Multi100 Project: Does Warfare Matter? Severity, Duration, and Outcomes of Civil Wars

Danilo Freire*

10 July 2022

1 Introduction

This Markdown file contains the code I used to replicate Figure 01 (p. 1399) and Table 01 (p. 1400-01) in "*Does Warfare Matter? Severity, Duration, and Outcomes of Civil Wars*" by Laia Balcells and Stathis N. Kalyvas (2014). The paper was originally published in the *Journal of Conflict Resolution*. The claim to be evaluated here is that "SNC ["symmetric nonconventional"] conflicts are likely to last longer than conventional conflicts but likely to be shorter than irregular ones (Hypothesis 1)." The article files are available at: https://osf.io/3j72h/?view_only=5b822a6c354940c5a5127c97d2c588fd.

The replication indicates that the result is indeed robust and supports the authors' main claims. I was able to successfully replicate both Figure 01 and Table 01 using the authors' Stata do files. I did not find any errors in their scripts. The analyses were conducted using Stata 15 and the `Statamarkdown` package for `R` (R Core Team 2022).

This replication is part of the Multi100 Project, a crowdsourced project whose goal is to assess the robustness of findings in the social and behavioural sciences. My researcher ID is 9EFM2 and all files required to reproduce the contents of this document are available at: https://github.com/danilofreire/multi100-project. Feel free to contact me at danilofreire@gmail.com or dfreire@lincoln.ac.uk if you have any questions.

*Senior Lecturer, School of Social and Political Sciences, University of Lincoln, Lincoln, UK, dfreire@lincoln.ac.uk, https://danilofreire.github.io.

2 Figure 01 - Kaplan Meier Survival Estimates

In this section, I replicate Figure 01 in Balcells and Kalyvas (2014, p. 1399). First, I install the packages required for the analysis.

```r
# Install necessary packages
r <- getOption("repos")
r["CRAN"] <- "https://cran.rstudio.com/"
options(repos = r)

# List of packages
packages <- c("devtools", "knitr", "rmarkdown")

installed_packages <- packages %in% rownames(installed.packages())
if (any(installed_packages == FALSE)) {
  install.packages(packages[!installed_packages])
}
invisible(lapply(packages, library, character.only = TRUE))

# Install and load Statamarkdown
devtools::install_github("Hemken/Statamarkdown")
library(Statamarkdown)
```

Then I use the authors' Stata script to create some of the variables included in the models.

```
// Load the dataset
sysuse ./TR_panelformat_Replication.dta

// Create variable warend
sort id year
gen warend=0
replace warend=1 if year==yrend
```

```
// Some months have 31 days, others 30, 29, 28.
gen curyear=mdy(moend,31,yrend) if yrend!=. & year==yrend & moend!=.
format %td curyear
replace curyear=mdy(moend,30,yrend) if yrend!=. & year==yrend & moend!=. & curyear==.
replace curyear=mdy(moend,29,yrend) if yrend!=. & year==yrend & moend!=. & curyear==.
replace curyear=mdy(moend,28,yrend) if yrend!=. & year==yrend & moend!=. & curyear==.
replace curyear=date("3112"+string(year),"DMY") if curyear==.

// Create variable warbegin
gen warbegin=0
sort id year
by id: replace warbegin=1 if _n==1

// Create variable origyear
gen origyear=.
format %td origyear
replace origyear=mdy(most,1,yrst) if year==yrst & most!=.
replace origyear=mdy(1,1,yrst) if year==yrst & most==.
replace origyear=date("3112"+string(year-1), "DMY") if year!=yrst

// Kaplan Meier estimates
stset curyear, id(id) failure(warend==1) time0(origyear) origin(time origyear) scale(30.41667)

// Plot survival curve
sts graph, by(technologyrebellion) ///
legend(label(1 "Conventional") label (2 "Irregular") label(3 "SNC"))

// Export graph
quietly graph export kaplan-meier.pdf, replace
```

```
// Save dataset
save df.dta
```

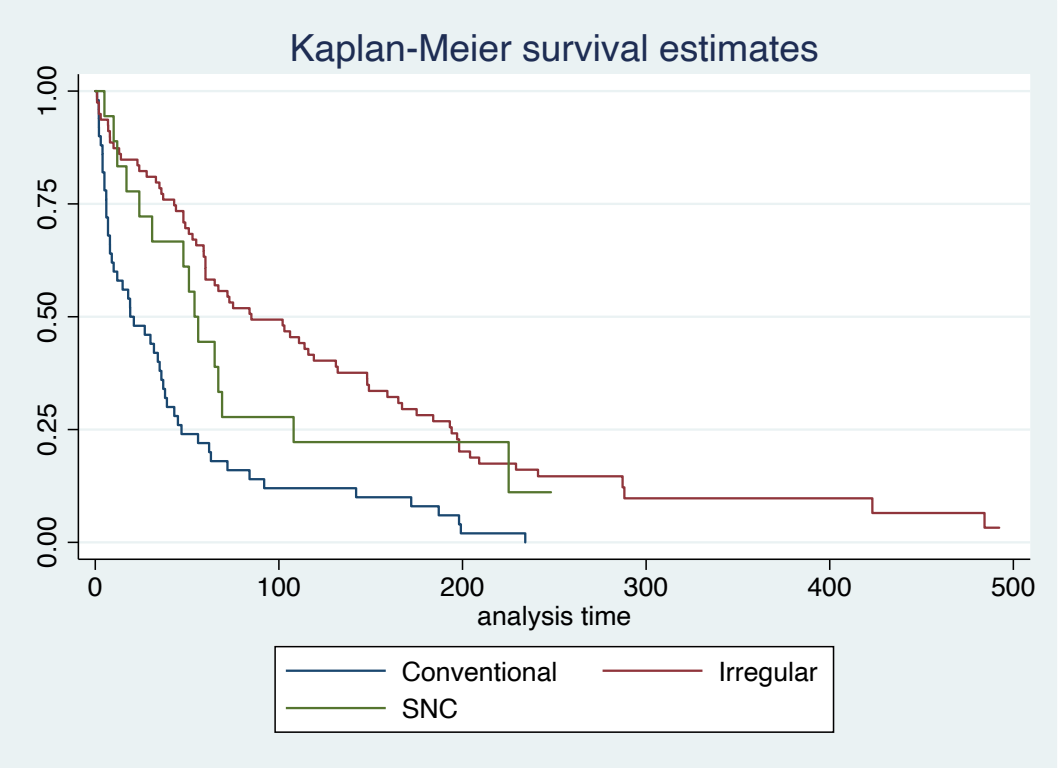

Figure 1: Kaplan Meier Survival Estimates

3 Table 01 - Weibull Regression on Civil War Duration

The code below replicates models 1 to 4 reported in Table 01 (Balcells and Kalyvas 2014, 1400). All estimates are identical to those included in the main paper.

```
// Load dataset
sysuse ./df.dta

// Set data
stset curyear, id(id) failure(warend==1) time0(origyear) origin(time origyear) scale(30.41667)

// Model 01
streg i.technologyrebellion, d(w) nolog vce(robust) time
```

```
// Model 02
streg i.technologyrebellion post1990 lmtnest ///
lpop gdpenl oil ethfrac deml, d(w) nolog  vce(robust) time

// Model 03
streg i.technologyrebellion post1990 lmtnest ///
lpop gdpenl oil ethfrac deml milper Extsupp_gov_bi ///
Extsupp_reb_bi, d(w) nolog  vce(robust) time

// Model 04
streg i.technologyrebellion post1990 lmtnest ///
lpop gdpenl oil ethfrac deml milper Extsupp_gov_bi ///
Extsupp_reb_bi western eeurop asia ssafrica lamerica, ///
d(w) nolog  vce(robust) time
```

```
                id:  id
     failure event:  warend == 1
obs. time interval:  (origyear, curyear]
 exit on or before:  failure
    t for analysis:  (time-origin)/30.41667
            origin:  time origyear

------------------------------------------------------------------------------
      1,206  total observations
          0  exclusions
------------------------------------------------------------------------------
      1,206  observations remaining, representing
        147  subjects
        135  failures in single-failure-per-subject data
 13,237.675  total analysis time at risk and under observation
```

```
                                          at risk from t =         0
                                earliest observed entry t =         0
                                     last observed exit t =  492.2958

         failure _d:  warend == 1
   analysis time _t:  (curyear-origin)/30.41667
             origin:  time origyear
                 id:  id

Weibull AFT regression

No. of subjects      =          147             Number of obs    =       1,206
No. of failures      =          135
Time at risk         =  13237.67526
                                                 Wald chi2(2)     =       26.36
Log pseudolikelihood =   -251.07231             Prob > chi2      =      0.0000

                                       (Std. Err. adjusted for 147 clusters in id)
------------------------------------------------------------------------------
                    |               Robust
                 _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------------
technologyrebellion |
         Irregular  |   1.203136   .2343467     5.13   0.000     .7438244    1.662447
               SNC  |   .8298717   .3767133     2.20   0.028     .0915272    1.568216
                    |
             _cons  |   3.680266   .1943823    18.93   0.000     3.299284    4.061249
--------------------+---------------------------------------------------------------
              /ln_p |  -.1656716   .0614201    -2.70   0.007    -.2860527   -.0452905
--------------------+---------------------------------------------------------------
```

```
                 p |   .8473245   .0520427                      .751223    .9557198
               1/p |   1.180185   .0724871                     1.046332    1.331163
------------------------------------------------------------------------------

         failure _d:  warend == 1
   analysis time _t:  (curyear-origin)/30.41667
             origin:  time origyear
                 id:  id

Weibull AFT regression

No. of subjects      =          131             Number of obs    =         906
No. of failures      =          104
Time at risk         =  9802.223583
                                                 Wald chi2(9)     =       33.37
Log pseudolikelihood =   -205.20084             Prob > chi2      =      0.0001

                                       (Std. Err. adjusted for 131 clusters in id)
------------------------------------------------------------------------------
                   |               Robust
                _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------------+----------------------------------------------------------------
technologyrebellion |
         Irregular  |   .8464113   .3308443     2.56   0.011     .1979683    1.494854
               SNC  |    .505669   .3995707     1.27   0.206    -.2774751    1.288813
                    |
          post1990 |  -.5473025   .3089441    -1.77   0.076    -1.152822    .0582169
            lmtnest |   .0389837   .1010648     0.39   0.700    -.1590996     .237067
               lpop |   .0779611   .1023826     0.76   0.446    -.1227052    .2786273
          gdpenl_fl |   .1626186   .1304021     1.25   0.212    -.0929648     .418202
```

```
             oil_fl |  -.2760041   .3886023    -0.71   0.478    -1.037651    .4856423
            ethfrac |   .8399731   .4583639     1.83   0.067    -.0584036     1.73835
               deml |   .1196862   .3832193     0.31   0.755    -.6314099    .8707823
              _cons |   2.552892   .9921166     2.57   0.010     .6083788    4.497404
--------------------+------------------------------------------------------------------
              /ln_p |  -.1902052   .0680165    -2.80   0.005    -.3235151   -.0568954
--------------------+------------------------------------------------------------------
                  p |   .8267894   .0562353                      .7236011    .9446928
                1/p |   1.209498   .0822658                      1.058545    1.381977
------------------------------------------------------------------------------------------

         failure _d:  warend == 1
   analysis time _t:  (curyear-origin)/30.41667
             origin:  time origyear
                 id:  id

Weibull AFT regression

No. of subjects      =          131             Number of obs    =         906
No. of failures      =          104
Time at risk         =  9802.223583
                                                 Wald chi2(12)    =       56.65
Log pseudolikelihood =   -196.71922             Prob > chi2      =      0.0000

                                        (Std. Err. adjusted for 131 clusters in id)
------------------------------------------------------------------------------------------
                   |               Robust
                _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------------------
technologyrebellion |
```

```
         Irregular  |   .9331167    .315459     2.96   0.003     .3148284    1.551405
               SNC  |   .6323871   .4241451     1.49   0.136    -.1989219    1.463696
                    |
          post1990  |  -.3287936   .3069017    -1.07   0.284      -.93031    .2727227
           lmtnest  |     .11468   .0978604     1.17   0.241    -.0771228    .3064827
              lpop  |   .1109128   .1051042     1.06   0.291    -.0950877    .3169133
         gdpenl_fl  |   .1965522   .1159414     1.70   0.090    -.0306887    .4237932
            oil_fl  |  -.1338352   .3576492    -0.37   0.708    -.8348148    .5671444
           ethfrac  |   .7465131   .4094533     1.82   0.068    -.0560005    1.549027
              deml  |   .1408604   .3480787     0.40   0.686    -.5413614    .8230822
           milper  |  -.0000533   .0000615    -0.87   0.386    -.0001738    .0000672
    Extsupp_gov_bi  |   .5323196   .2973185     1.79   0.073     -.050414    1.115053
    Extsupp_reb_bi  |   .7807511   .2709429     2.88   0.004     .2497128    1.311789
             _cons  |   1.048939   1.001576     1.05   0.295    -.9141136    3.011991
--------------------+------------------------------------------------------------------
             /ln_p  |  -.1114433   .0676577    -1.65   0.100      -.24405    .0211635
--------------------+------------------------------------------------------------------
                 p  |   .8945421   .0605227                     .7834484    1.021389
               1/p  |    1.11789   .0756339                     .9790589    1.276408
------------------------------------------------------------------------------------

        failure _d:  warend == 1
  analysis time _t:  (curyear-origin)/30.41667
            origin:  time origyear
                id:  id

Weibull AFT regression

No. of subjects      =          131              Number of obs    =         906
No. of failures      =          104
```

```
Time at risk         =  9802.223583

                                          Wald chi2(16)    =       76.43
Log pseudolikelihood =   -195.18854       Prob > chi2      =      0.0000

                                    (Std. Err. adjusted for 131 clusters in id)
------------------------------------------------------------------------------
                    |               Robust
                 _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------------
technologyrebellion |
         Irregular  |    .868049   .3456335     2.51   0.012     .1906197    1.545478
               SNC  |   .5774997   .4547154     1.27   0.204    -.3137261    1.468725
                    |
          post1990 |  -.2930918   .3215899    -0.91   0.362    -.9233963    .3372128
           lmtnest |   .1554951   .0994134     1.56   0.118    -.0393516    .3503418
              lpop |   .0773637    .105569     0.73   0.464    -.1295478    .2842751
         gdpenl_fl |   .2938091   .1516979     1.94   0.053    -.0035134    .5911316
            oil_fl |  -.1133936   .3515312    -0.32   0.747    -.8023821    .5755949
           ethfrac |   .5641797   .4883649     1.16   0.248    -.3929979    1.521357
              deml |   .1201555   .3710263     0.32   0.746    -.6070427    .8473537
            milper |   .0000139   .0000696     0.20   0.842    -.0001226    .0001503
     Extsupp_gov_bi |   .5362917   .3090616     1.74   0.083     -.069458    1.142041
     Extsupp_reb_bi |   .7395934    .301807     2.45   0.014     .1480625    1.331124
           western |     -.6284   .5737524    -1.10   0.273    -1.752934     .496134
            eeurop |  -.2630128   .5017461    -0.52   0.600    -1.246417    .7203914
              asia |   .5034296   .4780022     1.05   0.292    -.4334375    1.440297
          ssafrica |   .4705192   .4778814     0.98   0.325    -.4661111     1.40715
          lamerica |   .2331925   .5420171     0.43   0.667    -.8291415    1.295526
             _cons |    .930519    1.12554     0.83   0.408    -1.275499    3.136537
--------------------+---------------------------------------------------------------
```

```
              /ln_p |    -.10289    .070159    -1.47   0.143    -.2403991    .0346191
--------------------+------------------------------------------------------------------
                  p |   .9022262   .0632993                      .786314    1.035225
                1/p |   1.108369   .0777621                     .9659733    1.271757
------------------------------------------------------------------------------
```

4 Session Information

```
sessionInfo()
```

```
R version 4.2.1 (2022-06-23)
Platform: x86_64-apple-darwin17.0 (64-bit)
Running under: macOS Monterey 12.4

Matrix products: default
BLAS:   /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libRblas.0.dylib
LAPACK: /Library/Frameworks/R.framework/Versions/4.2/Resources/lib/libRlapack.dylib

locale:
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8

attached base packages:
[1] stats     graphics  grDevices utils
[5] datasets  methods   base

other attached packages:
[1] Statamarkdown_0.7.1 knitr_1.39
[3] devtools_2.4.3      usethis_2.1.6
[5] rmarkdown_2.14      nvimcom_0.9-131

loaded via a namespace (and not attached):
```

```
 [1] magrittr_2.0.3    pkgload_1.3.0
 [3] R6_2.5.1          rlang_1.0.3
 [5] fastmap_1.1.0     stringr_1.4.0
 [7] tools_4.2.1       pkgbuild_1.3.1
 [9] xfun_0.31         sessioninfo_1.2.2
[11] cli_3.3.0         remotes_2.4.2
[13] htmltools_0.5.2   ellipsis_0.3.2
[15] yaml_2.3.5        digest_0.6.29
[17] lifecycle_1.0.1   crayon_1.5.1
[19] processx_3.6.1    purrr_0.3.4
[21] callr_3.7.0       codetools_0.2-18
[23] fs_1.5.2          ps_1.7.1
[25] curl_4.3.2        memoise_2.0.1
[27] glue_1.6.2        cachem_1.0.6
[29] evaluate_0.15     stringi_1.7.6
[31] compiler_4.2.1    prettyunits_1.1.1
```

References

Balcells, L. and Kalyvas, S. N. (2014). Does Warfare Matter? Severity, Duration, and Outcomes of Civil Wars. *Journal of Conflict Resolution*, 58(8):1390–1418.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.