SOCY7706: Longitudinal Data Analysis
Instructor: Natasha Sarkisian

Panel Data Analysis: Fixed Effects Models

Fixed effects models are similar to the first difference model we considered for two wave data—they also focus on the change component and get rid of any stable inter-individual differences. In fact, for two wave data, a fixed effects model is the same thing as a first difference model, while when there are more than two waves, these are considered to be two alternative ways to estimate a model focusing on change only, although fixed effects are used much more often and they perform much better when the data are unbalanced.

Overall, there are two kinds of information in panel data, regardless of the type: the cross-sectional information reflected in the differences between subjects, and the time-series or within-subject information reflected in the changes within subjects over time. Fixed effects models as well as first difference models focus on within-subject change only, but they control for differences across subjects. The key distinction is that in a first difference model, we focus on the change component by subtracting the previous wave observation from the current wave, while in the fixed effects model, we subtract the overall mean for that subject over time (that is, it's difference from the previous wave vs. difference from the overall mean over time).

We will continue using the same data for our example, using the already reshaped version where the empty rows have been dropped.

```
. xtset hhidpn wave
       panel variable:  hhidpn (unbalanced)
        time variable:  wave, 1 to 9, but with gaps
                delta:  1 unit
```

We will focus on predicting hours of help given to parents. Note that at this point, before proceeding to multivariate analyses, you should start with examining your variables for normality (use histogram, qnorm, and ladder, gladder, and qladder commands) and check the relationships between your dependent variable and each continuous predictor for linearity (lowess is a good tool for that). When necessary, apply transformations and proceed with transformed variables, but be aware of the balance between finding perfect transformations and having interpretable results.

While it is possible to use ordinary multiple regression techniques on panel data, they are usually not appropriate because of non-independence of observations (multiple observations that come from the same person have something in common), heteroskedasticity (both across time and across units), and autocorrelation. To avoid the problems of heteroscedasticity across units, we estimate a model that allows for each person to have its own intercept – a fixed effects model:

```
. xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg
female age   minority raedyrs, fe
note: female omitted because of collinearity
note: age omitted because of collinearity
note: minority omitted because of collinearity
note: raedyrs omitted because of collinearity
Fixed-effects (within) regression               Number of obs      =     30541
Group variable: hhidpn                          Number of groups   =      6243
```

```
R-sq:  within  = 0.0243                         Obs per group: min =         1
       between = 0.0067                                        avg =       4.9
       overall = 0.0134                                        max =         9
                                                F(6,24292)         =    100.87
corr(u_i, Xb)  = -0.1592                        Prob > F           =    0.0000
------------------------------------------------------------------------------
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0193467   .0014772   -13.10   0.000    -.0222421   -.0164512
 rpoorhealth |   .0792176   .0798801     0.99   0.321    -.0773524    .2357876
    rmarried |  -.6578103   .1342641    -4.90   0.000    -.9209763   -.3946443
   rtotalpar |    -.52481   .0384257   -13.66   0.000    -.6001268   -.4494933
     rsiblog |  -.5767981   .1841559    -3.13   0.002    -.9377549   -.2158412
    hchildlg |   .3859163   .1720502     2.24   0.025     .0486873    .7231454
      female |  (omitted)
         age |  (omitted)
    minority |  (omitted)
     raedyrs |  (omitted)
       _cons |   3.786918   .3755791    10.08   0.000      3.05076    4.523076
-------------+----------------------------------------------------------------
     sigma_u |  2.6483618
     sigma_e |  3.5375847
         rho |  .35916136   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(6242, 24292) =     2.37         Prob > F = 0.0000
```

Note that all time-invariant variables were automatically omitted.
Since we have multiple lines of data for each person, we should also adjust standard errors for clustering – that will take care of non-independence of observation. In this case, we also have multiple individuals in the same household, so we will adjust for the household (if there are multiple levels of clustering, we pick the higher one):

```
. xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg, fe
cluster(hhid)
Fixed-effects (within) regression               Number of obs      =     30546
Group variable: hhidpn                          Number of groups   =      6246
R-sq:  within  = 0.0243                         Obs per group: min =         1
       between = 0.0067                                        avg =       4.9
       overall = 0.0134                                        max =         9
                                                F(6,4637)          =     51.22
corr(u_i, Xb)  = -0.1593                        Prob > F           =    0.0000
                                (Std. Err. adjusted for 4638 clusters in hhid)
------------------------------------------------------------------------------
             |               Robust
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0193476   .0017652   -10.96   0.000    -.0228081   -.0158871
 rpoorhealth |   .0790409   .0867095     0.91   0.362     -.090951    .2490328
    rmarried |   -.657813   .1811515    -3.63   0.000    -1.012956   -.3026699
   rtotalpar |  -.5247729   .0573825    -9.15   0.000    -.6372698   -.4122759
     rsiblog |  -.5768106   .2257471    -2.56   0.011    -1.019382   -.1342388
    hchildlg |   .3856857   .1859452     2.07   0.038     .0211446    .7502268
       _cons |   3.786839   .4569993     8.29   0.000     2.890903    4.682775
-------------+----------------------------------------------------------------
     sigma_u |  2.6480962
     sigma_e |  3.5374433
         rho |  .35913361   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Although the person-level intercepts are not presented in the output, we would get the same model if we ran a regular OLS model with a dummy variable for each person -- it will not run in

Stata IC, however, because of too many dummy variables (over 6000). These individual-specific intercepts can also be viewed as part of the decomposed residuals:
$Y_{it} = \alpha + Xit\beta + u_i + e_{it}$ where u_i is the effect of person i and e_{it} is the residual effect for time point t within that person. In a fixed effects model, each of person residuals u_i is assigned a specific value – it's a fixed intercept for each individual. Because person-level intercepts are essentially separate independent variables in a fixed effects models, these intercepts are allowed to be correlated with the independent variables in the model –e.g., in our output we have

```
corr(u_i, Xb)  = -0.1593
```

What this means is that we do not use our independent variables to explain person-specific effects – they are just set aside and we focus on explaining change over time. One big advantage of doing this is that we eliminate all person-specific effects, including those that we could not explicitly model with the variables at hand. So that way, we control for the influence of both observable and unobservable individual-level factors, and we can focus explicitly on change over time. A disadvantage, however, is that the data on cross-sectional variation are available but not used in estimating independent variables' effects.

As a preliminary step to estimating a fixed effects model, it is usually helpful to estimate a fully unconditional model:

```
. xtreg rallparhelptw, fe
Fixed-effects (within) regression               Number of obs      =     32727
Group variable: hhidpn                          Number of groups   =      6588

R-sq:  within  = 0.0000                         Obs per group: min =         1
       between = 0.0009                                        avg =       5.0
       overall =      .                                        max =         9
                                                F(0,26139)         =      0.00
corr(u_i, Xb)  =      .                         Prob > F           =         .
------------------------------------------------------------------------------
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   1.652933   .0199706    82.77   0.000      1.61379    1.692076
-------------+----------------------------------------------------------------
     sigma_u |  2.6511079
     sigma_e |  3.6127964
         rho |  .35000687   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(6587, 26139) =     2.44         Prob > F = 0.0000
```

Its most important function is to provide information about outcome variability at each of the two levels. Sigma_e will provide information about level-1 (across time) variability, and sigma_u will provide information on level-2 (across individuals) variability. So running this model allows us to decompose the variance in the dependent variable into variance components -- into within-group and between-group variance (although they are expressed as standard deviations – to get variances, we'd have to square them). This model does not explain anything, but it allows us to evaluate whether there is variation in group means (here, person-specific means), and how much of it. That's why it is always a good idea to run this basic model when starting the analyses – it's the null model of our regression analysis. If we find that there is no significant variation across individuals, then there is no need for a fixed effects model because individuals are pretty much the same. That significance test is the F test below the model.

As we already learned earlier, the proportion of variance due to group-level variation in means is also known as the intraclass correlation coefficient (ICC) and can be calculated as
ρ = sigma_u^2 / (sigma_u^2 + sigma_e^2)

```
. di 2.6511079^2 / (2.6511079^2 + 3.6127964^2)
.35000689
```

which is the rho number in the xtreg table. So 35% of the total variance in hours of help to parents is due to person-specific effects.

<u>Diagnostics</u>

Predict command after xtreg, fe allows us to get predicted values and residuals. It allows the following options:

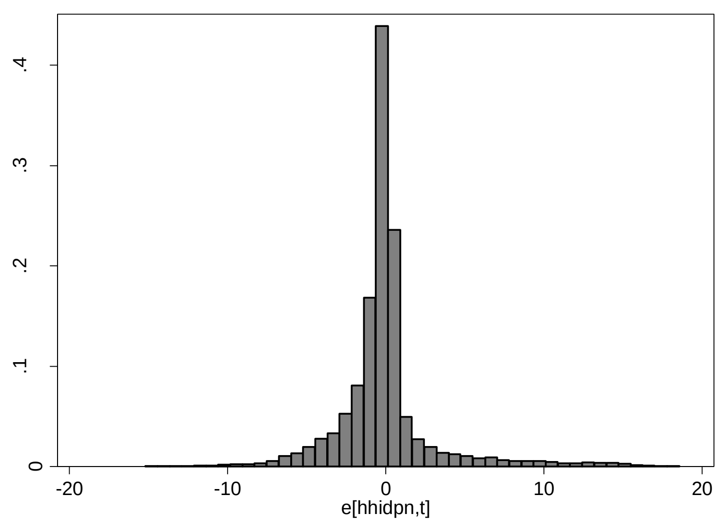
```
      xb            xb, fitted values; the default
      stdp          standard error of the fitted values
      ue            u_i + e_it, the combined residual
      xbu           xb + u_i, prediction including effect
      u             u_i, the fixed- or random-error component
      e             e_it, the overall error component
```

So to obtain two sets of residuals, level 1 (e) and level 2 (u), we run:

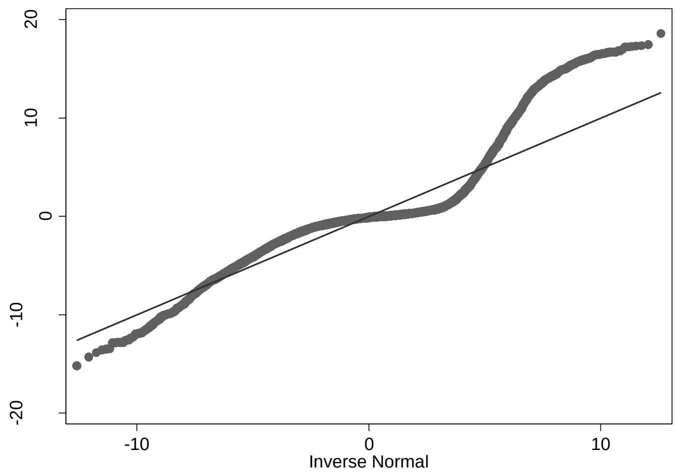
```
. qui xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog
hchildlg, fe cluster(hhid)

. predict level1, e
(24130 missing values generated)

. predict level2, u
(24130 missing values generated)
```

We can use these residuals to conduct regression diagnostics – e.g., examine normality:

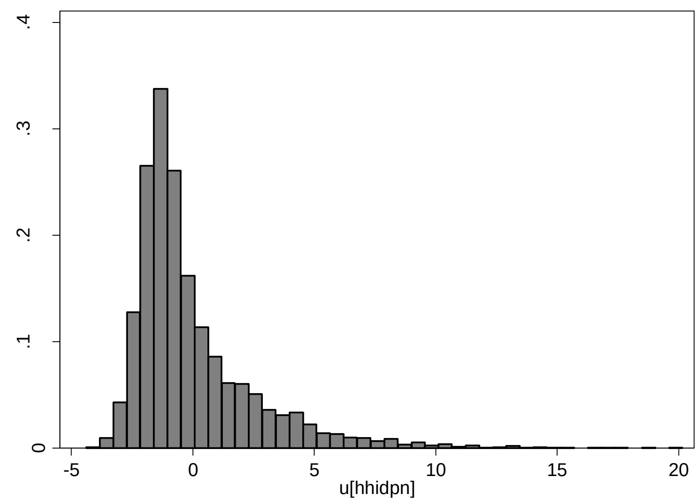
```
. histogram level1
(bin=44, start=-15.196963, width=.76734705)
```



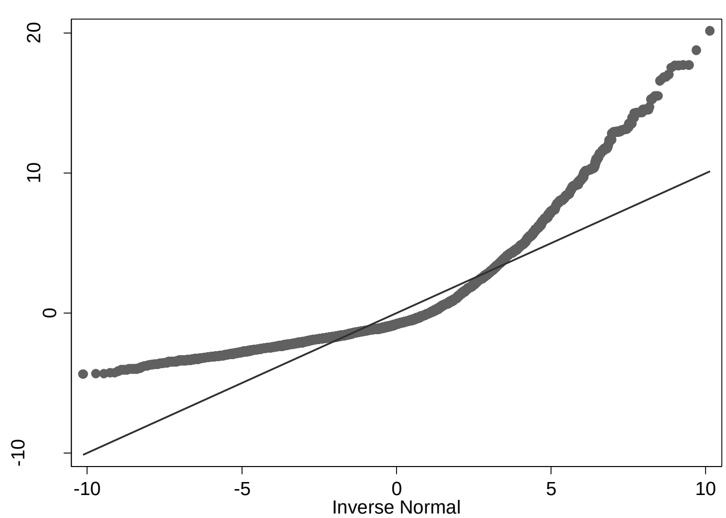
```
. qnorm level1
```




```
. histogram level2
(bin=44, start=-4.3855634, width=.55776229)
```

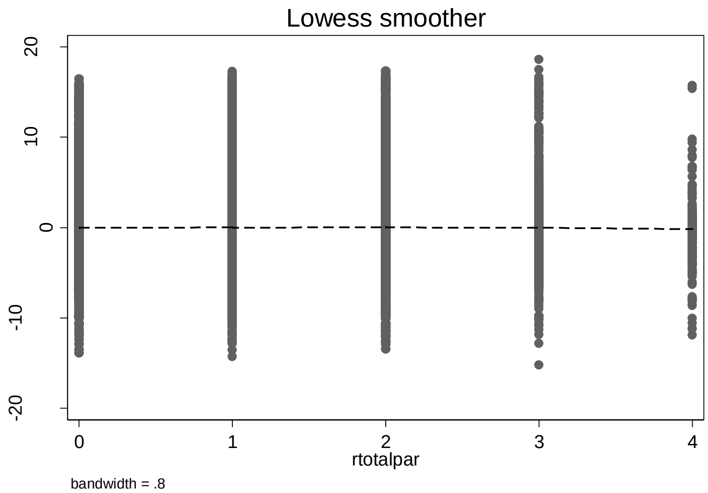



```
. qnorm level2
```

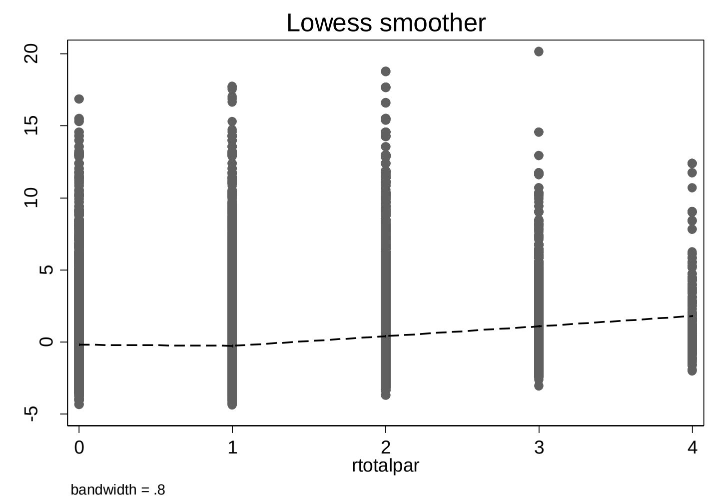


Next, let's look at linearity; we should do this for each continuous predictor:

```
. lowess level1 rtotalpar
```




```
. lowess level2 rtotalpar
```



If you find that you need to introduce a quadratic effect to better model nonlinear (quadratic) relationships, you would need to first group mean center the variable and then generate a quadratic term:

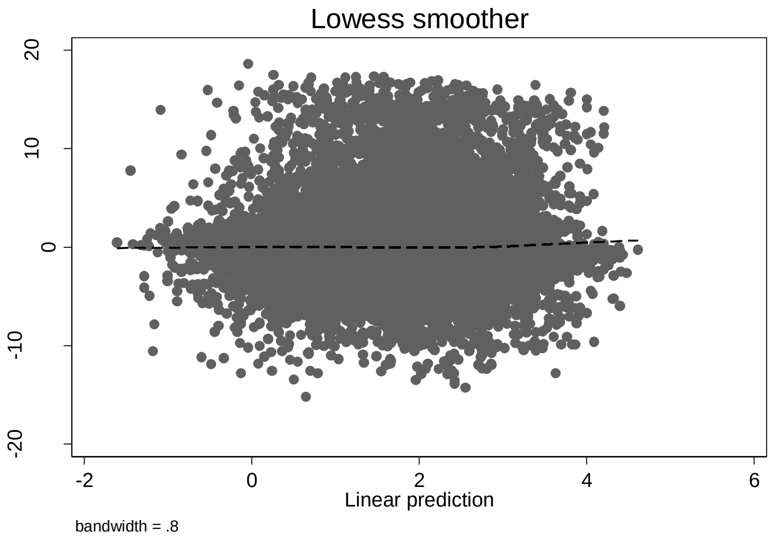
```
bysort hhidpn: egen rworkhours80_m=mean(rworkhours80) \ gen rworkhours80_diff=
rworkhours80- rworkhours80_m
gen rworkhours80_diff2= rworkhours80_diff^2
```

Then both rworkhours80_diff and rworkhours80_diff2 will be used in the model simultaneously to model the quadratic relationship. (See "Identifying Non-linearities In Fixed Effects Models" article by Craig T. McIntosh and Wolfram Schlenker.)

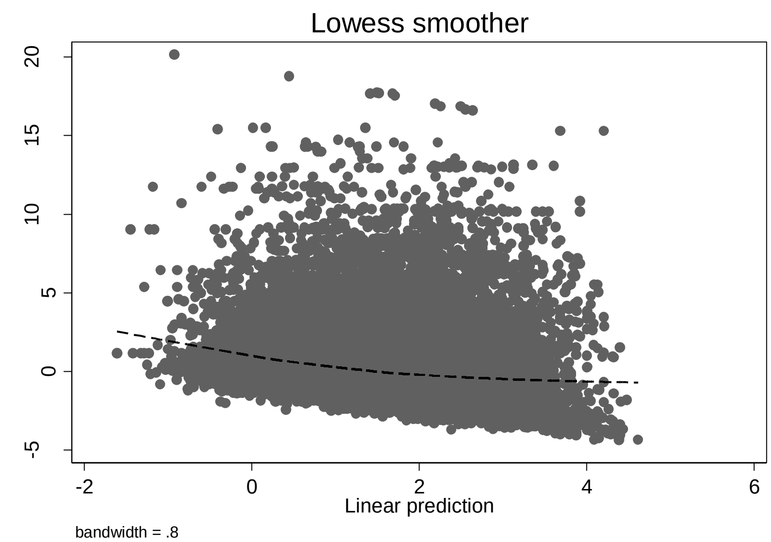
We can also obtain predicted values and examine distribution of residuals against these values; this allows us to assess whether there is heteroskedasticity.

```
. predict predval, xb
(11215 missing values generated)
```

```
. lowess level1 predval
```




```
. lowess level2 predval
```



In fixed effects model, level 2 residuals are not a random variable, they are all individual dummies in a sense, so we do not make much of an assumption about them – in fact, they can be correlated with independent variables, and we are not concerned about heteroskedasticity. This will be more of an issue for random effects models, however.

We can apply all the OLS diagnostic tools to the model with many dummies if we have enough system resources to estimate it – which would be easier if our dataset contained fewer units, of course. For more information on OLS diagnostics, see SOCY7704 class notes at http://www.sarkisian.net/socy7704.

Fixed Effects Model versus First Differences Model

Let's compare this fixed effects model to a first differences model.

First differences model:

$$\boxed{X_{it} - X_{i,t-1}} \longrightarrow \boxed{Y_{it} - Y_{i,t-1}}$$

Fixed effects model:

$$\boxed{X_{it} - \bar{X}_{i}} \longrightarrow \boxed{Y_{it} - \bar{Y}_{i}}$$

```
. reg D.(rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg),
cluster(hhid)
Linear regression                                      Number of obs =   23022
                                                       F(  6,  4221) =    1.63
                                                       Prob > F      =  0.1348
                                                       R-squared     =  0.0004
                                                       Root MSE      =  4.4523
                                (Std. Err. adjusted for 4222 clusters in hhid)
------------------------------------------------------------------------------
D.           |               Robust
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |
         D1. |  -.0045044   .0018505    -2.43   0.015    -.0081324   -.0008764
             |
 rpoorhealth |
         D1. |   .0978106   .0908495     1.08   0.282    -.0803022    .2759233
             |
    rmarried |
         D1. |  -.1340606   .1854965    -0.72   0.470    -.4977313    .2296101
             |
   rtotalpar |
         D1. |   .0074517   .0815175     0.09   0.927    -.1523654    .1672688
             |
     rsiblog |
         D1. |  -.0346379   .2086947    -0.17   0.868    -.4437894    .3745136
             |
    hchildlg |
         D1. |   .3036416   .2236533     1.36   0.175    -.1348365    .7421198
             |
       _cons |   .3333509   .0245212    13.59   0.000     .2852765    .3814252
------------------------------------------------------------------------------
```

Since the data are not balanced, however, we would prefer the fixed effects model.

Let's compare first difference and FE for two waves – that is when we expect them to be identical.

```
. preserve
. keep if wave<3
```

```
(41753 observations deleted)
. xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchil
> dlg , fe
Fixed-effects (within) regression               Number of obs      =     11327
Group variable: hhidpn                          Number of groups   =      6098

R-sq:  within  = 0.0021                         Obs per group: min =         1
       between = 0.0027                                        avg =       1.9
       overall = 0.0033                                        max =         2
                                                F(6,5223)          =      1.87
corr(u_i, Xb)  = -0.0706                        Prob > F           =    0.0829
------------------------------------------------------------------------------
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0091473   .0032143    -2.85   0.004    -.0154486    -.002846
 rpoorhealth |   .0261088    .159492     0.16   0.870    -.2865623    .3387798
    rmarried |  -.0947894   .3190668    -0.30   0.766    -.7202937    .5307149
   rtotalpar |  -.1342923   .1066734    -1.26   0.208    -.3434168    .0748321
     rsiblog |  -.3294339   .6821365    -0.48   0.629    -1.666707    1.007839
    hchildlg |   .2345638   .3881181     0.60   0.546      -.52631    .9954377
       _cons |   1.745728   1.254244     1.39   0.164     -.713115    4.204571
-------------+----------------------------------------------------------------
     sigma_u |  2.5509352
     sigma_e |  2.8203131
         rho |  .44997401   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(6097, 5223) =     1.44          Prob > F = 0.0000

. reg D.(rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchi
> ldlg), nocons
      Source |       SS       df       MS              Number of obs =    5229
-------------+------------------------------           F(  6,  5223) =    1.87
       Model |  178.031928     6  29.6719879           Prob > F      =  0.0829
    Residual |  83089.2185  5223  15.9083321           R-squared     =  0.0021
-------------+------------------------------           Adj R-squared =  0.0010
       Total |  83267.2505  5229  15.9241252           Root MSE      =  3.9885
------------------------------------------------------------------------------
D.           |
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |
         D1. |  -.0091473   .0032143    -2.85   0.004    -.0154486    -.002846
             |
 rpoorhealth |
         D1. |   .0261088    .159492     0.16   0.870    -.2865623    .3387798
             |
    rmarried |
         D1. |  -.0947894   .3190668    -0.30   0.766    -.7202937    .5307149
             |
   rtotalpar |
         D1. |  -.1342923   .1066734    -1.26   0.208    -.3434168    .0748321
             |
     rsiblog |
         D1. |  -.3294339   .6821365    -0.48   0.629    -1.666707    1.007839
             |
    hchildlg |
         D1. |   .2345639   .3881181     0.60   0.546      -.52631    .9954377
------------------------------------------------------------------------------
. restore
```

Replicating a fixed effects model by subtracting "group means"

Since fixed effects is based on "mean-differencing" the data (that's why it is also called the "within" estimator), we can replicate the results by subtracting person-specific means:

```
. xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg ,
fe cluster(hhid)
Fixed-effects (within) regression               Number of obs      =     30546
Group variable: hhidpn                          Number of groups   =      6246
R-sq:  within  = 0.0243                         Obs per group: min =         1
       between = 0.0067                                        avg =       4.9
       overall = 0.0134                                        max =         9
                                                F(6,4637)          =     51.22
corr(u_i, Xb)  = -0.1593                        Prob > F           =    0.0000
                                (Std. Err. adjusted for 4638 clusters in hhid)
------------------------------------------------------------------------------
             |               Robust
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0193476   .0017652   -10.96   0.000    -.0228081   -.0158871
 rpoorhealth |   .0790409   .0867095     0.91   0.362     -.090951    .2490328
    rmarried |   -.657813   .1811515    -3.63   0.000    -1.012956   -.3026699
   rtotalpar |  -.5247729   .0573825    -9.15   0.000    -.6372698   -.4122759
     rsiblog |  -.5768106   .2257471    -2.56   0.011    -1.019382   -.1342388
    hchildlg |   .3856857   .1859452     2.07   0.038     .0211446    .7502268
       _cons |   3.786839   .4569993     8.29   0.000     2.890903    4.682775
-------------+----------------------------------------------------------------
     sigma_u |  2.6480962
     sigma_e |  3.5374433
         rho |  .35913361   (fraction of variance due to u_i)
------------------------------------------------------------------------------

. for var rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg:
bysort hhidpn: egen Xm=mean(X) if e(sample) \ gen Xdiff=X-Xm

->  bysort hhidpn: egen rallparhelptwm=mean(rallparhelptw) if e(sample)
(24130 missing values generated)

->  gen rallparhelptwdiff=rallparhelptw-rallparhelptwm
(24130 missing values generated)

->  bysort hhidpn: egen rworkhours80m=mean(rworkhours80) if e(sample)
(24130 missing values generated)

->  gen rworkhours80diff=rworkhours80-rworkhours80m
(24130 missing values generated)

->  bysort hhidpn: egen rpoorhealthm=mean(rpoorhealth) if e(sample)
(24130 missing values generated)

->  gen rpoorhealthdiff=rpoorhealth-rpoorhealthm
(24130 missing values generated)

->  bysort hhidpn: egen rmarriedm=mean(rmarried) if e(sample)
(24130 missing values generated)

->  gen rmarrieddiff=rmarried-rmarriedm
(24130 missing values generated)

->  bysort hhidpn: egen rtotalparm=mean(rtotalpar) if e(sample)
(24130 missing values generated)
```

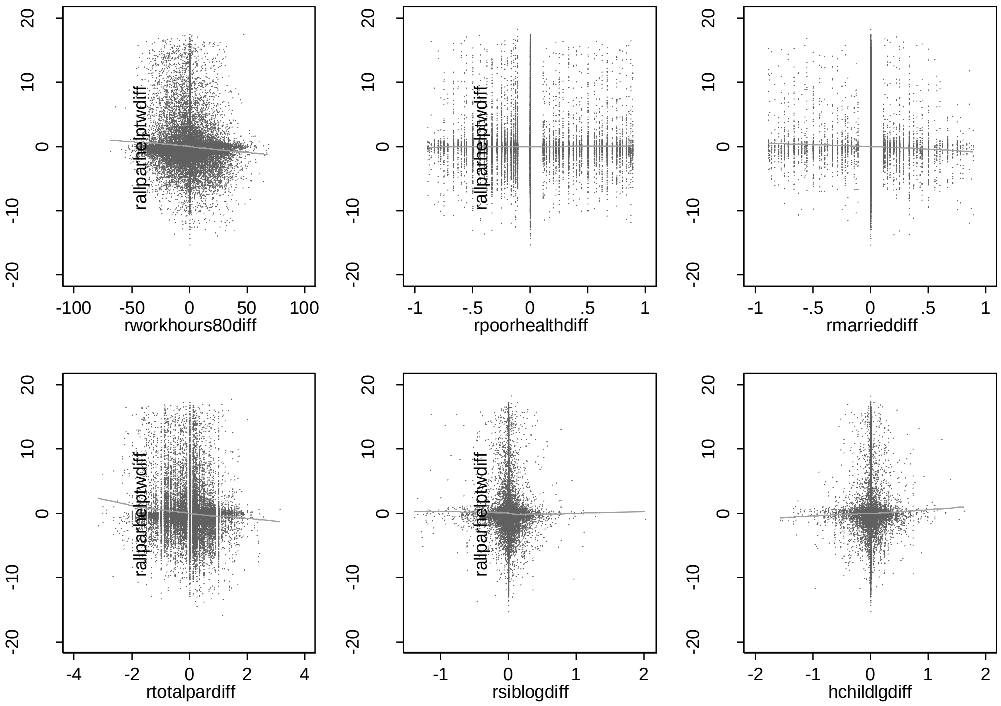
```
-> gen rtotalpardiff=rtotalpar-rtotalparm
(24130 missing values generated)

-> bysort hhidpn: egen rsiblogm=mean(rsiblog) if e(sample)
(24130 missing values generated)

-> gen rsiblogdiff=rsiblog-rsiblogm
(24130 missing values generated)

-> bysort hhidpn: egen hchildlgm=mean(hchildlg) if e(sample)
(24130 missing values generated)

-> gen hchildlgdiff=hchildlg-hchildlgm
(24130 missing values generated)

. reg rallparhelptwdiff rworkhours80diff rpoorhealthdiff rmarrieddiff rtotalpardiff
rsiblogdiff hchildlgdiff , cluster(hhid)

Linear regression                                      Number of obs =   30546
                                                       F(  6,  4637) =   51.22
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.0243
                                                       Root MSE      =  3.1551

                                (Std. Err. adjusted for 4638 clusters in hhid)
------------------------------------------------------------------------------
             |               Robust
rallparhel~f |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours~f |  -.0193476   .0017652   -10.96   0.000    -.0228081   -.0158871
rpoorhealt~f |   .0790409   .0867095     0.91   0.362     -.090951    .2490328
rmarrieddiff |   -.657813   .1811515    -3.63   0.000    -1.012956     -.30267
rtotalpard~f |  -.5247729   .0573825    -9.15   0.000    -.6372698   -.4122759
 rsiblogdiff |  -.5768106   .2257471    -2.56   0.011    -1.019382   -.1342388
hchildlgdiff |   .3856857   .1859452     2.07   0.038     .0211446    .7502268
       _cons |  -3.10e-09   1.28e-09    -2.42   0.016    -5.62e-09   -5.86e-10
------------------------------------------------------------------------------
```

One advantage of this specification is that we are using standard OLS with these variables – therefore, any diagnostics available with OLS could be used here as well (again, see SC704 notes for more detail), e.g. multicollinearity:

```
. vif

    Variable |       VIF       1/VIF
-------------+----------------------
rtotalpard~f |      1.13    0.886661
rworkhours~f |      1.12    0.894452
rmarrieddiff |      1.03    0.968216
rpoorhealt~f |      1.02    0.980188
hchildlgdiff |      1.01    0.986749
 rsiblogdiff |      1.00    0.997080
-------------+----------------------
    Mean VIF |      1.05
```

Or linearity:

```
. mrunning rallparhelptwdiff rworkhours80diff rpoorhealthdiff rmarrieddiff
rtotalpardiff rsiblogdiff hchildlgdiff
```



Note, however, that any transformations would have to be applied prior to mean-differencing the variables. Thus, diagnostics that rely on transforming variables (e.g., boxtid command) or testing interactions (fitint) won't always produce accurate results.

Just like we manually created mean-differenced variables, we can ask Stata to create a mean-differenced dataset for us using xtdata command. Make sure to save your dataset before doing that, though, because the data stored in memory will be lost once you transform the dataset:

```
. xtdata rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg,
fe clear

. reg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg

      Source |       SS       df       MS              Number of obs =   30546
-------------+------------------------------           F(  6, 30539) =  126.81
       Model |  7573.82015     6  1262.30336           Prob > F      =  0.0000
    Residual |   304003.09 30539   9.9545856           R-squared     =  0.0243
-------------+------------------------------           Adj R-squared =  0.0241
       Total |   311576.91 30545  10.2005863           Root MSE      =  3.1551

------------------------------------------------------------------------------
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0193476   .0013175   -14.69   0.000    -.0219299   -.0167653
 rpoorhealth |   .0790409   .0712347     1.11   0.267     -.060582    .2186638
    rmarried |   -.657813    .119747    -5.49   0.000    -.8925222   -.4231038
   rtotalpar |  -.5247729   .0342702   -15.31   0.000    -.5919439   -.4576019
     rsiblog |  -.5768106   .1642443    -3.51   0.000    -.8987362   -.2548849
    hchildlg |   .3856857   .1534408     2.51   0.012     .0849353    .6864361
       _cons |   3.786839   .3349614    11.31   0.000     3.130301    4.443377
------------------------------------------------------------------------------
```

Assymmetric Fixed Effects

So far, we have been assuming that an increase and a decrease in a given predictor's value would produce a symmetric response – that's the same assumption that we considered for a two time period example. If we are not willing to make that assumption or would like to test it, we can separate a given predictor into two variables – one including only increases (with the other values set to 0) and the other one only decreases. We do, however, need some special approaches for estimating such a model – e.g., see "Asymmetric Fixed-effects Models for Panel Data" article by Paul D. Allison (Socius 2019).

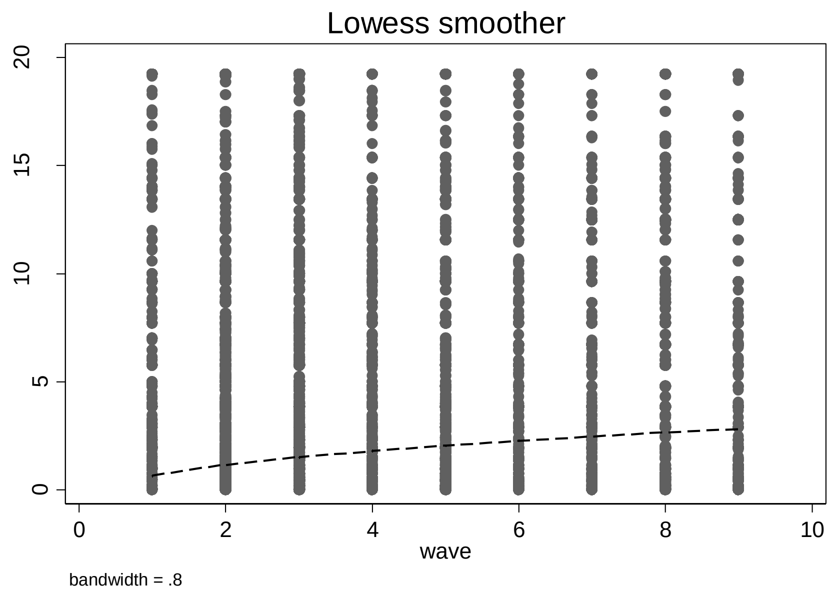
Two-Way Fixed Effects

In addition to the one-way fixed effects model that we just estimated, we could also consider estimating a two-way fixed-effects model. It is a good idea in most cases to include time into the model when estimating a fixed-effects model. Unfortunately, Stata does not automatically estimated two-way FE models – we have to introduce wave dummies:

```
. xi: xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg
i.wave, fe cluster(hhid)
i.wave            _Iwave_1-9          (naturally coded; _Iwave_1 omitted)

Fixed-effects (within) regression               Number of obs      =     30546
Group variable: hhidpn                          Number of groups   =      6246

R-sq:  within  = 0.0435                         Obs per group: min =         1
       between = 0.0288                                        avg =       4.9
       overall = 0.0365                                        max =         9

                                                F(14,4637)         =     44.56
corr(u_i, Xb)  = -0.0199                        Prob > F           =    0.0000

                               (Std. Err. adjusted for 4638 clusters in hhid)
------------------------------------------------------------------------------
             |               Robust
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0074452   .0017754    -4.19   0.000    -.0109258   -.0039646
 rpoorhealth |  -.0455057   .0853449    -0.53   0.594    -.2128224     .121811
    rmarried |  -.5600425   .1773223    -3.16   0.002    -.9076785   -.2124064
   rtotalpar |    .021109   .0663923     0.32   0.751    -.1090516    .1512695
     rsiblog |  -.3169884   .2216861    -1.43   0.153    -.7515985    .1176218
    hchildlg |    .122697   .1934373     0.63   0.526    -.2565321    .5019261
    _Iwave_2 |   .5832997   .0620318     9.40   0.000      .461688    .7049115
    _Iwave_3 |   .9157041   .0740675    12.36   0.000     .7704966    1.060912
    _Iwave_4 |   1.266869   .0896387    14.13   0.000     1.091134    1.442603
    _Iwave_5 |   1.202117   .0981853    12.24   0.000     1.009627    1.394607
    _Iwave_6 |   1.704193   .1215176    14.02   0.000      1.46596    1.942425
    _Iwave_7 |   2.032625   .1416203    14.35   0.000     1.754982    2.310268
    _Iwave_8 |   2.171453   .1639864    13.24   0.000     1.849961    2.492944
    _Iwave_9 |   2.145707   .1811296    11.85   0.000     1.790607    2.500807
       _cons |   1.613513   .4600734     3.51   0.000       .71155    2.515475
-------------+----------------------------------------------------------------
     sigma_u |     2.5653
     sigma_e |  3.5030799
         rho |  .34906895   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

It looks like on average, there is an increase in number of hours of help over time, so we could consider modeling it as a linear trend. Let's examine a lowess plot:

```
.lowess rallparhelptw wave
```




```
. xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog hchildlg
wave, fe cluster(hhid)

Fixed-effects (within) regression               Number of obs      =     30546
Group variable: hhidpn                          Number of groups   =      6246

R-sq:  within  = 0.0411                         Obs per group: min =         1
       between = 0.0257                                        avg =       4.9
       overall = 0.0338                                        max =         9

                                                F(7,4637)          =     66.59
corr(u_i, Xb)  = -0.0235                        Prob > F           =    0.0000

                                (Std. Err. adjusted for 4638 clusters in hhid)
------------------------------------------------------------------------------
             |               Robust
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0072974   .0017707    -4.12   0.000    -.0107689   -.0038259
 rpoorhealth |  -.0389903   .0853144    -0.46   0.648    -.2062471    .1282664
    rmarried |  -.5673026   .1764431    -3.22   0.001     -.913215   -.2213903
   rtotalpar |   .0036739   .0661762     0.06   0.956    -.1260631    .1334108
     rsiblog |  -.2868556   .2211379    -1.30   0.195    -.7203912      .14668
    hchildlg |   .1460003   .1932094     0.76   0.450     -.232782    .5247826
        wave |   .2868391   .0185104    15.50   0.000     .2505499    .3231283
       _cons |   1.481042   .4620466     3.21   0.001     .5752114    2.386874
-------------+----------------------------------------------------------------
     sigma_u |  2.5692546
     sigma_e |  3.5069475
         rho |  .34926754   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

To test whether it is appropriate to assume a linear trend, we test this model against the previous one in terms of its fit. We will use Bayesian Information Criterion (BIC) to compare models:

```
. estat ic

------------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+----------------------------------------------------------------
           . |  30546    -78813.1   -78172.16      7     156358.3    156416.6
------------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note

. qui xi: xtreg rallparhelptw rworkhours80 rpoorhealth rmarried rtotalpar rsiblog
hchildlg i.wave, fe cluster(hhid)

. estat ic

------------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+----------------------------------------------------------------
           . |  30546    -78813.1   -78134.05     14     156296.1    156412.7
------------------------------------------------------------------------------
               Note:  N=Obs used in calculating BIC; see [R] BIC note
```

BIC difference:

```
. di 156416.6-156412.7
3.9
```

The model with smaller BIC has better fit, and the strength of evidence in its favor is evaluated as follows:

BIC Difference	Evidence
0-2	Weak
2-6	Positive
6-10	Strong
>10	Very strong

So in this case, the model with dummies has a somewhat better fit as it has smaller BIC and the difference is 3.9, but the evidence in its favor is not strong. So linear trend could still be a reasonable choice.

<u>Autocorrelation</u>

So far, we have dealt with two problems of panel data -- heteroskedasticity across units and non-independence of observations. One problem that might be remaining is autocorrelation, that is, correlation between residuals at a given wave and the ones for the previous one. To test for autocorrelation:

```
. net search xtserial
```

Click on st0039 from http://www.stata-journal.com/software/sj3-2 and install

```
. xtserial rallparhelptw  rworkhours80  rpoorhealth rmarried rtotalpar rsiblog
hchildlg

Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
    F(  1,    4558) =     34.757
           Prob > F =      0.0000
```

This test focuses on residuals from a first difference model (ΔY regressed on ΔX). Here, the hypothesis of no first order autocorrelation is rejected; therefore, we would want a model explicitly accounting for autoregressive error term.

We can use xtregar models that assume that:
y_it = a + x_it * B + u_i + e_it
where e_it = rho * e_i,t-1 + z_it with |rho| < 1

```
. xtregar rallparhelptw  rworkhours80  rpoorhealth rmarried rtotalpar rsiblog
hchildlg, fe lbi

FE (within) regression with AR(1) disturbances  Number of obs      =     24300
Group variable: hhidpn                          Number of groups   =      5800

R-sq:  within  = 0.0079                         Obs per group: min =         1
       between = 0.0000                                        avg =       4.2
       overall = 0.0021                                        max =         8

                                                F(6,18494)         =     24.42
corr(u_i, Xb)  = -0.1672                        Prob > F           =    0.0000

------------------------------------------------------------------------------
rallparhel~w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
rworkhours80 |  -.0124486   .0018924    -6.58   0.000    -.0161579   -.0087393
 rpoorhealth |   .1305854   .0915099     1.43   0.154    -.0487824    .3099532
    rmarried |   -.444861   .1778864    -2.50   0.012    -.7935348   -.0961872
   rtotalpar |  -.3642803   .0512597    -7.11   0.000    -.4647541   -.2638066
     rsiblog |   .0112739   .2059205     0.05   0.956    -.3923493    .4148972
    hchildlg |   .6969819   .2383535     2.92   0.003      .229787    1.164177
       _cons |   2.193055   .3245314     6.76   0.000     1.556943    2.829166
-------------+----------------------------------------------------------------
      rho_ar |  .24444167
     sigma_u |  3.0974642
     sigma_e |  3.6788507
     rho_fov |  .41483009   (fraction of variance because of u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(5799,18494) =     1.70         Prob > F = 0.0000
modified Bhargava et al. Durbin-Watson = 1.5724782
Baltagi-Wu LBI = 2.0213388
```

Xtregar also offers additional tests for autocorrelation, based on Durbin-Watson statistic—we used lbi option to obtain those. A value of the modified Durbin-Watson statistic or Baltagi-Wu LBI-statistic of 2 indicates no autocorrelation (the values can be between 0 and 4). As a rough rule of thumb, values below 1 mean you should definitely correct for serial correlation. Small values indicate successive error terms are positively correlated. You can also find critical values for some specific numbers of cases (N), time points (T), and number of estimated parameters (k) here: http://www.stata.com/statalist/archive/2010-08/msg00542.html. In contrast, with the values of such a statistic >2, successive error terms are, on average, different in value from one another, i.e., negatively correlated. This is much less common, however. In regressions, this can lead to an underestimation of the level of statistical significance.