



Teaching with TwoRavens: Replicating Fearon and Laitin (2003)

Vito D'Orazio*

March 14, 2017[†]

Fearon and Laitin (2003) is a landmark study on civil wars. It shows that many of the post-Cold War civil conflicts were not due to ethnic differences, but rather to factors that enabled insurgencies to thrive. These factors include mountainous terrain, low GDP, high population, and political instability.

This document will guide you through a replication of the statistical results in Fearon and Laitin (2003) using TwoRavens.

*School of Economic, Political, and Policy Sciences at the University of Texas at Dallas; vjdorazio@gmail.com

[†] Current version of this document available at <http://2ra.vn/teaching>

1 Access the data

Follow the link to the [Teaching with TwoRavens dataverse](#). Click on the Fearon and Laitin replication dataset. Next to the “Download” button for the file `repdata2.tab`, click “Explore.” This should open a page, as shown in figure 1.¹

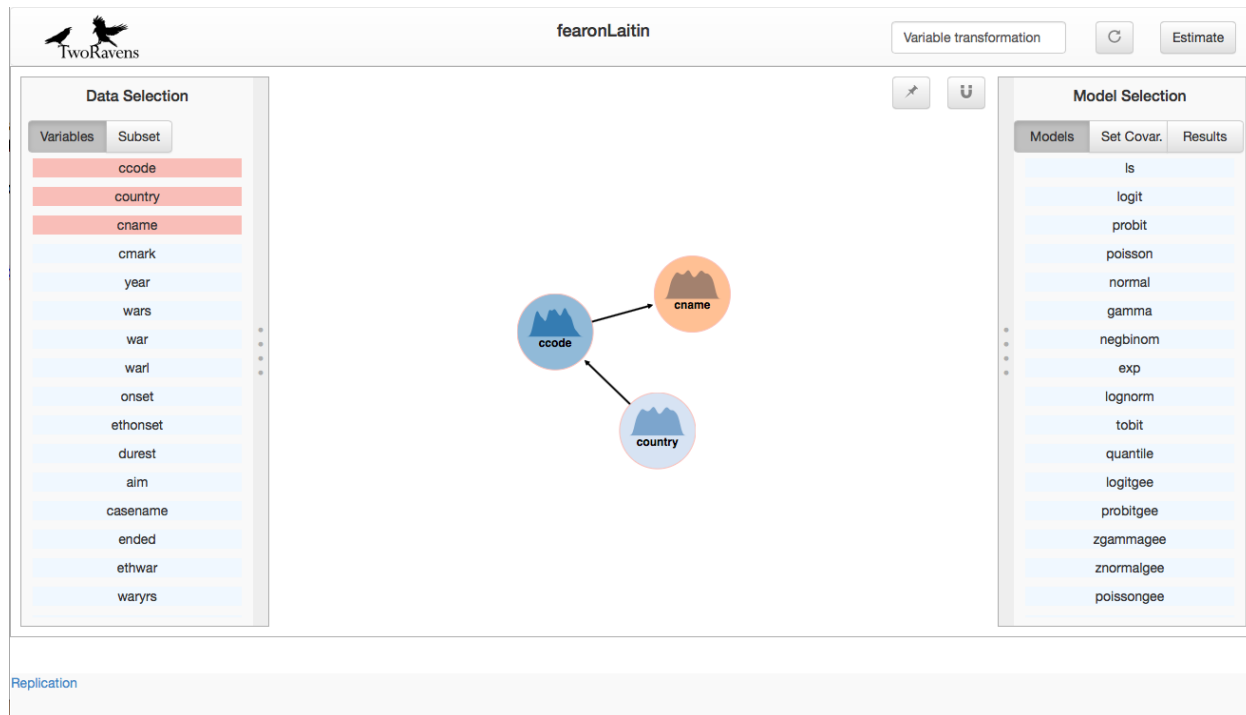


Figure 1: TwoRavens loaded with the Fearon and Laitin data.

In this module we will replicate the statistical analyses presented in Fearon and Laitin (2003). Their replication file, retrieved from <http://web.stanford.edu/group/ethnic/publicdata/publicdata.html>, reads as follows:

```
/* .do File to produce Table 1 results from replication data set */  
  
use repdata ,clear  
  
/* Model #1 */  
logit onset warl gdpnl lpopl lmtnest ncontig Oil nwstate instab polity2l ethfrac relfrac ,nolog  
  
/* Model #2 */  
logit ethonset warl gdpnl lpopl lmtnest ncontig Oil nwstate instab polity2l ethfrac relfrac if second > .049999,nolog  
  
/* Model #3 */  
logit onset warl gdpnl lpopl lmtnest ncontig Oil nwstate instab anocl deml ethfrac relfrac ,nolog  
  
/* Model #4 */  
logit emponset empwarl empdpnl emplpopl emlmtnest empncontig Oil nwstate instab empethfrac ,nolog  
  
/* Model #5 */  
logit cowonset cowwarl gdpnl lpopl lmtnest ncontig Oil nwstate instab anocl deml ethfrac relfrac ,nolog
```

¹Please refer to the [User Guide](#) for information about using TwoRavens (D’Orazio and Honaker, 2016). Fuller details are also found in Honaker and D’Orazio (2014). Please refer to Fearon and Laitin (2003) for additional descriptions of the data and variables.

This is a replication script for Stata, so some knowledge of Stata syntax is required. `logit` means that we are using the logistic regression. The first variable name after `logit` is the dependent variable, which in this case is either `onset`, `ethonset`, `emponset`, or `cowonset`. In Stata syntax, the variable names following the dependent variable are the independent variables.

All of these variables are listed in the Data Selection panel of TwoRavens, under the Variables tab. If you mouseover these variable names, or click on them and mouseover the pebble in the center panel, you can view additional summary statistics, a brief description of the variable, and see its distribution.

Note that for Model #2, Fearon and Laitin subset the data for all observations where `second > 0.49999`, and estimates the model on that subset. Finally, the `no log` at the end of each line suppresses some of the information that Stata automatically prints to the console, and is not relevant for us here.

2 Instructions for Replicating Fearon and Laitin Models

Begin with a clear modeling space. This means no variables are selected and thus no pebbles are in the center panel.

- Click the magnet button in the top right of the center panel
- Alternatively, click each selected variable name in the variable list in the left panel

2.1 Model #1

- Click the names of all the variables in Model #1 in the left panel
 - `onset warl gdpenl lpopl lmtnest ncontig Oil nwstate instab polity2l ethfrac relfrac`
- Mouseover `onset` and select **Dep Var** to tag `onset` as the dependent variable
 - The `onset` pebble should have a light blue outline and no fill color, signaling that this pebble has been tagged as the dependent variable
- Right click on each independent variable to create an arrow, and left click on `onset` to draw an arrow from the IV to the DV
 - These arrows signify the direction of the relationship

The resulting image in the center panel should appear like figure 2. Now that the relationships among variables have been specified, we now select the appropriate statistical model:

- Click on `logit` in the Model Selection panel on the right, under the **Models** tab

At this point, the variables have been selected, the relationships among them specified, and statistical model has been selected. Now, estimate the model:

- Click the **Estimate** button in the upper right

Once the model is estimated, the results will appear in the right panel under the **Results** tab. Click the **Models** tab to return the right panel to the model list.

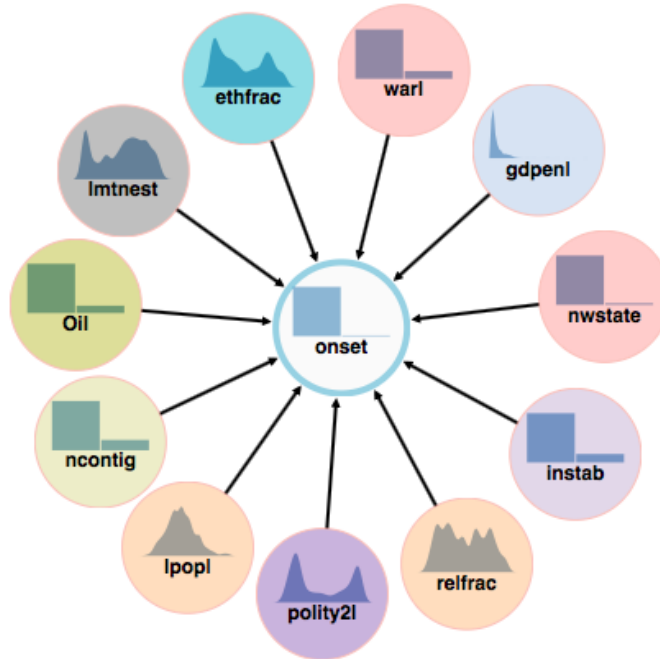


Figure 2: Relationships in Fearon and Laitin's Model #1.

2.2 Model #2

- Click the `onset` variable in the left panel
 - Removes this pebble from the center panel
- Click `ethonset` and make sure the names of all the variables in Model #2 are in the center panel
 - `ethonset warl gdpenl lpopl lmtnest ncontig Oil nwstate instab polity2l ethfrac relfrac`
- Mouseover `ethonset` and select `Dep Var` to tag `ethonset` as the dependent variable
- Right click on each independent variable to create an arrow, and left click on `ethonset` to draw an arrow from the IV to the DV

Notice that we need to estimate this model on a subset of the data. Specifically, on the subset where `second > .049999`.

- Click the `second` variable in the left panel
- Click `Subset` tab in the left panel
- Brush the region of the `second` plot greater than `.049999`
 - This tells us that we want to estimate the model on a subset of the data
- Click `Estimate` in the upper right

Once the model is estimated, the results will appear in the right panel under the `Results` tab. You may toggle between `Model 1` and `Model 2` to see different sets of results. Click the `Models` tab to return the right panel to the model list.

2.3 Model #3

- Click `ethonset` to drop this pebble from the modeling space
- Click `onset` to add this pebble
- Mouseover `onset` and click `Dep Var`
- Add all variables in Model #3 to the center panel
 - `onset warl gdpenl lpopl lmtnest ncontig Oil nwstate instab anocl deml ethfrac relfrac`
 - It is fine to have extra variables in the center panel, so don't feel like you have to drop those you aren't using specifically for Model #3
- Right click on all independent variables and draw arrows to the dependent variable, `onset`
- Ensure that `logit` is still selected in the right panel under the `Models` tab
- Click `Estimate` in the upper right

Once the model is estimated, the results will appear in the right panel under the `Results` tab. Click the `Models` tab to return the right panel to the model list.

2.4 Model #4

- Click the magnet button in the upper right of the center panel
 - Note that most of the variables in Model #4 are different from Model #3
- Click on the names of all variables in Model #4
 - `emponset empwarl empgdpenl emplpopl emplmtnest empncontig Oil nwstate instab em-pethfrac`
- Right click on all independent variables and draw arrows to the dependent variable, `emponset`
- Ensure that `logit` is still selected in the right panel under the `Models` tab
- Click `Estimate` in the upper right

Once the model is estimated, the results will appear in the right panel under the `Results` tab. Click the `Models` tab to return the right panel to the model list.

2.5 Model #5

- Click the magnet button in the upper right of the center panel
 - Most of the variables in Model #5 are different from Model #4
- Click on the names of all variables in Model #5
 - `cowonset cowwarl gdpenl lpopl lmtnest ncontig Oil nwstate instab anocl deml ethfrac relfrac`
- Right click on all independent variables and draw arrows to the dependent variable, `cowonset`

- Ensure that `logit` is still selected in the right panel under the `Models` tab
- Click `Estimate` in the upper right

Once the model is estimated, the results will appear in the right panel under the `Results` tab.

3 Results and Replication

In the `Results` tab, all five models will be listed, as seen in figure 3. Click on the models to toggle through the results. For a complete replication script in R, including all system information necessary for a complete replication of your results, click on the `Replication` link in the lower left.

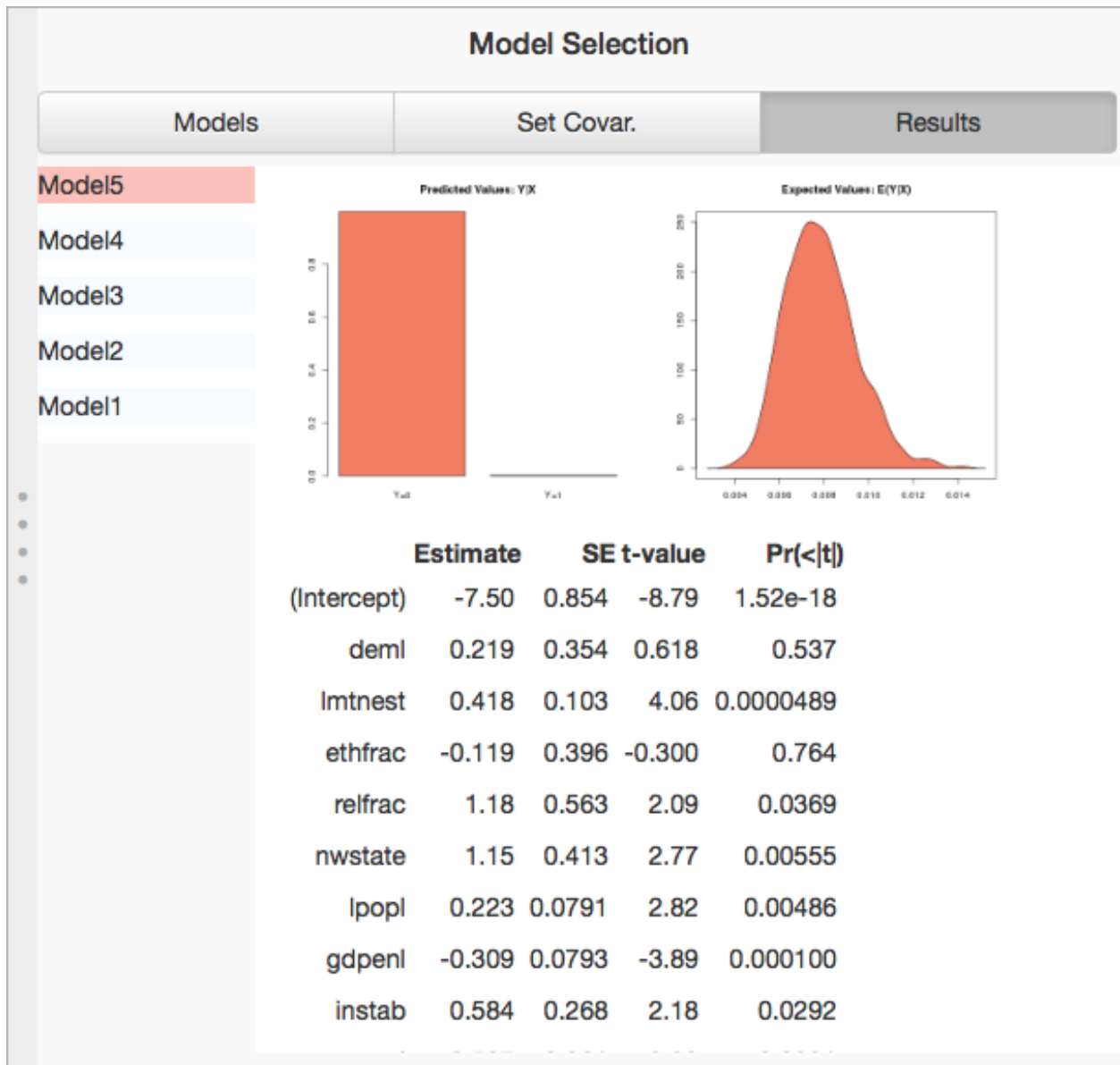


Figure 3: Results after estimating Fearon and Laitin models.

These results should be identical to the results in Fearon and Laitin's Table 1, which appears below:

	Model				
	(1) Civil War	(2) "Ethnic" War	(3) Civil War	(4) Civil War (Plus Empires)	(5) Civil War (COW)
Prior war	-0.954** (0.314)	-0.849* (0.388)	-0.916** (0.312)	-0.688** (0.264)	-0.551 (0.374)
Per capita income ^{a,b}	-0.344*** (0.072)	-0.379*** (0.100)	-0.318*** (0.071)	-0.305*** (0.063)	-0.309*** (0.079)
log(population) ^{a,b}	0.263*** (0.073)	0.389*** (0.110)	0.272*** (0.074)	0.267*** (0.069)	0.223** (0.079)
log(% mountainous)	0.219** (0.085)	0.120 (0.106)	0.199* (0.085)	0.192* (0.082)	0.418*** (0.103)
Noncontiguous state	0.443 (0.274)	0.481 (0.398)	0.426 (0.272)	0.798** (0.241)	-0.171 (0.328)
Oil exporter	0.858** (0.279)	0.809* (0.352)	0.751** (0.278)	0.548* (0.262)	1.269*** (0.297)
New state	1.709*** (0.339)	1.777*** (0.415)	1.658*** (0.342)	1.523*** (0.332)	1.147** (0.413)
Instability ^a	0.618** (0.235)	0.385 (0.316)	0.513* (0.242)	0.548* (0.225)	0.584* (0.268)
Democracy ^{a,c}	0.021 (0.017)	0.013 (0.022)			
Ethnic fractionalization	0.166 (0.373)	0.146 (0.584)	0.164 (0.368)	0.490 (0.345)	-0.119 (0.396)
Religious fractionalization	0.285 (0.509)	1.533* (0.724)	0.326 (0.506)		1.176* (0.563)
Anocracy ^a			0.521* (0.237)		0.597* (0.261)
Democracy ^{a,d}			0.127 (0.304)		0.219 (0.354)
Constant	-6.731*** (0.736)	-8.450*** (1.092)	-7.019*** (0.751)	-6.801*** (0.681)	-7.503*** (0.854)
<i>N</i>	6327	5186	6327	6360	5378

Note: The dependent variable is coded "1" for country years in which a civil war began and "0" in all others. Standard errors are in parentheses. Estimations performed using Stata 7.0. * $p < .05$; ** $p < .01$; *** $p < .001$.

^a Lagged one year.
^b In 1000's.
^c Polity IV; varies from -10 to 10.
^d Dichotomous.

Figure 4: Fearon and Laitin (2003) Table 1

4 Extensions

At this point, it is straightforward to perform additional robustness checks or extensions on any of these five models using variables in the data.

For example, notice that there is a dummy variable in the data called `colfra`. This variable indicates whether the state is a former French colony. On a different sample, Collier and Hoeffler (2002) investigate whether being a former French colony makes one less likely to experience a civil war. They find that the coefficient is negative, but not statistically significant at 0.05.

Using the Fearon and Laitin data, we can test the hypothesis that being a former French colony makes one less likely to experience a civil war by including the `colfra` variable in Fearon and Laitin's model.

- Replicate the construction of Model #1
- Add `colfra` as an independent variable
- Click Estimate

As we can see, the coefficient is negative, but not statistically significant at 0.05. This is consistent with the findings in Collier and Hoeffler (2002).

References

- Collier, Paul and Anke Hoeffler. 2002. “On the incidence of civil war in Africa.” *Journal of conflict resolution* 46(1):13–28.
- D’Orazio, Vito and James Honaker. 2016. *A User Guide to TwoRavens: An overview of features and capabilities*.
URL: <http://2ra.vn/papers/tworavens-guide.pdf>
- Fearon, James D and David D Laitin. 2003. “Ethnicity, insurgency, and civil war.” *American political science review* 97(01):75–90.
- Honaker, James and Vito D’Orazio. 2014. Statistical Modeling by Gesture: A Graphical, Browser-based Statistical Interface for Data Repositories. In *Extended Proceedings of the 25th ACM Conference on Hypertext and Social Media*. Vol. 1210 of *DataWiz 2014* ACM.
URL: http://ceur-ws.org/Vol-1210/datawiz2014_05.pdf