FLS 6415 - Causal Inference for the Political Economy of Development

Week 10 - Various & Matching

Jonathan Phillips

October 2017

Matching

- An alternative solution to the confounding problem

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals
  - Matching should really be called **'pruning'**

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals
  - Matching should really be called **'pruning'**
  - That slightly changes the causal effect we're estimating (eg. 'Average Treatment Effect on the Treated')

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals
  - Matching should really be called **'pruning'**
  - That slightly changes the causal effect we're estimating (eg. 'Average Treatment Effect on the Treated')
  - But allows us to have more confidence our effects are causal

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals
  - Matching should really be called **'pruning'**
  - That slightly changes the causal effect we're estimating (eg. 'Average Treatment Effect on the Treated')
  - But allows us to have more confidence our effects are causal
- As with regression, it succeeds only to the extent we **match on all confounders**

Matching

- An alternative solution to the confounding problem
- Matching **forces balance** between treated and control units
- It does so by **dropping units** that don't have good counterfactuals
  - Matching should really be called **'pruning'**
  - That slightly changes the causal effect we're estimating (eg. 'Average Treatment Effect on the Treated')
  - But allows us to have more confidence our effects are causal
- As with regression, it succeeds only to the extent we **match on all confounders**
- Unmeasured confounders are a big problem still

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance
- Matching is similar:

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance
- Matching is similar:

1. For each treated unit, find a control unit with very close values of confounding variables, and keep both

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance
- Matching is similar:

1. For each treated unit, find a control unit with very close values of confounding variables, and keep both
2. Repeat for every treated unit

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance
- Matching is similar:

1. For each treated unit, find a control unit with very close values of confounding variables, and keep both
2. Repeat for every treated unit
3. Drop all the unmatched units (eg. ’extra’ control units that are ’far away’ from treated units)

Matching

- Controlling in a regression is like separating our dataset into male and female, and comparing treated and control groups separately for each, so we are guaranteed balance
- Matching is similar:

1. For each treated unit, find a control unit with very close values of confounding variables, and keep both
2. Repeat for every treated unit
3. Drop all the unmatched units (eg. 'extra' control units that are 'far away' from treated units)
4. Assess balance - re-run the matching process as many times as you can to maximize balance!

Matching

- Matching is **not** an analysis method, it's a **pre-processing stage**

Matching

- Matching is **not** an analysis method, it's a **pre-processing stage**
- After matching, we can either:

1. Calculate the difference in means between treated and control groups

Matching

- Matching is **not** an analysis method, it's a **pre-processing stage**
- After matching, we can either:

1. Calculate the difference in means between treated and control groups
2. Conduct the normal regression: $Y \sim D$
    - Option to include all our matching variables as controls
    - This will help control for any **residual imbalance** (esp. for continuous variables)

Matching

- Which variables to match on?

Matching

- Which variables to match on?
  - Treatment variable?

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable?

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable? **No!** That's selecting on the dependent variable - biased!

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable? **No!** That's selecting on the dependent variable - biased!
  - Post-treatment variables?

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable? **No!** That's selecting on the dependent variable - biased!
  - Post-treatment variables? **No!** This will bias our causal effect, just as in regression

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable? **No!** That's selecting on the dependent variable - biased!
  - Post-treatment variables? **No!** This will bias our causal effect, just as in regression
  - Confounders?

Matching

- Which variables to match on?
  - Treatment variable? **No!** We need treated and control units who are both male
  - Outcome variable? **No!** That's selecting on the dependent variable - biased!
  - Post-treatment variables? **No!** This will bias our causal effect, just as in regression
  - Confounders? **Yes!** We want to remove imbalance due to confounders

Matching

- Matching's advantage over (only) regression is that it is **non-parametric**
  - We don't need to make any assumptions about linearity, quadratic relationship etc.

Matching

- Matching's advantage over (only) regression is that it is **non-parametric**
  - We don't need to make any assumptions about linearity, quadratic relationship etc.
    - I.e. Reduced 'Model Dependence'

Matching

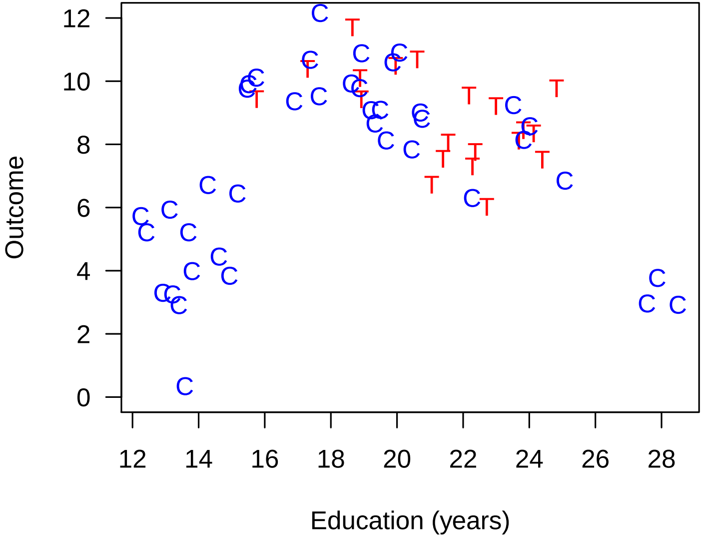
- Matching's advantage over (only) regression is that it is **non-parametric**
  - We don't need to make any assumptions about linearity, quadratic relationship etc.
    - I.e. Reduced 'Model Dependence'
  - And we get better overlap because we're not extrapolating outside of the data

Matching

- Matching's advantage over (only) regression is that it is **non-parametric**
  - We don't need to make any assumptions about linearity, quadratic relationship etc.
    - I.e. Reduced 'Model Dependence'
  - And we get better overlap because we're not extrapolating outside of the data
  - True, there are lots of choices in matching, but our aim is just to increase balance, unlike regression which has no success measure
- The disadvantages are:
  - We may change our definition of the causal effect a little
  - We might lose statistical power by discarding too many units
  - A tricky trade-off between number of units and balance

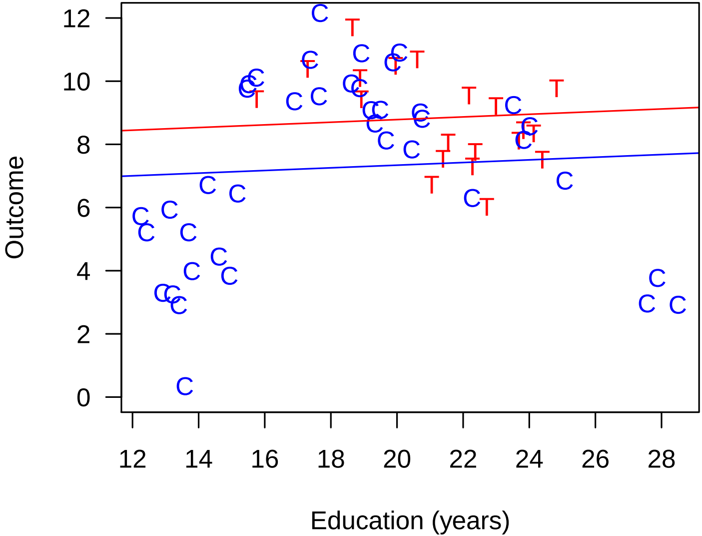
Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



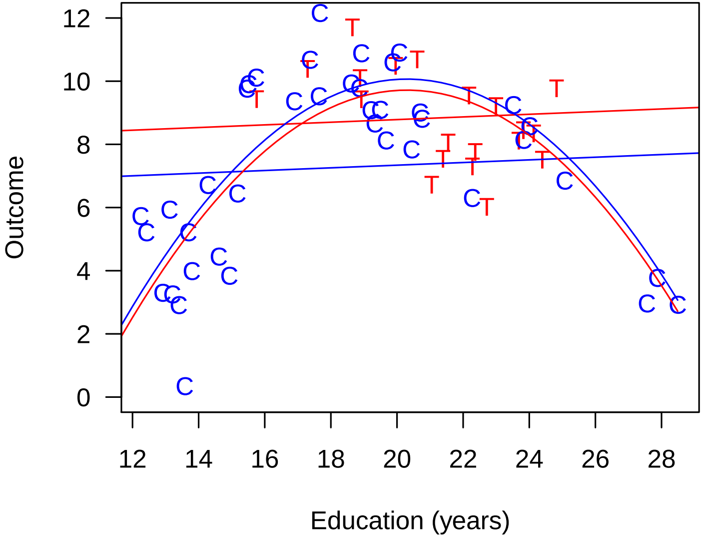
Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



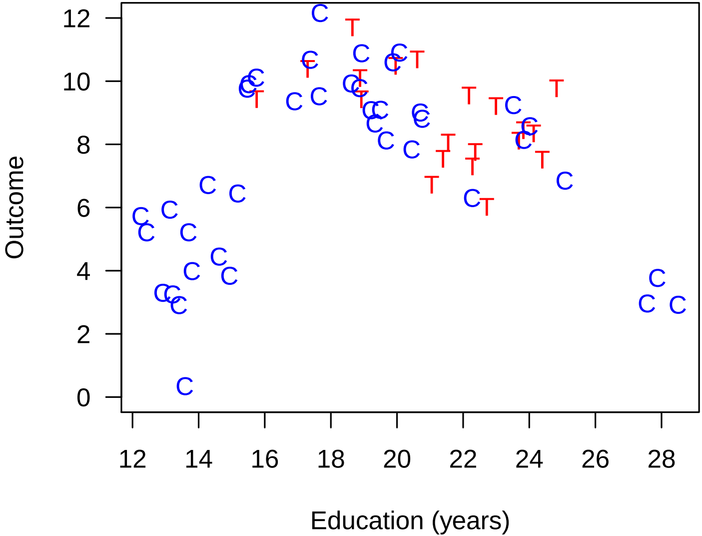
Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



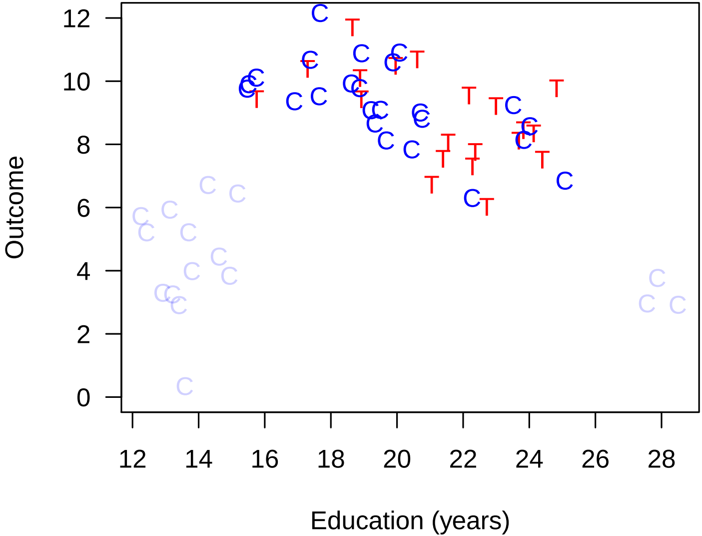
Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



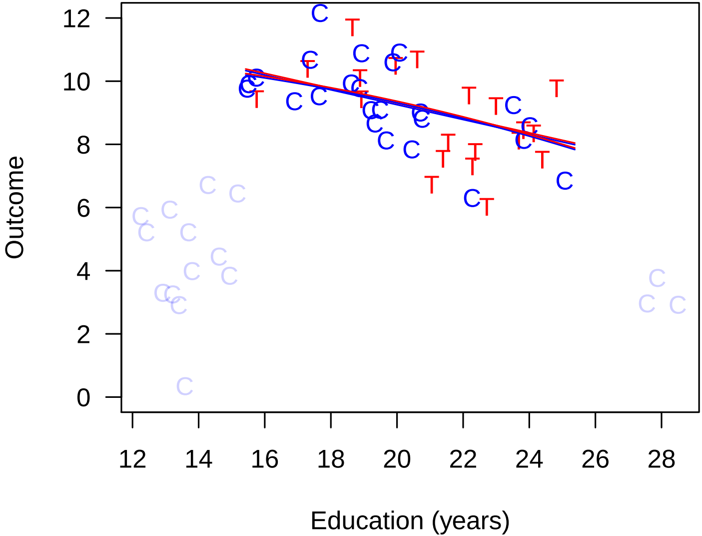
Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



Matching to Reduce Model Dependence

(Ho, Imai, King, Stuart, 2007: fig.1, *Political Analysis*)



Matching

- To identify ’close’ matches we need some measure of distance between units’ covariates

Matching

- To identify ’close’ matches we need some measure of distance between units’ covariates

1. Matching on few categorical variables: **Exact Matching**

Matching

- To identify ’close’ matches we need some measure of distance between units’ covariates

1. Matching on few categorical variables: **Exact Matching**
2. Matching on continuous variables (sequential): **Nearest-Neighbour Matching**

Matching

- To identify ’close’ matches we need some measure of distance between units’ covariates

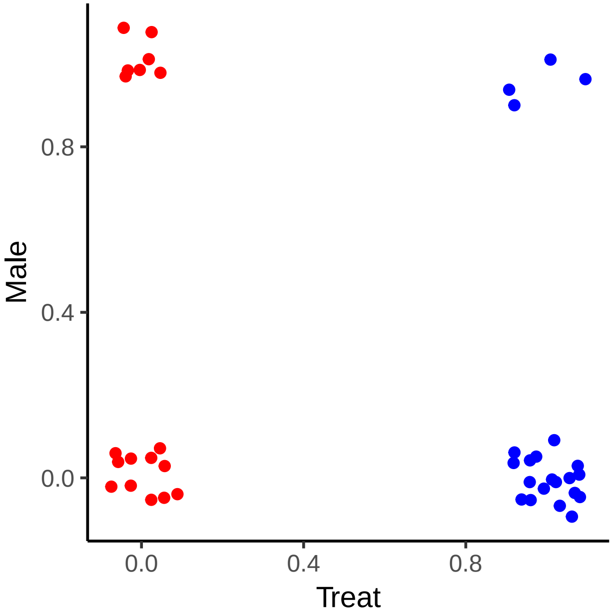
1. Matching on few categorical variables: **Exact Matching**
2. Matching on continuous variables (sequential): **Nearest-Neighbour Matching**
3. Matching to maximize balance: **Optimal/Genetic Matching**

Matching

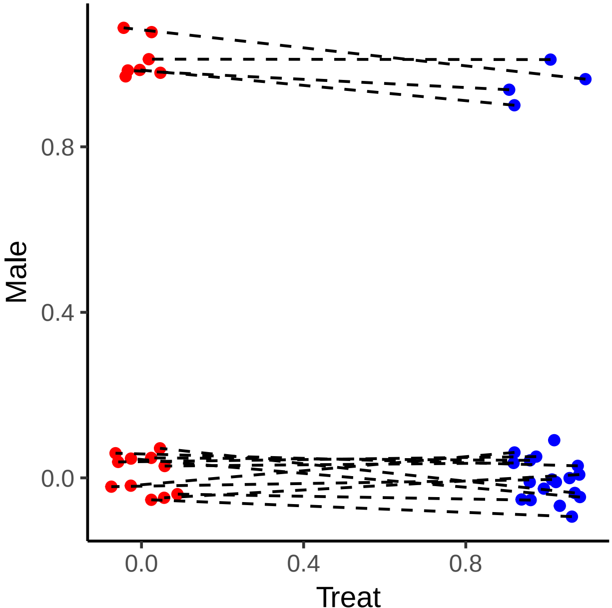
- To identify ’close’ matches we need some measure of distance between units’ covariates

1. Matching on few categorical variables: **Exact Matching**
2. Matching on continuous variables (sequential): **Nearest-Neighbour Matching**
3. Matching to maximize balance: **Optimal/Genetic Matching**
4. Matching on the probability of treatment: **Propensity Score Matching**

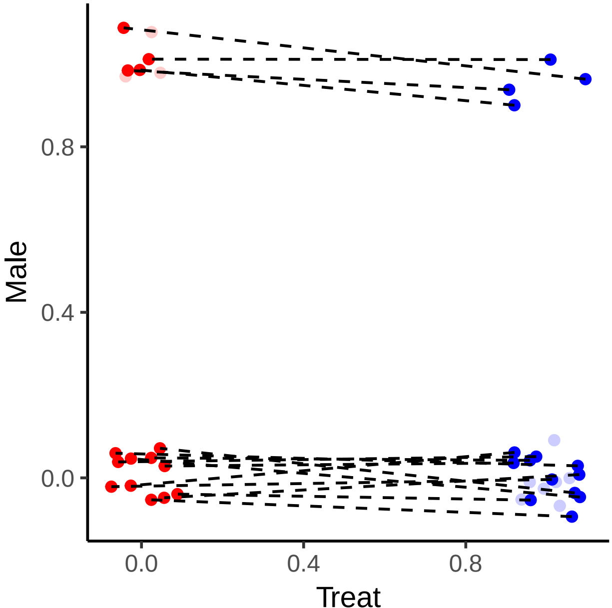
Exact Matching



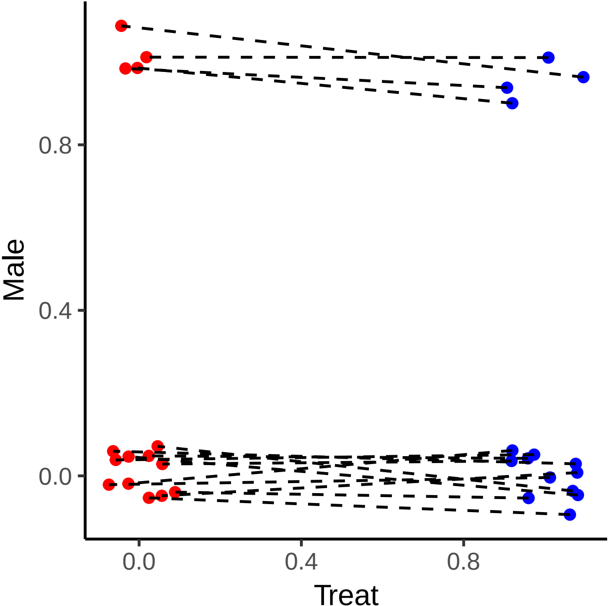
Exact Matching



Exact Matching



Exact Matching



Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**

Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**
- After matching, we **prune/remove** unmatched units

Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**
- After matching, we **prune/remove** unmatched units
- **Then delete the link between the paired units, we don't need it any more**

Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**
- After matching, we **prune/remove** unmatched units
- **Then delete the link between the paired units, we don't need it any more**
- Then compare the outcome of the **remaining** treated and control units

Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**
- After matching, we **prune/remove** unmatched units
- **Then delete the link between the paired units, we don't need it any more**
- Then compare the outcome of the **remaining** treated and control units
  - Difference in means

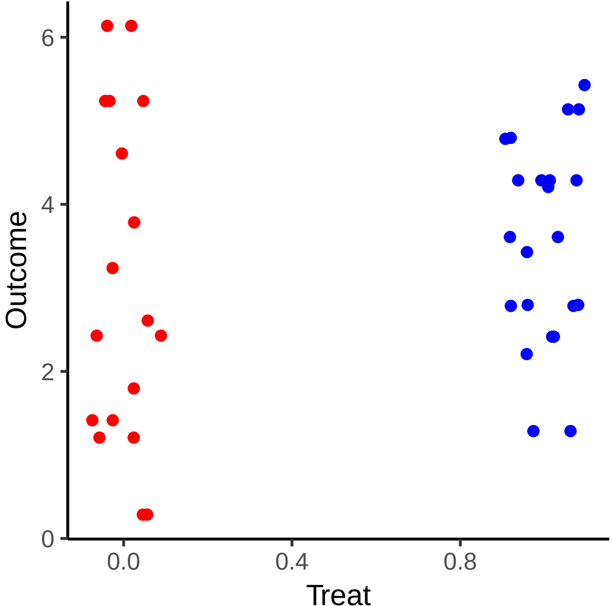
Exact Matching

- Exact matching defines clear counterfactuals:
  - What is the difference in the outcome between treated and control units **for units of the same gender**
- After matching, we **prune/remove** unmatched units
- **Then delete the link between the paired units, we don't need it any more**
- Then compare the outcome of the **remaining** treated and control units
  - Difference in means
  - Or regression of outcome on treatment

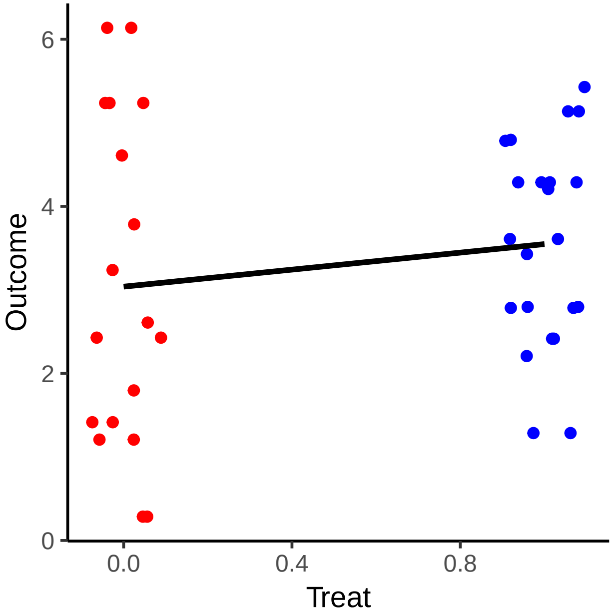
Exact Matching

	Units	Means Treated	Means Control	Mean Diff
1	All	0.18	0.39	-0.21
2	Matched	0.27	0.27	0.00

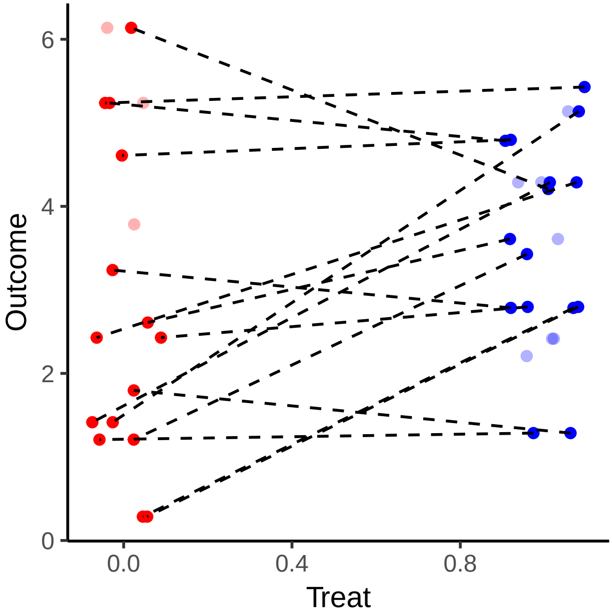
Exact Matching Analysis



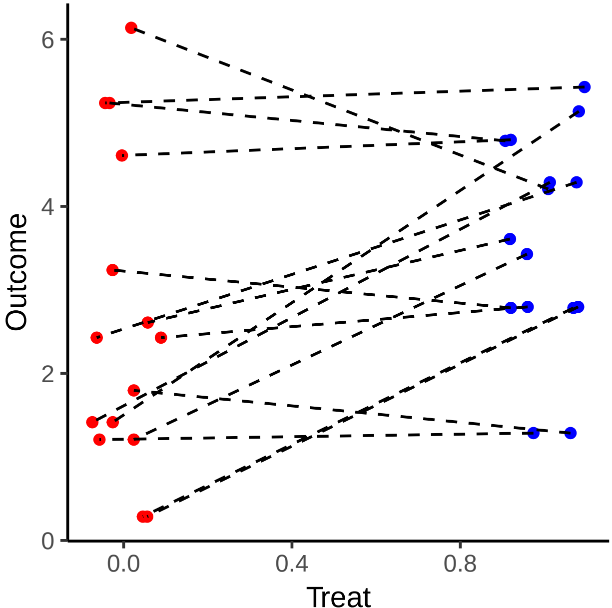
Exact Matching Analysis



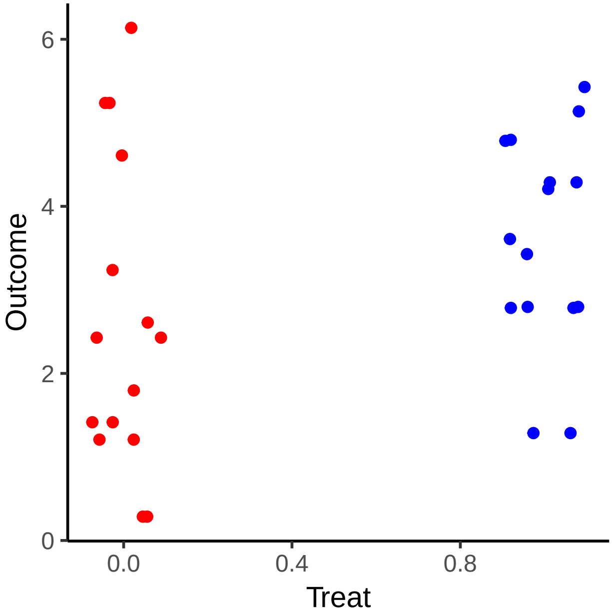
Exact Matching Analysis



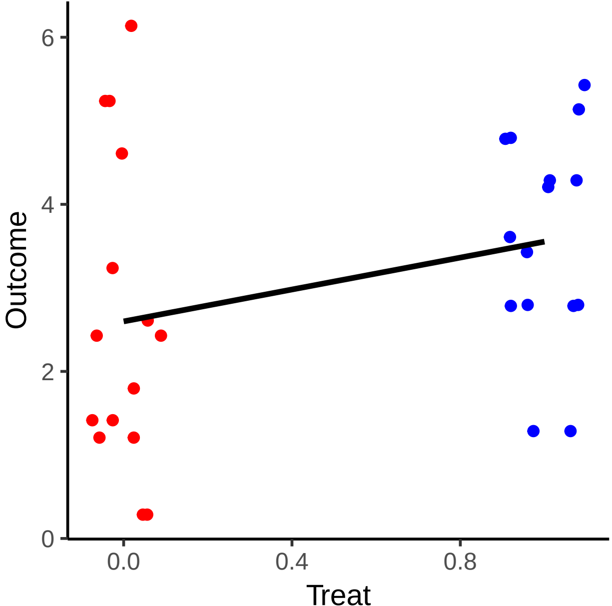
Exact Matching Analysis



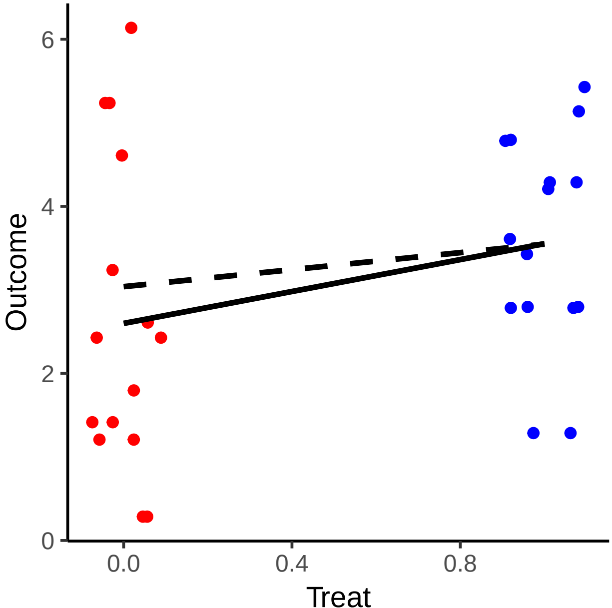
Exact Matching Analysis



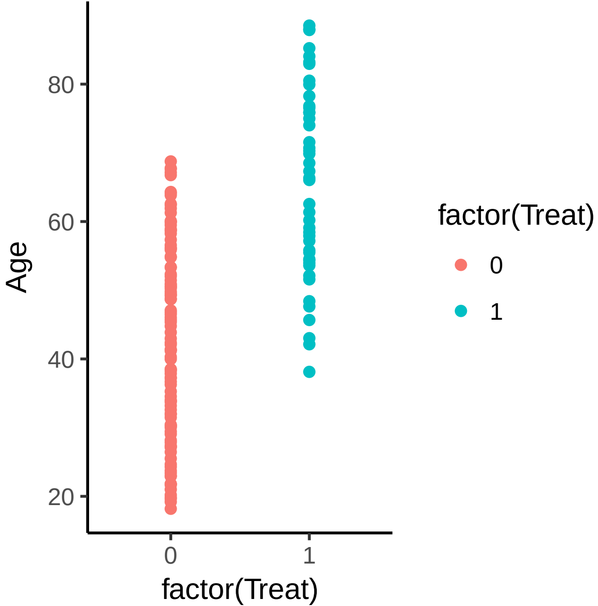
Exact Matching Analysis



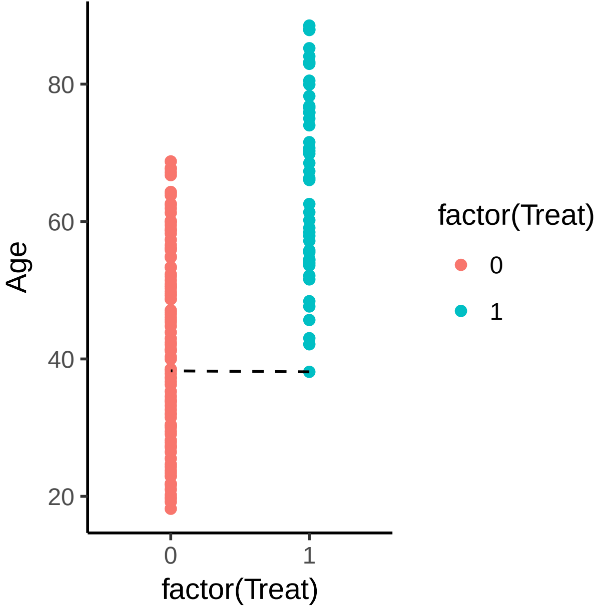
Exact Matching Analysis



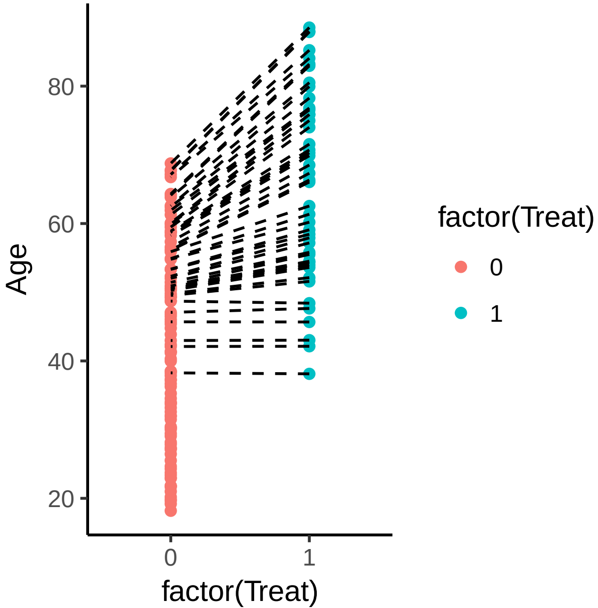
Nearest Neighbour Matching



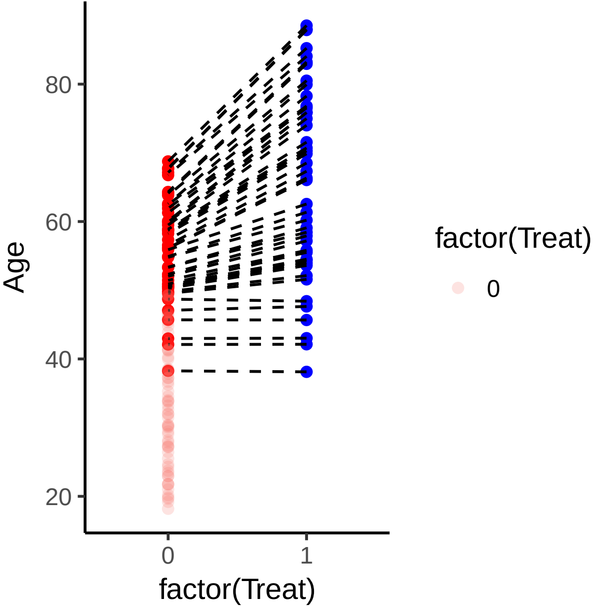
Nearest Neighbour Matching



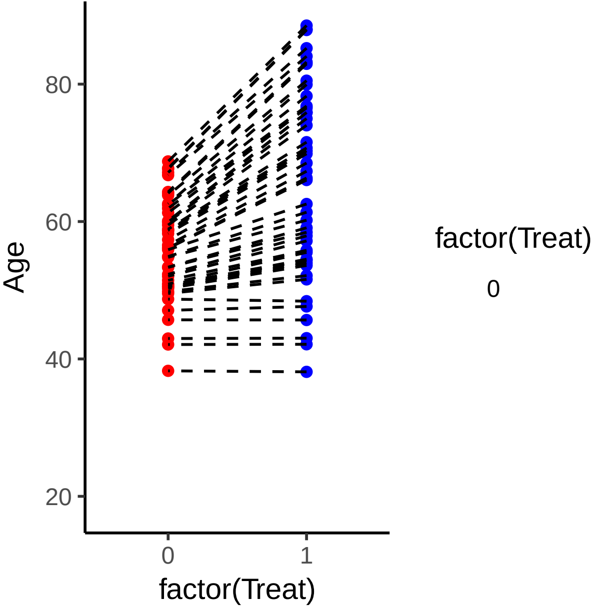
Nearest Neighbour Matching



Nearest Neighbour Matching



Nearest Neighbour Matching



Nearest Neighbour Matching

	Units	Means Treated	Means Control	Mean Diff
1	All	65.70	42.67	23.03
2	Matched	65.70	56.09	9.61

Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:

Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
  1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units

Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
  1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units
     - We need some **absolute** limits on the distance we can match units within

Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
  1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units
     - We need some **absolute** limits on the distance we can match units within
     - We can add 'calipers' to matching to match only within a fixed range

Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
  1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units
     - We need some **absolute** limits on the distance we can match units within
     - We can add ’calipers’ to matching to match only within a fixed range
  2. **The order of matching matters:** The first matches use up units that might make better matches for later treated units

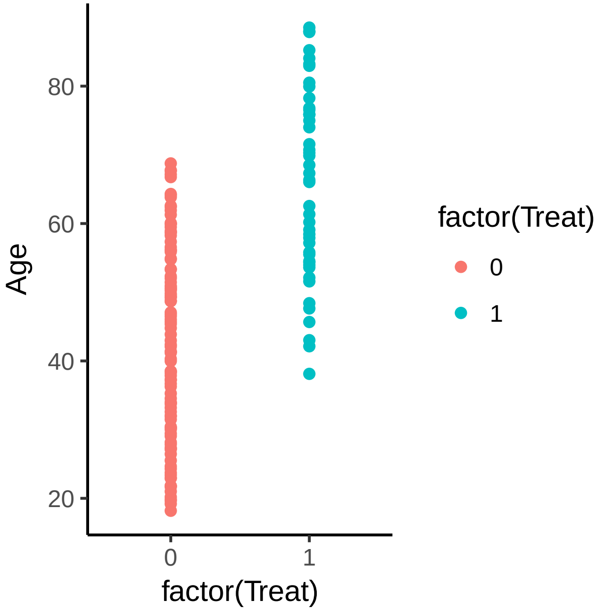
Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
    1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units
        - We need some **absolute** limits on the distance we can match units within
        - We can add 'calipers' to matching to match only within a fixed range
    2. **The order of matching matters:** The first matches use up units that might make better matches for later treated units
        - To maximize balance we need to 'look ahead' and match in the right order

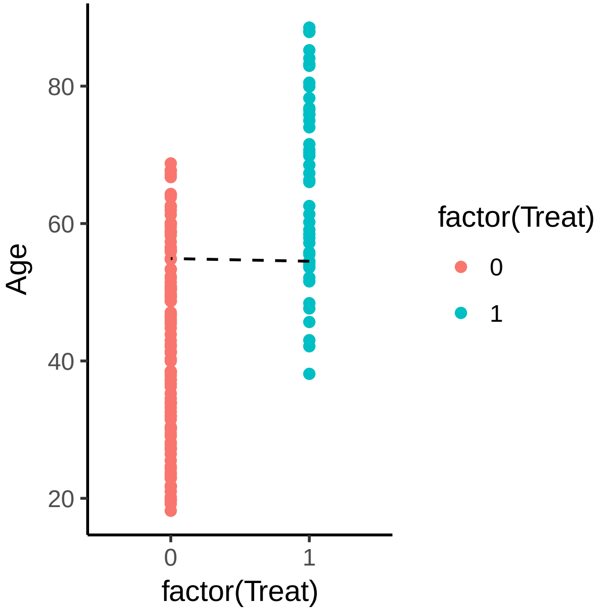
Nearest Neighbour Matching

- Two potential problems with nearest neighbour matching:
  1. **Nearest does not mean close:** The oldest treated units are matched with, but very different to, the oldest control units
     - We need some **absolute** limits on the distance we can match units within
     - We can add ’calipers’ to matching to match only within a fixed range
  2. **The order of matching matters:** The first matches use up units that might make better matches for later treated units
     - To maximize balance we need to ’look ahead’ and match in the right order
     - For this we can use optimal or genetic matching, which is fully automated

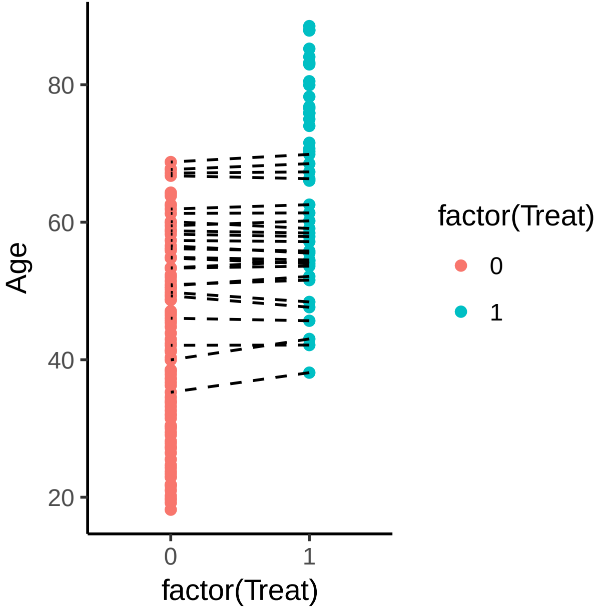
Nearest Neighbour Matching with Caliper



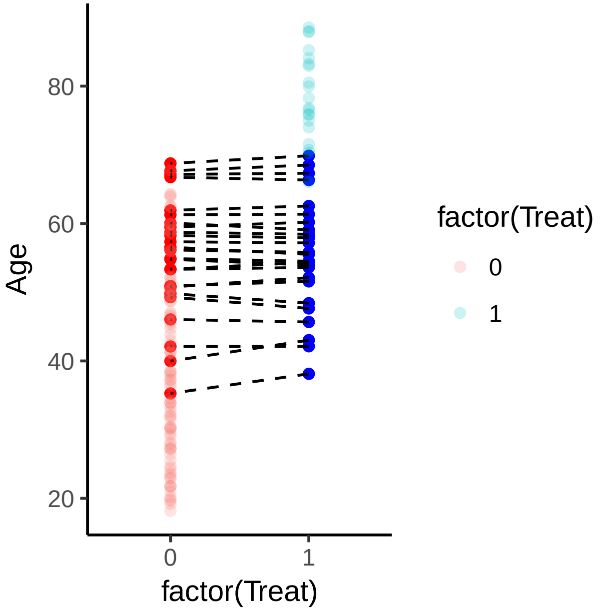
Nearest Neighbour Matching with Caliper



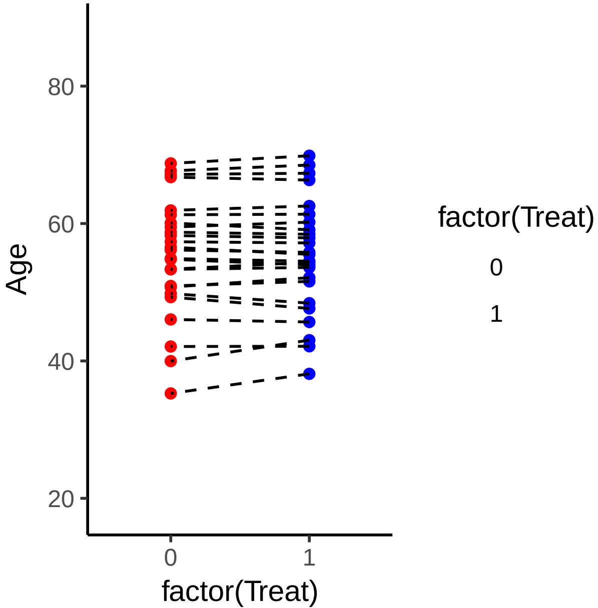
Nearest Neighbour Matching with Caliper



Nearest Neighbour Matching with Caliper



Nearest Neighbour Matching with Caliper



Nearest Neighbour Matching with Caliper

	Units	Means Treated	Means Control	Mean Diff
1	All	65.70	42.67	23.03
2	Matched	55.41	55.23	0.18

- Note: p-values don't mean so much for balance tests

Nearest Neighbour Matching with Caliper

	Units	Means Treated	Means Control	Mean Diff
1	All	65.70	42.67	23.03
2	Matched	55.41	55.23	0.18

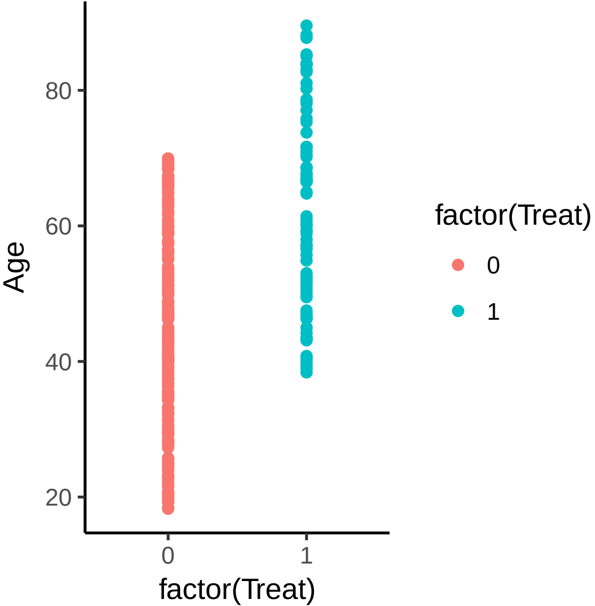
- Note: p-values don't mean so much for balance tests
- We always want to improve balance as much as possible

Nearest Neighbour Matching with Caliper

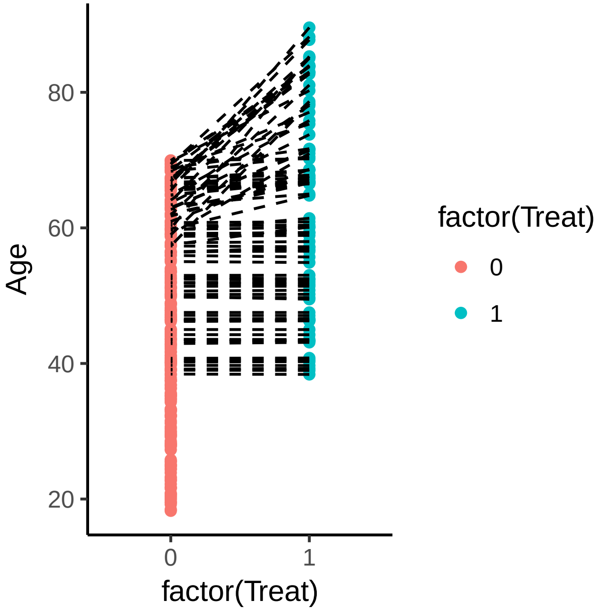
	Units	Means Treated	Means Control	Mean Diff
1	All	65.70	42.67	23.03
2	Matched	55.41	55.23	0.18

- Note: p-values don’t mean so much for balance tests
- We always want to improve balance as much as possible
- Better to compare (standardized) difference in means

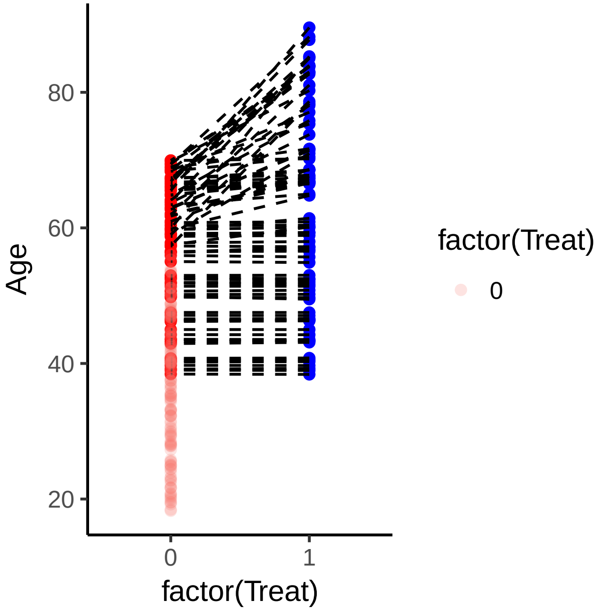
Optimal Matching



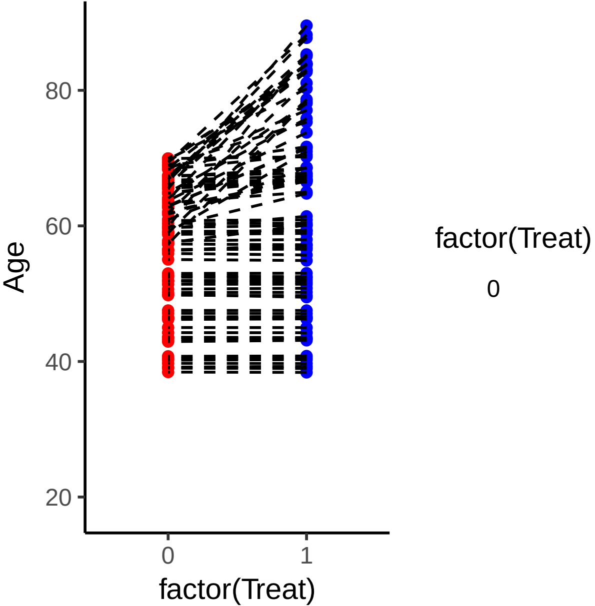
Optimal Matching



Optimal Matching



Optimal Matching



Optimal Matching

	Units	Means Treated	Means Control	Mean Diff
1	All	62.60	44.64	17.96
2	Matched	62.60	57.57	5.03

Propensity Score Matching

- With many covariates we have a dimensionality challenge

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero
  - Counterfactuals are impossible to define
- The propensity score collapses matching to a single dimension

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero
  - Counterfactuals are impossible to define
- The propensity score collapses matching to a single dimension
  - Confounders only matter to the extent they affect treatment

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero
  - Counterfactuals are impossible to define
- The propensity score collapses matching to a single dimension
  - Confounders only matter to the extent they affect treatment
  - So let's use the confounders to **predict treatment**

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero
  - Counterfactuals are impossible to define
- The propensity score collapses matching to a single dimension
  - Confounders only matter to the extent they affect treatment
  - So let’s use the confounders to **predict treatment**
  - That’s different to actual treatment status, with the remainder due to ’random’ factors (if we include all confounders)

Propensity Score Matching

- With many covariates we have a dimensionality challenge
  - Overlap is almost zero
  - Counterfactuals are impossible to define
- The propensity score collapses matching to a single dimension
  - Confounders only matter to the extent they affect treatment
  - So let's use the confounders to **predict treatment**
  - That's different to actual treatment status, with the remainder due to 'random' factors (if we include all confounders)
- Then use the propensity score (probability 0-1) to match treated and control units which have the same ex ante probability of treatment

Propensity Score Matching

- But some concerns about drawbacks of propensity score matching

Propensity Score Matching

- But some concerns about drawbacks of propensity score matching
- May have poor balance on individual confounders

Propensity Score Matching

- But some concerns about drawbacks of propensity score matching
- May have poor balance on individual confounders
- Balance may get worse as we remove more units

Propensity Score Matching

- But some concerns about drawbacks of propensity score matching
- May have poor balance on individual confounders
- Balance may get worse as we remove more units
- We have to get the functional form of the treatment explanation right (linear, quadratic etc.) so we remain vulnerable to model dependence!

Propensity Score Matching

- Treatment: 1/0
- Confounder: Age
- Logit model predicting treatment:

$$Treat_i = \alpha + \beta Age_i + \epsilon_i$$

Propensity Score Matching

- Treatment: 1/0
- Confounder: Age
- Logit model predicting treatment:

$$Treat_i = \alpha + \beta Age_i + \epsilon_i$$

$$Predicted\_Treat_i = -7.19 + 0.116 Age_i + \epsilon_i$$

Propensity Score Matching

- Treatment: 1/0
- Confounder: Age
- Logit model predicting treatment:

$$Treat_i = \alpha + \beta Age_i + \epsilon_i$$

$$Predicted\_Treat_i = -7.19 + 0.116 Age_i + \epsilon_i$$

- Match on the values of $Predicted\_Treat_i$ (fitted values of the regression)

Propensity Score Matching

- Treatment: 1/0
- Confounder: Age
- Logit model predicting treatment:

$$Treat_i = \alpha + \beta Age_i + \epsilon_i$$

$$Predicted\_Treat_i = -7.19 + 0.116 Age_i + \epsilon_i$$

- Match on the values of $Predicted\_Treat_i$ (fitted values of the regression)
- I.e. match units with a similar *probability* of treatment

Propensity Score Matching

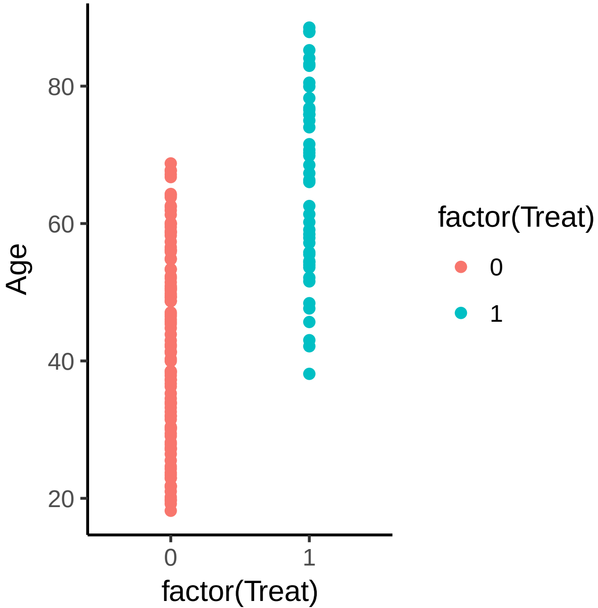
- Treatment: 1/0
- Confounder: Age
- Logit model predicting treatment:

$$Treat_i = \alpha + \beta Age_i + \epsilon_i$$

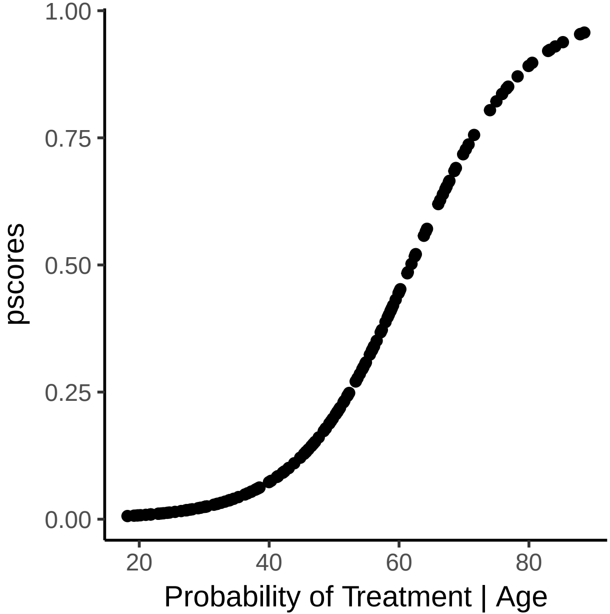
$$Predicted\_Treat_i = -7.19 + 0.116 Age_i + \epsilon_i$$

- Match on the values of $Predicted\_Treat_i$ (fitted values of the regression)
- I.e. match units with a similar *probability* of treatment
- ...Regardless of whether they actually get treated

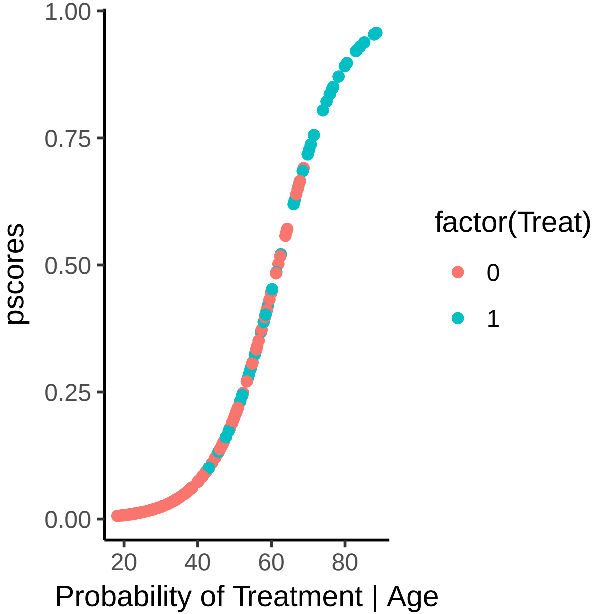
Propensity Score Matching



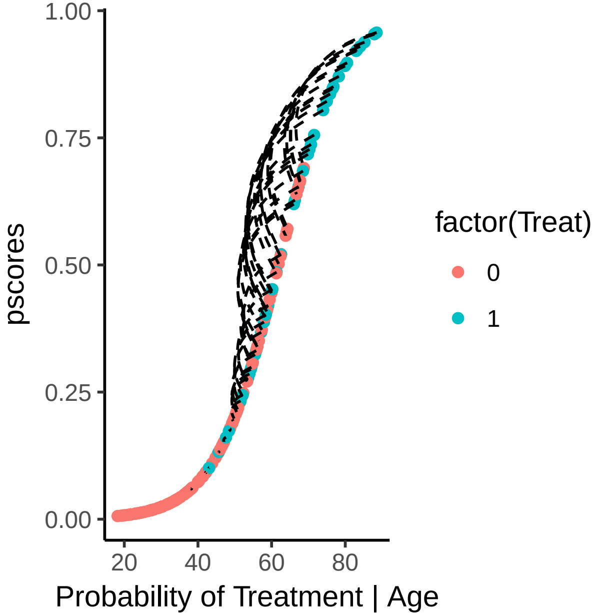
Propensity Score Matching



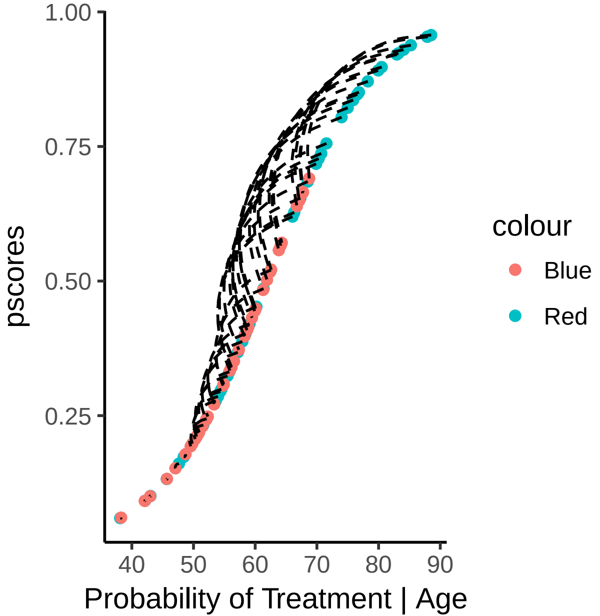
Propensity Score Matching



Propensity Score Matching



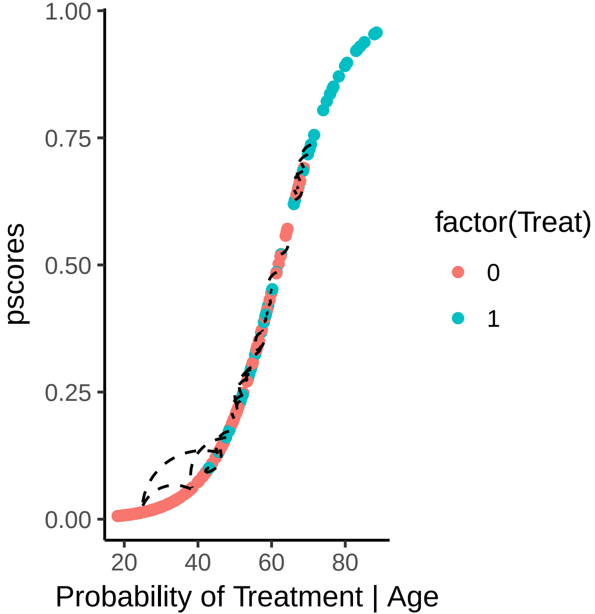
Propensity Score Matching



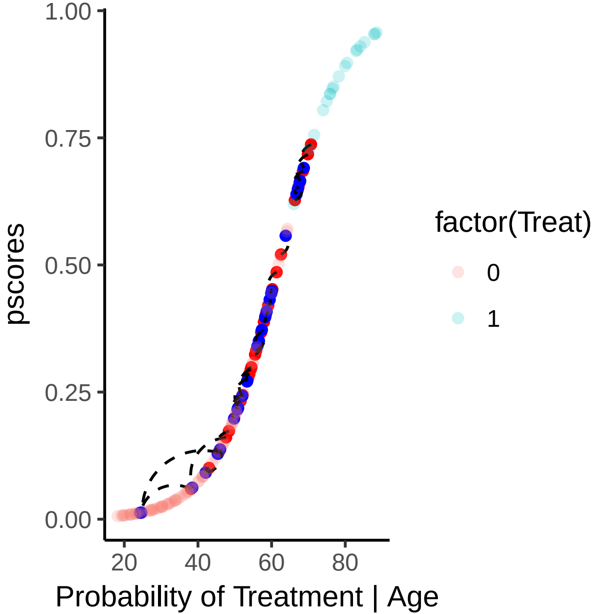
Propensity Score Matching

	Units	Means Treated	Means Control	Mean Diff
1	All	0.57	0.18	0.39
2	Matched	0.57	0.36	0.21

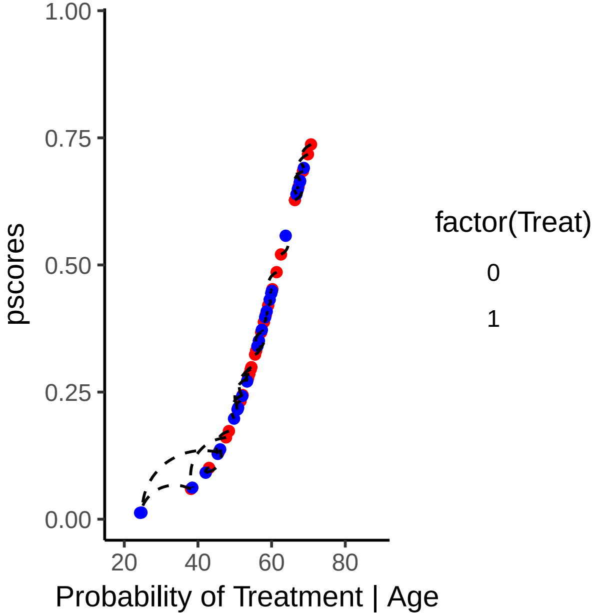
Propensity Score Matching with Caliper



Propensity Score Matching with Caliper



Propensity Score Matching with Caliper



Propensity Score Matching with Caliper

	Units	Means Treated	Means Control	Mean Diff
1	All	0.57	0.18	0.39
2	Matched	0.36	0.34	0.02

Matching

- How much matching should we undertake?

Matching

- How much matching should we undertake?
- We can always enforce **stricter** matching (eg. narrower calipers, more exact matching) to get better balance

Matching

- How much matching should we undertake?
- We can always enforce **stricter** matching (eg. narrower calipers, more exact matching) to get better balance
- But our N will approach zero, so little statistical power

Matching

- How much matching should we undertake?
- We can always enforce **stricter** matching (eg. narrower calipers, more exact matching) to get better balance
- But our N will approach zero, so little statistical power
- A Bias-variance trade-off

Matching

- How much matching should we undertake?
- We can always enforce **stricter** matching (eg. narrower calipers, more exact matching) to get better balance
- But our N will approach zero, so little statistical power
- A Bias-variance trade-off
- Try alternative specifications

Matching

- Matching preferred to regression where:

Matching

- Matching preferred to regression where:
  - Never! Do both!
- Matching makes a big contribution where there's poor overlap

Matching

- Matching preferred to regression where:
  - Never! Do both!
- Matching makes a big contribution where there's poor overlap
- Matching + Regression = "Doubly Robust"

Matching

- Matching preferred to regression where:
  - Never! Do both!
- Matching makes a big contribution where there's poor overlap
- Matching + Regression = "Doubly Robust"
  - If **either** matching produces balance **OR** we have the correct functional form for regression, we can make causal inference

Matching

- Arceneaux, Gerber and Green (2005)

Matching

- Arceneaux, Gerber and Green (2005)
- How does matching work on experimental (IV) data? (eg. for how to get voters to vote)

Matching

- Arceneaux, Gerber and Green (2005)
- How does matching work on experimental (IV) data? (eg. for how to get voters to vote)
- Matching is biased compared to the experimental results

Matching

- Arceneaux, Gerber and Green (2005)
- How does matching work on experimental (IV) data? (eg. for how to get voters to vote)
- Matching is biased compared to the experimental results
- Lots of controls

Matching

- Arceneaux, Gerber and Green (2005)
- How does matching work on experimental (IV) data? (eg. for how to get voters to vote)
- Matching is biased compared to the experimental results
- Lots of controls
- But unobserved confounders mean matching can't recover causal estimates

Matching

- Bias was due to whether people actually answered phone calls

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**
- Experimental measure: 0.4

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**
- Experimental measure: 0.4
- OLS estimate: 2.7

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**
- Experimental measure: 0.4
- OLS estimate: 2.7
- Matching estimate: 2.8

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**
- Experimental measure: 0.4
- OLS estimate: 2.7
- Matching estimate: 2.8
- We can't control for likelihood of answering the phone using the (many) covariates they have

Matching

- Bias was due to whether people actually answered phone calls
- Huge N, **Perfect balance**
- Experimental measure: 0.4
- OLS estimate: 2.7
- Matching estimate: 2.8
- We can't control for likelihood of answering the phone using the (many) covariates they have
- Matching still relies on **measuring all confounders**

Political Economy

- Boas and Hidalgo (2011)

Political Economy

- Boas and Hidalgo (2011)
- They already used an RDD on close electoral victories to show incumbents are more likely to get media licences

Political Economy

- Boas and Hidalgo (2011)
- They already used an RDD on close electoral victories to show incumbents are more likely to get media licences
- But how do media licences affect performance in the next election?

Political Economy

- Boas and Hidalgo (2011)
- They already used an RDD on close electoral victories to show incumbents are more likely to get media licences
- But how do media licences affect performance in the next election?
- No ’as-if’ random variation in treatment

Political Economy

- Boas and Hidalgo (2011)
- They already used an RDD on close electoral victories to show incumbents are more likely to get media licences
- But how do media licences affect performance in the next election?
- No ’as-if’ random variation in treatment
- So they use an observational study with matching to create plausible counterfactuals

Political Economy

- **Population:**

Political Economy

- **Population:** Brazilian councillors

Political Economy

- **Population:** Brazilian councillors
- **Sample:**

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:**

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:**

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:** No approval before the election campaign (rejection or no decision)

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:** No approval before the election campaign (rejection or no decision)
- **Treatment Assignment:**

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:** No approval before the election campaign (rejection or no decision)
- **Treatment Assignment:** Ministry's decision process (unknown)
  - We know it's influenced by incumbency, for example

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:** No approval before the election campaign (rejection or no decision)
- **Treatment Assignment:** Ministry's decision process (unknown)
  - We know it's influenced by incumbency, for example
- **Outcome:**

Political Economy

- **Population:** Brazilian councillors
- **Sample:** Brazilian councillors who apply for media licences (not in cities >2 million)
- **Treatment:** Approval for a licence before the election campaign
- **Control:** No approval before the election campaign (rejection or no decision)
- **Treatment Assignment:** Ministry's decision process (unknown)
    - We know it's influenced by incumbency, for example
- **Outcome:** Vote Share in next election

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation
- Try to balance treatment and control units on confounders using matching

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation
- Try to balance treatment and control units on confounders using matching
- Seek to maximize balance, using genetic matching

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation
- Try to balance treatment and control units on confounders using matching
- Seek to maximize balance, using genetic matching
- Units in dataset before matching: 1455

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation
- Try to balance treatment and control units on confounders using matching
- Seek to maximize balance, using genetic matching
- Units in dataset before matching: 1455
- Units in dataset after matching: 622 (311 treated, 311 control)

Political Economy

- Confounders with imbalance: Application timing, competition, incumbency, municipality type, political bias, occupation
- Try to balance treatment and control units on confounders using matching
- Seek to maximize balance, using genetic matching
- Units in dataset before matching: 1455
- Units in dataset after matching: 622 (311 treated, 311 control)
- Clear improvement in balance after matching

Political Economy

- How matching changes our units may affect the definition of our treatment effect

Political Economy

- How matching changes our units may affect the definition of our treatment effect
  - Eg. if we keep all treated units and throw away some control units, this is an average treatment effect on the treated

Political Economy

- How matching changes our units may affect the definition of our treatment effect
  - Eg. if we keep all treated units and throw away some control units, this is an average treatment effect on the treated
  - If we keep all control units and throw away some treated units, this is an average treatment effect on the untreated

Political Economy

- How matching changes our units may affect the definition of our treatment effect
  - Eg. if we keep all treated units and throw away some control units, this is an average treatment effect on the treated
  - If we keep all control units and throw away some treated units, this is an average treatment effect on the untreated
- Simple comparison of means between treated and control group

Political Economy

- How matching changes our units may affect the definition of our treatment effect
  - Eg. if we keep all treated units and throw away some control units, this is an average treatment effect on the treated
  - If we keep all control units and throw away some treated units, this is an average treatment effect on the untreated
- Simple comparison of means between treated and control group
- ATT causal effect estimate: 0.39% points (17%) increase in vote share if you have a media licence

Political Economy

- Zucco (2013)

Political Economy

- Zucco (2013)
- How does Bolsa Familia affect voting?

Political Economy

- Zucco (2013)
- How does Bolsa Familia affect voting?
- Similar to De La O (2010) but without the natural experiment

Political Economy

- Zucco (2013)
- How does Bolsa Familia affect voting?
- Similar to De La O (2010) but without the natural experiment
- "There were no randomized pilot programs, there are no obvious discontinuities to be exploited, and CCT eligibility and actual coverage are highly correlated with several other socioeconomic variables."

Political Economy

- **Population:**

Political Economy

- **Population:** Brazilian municipalities

Political Economy

- **Population:** Brazilian municipalities
- **Sample:**

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:**

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:**

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:** Low % Families in municipality receiving Bolsa Familia

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:** Low % Families in municipality receiving Bolsa Familia
- **Treatment Assignment:**

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:** Low % Families in municipality receiving Bolsa Familia
- **Treatment Assignment:** Complex, based on poverty, geography

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:** Low % Families in municipality receiving Bolsa Familia
- **Treatment Assignment:** Complex, based on poverty, geography
- **Outcome:**

Political Economy

- **Population:** Brazilian municipalities
- **Sample:** Brazilian municipalities
- **Treatment:** High % Families in municipality receiving Bolsa Familia
- **Control:** Low % Families in municipality receiving Bolsa Familia
- **Treatment Assignment:** Complex, based on poverty, geography
- **Outcome:** Vote Share to Incumbent President in next election

Political Economy

- **Overlap Problem:** If MDS used fixed rules to allocate Bolsa Familia, there would be no overlap and no plausible counterfactuals

Political Economy

- **Overlap Problem:** If MDS used fixed rules to allocate Bolsa Familia, there would be no overlap and no plausible counterfactuals
- Luckily, implementation does not follow perfectly fixed rules - so matching is feasible

Political Economy

- **Overlap Problem:** If MDS used fixed rules to allocate Bolsa Familia, there would be no overlap and no plausible counterfactuals
- Luckily, implementation does not follow perfectly fixed rules - so matching is feasible
- But that also suggests other informal/unobserved factors affect treatment assignment - confounding!

Political Economy

- Matching Stage:

Political Economy

- Matching Stage:
- (Generalized) Propensity Score Matching

Political Economy

- Matching Stage:
- (Generalized) Propensity Score Matching
  1. Regression of treatment on covariates ($Bolsa\_Familia \sim HDI + Population + Target\_BF\_Coverage...$)

Political Economy

- Matching Stage:
- (Generalized) Propensity Score Matching
  1. Regression of treatment on covariates ($Bolsa\_Familia \sim HDI + Population + Target\_BF\_Coverage...$)
  2. Predict probability of treatment (propensity score) for each unit

Political Economy

- Matching Stage:
- (Generalized) Propensity Score Matching
  1. Regression of treatment on covariates ($Bolsa\_Familia \sim HDI + Population + Target\_BF\_Coverage...$)
  2. Predict probability of treatment (propensity score) for each unit
  3. Match treatment and control units on propensity score (in strata)

Political Economy

- Matching Stage:
- (Generalized) Propensity Score Matching
  1. Regression of treatment on covariates ($Bolsa\_Familia \sim HDI + Population + Target\_BF\_Coverage...$)
  2. Predict probability of treatment (propensity score) for each unit
  3. Match treatment and control units on propensity score (in strata)
  4. Regression of Outcome on propensity score by strata

Political Economy

- Results:
  - 1% point more families covered by Bolsa Familia increases vote share by 0.12-0.18% points

Political Economy

- Results:
  - 1% point more families covered by Bolsa Familia increases vote share by 0.12-0.18% points
  - Spending extra R$100 per person through Bolsa Familia increases vote share by 7-15% points

Political Economy

- Results:
  - 1% point more families covered by Bolsa Familia increases vote share by 0.12-0.18% points
  - Spending extra R$100 per person through Bolsa Familia increases vote share by 7-15% points
  - Consistent estimates from matching in survey data