

\*\* Matching for Chapter 6 in Finding Pathways. Data is from Collier and Hoeffler's 2004 article  
"Greed and Grievance in Civil Wars" accessed here: <http://users.ox.ac.uk/~ball0144/research.htm>

```
**install the Mahapick package developed by David Kantor
**ssc install mahapick
```

```
**Begin by examining the distribution of the treatment/key IV
histogram sxp, percent
```

```
** Begin by computing the pairwise Mahalanobis distance for ALL possible dyads in the Collier and
Hoeffler data. Use the
** model from their regression, except exclude the level of primary exports and its square from the
matching
```

```
** This code creates an ideal match between a war case and a non-war case. It does not create a
unique match.
** for each war case this will give us the best match based on Mahalanobis distance between the war
case and one of the non-war cases.
**use Chapter6replication.dta
```

```
mahapick coldwar secm gyl peace prevwar mount georgia frac lnpop, idvar(id_year) treated(warsa)
genfile(CollierMatchedPairs) score
```

```
** Close Chapter6replication and open the new matched dataset (CollierMatchedPairs.dta)
```

```
** keep only the observations for which there is a match
keep if _matchnum==1
```

```
** Rename the _prime_id variable and the id_year variable to remind us of their meaning
ren id_year nonwar_id_year
ren _prime_id war_id_year
```

```
**save this dataset "CollierMatchedPairs"
```

```
**Go back to Chapter6replication to compute Mahalanobis Distance for the treatment variable and its
squared term
**using Chapter6replication.dta
mahascores sxp sxp2, idvar(id_year) compute_invcovarmat genfile(CollierTreatmentMaha)
ren id_year nonwar_id_year
ren _refid war_id_year
**save dataset
```

```
** Going to now add the original data from "Chapter6replication" to the Matched Data for the 46
observations.
** This requires first renaming all the variables with the prefix "war" and another dataset with
"non-war"
** the new dataset will consist of 46 matched pairs. Each row is a "matched pair" of two cases and
each column is a variable for either the war or the non-war case.
```

```
**to add a "war" prefix to all variables. This data will then be merged, only for the war cases, to
the matched dataset
```

```
**Use the Chapter6replication data
foreach var of varlist * {
    rename `var' war_`var'
}
*save new dataset CollierForMatching_War
```

```
** add non-war for non-war cases use original Chapter6replication
foreach var of varlist * {
    rename `var' nonwar_`var'
}
```

```
**save new dataset CollierForMatching_NonWar
```

```

**merge the three datasets together using the unique id_year variables names. Will end with a file
that has only the matched pairs and
**relevant data for each country in the matched pair

**also need to merge in the dataset that has the distance on the Key IV for each pair of cases
(data: CollierTreatmentMaha)
**first merge the CollierMatchedPairs dataset with the CollierTreatmentMaha dataset

**sort both the CollierForMatchingWar dataset and the CollierMatchedPairs dataset using war_id_year.
Merge the two datasets
**open the CollierMatchedPairs dataset
merge war_id_year using "CollierForMatching_War"
table _merge

**should be 46 observations that are consistent in the 2 datasets
** drop the observations that are not in both dataset (i.e. keep the 46 matches)

**repeat the same to merge the NonWar dataset, sort using the nonwar_id_year
table _merge

**should be 46 observations that are consistent in the 2 datasets
** drop the observations that are not in both dataset (i.e. keep the 46 matches)

**Now merge the data that identifies the distance between the cases
**sort both the Matched Pairs dataset and the Treatment Distance Dataset using the war and nonwar
ids
**merge in the dataset with the treatment distance between pairs
**should be 46 observations consistent in the two datasets

**This now gives us the final dataset we need for visualization.
** in this new dataset the "_score" variable is the Mahalanobis distance between the two cases
based on the
** various Xs from Collier and HOeffler
** rename as CovariateDistance

```