Practical Implementation of TSLSFECLUS Algorithm

Siyi Luo, Wenjia Zhu, Randall P. Ellis

Our algorithm is programmed in SAS for ease of implementation. The main macro that performs the iterative procedure is TSLSFECLUS. This macro can accommodate a wide range of model features such as endogeneity, cluster standard error correction, and multiple high-dimensional fixed effects. In addition, it allows multiple specifications that differ only in their dependent variables to be estimated in a single call. Finally, the macro automatically outputs the number of iterations needed for model convergence together with the model estimates. The macro mainly contains the following four steps.

1): Given model specification, identify multiple high-dimensional fixed effects to absorb. Set the values of maximum number of iteration $Maxiter$ and the tolerance level tol.

2): Absorb fixed effects from all dependent and explanatory (including instrumental) variables, one by one until all the fixed effects are absorbed once. Save standardized data S_1, and the estimated parameters of interest from model S_1, labeled as $\{\hat{\beta}^{(1)}{}_k\}_{k=1,2,\dots,K}$.

3): Repeat step 2) and record S_2, and obtain $\{\hat{\beta}^{(2)}{}_k\}_{k=1,2,\dots,K}$ from estimating the model using S_2.

4): Calculate $|\Delta_2| = \max\{|\frac{\hat{\beta}^{(2)}{}_k - \hat{\beta}^{(1)}{}_k}{\hat{\beta}^{(1)}{}_k}|\}_{k=1,2,\dots,K}$, the maximum absolute value of percentage difference between adjacent iterations among *K* estimated parameters of interest. If $|\Delta_2| < tol$, then stop here and report coefficient estimates $\{\hat{\beta}^{(2)}{}_k\}_{k=1,2,\dots,K}$; otherwise, repeat step 2) until $|\Delta_i| = \max\{|\frac{\hat{\beta}^{(i)}{}_k - \hat{\beta}^{(i-1)}{}_k}{\hat{\beta}^{(i-1)}{}_k}|\}_{k=1,2,\dots,K} < tol$ or the maximum number of iterations have been reached. The reported number of iteration = $\min\{i = \{i||\Delta_i| < tol\}, Maxiter\}$.

For more details of the algorithm, please refer to our working paper "An Iterative Approach to Estimation with Multiple High-Dimensional Fixed Effects: Controlling Simultaneously for Patients, Providers and Counties" (Luo, Zhu and Ellis, Boston University Working Paper, 2017).

Below is a sample call of our two macros in SAS. The second macro can be called directly if only one iteration is desired, such as if there is only one high-dimensional fixed effect:

Libname junk “directory for storing temporary data sets”;

```
%auto_iter(
indsn = in_data,                /* input data */
tol = 0.0001,                   /* tolerance level for convergence */
maxiter = 10,                   /* maximum number of iteration */
betasefinal = out_data,         /* output data for storing estimates from all iterations */
fevarcount = 3,                 /* number of absorbed fixed effects */
auto_tempdir = junk,            /* directory for storing temporary data sets */
auto_depvar = y,                /* dependent variable */
auto_endog = x,                 /* endogenous variable */
auto_inst = z,                  /* instrumental variable */
auto_exog = ,                   /* exogenous variable */
auto_fe = i c t,                /* variables defining absorbed fixed effects */
auto_FE_iter = 1,               /* incremental on iteration number */
auto_cluster = c,               /* variable defining cluster level */
auto_othervar = ,               /* other variables to be carried along to final dataset
                                   for final analysis */
auto_regtype = TSLS,            /* TSLS or OLS */
auto_showmeans = no,            /* yes or no to showing sample summary statistics */
auto_showrf = no,               /* yes or no to showing reduced form results of TSLS
                                   model */
auto_showols = no,              /* yes or no to OLS without cluster correction */
auto_dosurveyreg = no,          /* yes or no to doing PROC SURVEYREG */
auto_wide = no                  /* yes or no to wide format table */
);

*which calls the following core macro iteratively;
%TSLSCLUS_iterFE(
runtitle = "TSLS: one fixed effect",        /* running title */
indata = in_data,               /* input data for each iteration: &indsn. for the first
                                   iteration, standardized data for subsequent iterations */
depvar = y,                     /* dependent variable */
endog = x,                      /* endogenous variable */
inst = z,                       /* instrumental variable */
exog = ,                        /* exogenous variable */
fe = i c t,                     /* variables defining absorbed fixed effects */
FE_iter = 1,                    /* incremental on iteration number */
cluster = c,                    /* variable defining cluster level */
othervar = ,                    /* other variables to be carried along to final dataset
                                   for final analysis */
tempdir = junk,                 /* directory for storing temporary data sets */
```

```
regtype = TSLS,                 /* TSLS or OLS */
showmeans = no,                 /* yes or no to showing sample summary statistics */
showrf = no,                    /* yes or no to showing reduced form results of TSLS
                                   model */
showols = no,                   /* yes or no to OLS without cluster correction */
dosurveyreg = no,               /* yes or no to doing PROC SURVEYREG */
wide = no,                      /* yes or no to wide format table */
estresult = out_data            /* data set for outputting estimates */
);
```