Final Project

Nikita Kotsehub, Aniket Das

CS112 Fall 2020

Introduction

PTAs (Preferential Trade Agreements) are international trading arrangements that outline permitted policy changes by states. Signing a PTA commits the government to a set of policies, thus establishing transparency and reducing policy uncertainty. Hollyer & Rosendorff (2012) tested a hypothesis that governments that sign Preferential trade agreements (PTAs) survive longer in office than similar governments that do not sign them. In "Leadership Survival, Regime Type, Policy Uncertainty, and PTA Accession," they account for the self-selection bias arising from observable or non-observable covariates. For example, if a government already has much support, it is more likely to sign a PTA, making their position even stronger. To mitigate the self-selection bias and balance out the covariates, the authors employ the nearest-neighbor algorithm for propensity score matching (PSM). However, in their nearest-neighbor matching, they set parameter replace to True, which introduces potential bias. Therefore, in this paper, we aim to replicate the matching process used in the paper and improve upon the procedure by employing the genetic matching algorithm.

Replication & Extension

We successfully replicated the matching process in the paper and accompanying diagnostics using the Nearest-neighbor algorithm with propensity scores (Appendix A). We improved the process by using Genetic matching instead of propensity scores and changing the replace parameter to true.

The PSM is an outdated matching model that attempts to squeeze all the covariates into one single variable, and then match on this variable. This introduces bias, imbalance, inefficiency, and model dependence. Towards the end when propensity scores are very similar, the model is essentially pruning at random which increases bias and imbalance. The model cannot distinguish between good matches and excellent matches, leading to inefficiency. Unnecessary imbalance and bias increase model dependence due to suboptimal matching (King & Nielsen, 2018).

Fully blocked randomized experiments, which are approximated by other matching methods, dominate over completely randomized experiments, which are approximated by PSM (King & Nielsen, 2018). Genetic matching is a multivariate matching method that finds optimal balance by giving a weight to each covariate, choosing the weights that give the best balance, using these selected weights to produce another generation, and repeating the process. Since it is not blind to covariates, it can distinguish between good matches and excellent matches. Additionally, it doesn't prune at random because it isn't blind to the covariate distance between the units, unlike PSM. Lastly, it has been proven to perform better than commonly used models such as Mahalanobis distance matching and PSM in simulations (Tsai & Peace, 2012).

The paper, without specifying any reason, matched without replacement, which introduces bias because the order of matches matters: the first treated unit will be matched first, followed by the second unit, and so on. Matching with replacement decreases bias because a control that resembles multiple treated units will be matched to all of them and the order of the units doesn't matter.

Genetic matching produced a 42% better average balance on the covariates than nearest-neighbor matching using propensity scores when replacement was allowed (Appendix B and C, respectively) and 1.66% better on average than PSM without replacement (Figure 1), but while PSM dropped 3 treated units. Genetic matching dropped only 1 (Appendix A and B, respectively).
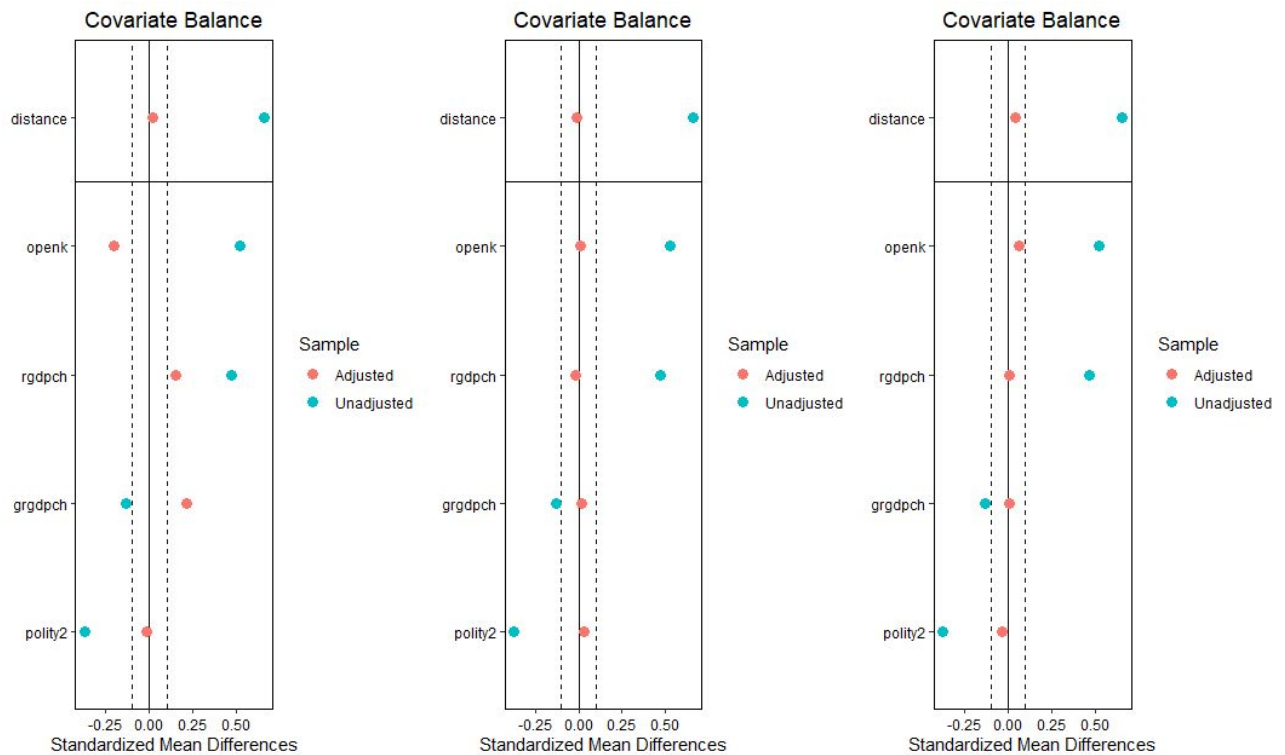
Fig. 1. Covariate balance plot comparison between PSM with replacement on the left, genetic matching in the center, and PSM without replacement on the right. The X-axis is the standardized mean differences between the treatment and control group and Y-axis describes the different covariates. Genetic matching gives the best results because it's adjusted standardized mean differences are closest to zero.

Conclusion

In this paper, we

1. identified several concerns about the application of propensity score matching in the Hollyer & Rosendorff paper,

2. explained the issues with PSM, justifying them with researches from leading experts in matching and causal inference, and

3. extended the matching technique by implementing a genetic matching with replacement.

Our extension significantly improves the covariate balance between treatment and control units, thus making causal inferences about PTA's effect on governmental stability more accurate and less biased. We encourage other scientists and researchers to move away from the propensity score matching and adopt efficient methods such as genetic matching. One can find a link to the R code replication file and the description of our contributions in Appendix D.

References

Hollyer, J., & Rosendorff, B. (2012). Leadership Survival, Regime Type, Policy Uncertainty and
  PTA Accession. International Studies Quarterly, 56(4), 748-764. Retrieved November 14,
  2020, from http://www.jstor.org/stable/41804829

Hollyer, James R.; Rosendorff, B. Peter, 2018, "Leadership Survival, Regime Type, Policy
  Uncertainty and PTA Accession", https://doi.org/10.7910/DVN/O1PMGO, Harvard
  Dataverse, V1, UNF:6:0wsSEb3aAxcBOfZtGV0Khg== [fileUNF]

King, G., & Nielsen, R. (2019). Why Propensity Scores Should Not Be Used for Matching.
  Political Analysis, 27(4), 435-454. doi:10.1017/pan.2019.11. Retrieved December 17,
  2020, from

  https://gking.harvard.edu/files/gking/files/psnot.pdf

Tsai, Kao-Tai & Peace, Karl. (2012). Genetic Matching: An Efficient Algorithm to Adjust
  Covariate Imbalance for Data Analysis and Modeling. Retrieved December 17, 2020,
  from

  https://www.researchgate.net/publication/267786620_Genetic_Matching_An_Efficient_
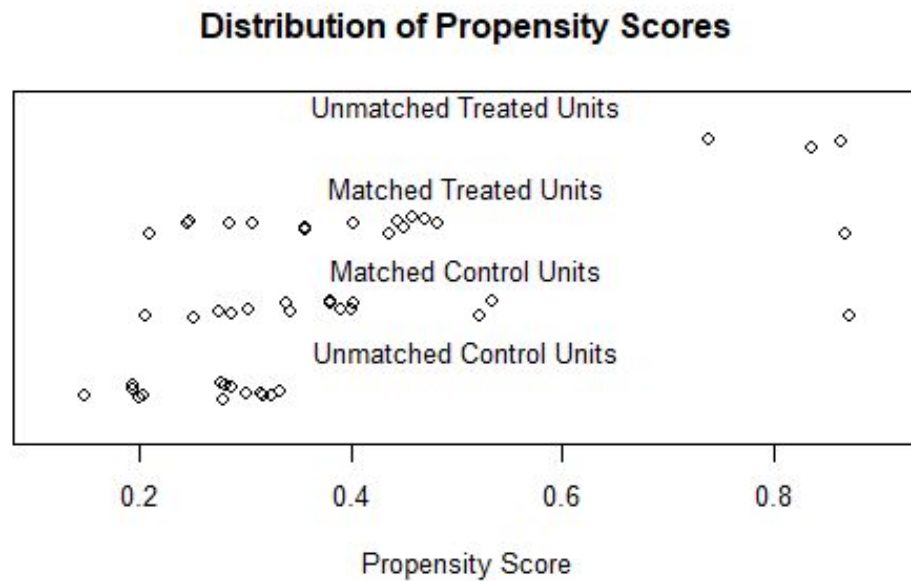 Algorithm_to_Adjust_Covariate_Imbalance_for_Data_Analysis_and_Modeling

Appendices

Appendix A (Replication)

Propensity score Nearest-neighbor matching results from the paper (without replacement)
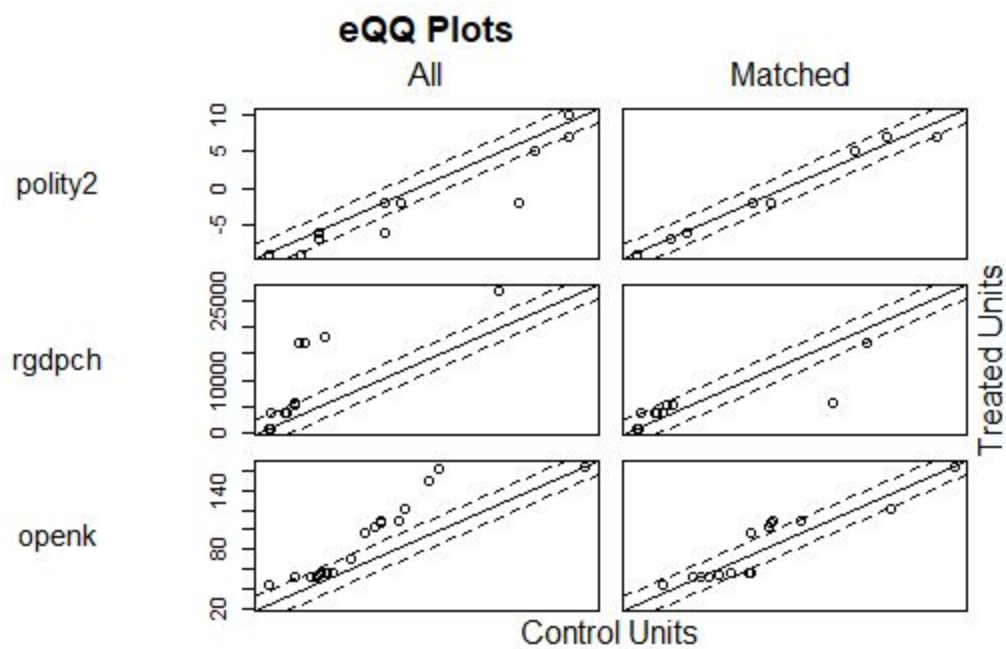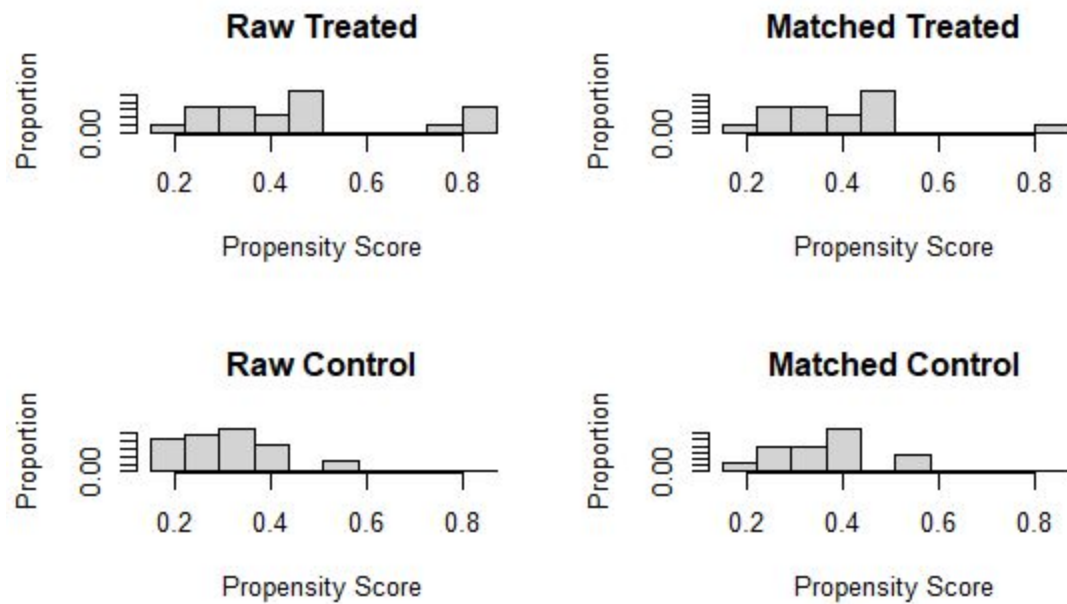
```
Summary of Balance for Matched Data:
         Means Treated Means Control Std. Mean Diff. Var. Ratio eCDF Mean
distance        0.4016        0.3924           0.0427     0.9696    0.0582
polity2        -1.3333       -1.1333          -0.0323     0.8977    0.0303
rgdpch       4019.8366     3956.6506           0.0084     0.4567    0.1745
openk          81.8127       79.1040           0.0666     1.1381    0.1035
grgdpch         0.5033        0.4580           0.0101     0.9280    0.0986
         eCDF Max Std. Pair Dist.
distance   0.3333        0.1319
polity2    0.1333        0.8717
rgdpch     0.4667        0.3813
openk      0.3333        0.5874
grgdpch    0.2667        1.1813


Percent Balance Improvement:
         Std. Mean Diff. Var. Ratio eCDF Mean eCDF Max
distance            93.5        96.5      74.7     34.3
polity2             91.4      -134.0      72.8     35.0
rgdpch              98.2        29.3      18.3     15.1
openk               87.3        77.6      38.2     15.9
grgdpch             92.3       -46.3     -60.8    -65.7


Sample Sizes:
          Control Treated
All            29      18
Matched        15      15
Unmatched      14       3
Discarded       0       0
```
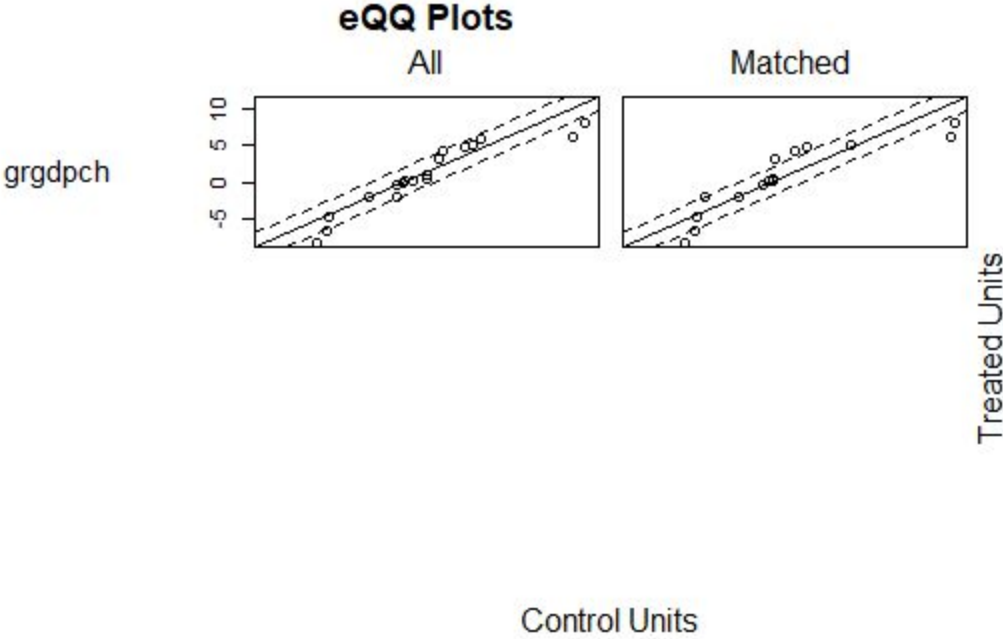
Accompanying plots distinguishing matched units on propensity score and covariate scales

(Code replication from the paper):

Distribution of Propensity Scores

Unmatched Treated Units

Matched Treated Units

Matched Control Units

Unmatched Control Units

0.2   0.4   0.6   0.8

Propensity Score

Appendix B (Genetic matching)

Results from using Genetic matching with replacement

```
Summary of Balance for Matched Data:
         Means Treated Means Control Std. Mean Diff. Var. Ratio eCDF Mean
distance        0.4544        0.4566         -0.0101      0.8583    0.0688
polity2        -2.2353       -2.4118          0.0285      0.8836    0.0267
rgdpch       5612.0470     5726.1371         -0.0151      0.6627    0.1164
openk          90.5076       90.0335          0.0117      0.8171    0.0926
grgdpch         0.8476        0.7624          0.0190      1.2821    0.0818
         eCDF Max Std. Pair Dist.
distance   0.2353          0.1678
polity2    0.1176          0.4843
rgdpch     0.2353          0.3190
openk      0.2353          0.4881
grgdpch    0.2353          0.3443

Percent Balance Improvement:
         Std. Mean Diff. Var. Ratio eCDF Mean eCDF Max
distance            98.5        82.6      70.0     53.7
polity2             92.4      -168.5      76.0     42.6
rgdpch              96.8        62.9      45.5     57.2
openk               97.8        65.1      44.7     40.7
grgdpch             85.5      -387.0     -33.5    -46.2

Sample Sizes:
              Control Treated
All              29.       18
Matched (ESS)   8.26       17
Matched          10.       17
Unmatched        19.        1
Discarded         0.        0
```

Appendix C (PSM Nearest Neighbor matching with replacement)

Results from Nearest neighbor matching using propensity scores with replacement

```
Summary of Balance for Matched Data:
          Means Treated Means Control Std. Mean Diff. Var. Ratio eCDF Mean
distance         0.4544        0.4504           0.0186      0.8118    0.0413
polity2         -2.2353       -2.1176          -0.0190      1.1224    0.0588
rgdpch        5612.0470     4476.6188           0.1502      0.8397    0.1927
openk           90.5076       98.7100          -0.2017      0.9995    0.1489
grgdpch          0.8476       -0.1194           0.2157      0.8036    0.1598
          eCDF Max Std. Pair Dist.
distance    0.2941           0.1030
polity2     0.1176           0.6647
rgdpch      0.4118           0.2296
openk       0.4118           0.5278
grgdpch     0.3529           0.9872


Percent Balance Improvement:
          Std. Mean Diff. Var. Ratio eCDF Mean eCDF Max
distance             97.2        76.3      82.0     42.1
polity2              94.9      -150.5      47.1     42.6
rgdpch               68.1        84.2       9.7     25.1
openk                61.6        99.9      11.1     -3.8
grgdpch             -64.3      -328.6    -160.7   -119.3


Sample Sizes:
              Control Treated
All               29.      18
Matched (ESS)     5.9      17
Matched           8.       17
Unmatched        21.        1
Discarded         0.        0
```

Appendix D (Description of contributions and GitHub link to R replication & extension code)

Aniket Das and Nikita Kotsehub worked collaboratively on identifying issues with the Hollyer & Rosendorff paper and brainstorming on extension methods that would improve the application of matching techniques. Nikita Kotsehub wrote the Introduction, Conclusion, and produced the plots and figures in the paper. Aniket Das coded the extension by doing the genetic matching with replacement and described the previous issues and the improvements in the Extension & Replication section, providing evidence in the Appendices. One can find our Replication & Extension R Code file on GitHub by following the link:

https://github.com/anik8das/Matching-replication-and-improvement/tree/main