

Causal Inference
Jose Alvarez Cabrera
CS112 Fall 2018

(1) Debugging

Debugging--in the 3 cases below (a through c), identify the major coding error in each case and explain how to fix it, in 1-2 sentences. DO NOT actually copy/paste corrected code:

(a) <https://gist.github.com/diamonaj/2e5d5ba5226b7b9760f5d1bf1e7bf765>

The code inputs a GenMatch object (mainly a set of weights) into the match.out parameter (which should be mainly a match dataset) of MatchBalance().

(b) <https://gist.github.com/diamonaj/3b6bc83d040098486634184d99fc4c55>

The code runs GenMatch with estimand="ATE" and does not run Match() with the estimand parameter specified. As the default is "ATT", the properties of the matching originally (and "internally") induced by GenMatch() will not translate to the output of Match().

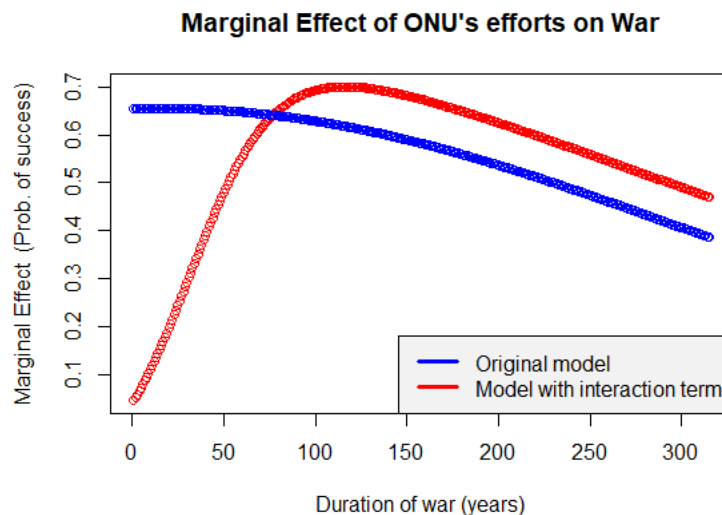
(c) <https://gist.github.com/diamonaj/a88cb40132ed8584e5182b585e1c84c8>

The code runs Match() with M=2 (number of matched (control/treated) units per (control/treated) unit), but does not initially run GenMatch() in the same way. In general, GenMatch() should be run with the same specifications as Match() if one desires the matching internally induced by GenMatch() to translate to the output of Match() as well.

Peacekeeping Data set

(2) Replicate figure 8.

Link to my code: <https://gist.github.com/josealvarez97/1a8e03d990000082997bd02437f0c9e8>



(3) Define treatment as below:

```
Tr <- rep(0, length(foo$untype))
Tr[which(foo$untype != "None")] <- 1
```

What does this mean? What is “treatment”?

As the **codebook** by Doyle & Sambanis (2000) clarifies, untype refers to the “type of UN peace operation, distinguishing between traditional and multidimensional peacekeeping/peacebuilding operations.” So, in general, when it is ‘None’, it means that there was not any kind of UN peace operation (the treatment) in the given unit.

In other words, the treatment¹ is *any* type of UN peace operation. That is why we define treatment = 1 as any value of untype different than ‘None’.

(4) Let’s pretend that you work for an NGO and your manager asks you to estimate the impact of the treatment identified above on lenient peacebuilding success 2 years and 5 years after the war.

(a) In no more than 1 sentence, articulate the causal question as best you can (being as clear as you can about treatment and control):

What is the effect of UN's peace keeping operations on the probability of *lenient peacebuilding success* 2 years, and 5 years, after war, where lenient peacebuilding success means that at least state sovereignty is not divided?

(b) In no more than 1 sentence, explain how/why SUTVA might be violated here. In no more than 1 additional sentence, explain how you could in theory use the "restrict" argument (in Match()/GenMatch()) to help address this potential problem.

As SUTVA can be violated when the treatment is not equivalent across units, it is arguably highly possible that UN's peacekeeping efforts differ largely (e.g., in terms of approaches, strategies, etc.) across units for even very small factors.

The restrict argument allows to restrict possible matches. If our expertise allows us to know that two wars are substantially different and would yield an inequivalent comparison of what would have happened if the UN had tried to intervene (because the treatment would have been inequivalent due to differences between the units), we may restrict such possible match(es).

(c) Use simple logistic regression, propensity score matching, and genetic matching to try to answer these questions.

Link to my code (same link as above, but start at line #194):

<https://gist.github.com/josealvarez97/1a8e03d990000082997bd02437f0c9e8>

¹ In the script, this variable is declared as *treatment2* to avoid any confusion (as the dataset contains other variables that refer to “other kinds” of treatment).

Table 1 - ATT Estimates of the effect on Lenient Success 2 years, and 5 years, after war.

Procedure / Method	Tmt Effect (bias adj)	Tmt Effect (no bias adj)	p-value (measures Balance)
Logistic Regression Len success 2 years	NA*	0.1139208	
Logistic Regression Len success 5 years	NA*	0.145964	
p-score Matching Len success 2 years	0.1658583	0.1666667	0.18803
p-score Matching Len success 5 years	0.1936361	0.1944444	0.18803
GenMatch Len success 2 years	0.12625	0.1666667	0.288
GenMatch Len success 5 years	0.1540287	0.1944444	0.288

*No need to provide bias-adjusted results for logistic regression--only for matching estimates.

Functional form of the propensity score model **for p-score matching**: $treatment2 \sim B_0 + B_1 * wartype + B_2 * logcost + B_3 * logcost^2 + B_4 * logcost^3 + B_5 * wardur + B_6 * wardur^2 + B_7 * wardur^3 + B_8 * factnum + B_9 * factnum2 + B_{10} * trnsfcap + B_{11} * develop + B_{12} * exp + B_{13} * decade + B_{14} * treaty$ (see codebook by Doyle & Sambanis (2000) for the precise definition of each² of this variables). **For genetic matching**, we matched on the same basic covariates, *in addition to a version of the propensity score based on a functional form that only includes the basic covariates and not any higher order terms*³, plus quadratic and cubic terms of logcost and wardur (believed to be important variables). **We measured balance** of the produced matched data sets only on the basic variables (the covariates themselves): wartype, logcost, wardur, factnum, factnum2, trnsfcap, develop, exp, decade, treaty.

² All variables except for *treatment2* will appear in the codebook. *treatment2* is exclusive of this paper, and it is defined as 1 for any value of *untype* other than 'None'. Refer back to section 3 for a detailed explanation of the meaning of *treatment2*.

³ Note: I would appreciate some feedback on this. Why was I able to achieve better balance in GenMatch with a simpler version of the propensity score functional form, if a more complex version of the propensity score functional form achieved better balance in p-score matching? It defies the theory in the claim that the better the balance in propensity score matching, the better that should have been the estimation of the propensity score model. In brief, it seems like if a less accurate version (in theory) of the propensity score model allowed GenMatch to achieve better balance, than a more accurate (but more complex) one.

MEMO: Advice in UN's Peace Building Efforts Policies

To: General Secretary of the UN's Nations

Executive Summary:

After comparing the results of different causal models and procedures for approximating the effect of an intervention, estimates **suggest a positive effect of UN's peace keeping efforts** in the *probability of* lenient success⁴ 2, and 5 years, after war. They also make obvious the fact that the effect on the probability of lenient success 5 years after war is *considerably* greater than the one of 2 years after war. Still, it matters to recall this fact, because it may support that the results are not due to chance and the units follow the expected behavior. In other words, UN's efforts do seem to have a systematic positive impact on the probability of lenient success, even if it takes a relatively long time to see the effect after the intervention.

Conclusion:

Recalling that reliable estimates from procedures that adjust for common pitfalls that induce biases yielded a similar positive approximation of the effect – an increase of approximately 0.17 and 0.19 respectively (these are highly **reliable** estimates given that several procedures that reduced bias yield similar values. (See Table 1) – **we suggest keeping the UN's efforts on peacebuilding across the world**, but *not without* continuing to seek building better models that allow to better predict when the intervention will be positive, and when it will not. **We have to recall that** the fact that UN's keeping efforts showed evidence of being positive in general, **does not mean – and does not support the idea – that every intervention should be as likely to yield a positive effect**. On this particular context, it is crucial to accompany causal models, that allow to understand the effect, with predictive models, that give greater insight into when it is more appropriate to apply the treatment (UN's interventions) *in practice*.

⁴ Assumes familiarity of the audience with the term lenient success.

References

- Doyle, M., & Sambanis, N. (2000). Data Sets Notes for International Peacebuilding: A Theoretical and Quantitative Analysis. *The American Political Science Review*, 94(4), 779-801.
doi:10.2307/2586208/. Retrieved from
<http://www.nyu.edu/gsas/dept/politics/faculty/cohen/codebook.pdf>