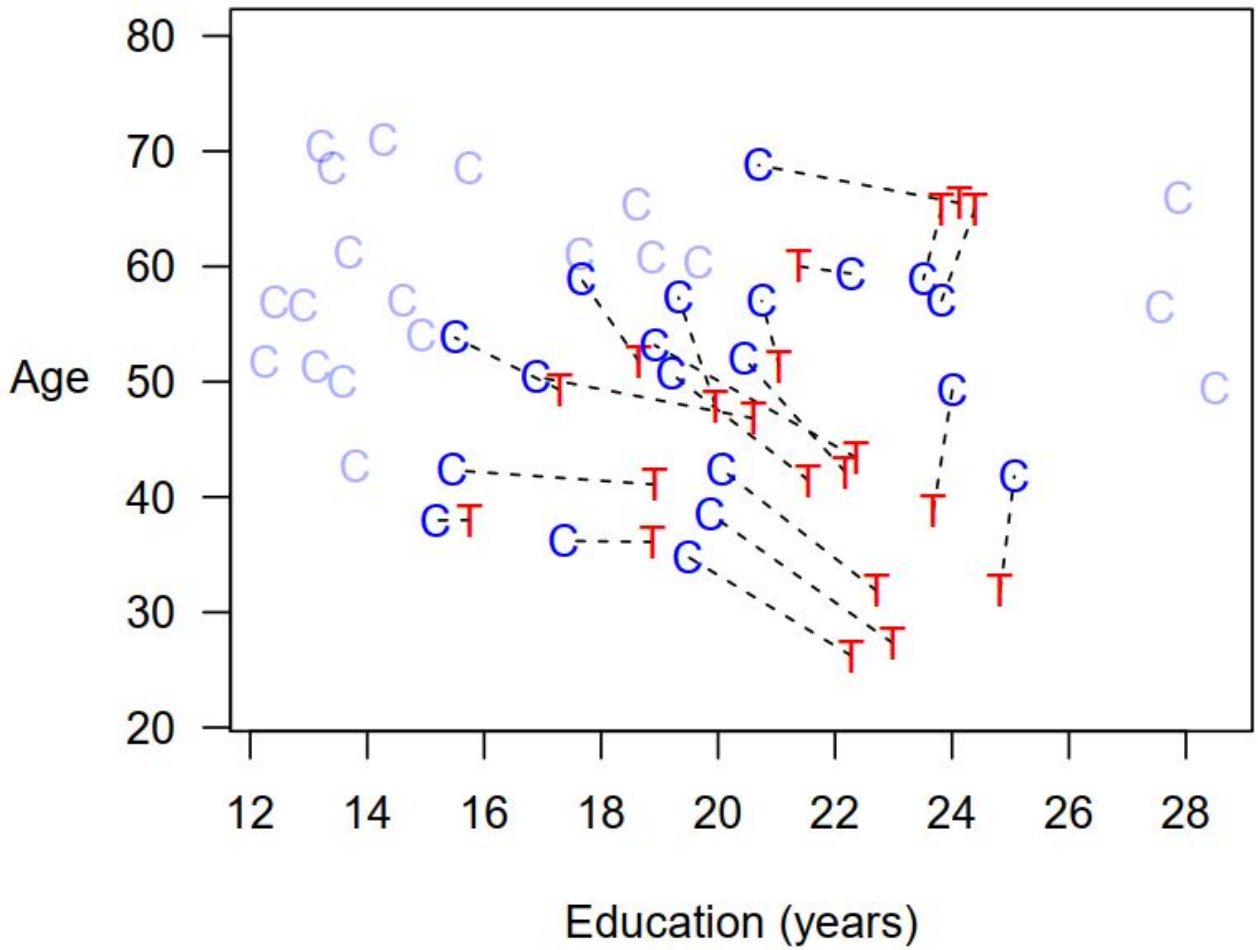
Source: King (2020)

Mahalanobis Distance Matching



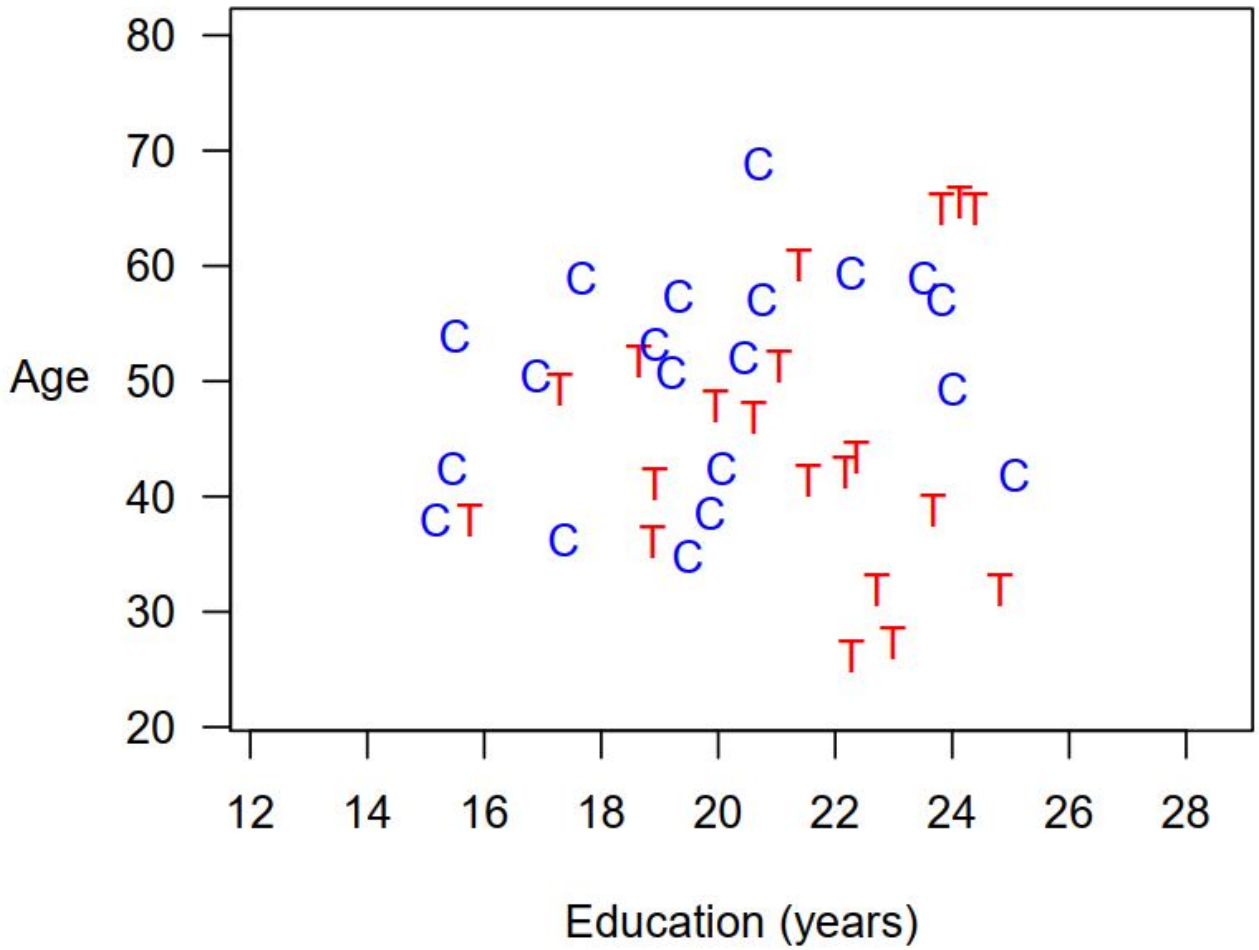
Source: King (2020)

Mahalanobis Distance Matching



Source: King (2020)

Mahalanobis Distance Matching



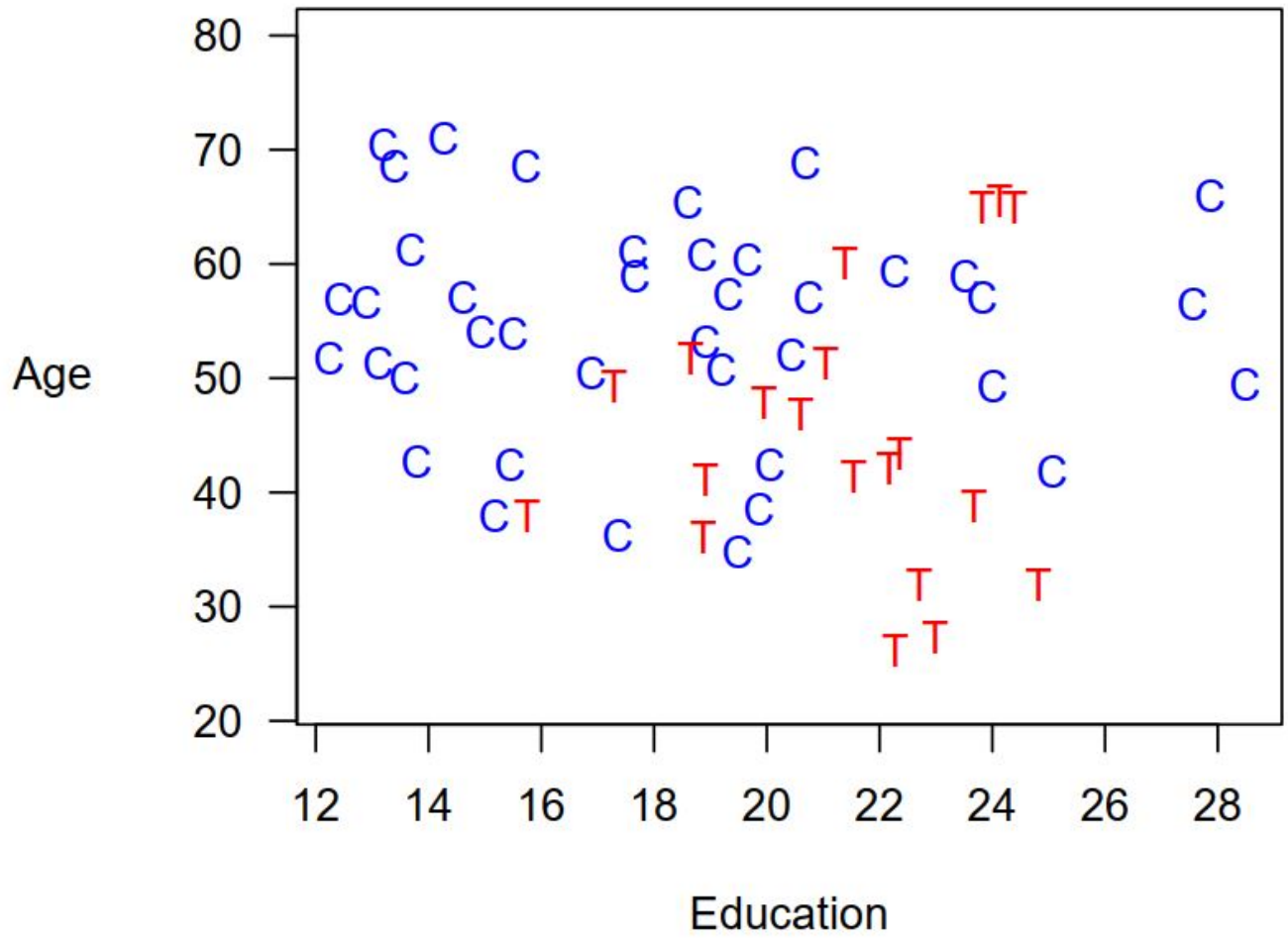
Source: King (2020)

Coarsened Exact Matching

Compare and preprocess observations on the coarsened values of covariates X

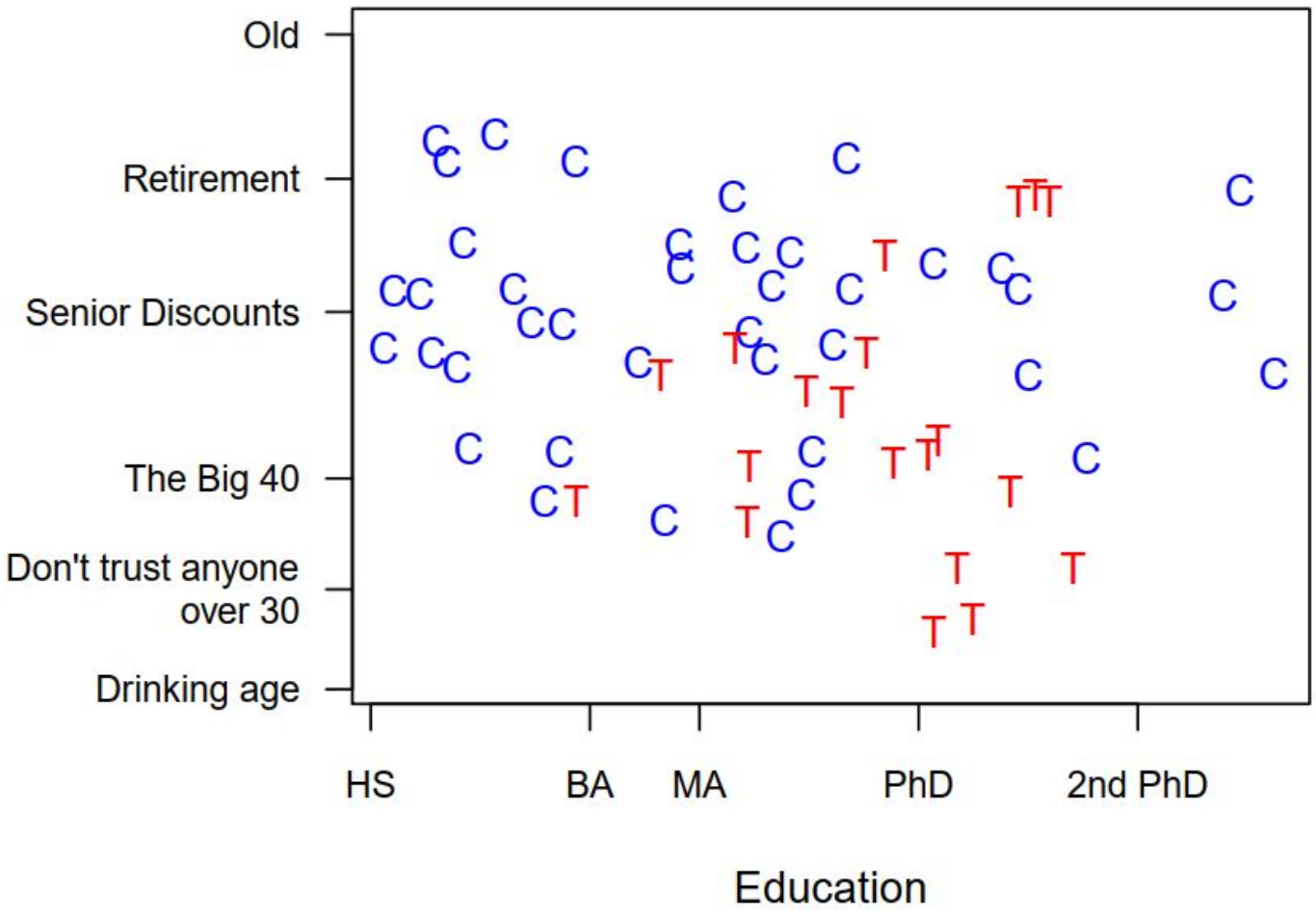
- Temporarily coarsen X to a proper level (into different strata)
  - household income value --→ low-, middle-, high-income
- Apply exact matching to the coarsened X
- Remove strata with 0 treated or 0 control units
- Pass on original values (except those pruned)
- Consider weighting controls in each stratum to equal treateds

Coarsened Exact Matching



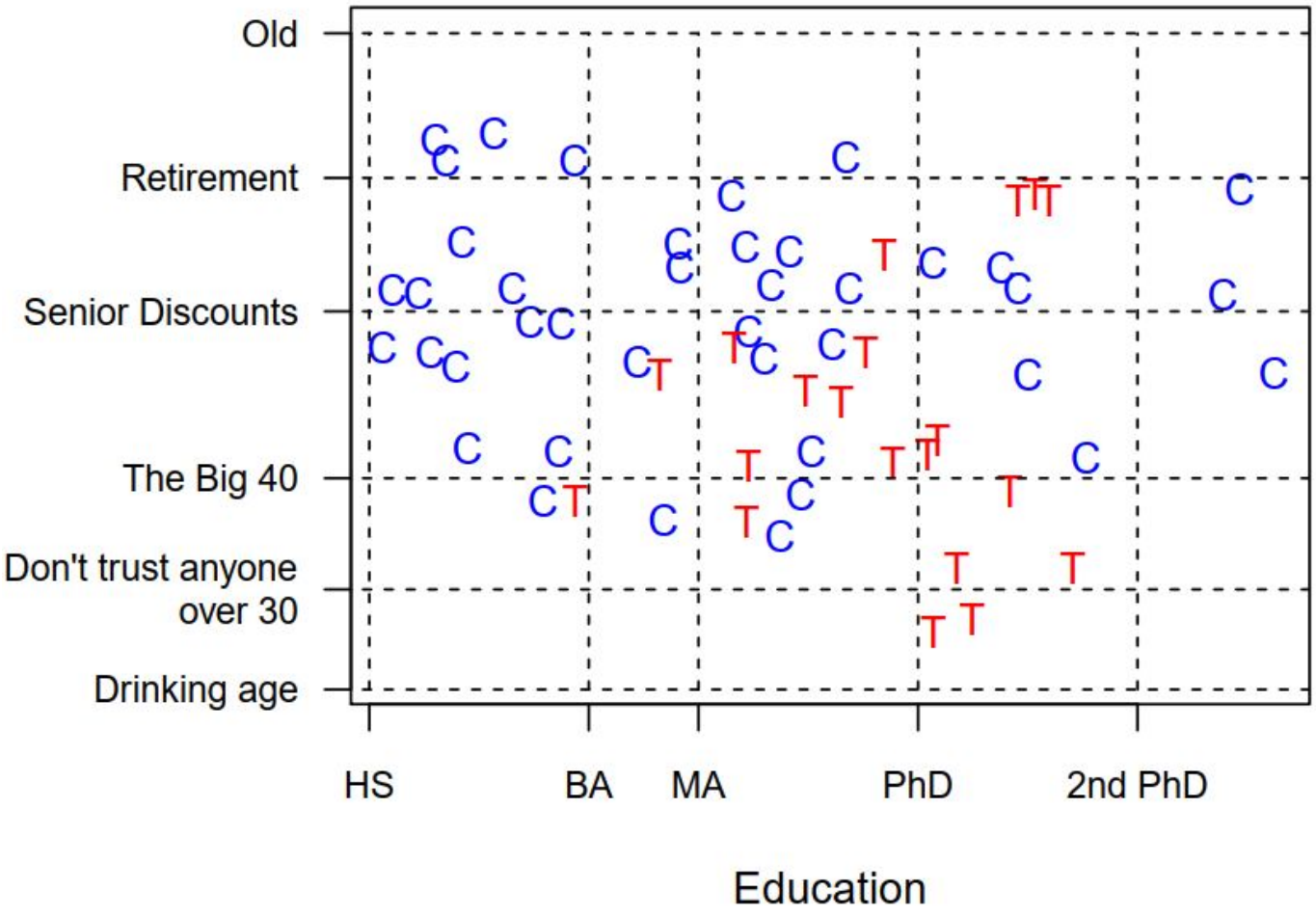
Source: King (2020)

Coarsened Exact Matching



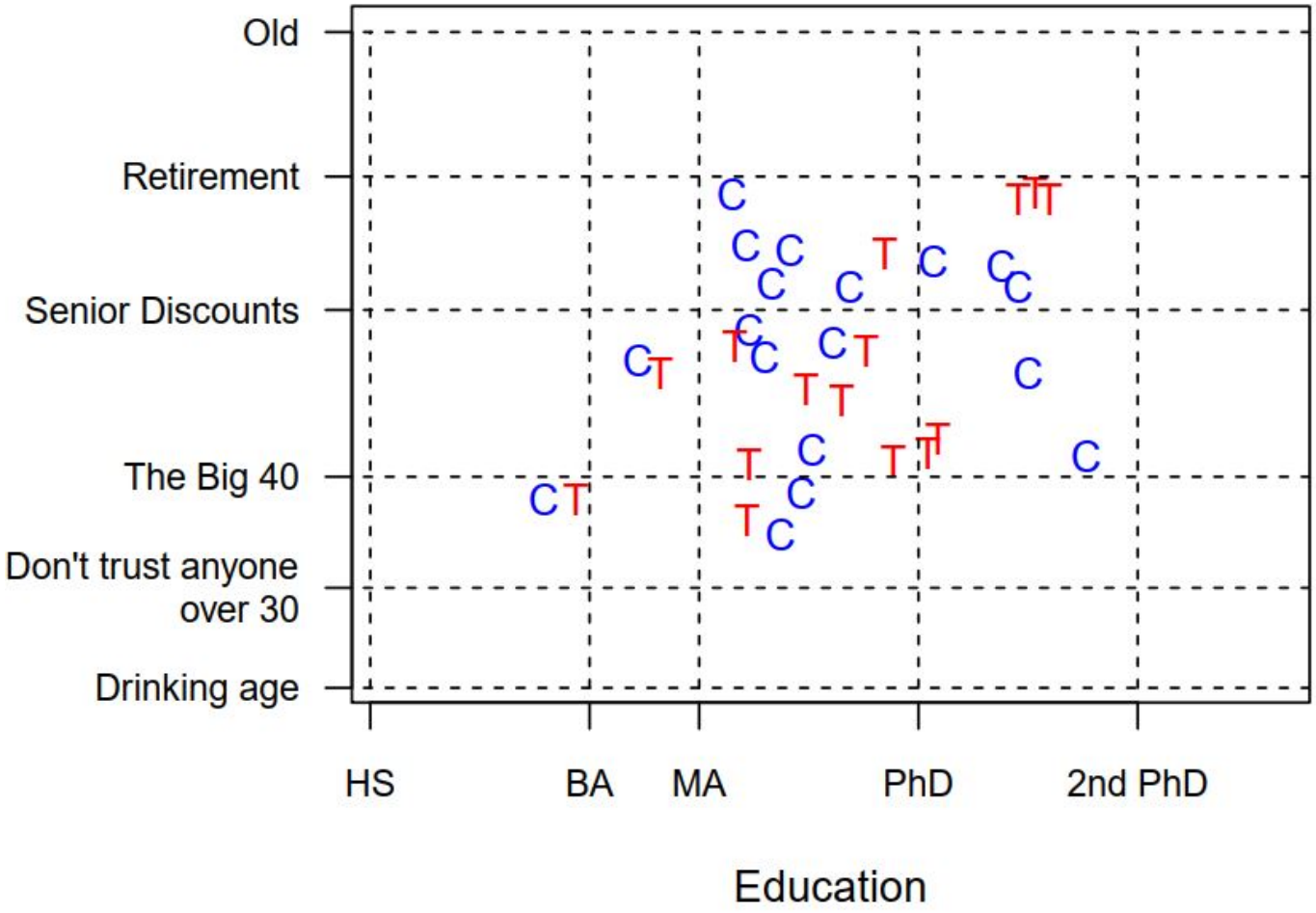
Source: King (2020)

Coarsened Exact Matching



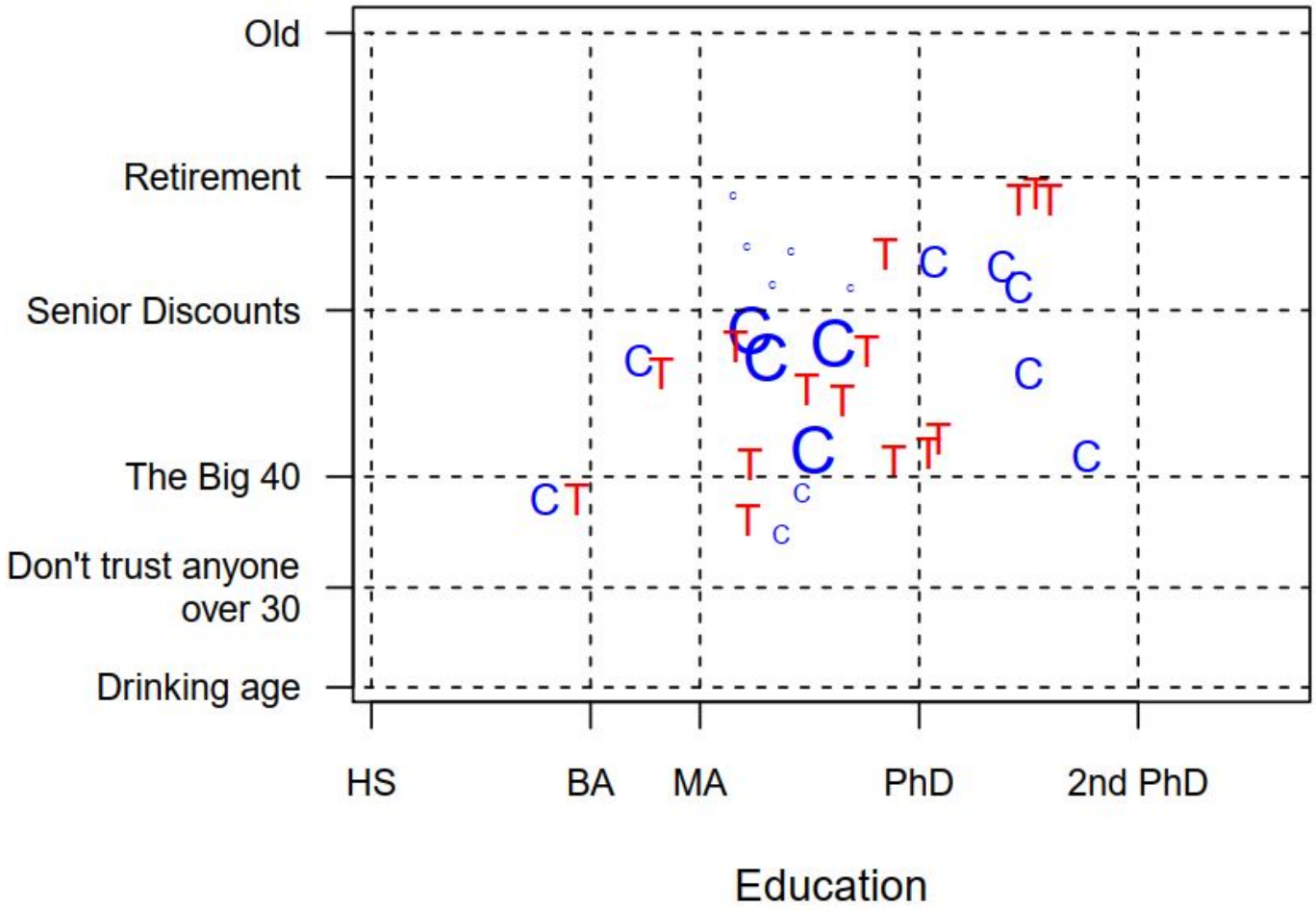
Source: King (2020)

Coarsened Exact Matching



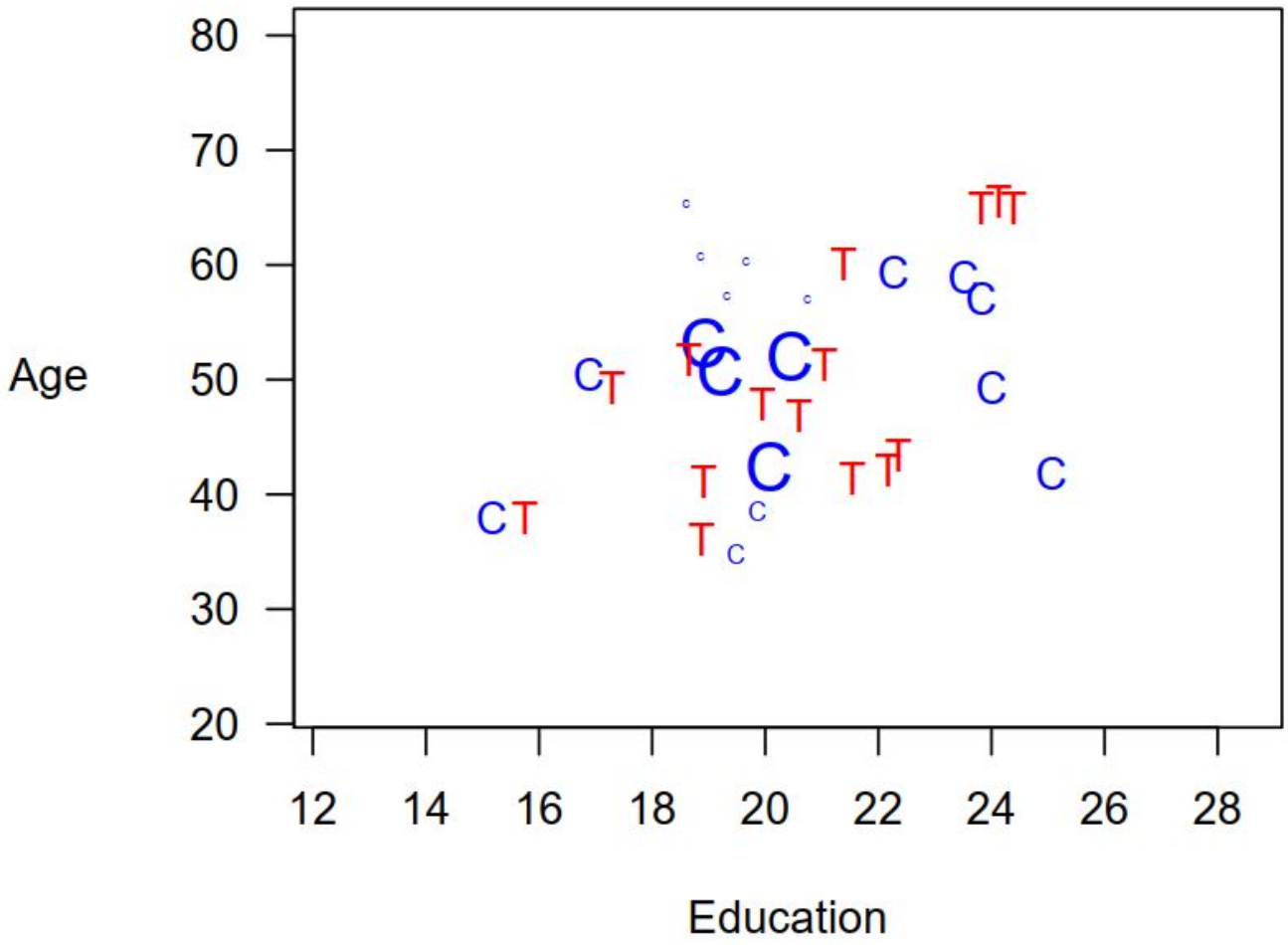
Source: King (2020)

Coarsened Exact Matching



Source: King (2020)

Coarsened Exact Matching



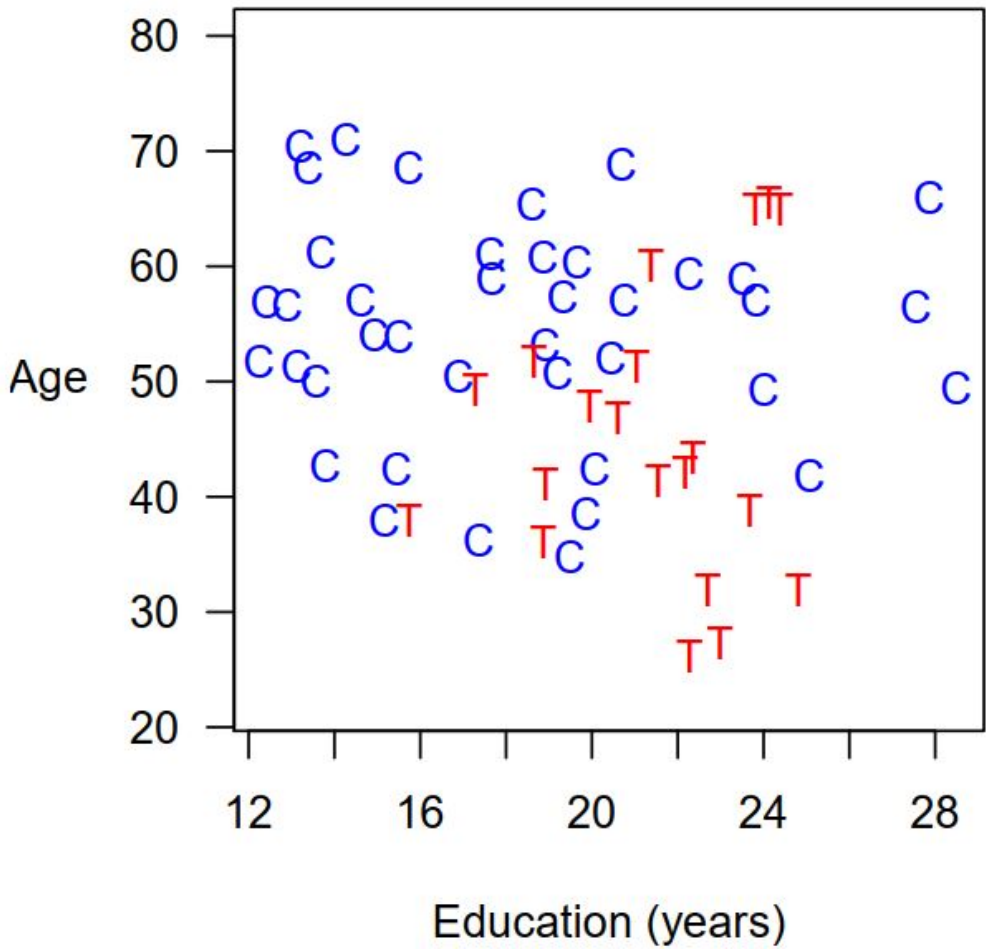
Source: King (2020)

Propensity Score Matching

Compare and preprocess observations on the propensity score distance based on X

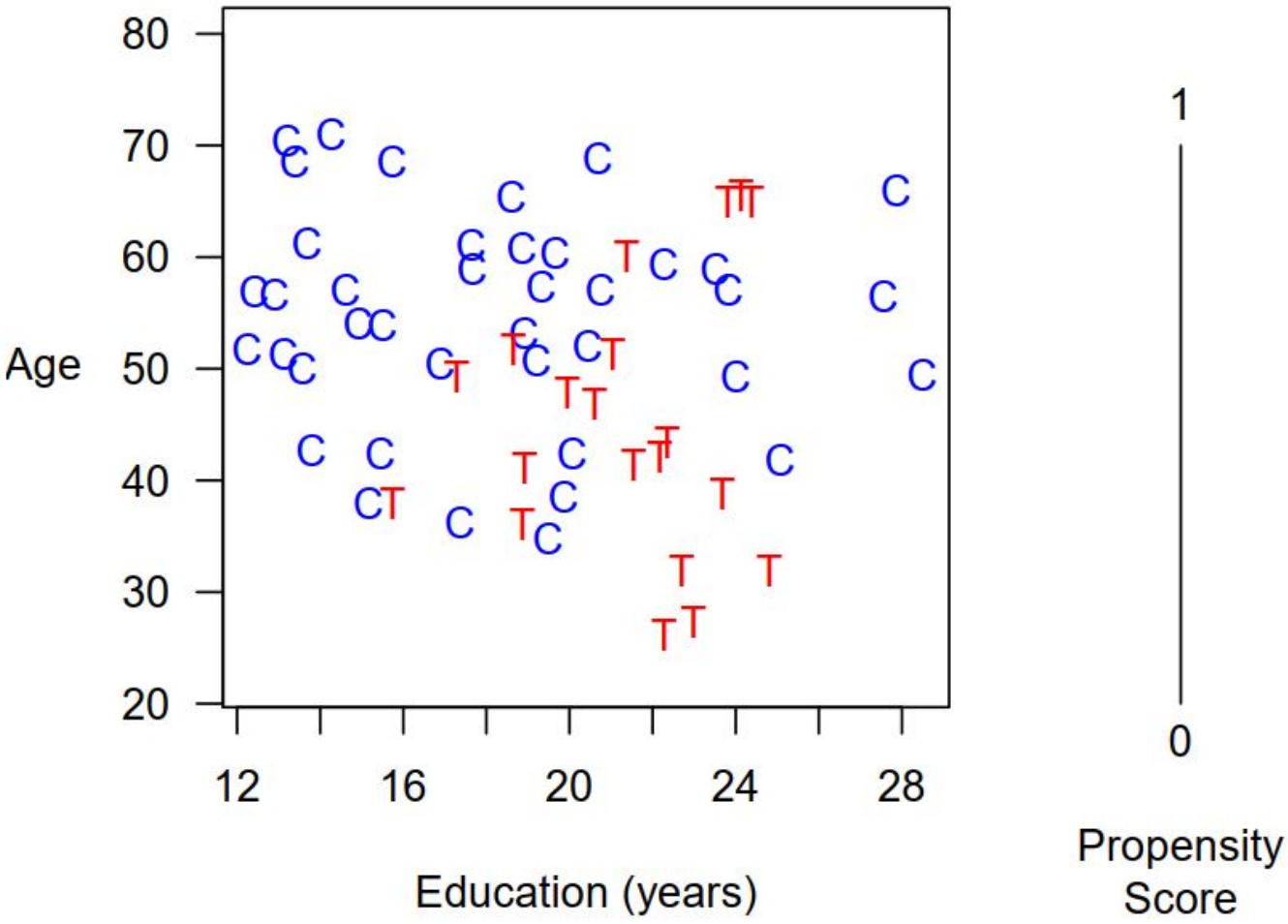
- Propensity score: $\pi = Pr(T_i = 1|X) = \frac{1}{1+e^{-X\beta}}$
  - it is a model-based probability that an observation gets treated
- Distance in propensity score: $|\pi_c - \pi_t|$
- Match each treated unit to the nearest control unit in propensity score
- Remove unused controls
- Set a caliper and remove matches if distance $>$ caliper

Propensity Score Matching



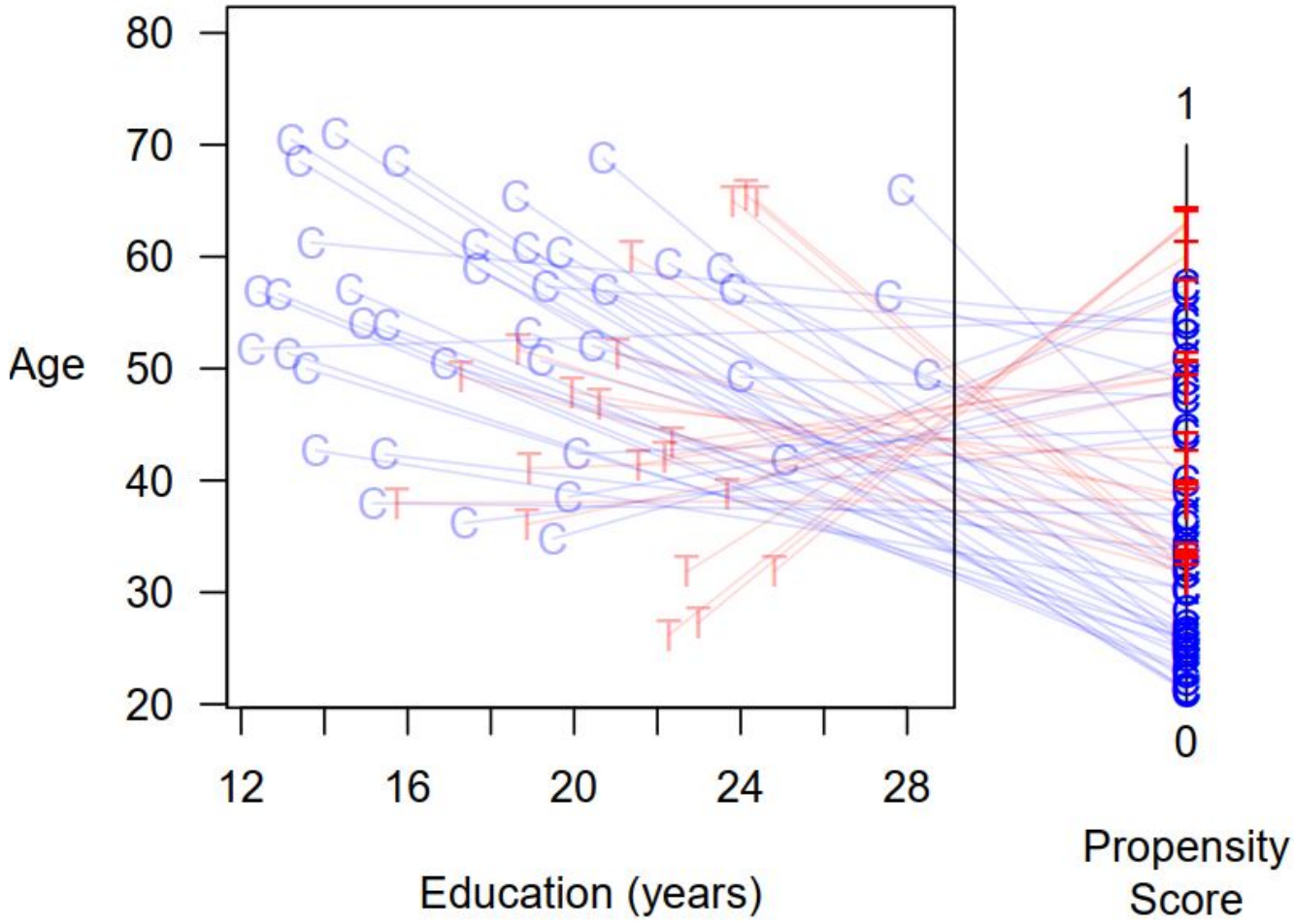
Source: King (2020)

Propensity Score Matching



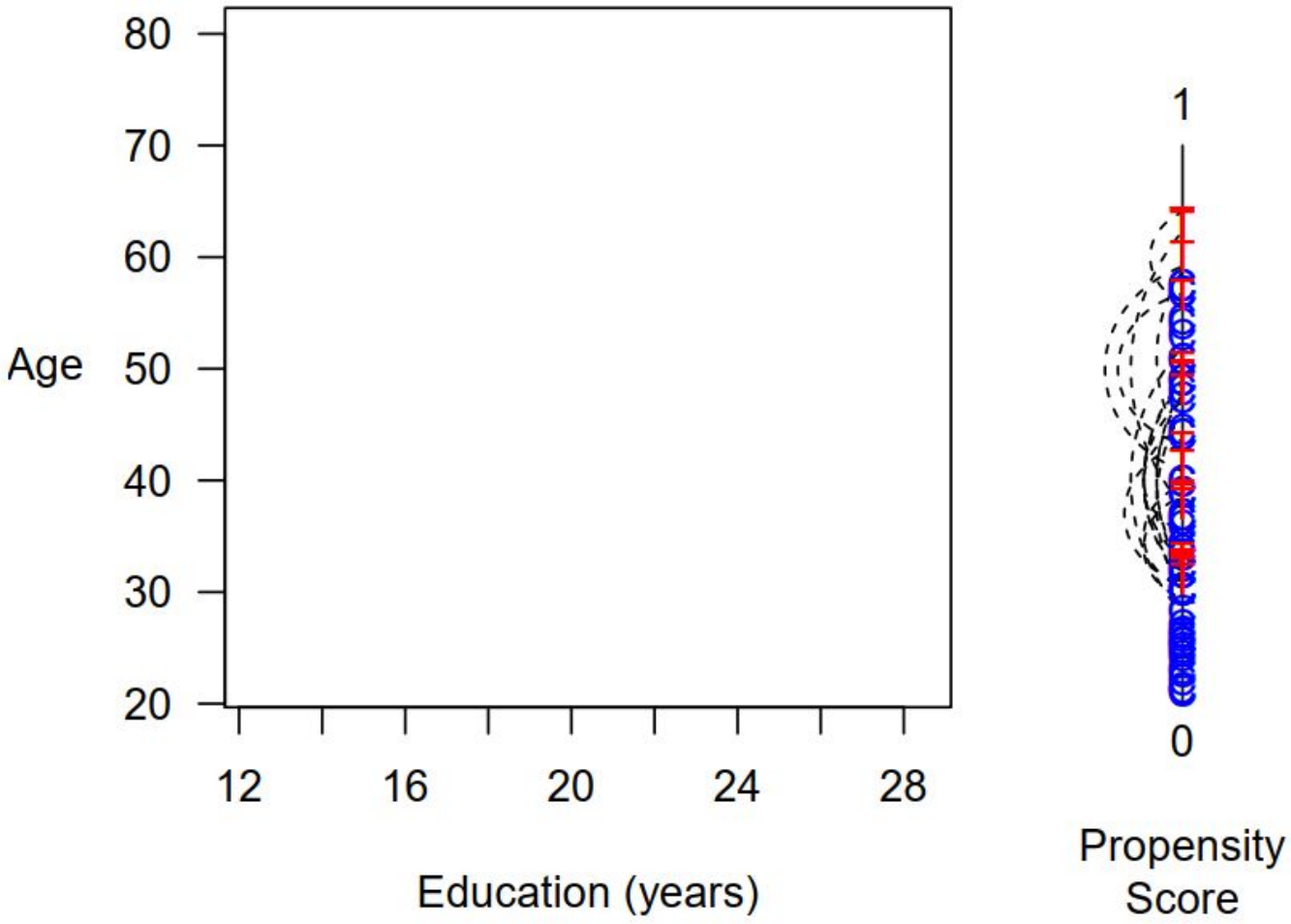
Source: King (2020)

Propensity Score Matching



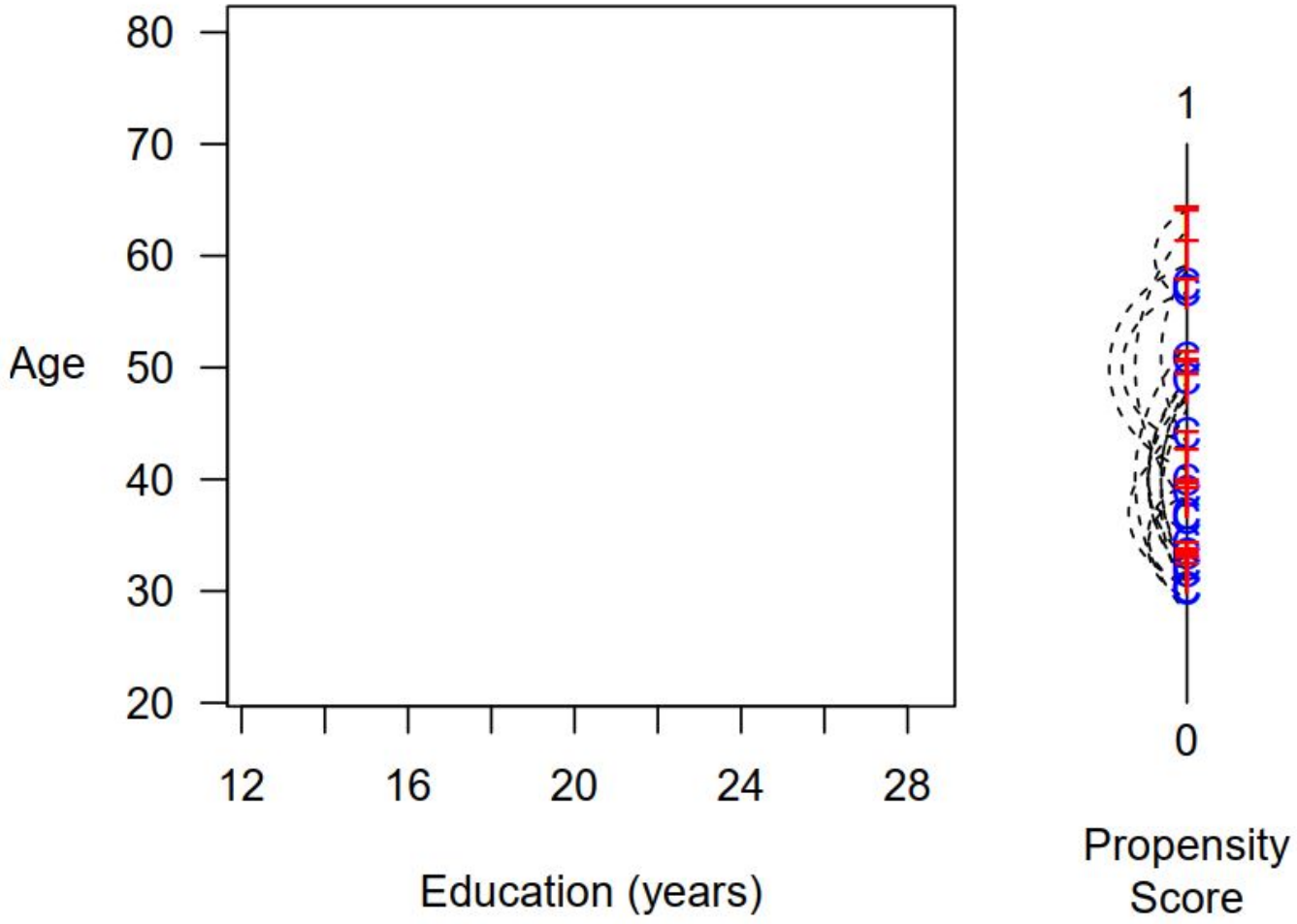
Source: King (2020)

Propensity Score Matching



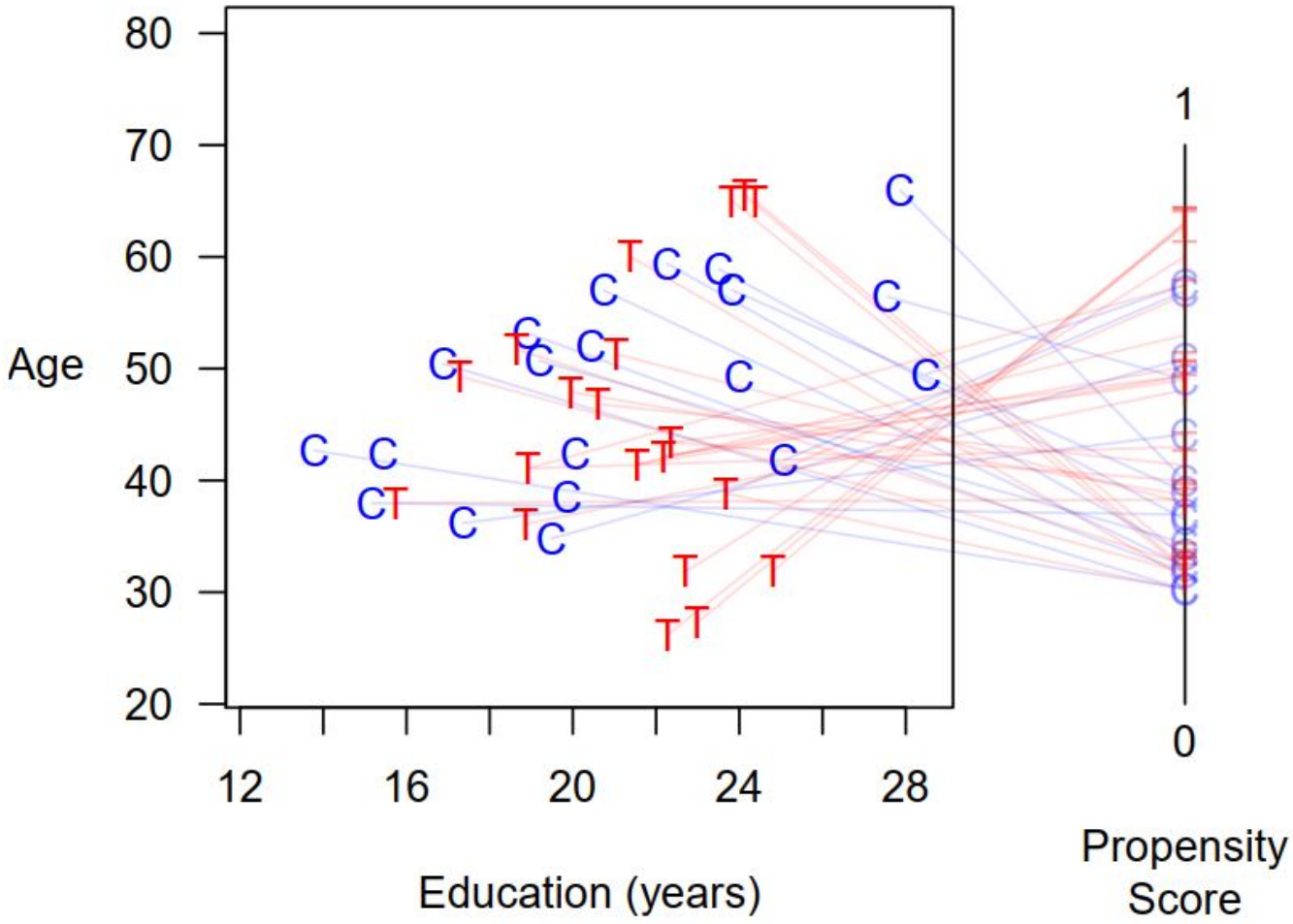
Source: King (2020)

Propensity Score Matching



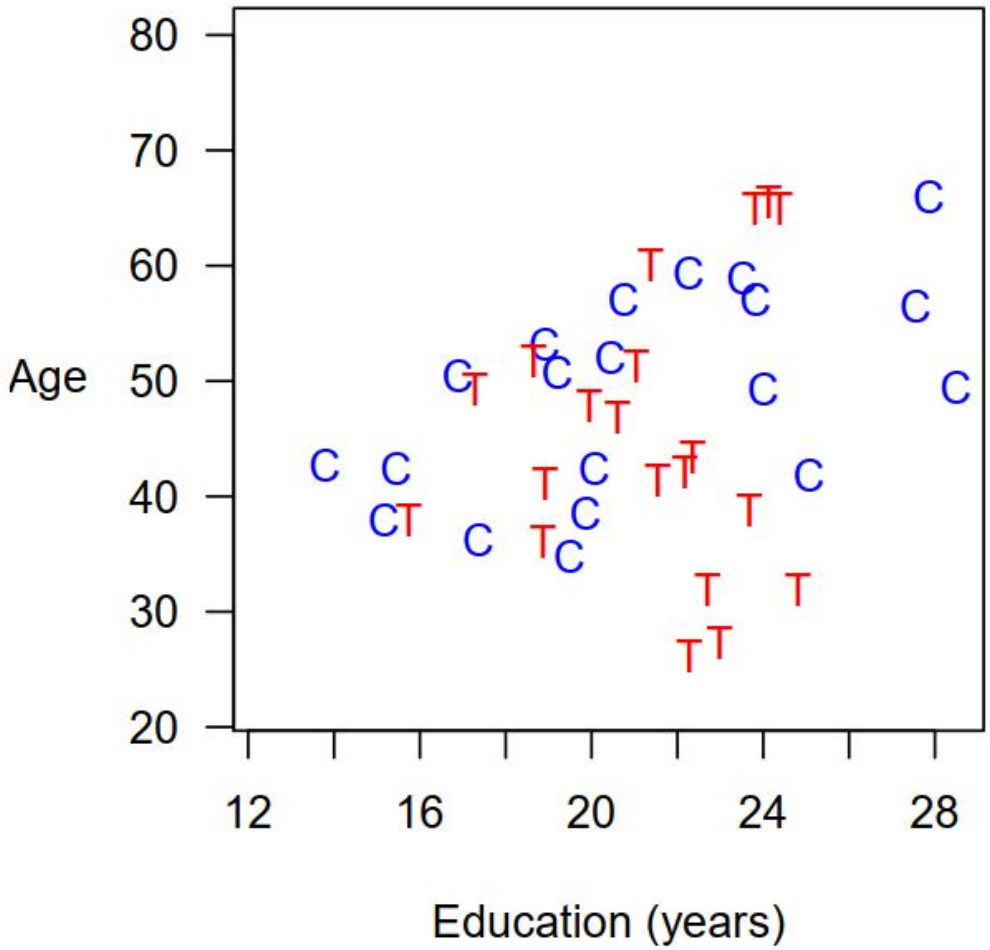
Source: King (2020)

Propensity Score Matching



Source: King (2020)

Propensity Score Matching

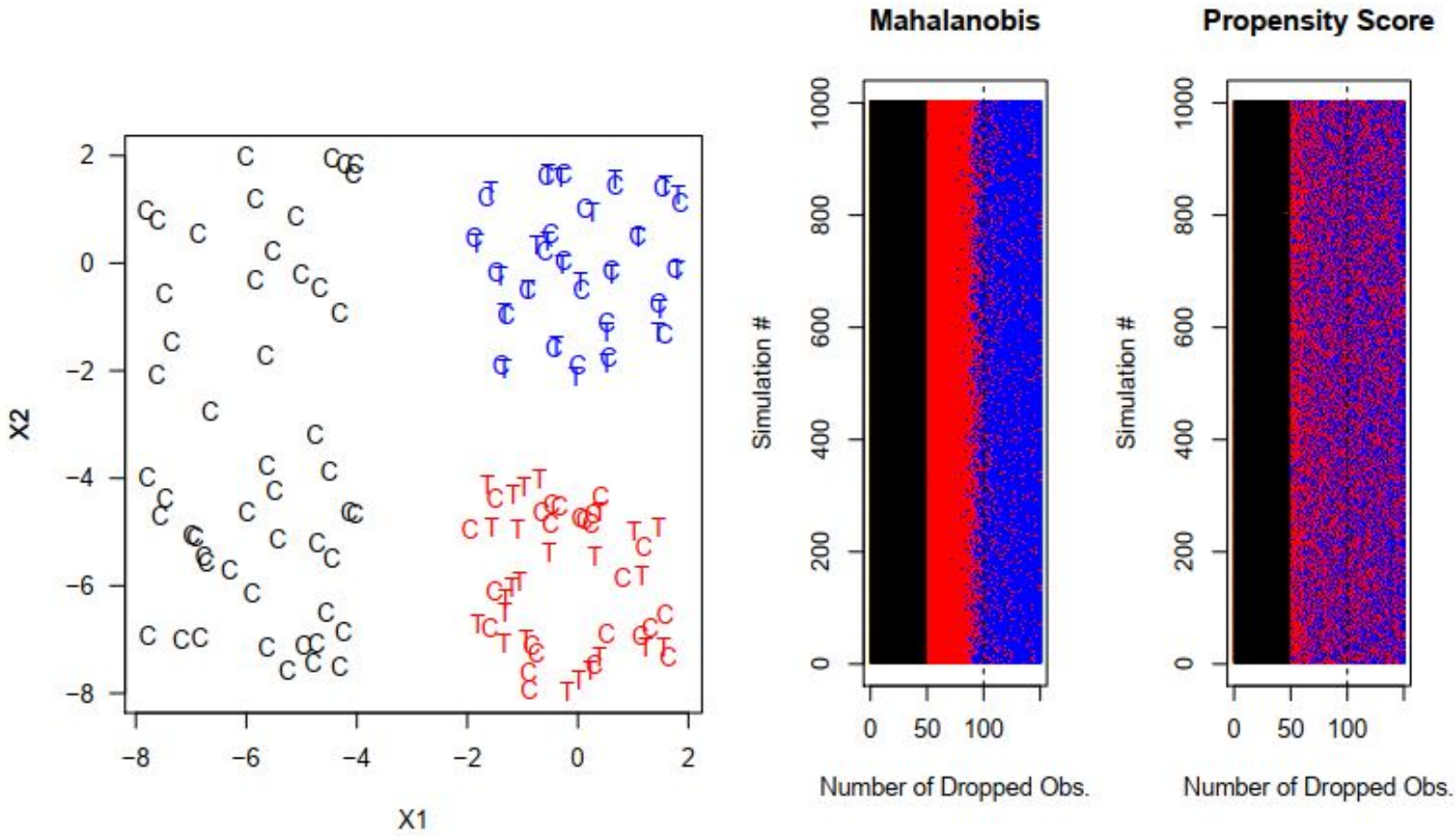


Source: King (2020)

Summary

- To mimic fully blocked experiment:
  - exact matching, Mahalanobis distance matching, coarsened exact matching
- To mimic randomized experiment:
  - propensity score matching

Propensity Score Matching Are Not Reliable



Source: King & Nielsen (2019)

Coming Up

- Lab tutorial on matching this week
- Hierarchical models in the next week (last topic)
- Online exam (Finals week, one-entry, fixed duration, flexible schedule)