csdid with RC/Panel

First. Some other tools. Please copy this file, run it, and type:

```
** just for data example
ssc install frause
** for installing csdid2
fra install fra
fra install csdid2
```

Now the example.

First Run csdid with panel data

```
frause mpdta, clear
set seed 1
csdid lemp, ivar( countyreal) gvar(first) time(year)
```

```
(Written by R.              )
.............
Difference-in-difference with Multiple Time Periods

                                                      Number of obs = 2,500
Outcome model  : regression adjustment
Treatment model: none
------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
g2004        |
 t_2003_2004 |  -.0105032    .023251    -0.45   0.651    -.0560744    .0350679
 t_2003_2005 |  -.0704232   .0309848    -2.27   0.023    -.1311522   -.0096941
 t_2003_2006 |  -.1372587   .0364357    -3.77   0.000    -.2086713   -.0658461
 t_2003_2007 |  -.1008114   .0343592    -2.93   0.003    -.1681542   -.0334685
-------------+----------------------------------------------------------------
```

```
g2006        |
 t_2003_2004 |   .0065201   .0233268     0.28   0.780    -.0391996    .0522398
 t_2004_2005 |  -.0027508   .0195586    -0.14   0.888    -.0410849    .0355833
 t_2005_2006 |  -.0045946   .0177552    -0.26   0.796    -.0393942    .0302049
 t_2005_2007 |  -.0412245   .0202292    -2.04   0.042    -.0808729    -.001576
-------------+----------------------------------------------------------------
g2007        |
 t_2003_2004 |   .0305067   .0150336     2.03   0.042     .0010414    .0599719
 t_2004_2005 |  -.0027259   .0163958    -0.17   0.868    -.0348611    .0294093
 t_2005_2006 |  -.0310871   .0178775    -1.74   0.082    -.0661264    .0039522
 t_2006_2007 |  -.0260544   .0166554    -1.56   0.118    -.0586985    .0065896
------------------------------------------------------------------------------
Control: Never Treated

See Callaway and Sant'Anna (2021) for details
```

Then, drop at randome 10% of the data, so technically have unbalanced panel. It can be estimated using repeated crossection

```
drop if runiform()<.1
** as Panel with unbalanced Data
csdid lemp lpop,  ivar( countyreal) gvar(first) time(year)
```

```
(239 observations deleted)
Panel is not balanced
Will use observations with Pair balanced (observed at t0 and t1)
.............
Difference-in-difference with Multiple Time Periods

                                                      Number of obs = 2,227
Outcome model  : least squares
Treatment model: inverse probability
------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
g2004        |
 t_2003_2004 |  -.0106321   .0213814    -0.50   0.619     -.052539    .0312747
 t_2003_2005 |  -.0997826   .0338647    -2.95   0.003    -.1661562   -.0334089
 t_2003_2006 |   -.164467   .0421525    -3.90   0.000    -.2470845   -.0818495
 t_2003_2007 |  -.1372957    .029445    -4.66   0.000    -.1950069   -.0795845
-------------+----------------------------------------------------------------
g2006        |
```

```
 t_2003_2004 |  -.0102468   .0250207    -0.41   0.682    -.0592865    .0387928
 t_2004_2005 |  -.0012623   .0190316    -0.07   0.947    -.0385636    .0360389
 t_2005_2006 |   .0018708   .0192118     0.10   0.922    -.0357835    .0395252
 t_2005_2007 |  -.0439769   .0200335    -2.20   0.028    -.0832418    -.004712
-------------+----------------------------------------------------------------
g2007        |
 t_2003_2004 |   .0313098    .015797     1.98   0.047     .0003482    .0622713
 t_2004_2005 |   .0023558   .0162905     0.14   0.885    -.0295731    .0342846
 t_2005_2006 |  -.0317725   .0198823    -1.60   0.110    -.0707412    .0071962
 t_2006_2007 |  -.0399225   .0163005    -2.45   0.014    -.0718709   -.0079741
------------------------------------------------------------------------------
Control: Never Treated

See Callaway and Sant'Anna (2021) for details
```

```
** As RC
csdid lemp lpop,  cluster( countyreal) gvar(first) time(year)
```

```
.............
Difference-in-difference with Multiple Time Periods

                                                     Number of obs = 2,261
Outcome model  : least squares
Treatment model: inverse probability
                          (Std. err. adjusted for 500 clusters in countyreal)
------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
g2004        |
 t_2003_2004 |  -.0126297   .0345604    -0.37   0.715    -.0803669    .0551075
 t_2003_2005 |  -.0913197   .0384773    -2.37   0.018    -.1667338   -.0159055
 t_2003_2006 |  -.2070359    .049318    -4.20   0.000    -.3036974   -.1103744
 t_2003_2007 |   -.136554   .0317776    -4.30   0.000     -.198837   -.0742711
-------------+----------------------------------------------------------------
g2006        |
 t_2003_2004 |  -.0098464   .0303239    -0.32   0.745    -.0692801    .0495873
 t_2004_2005 |  -.0381135   .0237243    -1.61   0.108    -.0846122    .0083852
 t_2005_2006 |  -.0141839   .0318463    -0.45   0.656    -.0766015    .0482337
 t_2005_2007 |  -.0230668   .0350484    -0.66   0.510    -.0917604    .0456267
-------------+----------------------------------------------------------------
g2007        |
```

```
 t_2003_2004 |   .0596817   .0309245     1.93   0.054    -.0009293    .1202926
 t_2004_2005 |  -.0467024   .0294592    -1.59   0.113    -.1044414    .0110365
 t_2005_2006 |  -.0388017    .034514    -1.12   0.261     -.106448    .0288445
 t_2006_2007 |  -.0431742   .0312701    -1.38   0.167    -.1044625     .018114
-------------------------------------------------------------------------------
Control: Never Treated

See Callaway and Sant'Anna (2021) for details
```

The second one produces odd results, which is not unexpected.

Now using the residuals Idea:

```
** Getting residuals wo year
qui:reghdfe lemp, abs(countyreal) resid
** AS RC but with residuals
csdid _reghdfe_resid lpop, cluster( countyreal) gvar(first) time(year)
```

```
.............
Difference-in-difference with Multiple Time Periods

                                                      Number of obs = 2,260
Outcome model  : least squares
Treatment model: inverse probability
                          (Std. err. adjusted for 499 clusters in countyreal)
-------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+-----------------------------------------------------------------
g2004        |
 t_2003_2004 |  -.0235323   .0248084    -0.95   0.343     -.072156    .0250913
 t_2003_2005 |  -.0901505   .0312695    -2.88   0.004    -.1514376   -.0288633
 t_2003_2006 |  -.1579142   .0380608    -4.15   0.000     -.232512   -.0833165
 t_2003_2007 |   -.119084   .0323402    -3.68   0.000    -.1824697   -.0556984
-------------+-----------------------------------------------------------------
g2006        |
 t_2003_2004 |  -.0141947    .024901    -0.57   0.569    -.0629997    .0346104
 t_2004_2005 |  -.0051773   .0197164    -0.26   0.793    -.0438208    .0334661
 t_2005_2006 |    .001378   .0197077     0.07   0.944    -.0372484    .0400043
 t_2005_2007 |  -.0374866   .0194234    -1.93   0.054    -.0755558    .0005825
-------------+-----------------------------------------------------------------
g2007        |
 t_2003_2004 |   .0218685   .0177405     1.23   0.218    -.0129023    .0566393
```

```
 t_2004_2005 |  -.0041205   .0160708    -0.26   0.798    -.0356186    .0273776
 t_2005_2006 |  -.0344927    .019066    -1.81   0.070    -.0718613     .002876
 t_2006_2007 |  -.0310214   .0164804    -1.88   0.060    -.0633224    .0012796
-------------------------------------------------------------------------------
Control: Never Treated

See Callaway and Sant'Anna (2021) for details
```

```
** Getting residuals
qui:reghdfe lemp, abs(countyreal year) resid
** AS RC but with residuals
csdid _reghdfe_resid lpop, cluster( countyreal) gvar(first) time(year)
```

```
.............
Difference-in-difference with Multiple Time Periods

                                                      Number of obs = 2,260
Outcome model  : least squares
Treatment model: inverse probability
                          (Std. err. adjusted for 499 clusters in countyreal)
-------------------------------------------------------------------------------
             | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
-------------+-----------------------------------------------------------------
g2004        |
 t_2003_2004 |   -.022436   .0249588    -0.90   0.369    -.0713544    .0264824
 t_2003_2005 |  -.0898449   .0312509    -2.87   0.004    -.1510956   -.0285942
 t_2003_2006 |  -.1577356   .0380741    -4.14   0.000    -.2323594   -.0831118
 t_2003_2007 |  -.1192472   .0322187    -3.70   0.000    -.1823947   -.0560998
-------------+-----------------------------------------------------------------
g2006        |
 t_2003_2004 |  -.0139274   .0249804    -0.56   0.577    -.0628881    .0350333
 t_2004_2005 |  -.0055821   .0197258    -0.28   0.777     -.044244    .0330798
 t_2005_2006 |   .0011665   .0196647     0.06   0.953    -.0373756    .0397087
 t_2005_2007 |  -.0369996   .0194188    -1.91   0.057    -.0750597    .0010606
-------------+-----------------------------------------------------------------
g2007        |
 t_2003_2004 |   .0215591   .0178184     1.21   0.226    -.0133643    .0564825
 t_2004_2005 |  -.0041803   .0160877    -0.26   0.795    -.0357116    .0273509
 t_2005_2006 |  -.0343154   .0190396    -1.80   0.071    -.0716322    .0030015
 t_2006_2007 |   -.030636   .0164794    -1.86   0.063    -.0629351    .0016631
-------------------------------------------------------------------------------
```

```
Control: Never Treated

See Callaway and Sant'Anna (2021) for details
```

I think this produces much better results! Not sure how would this translate in other frameworks, tho.