Panel data estimators

Christopher F Baum

ECON 8823: Applied Econometrics

Boston College, Spring 2021

The panel data estimation problem

Consider a balanced panel of $N \times T$ observations. If we subscripted each coefficient as β_{kit}, the model contains $K \times N \times T$ regression coefficients, it cannot be estimated from the data. We could ignore the nature of the panel data and apply pooled ordinary least squares, which would assume that $\beta_{kit} = \beta_k \; \forall \; k, i, t$, but that model might be viewed as overly restrictive and is likely to have a very complicated error process (e.g., heteroskedasticity across panel units, serial correlation within panel units, and so forth). Thus the pooled OLS solution is not often considered to be practical and can usually be rejected by the data.

One set of panel data estimators allow for heterogeneity across panel units (and possibly across time), but confine that heterogeneity to the intercept terms of the relationship. These techniques, the *fixed effects* and *random effects* models, we consider below. They impose restrictions on the model above of $\beta_{kit} = \beta_k \; \forall i, t, \; k > 1$, assuming that β_1 refers to the constant term in the relationship.

The fixed effects estimator

The general structure above may be restricted to allow for heterogeneity across units without the full generality (and infeasibility) that this equation implies. In particular, we might restrict the slope coefficients to be constant over both units and time, and allow for an intercept coefficient that varies by unit or by time. For a given observation, an intercept varying over units results in the structure:

$$y_{it} = \sum_{k=2}^{K} X_{kit}\beta_k + u_i + \epsilon_{it} \tag{1}$$

There are two interpretations of u_i in this context: as a parameter to be estimated in the model (a so-called *fixed effect*) or alternatively, as a component of the disturbance process, giving rise to a composite error term $[u_i + \epsilon_{it}]$: a so-called *random effect*. Under either interpretation, u_i is taken as a random variable.

If we treat it as a fixed effect, we assume that the u_i may be correlated with some of the regressors in the model. The fixed-effects estimator removes the fixed-effects parameters from the estimator to cope with this incidental parameter problem, which implies that all inference is conditional on the fixed effects in the sample.

Use of the random effects model implies additional orthogonality conditions—that the u_i are not correlated with the regressors—and yields inference about the underlying population that is not conditional on the fixed effects in our sample.

We could treat a time-varying intercept term similarly: as either a fixed effect (giving rise to an additional coefficient) or as a component of a composite error term. We concentrate here on so-called *one-way fixed (random) effects* models in which only the individual effect is considered in the "large N, small T" context most commonly found in economic and financial research.

Stata's set of `xt` commands include those which extend these panel data models in a variety of ways. For more information, see `help xt`.

One-way fixed effects: the within estimator

Rewrite the equation to express the individual effect u_i as

$$y_{it} = X_{it}^* \beta^* + Z_i \alpha + \epsilon_{it} \tag{2}$$

In this context, the X^* matrix does not contain a units vector. The heterogeneity or individual effect is captured by Z, which contains a constant term and possibly a number of other individual-specific factors. Likewise, β^* contains $\beta_2 \ldots \beta_K$ from the equation above, constrained to be equal over i and t. If Z contains only a units vector, then pooled OLS is a consistent and efficient estimator of $[\beta^* \; \alpha]$.

However, it will often be the case that there are additional factors specific to the individual unit that must be taken into account, and omitting those variables from Z will cause the equation to be misspecified.

The *fixed effects* model deals with this problem by relaxing the assumption that the regression function is constant over time and space in a very modest way. A one-way fixed effects model permits each cross-sectional unit to have its own constant term while the slope estimates (β^*) are constrained across units, as is the σ^2_ϵ.

This estimator is often termed the *LSDV* (least-squares dummy variable) model, since it is equivalent to including $(N-1)$ dummy variables in the OLS regression of y on X (including a units vector). The *LSDV* model may be written in matrix form as:

$$y = X\beta + D\alpha + \epsilon \tag{3}$$

where D is a $NT \times N$ matrix of dummy variables d_i (assuming a balanced panel of $N \times T$ observations).

The model has $(K-1)+N$ parameters (recalling that the β^* coefficients are all slopes) and when this number is too large to permit estimation, we rewrite the least squares solution as

$$b = (X' M_D X)^{-1}(X' M_D y) \tag{4}$$

where

$$M_D = I - D(D'D)^{-1}D' \tag{5}$$

is an idempotent matrix which is block–diagonal in $M_0 = I_T - T^{-1}\iota\iota'$ (ι a T–element units vector).

Premultiplying any data vector by M_0 performs the demeaning transformation: if we have a T–vector Z_i, $M_0 Z_i = Z_i - \bar{Z}_i \iota$. The regression above estimates the slopes by the projection of demeaned y on demeaned X without a constant term.

The estimates a_i may be recovered from $a_i = \bar{y}_i - b' \bar{X}_i$, since for each unit, the regression surface passes through that *unit's* multivariate point of means. This is a generalization of the OLS result that in a model with a constant term the regression surface passes through the *entire sample's* multivariate point of means.

The large-sample VCE of b is $s^2[X' M_D X]^{-1}$, with s^2 based on the least squares residuals, but taking the proper degrees of freedom into account: $NT - N - (K - 1)$.

This model will have explanatory power *if and only if* the variation of the individual's y above or below the individual's mean is significantly correlated with the variation of the individual's X values above or below the individual's vector of mean X values. For that reason, it is termed the *within estimator*, since it depends on the variation *within* the unit.

It does not matter if some individuals have, e.g., very high y values and very high X values, since it is only the within variation that will show up as explanatory power. This is the panel analogue to the notion that OLS on a cross-section does not seek to explain the mean of y, but only the variation around that mean.

This has the clear implication that any characteristic which does not vary over time for each *unit* cannot be included in the model: for instance, an individual's gender, or a firm's three-digit SIC (industry) code, or the nature of a country as landlocked.

The unit-specific intercept term absorbs all heterogeneity in y and X that is a function of the identity of the unit, and any variable constant over time for each unit will be perfectly collinear with the unit's indicator variable.

The one-way individual fixed effects model may be estimated by the Stata command `xtreg` using the `fe` (fixed effects) option. The command has a syntax similar to `regress`:

`xtreg` *depvar indepvars*`, fe [`*options*`]`

As with standard regression, options include `robust` and `cluster()`. The command output displays estimates of σ_u^2 (labeled `sigma_u`), σ_ϵ^2 (labeled `sigma_e`), and what Stata terms `rho`: the fraction of variance due to u_i. Stata estimates a model in which the u_i are taken as deviations from a single constant term, displayed as `_cons`; therefore testing that all u_i are zero is equivalent in our notation to testing that all α_i are identical. The empirical correlation between u_i and the regressors in X^* is also displayed as `corr(u_i, Xb)`.

The fixed effects estimator does not require a balanced panel. As long as there are at least two observations per unit, it may be applied. However, since the individual fixed effect is in essence estimated from the observations of each unit, the precision of that effect (and the resulting slope estimates) will depend on N_i.

We wish to test whether the individual-specific heterogeneity of α_i is necessary: are there distinguishable intercept terms across units? `xtreg, fe` provides an F-test of the null hypothesis that the constant terms are equal across units. If this null is rejected, pooled OLS would represent a misspecified model.

The one-way fixed effects model also assumes that the errors are not contemporaneously correlated across units of the panel. This hypothesis can be tested (provided $T > N$) by the Lagrange multiplier test of Breusch and Pagan, available as the author's `xttest2` routine (`findit xttest2`).

In this example, using the `traffic` dataset, we have 1982–1988 state-level data for 48 U.S. states on traffic fatality rates (deaths per 100,000). We model the highway fatality rates as a function of several common factors: `beertax`, the tax on a case of beer, `spircons`, a measure of spirits consumption and two economic factors: the state unemployment rate (`unrate`) and state per capita personal income, $000 (`perincK`). We present descriptive statistics for these variables of the `traffic.dta` dataset.

Try it out:

```
. bcuse traffic, clear
. summarize fatal beertax spircons unrate perincK
```

Variable	Obs	Mean	Std. Dev.	Min	Max
fatal	336	2.040444	.5701938	.82121	4.21784
beertax	336	.513256	.4778442	.0433109	2.720764
spircons	336	1.75369	.6835745	.79	4.9
unrate	336	7.346726	2.533405	2.4	18
perincK	336	13.88018	2.253046	9.513762	22.19345

Try it out:

```
. xtreg fatal beertax spircons unrate perincK, fe
Fixed-effects (within) regression               Number of obs      =       336
Group variable (i): state                       Number of groups   =        48
R-sq:  within  = 0.3526                         Obs per group: min =         7
       between = 0.1146                                        avg =       7.0
       overall = 0.0863                                        max =         7
                                                F(4,284)           =     38.68
corr(u_i, Xb)  = -0.8804                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       fatal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.4840728   .1625106    -2.98   0.003    -.8039508   -.1641948
    spircons |   .8169652   .0792118    10.31   0.000     .6610484    .9728819
      unrate |  -.0290499   .0090274    -3.22   0.001    -.0468191   -.0112808
     perincK |   .1047103   .0205986     5.08   0.000      .064165    .1452555
       _cons |   -.383783   .4201781    -0.91   0.362    -1.210841    .4432754
-------------+----------------------------------------------------------------
     sigma_u |  1.1181913
     sigma_e |  .15678965
         rho |  .98071823   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47, 284) =    59.77             Prob > F = 0.0000
```

All explanatory factors are highly significant, with the unemployment rate having a negative effect on the fatality rate (perhaps since those who are unemployed are income-constrained and drive fewer miles), and income a positive effect (as expected because driving is a normal good).

Note the empirical correlation labeled `corr(u_i, Xb)` of -0.8804. This correlation indicates that the unobserved heterogeneity term, proxied by the estimated fixed effect, is strongly correlated with a linear combination of the included regressors. That is not a problem for the fixed effects model, but as we shall see it is an important magnitude.

We have considered one-way fixed effects models, where the effect is attached to the individual. We may also define a two-way fixed effect model, where effects are attached to each unit and time period. Stata lacks a command to estimate two-way fixed effects models. If the number of time periods is reasonably small, you may estimate a two-way FE model by creating a set of time indicator variables and including all but one in the regression.

In Stata 11 onward, that is very easy to do using factor variables by specifying `i.year` in the regressor list. The joint significance of those variables can be assessed with `testparm`.

The joint test that all of the coefficients on those indicator variables are zero will be a test of the significance of time fixed effects. Just as the individual fixed effects (LSDV) model requires regressors' variation over time within each *unit*, a time fixed effect (implemented with a time indicator variable) requires regressors' variation over units within each *time period*.

If we are estimating an equation from individual or firm microdata, this implies that we cannot include a "macro factor" such as the rate of GDP growth or price inflation in a model with time fixed effects, since those factors do not vary across individuals.

```
. xtreg fatal beertax spircons unrate perincK i.year, fe

Fixed-effects (within) regression               Number of obs      =       336
Group variable: state                           Number of groups   =        48

R-sq:  within  = 0.4528                         Obs per group: min =         7
       between = 0.1090                                        avg =       7.0
       overall = 0.0770                                        max =         7

                                                F(10,278)          =     23.00
corr(u_i, Xb)  = -0.8728                        Prob > F           =    0.0000

------------------------------------------------------------------------------
       fatal |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     beertax |  -.4347195   .1539564    -2.82   0.005    -.7377878   -.1316511
    spircons |    .805857   .1126425     7.15   0.000     .5841163    1.027598
      unrate |  -.0549084   .0103418    -5.31   0.000    -.0752666   -.0345502
     perincK |   .0882636   .0199988     4.41   0.000     .0488953    .1276319
             |
        year |
       1983  |  -.0533713    .030209    -1.77   0.078    -.1128387    .0060962
       1984  |  -.1649828    .037482    -4.40   0.000    -.2387674   -.0911983
       1985  |  -.1997376   .0415808    -4.80   0.000    -.2815908   -.1178845
       1986  |  -.0508034   .0515416    -0.99   0.325    -.1522647     .050658
       1987  |  -.1000728     .05906    -1.69   0.091    -.2163345    .0161889
       1988  |   -.134057   .0677696    -1.98   0.049    -.2674638   -.0006503
             |
       _cons |   .1290568   .4310663     0.30   0.765    -.7195118    .9776253
-------------+----------------------------------------------------------------
     sigma_u |  1.0987683
     sigma_e |  .14570531
         rho |  .98271904   (fraction of variance due to u_i)
------------------------------------------------------------------------------
F test that all u_i=0:     F(47, 278) =    64.52             Prob > F = 0.0000
```

```
. testparm i.year
( 1)  1983.year = 0
( 2)  1984.year = 0
( 3)  1985.year = 0
( 4)  1986.year = 0
( 5)  1987.year = 0
( 6)  1988.year = 0
       F(  6,   278) =    8.48
            Prob > F =    0.0000
```

The four quantitative factors included in the one-way fixed effects model retain their sign and significance in the two-way fixed effects model. The time effects are jointly significant, suggesting that they should be included in a properly specified model. Otherwise, the model is qualitatively similar to the earlier model, with a sizable amount of variation explained by the individual (state) fixed effect.

The between estimator

Another estimator that may be defined for a panel data set is the *between estimator*, in which the group means of y are regressed on the group means of X in a regression of N observations. This estimator *ignores* all of the individual-specific variation in y and X that is considered by the within estimator, replacing each observation for an individual with their mean behavior.

This estimator is not widely used, but has sometimes been applied in cross-country studies where the time series data for each individual are thought to be somewhat inaccurate, or when they are assumed to contain random deviations from long-run means. If you assume that the inaccuracy has mean zero over time, a solution to this measurement error problem can be found by averaging the data over time and retaining only one observation per unit.

This could be done explicitly with Stata's `collapse` command. However, you need not form that data set to employ the between estimator, as the command `xtreg` with the `be` (between) option will invoke it. Use of the between estimator requires that $N > K$. Any macro factor that is constant over *individuals* cannot be included in the between estimator, since its average will not differ by individual.

We can show that the pooled OLS estimator is a matrix weighted average of the within and between estimators, with the weights defined by the relative precision of the two estimators. We might ask, in the context of panel data: where are the interesting sources of variation? In individuals' variation around their means, or in those means themselves? The within estimator takes account of only the former, whereas the between estimator considers only the latter.

Stata's `xtsum` command can be used to analyze the source of variation in panel data by identifying the within variation and between variation in each variable.

The random effects estimator

As an alternative to considering the individual-specific intercept as a "fixed effect" of that unit, we might consider that the individual effect may be viewed as a random draw from a distribution:

$$y_{it} = X_{it}^{*}\beta^{*} + [u_i + \epsilon_{it}] \tag{6}$$

where the bracketed expression is a composite error term, with the u_i being a single draw per unit. This model could be consistently estimated by OLS or by the between estimator, but that would be inefficient in not taking the nature of the composite disturbance process into account.

A crucial assumption of this model is that u_i is independent of X^*: individual i receives a random draw that gives her a higher wage. That u_i must be independent of individual i's measurable characteristics included among the regressors X^*. If this assumption is not sustained, the random effects estimator will yield inconsistent estimates since the regressors will be correlated with the composite disturbance term.

If the individual effects can be considered to be strictly independent of the regressors, then we can model the individual-specific constant terms (reflecting the unmodeled heterogeneity across units) as draws from an independent distribution. This greatly reduces the number of parameters to be estimated, and conditional on that independence, allows for inference to be made to the population from which the survey was constructed.

In a large survey, with thousands of individuals, a random effects model will estimate K parameters, whereas a fixed effects model will estimate $(K-1)+N$ parameters, with the sizable loss of $(N-1)$ degrees of freedom.

In contrast to fixed effects, the random effects estimator can identify the parameters on time-invariant regressors such as race or gender at the individual level.

Therefore, where its use can be warranted, the random effects model is more efficient and allows a broader range of statistical inference. The assumption of the individual effects' independence is testable using `hausman`, and should always be tested.

In actual empirical work, it is extremely unusual to find that the key assumption underlying the pure random effects model is satisfied. Beyond textbook examples, it is difficult to find instances where the unobserved random effect can plausibly be uncorrelated with all observable attributes of the unit.

For instance, if you applied the estimator to country-level data on GDP growth, you would attribute the country-specific random component of the error term to a draw from nature that is uncorrelated with all observable characteristics of the country's performance.

The correlated random effects estimator

Even if the unobserved effects, the u_i terms, are considered random, there is an approach that allows them to be correlated with the regressors if we model the correlations. As u_i is constant over time, it makes sense for it to be correlated with the time-average of the x_{it}.

Assume that

$$u_i = \alpha + \gamma \bar{x}_i + r_i \tag{7}$$

where we assume that the r_i terms are uncorrelated with x_{it}, implying that $Cov(\bar{x}_i, r_i) = 0$. In this model, u_i and $\bar{x}_i$ are correlated whenever $\gamma \neq 0$.

The *correlated random effects* (CRE) approach (Mundlak, *Econometrica* 1978) then defines the model to be estimated as

$$y_{it} = \alpha + \beta x_{it} + \gamma \bar{x}_i + r_i + \epsilon_{it} \quad (8)$$

The equation still has a composite error term, but by adding the time-averaged values of the regressors, we control for the correlation between u_i and x_{it}.

The β coefficients are identical to the fixed effects coefficients on those regressors, as adding the time-averages is equivalent to centering the regressors via the Frisch–Waugh–Lovell theorem.

The β coefficients are then measuring the effects of the regressors on the outcome, controlling for the average level of x_{it} when computing the partial effect. The γ coefficients allow us to include time-invariant variables in the model, which cannot be used in a fixed-effects context.

An additional advantage of the CRE approach is that it provides a formal way of choosing between the pure RE model and the augmented CRE model. If the u_i are indeed uncorrelated with x_{it}, the estimates of γ will not be distinguishable from zero.

Care must be taken in applying the CRE model to an unbalanced panel, as the time averages should be computed over those observations for which all variables are available.

In the balanced panel case, you need not include the averages of time dummies, as they will be the same for each variable. However, in an unbalanced panel, with different numbers of observations per year, that will not be the case.

To apply the CRE model, it is best to use Perales' `mundlak` command, available from `SSC`. A more general implementation of the CRE model is provided by Schunk and Perales' `xthybrid` (*Stata Journal*, 2017) with the `cre` option. However, the `SSC` version of that routine is the latest released, and it should be installed rather than the *SJ* version.

To illustrate the CRE model, consider a version of the state-level traffic fatalities model (`bcuse traffic`) with a set of `region` indicators included for the four Census regions. This model cannot be estimated with fixed effects, as `region` is constant and collinear with the fixed effects.

```
 . // use xi, as mundlak does not handle FVs
. xi: mundlak  fatal beertax spircons unrate perincK i.region, keep
i.region          _Iregion_1-4        (_Iregion_1 for region==E omitted)
The variable _Iregion_2 does not vary sufficiently within groups and will not be use
> dditional regressors.
0% of the total variance in _Iregion_2 is within groups.
The variable _Iregion_3 does not vary sufficiently within groups and will not be use
> dditional regressors.
0% of the total variance in _Iregion_3 is within groups.
The variable _Iregion_4 does not vary sufficiently within groups and will not be use
> dditional regressors.
0% of the total variance in _Iregion_4 is within groups.
...
```

Variable	RE	Mundlak
beertax	-0.133	-0.484
spircons	0.286	0.817
unrate	-0.050	-0.029
perincK	-0.015	0.105
_Iregion_2	0.274	-0.071
_Iregion_3	1.126	0.292
_Iregion_4	0.881	0.543
mean__beertax		0.541
mean__spircons		-0.602
mean__unrate		0.033
mean__perincK		-0.263
_cons	1.697	3.627

...

N	336	336
N_g	48.000	48.000
g_min	7.000	7.000
g_avg	7.000	7.000
g_max	7.000	7.000
rho	0.826	0.826
rmse	0.180	0.157
chi2	94.239	231.496
p	0.000	0.000
df_m	7.000	11.000
sigma	0.376	0.376
sigma_u	0.341	0.341
sigma_e	0.157	0.157
r2_w	0.234	0.353
r2_o	0.290	0.627
r2_b	0.302	0.657

```
. testparm mean_*
 ( 1)  mean__beertax = 0
 ( 2)  mean__spircons = 0
 ( 3)  mean__unrate = 0
 ( 4)  mean__perincK = 0
           chi2(  4) =  107.52
         Prob > chi2 =    0.0000

.
```

The `mundlak` routine automatically fits both the pure RE model and the CRE model. The joint test of the `mean*` coefficients strongly rejects the null hypothesis that those coefficients are jointly zero. This indicates that the pure RE model is not appropriate, and that the CRE model should be used.

The seemingly unrelated regression estimator

An alternative technique which may be applied to "small *N*, large *T*" panels is the method of *seemingly unrelated regressions* or SURE. The "small *N*, large *T*" setting refers to the notion that we have a relatively small number of panel units, each with a lengthy time series: for instance, financial variables of the ten largest U.S. manufacturing firms, observed over the last 40 calendar quarters, or annual data on the G7 countries for the last 30 years.

The SURE technique (implemented in Stata as `sureg`) requires that the number of time periods exceeds the number of cross-sectional units.

The concept of ‘seemingly unrelated’ regressions is that we have several panel units, for which we could separately estimate proper OLS equations: that is, there is no simultaneity linking the units’ equations. The units might be firms operating in the same industry, or industries in a particular economy, or countries in the same region.

We might be interested in estimating these equations jointly in order to take account of the likely correlation, across equations, of their error terms. These correlations represent common shocks. Incorporating those correlations in the estimation can provide gains in efficiency.

The SURE model is considerably more flexible than the fixed-effect model for panel data, as it allows for coefficients that may differ across units (but may be tested, or constrained to be identical) as well as separate estimates of the error variance for each equation. In fact, the regressor list for each equation may differ: for a particular country, for example, the price of an important export commodity might appear, but only in that country's equation. To use `sureg`, your data must be stored in the 'wide' format: the same variable for different units must be named for that unit.

Its limitation, as mentioned above, is that it cannot be applied to models in which $N > T$, as that will imply that the residual covariance matrix is singular. SURE is a generalized least squares (GLS) technique which makes use of the inverse of that covariance matrix.

A limitation of official Stata's `sureg` command is that it can only deal with balanced panels. This may be problematic in the case of firm-level or country-level data where firms are formed, or merged, or liquidated during the sample period, or when new countries emerge, as in Eastern Europe.

I wrote an extended version of `sureg`, named `suregub`, which will handle SURE in the case of unbalanced panels as long as the degree of imbalance is not too severe: that is, there must be some time periods in common across panel units. It is available in the SSC package `itsp_ado`.

One special case of note: if the equations contain exactly the same regressors (that is, numerically identical), SURE results will exactly reproduce equation-by-equation OLS results. This situation is likely to arise when you are working with a set of demand equations (for goods or factors) or a set of portfolio shares, wherein the explanatory variables should be the same for each equation.

Although SURE will provide no efficiency gain in this setting, you may still want to employ the technique on such a set of equations, as by estimating them as a system you gain the ability to perform hypothesis tests across equations, or estimate them subject to a set of linear constraints. The `sureg` command supports linear constraints, defined in the same manner as single-equation `cnsreg`.

We illustrate `sureg` with a macro example using the Penn World Tables (v9.0) dataset, `pwt90`. For simplicity, we choose three countries from that dataset: Spain, Italy, and Greece for 1960–2007. Our model considers the consumption share of real GDP per capita (`csh_c`) as a function of its lagged value and the shares of investment and government spending (`csh_i, csh_g`).

```
. bcuse pwt90, nodesc clear

. keep if tin(1960,)
(1,820 observations deleted)

. keep csh_c csh_i csh_g countrycode year

. keep if inlist(countrycode, "ITA", "ESP", "GRC")
(9,845 observations deleted)

. levelsof countrycode, local(ctylist)
`"ESP"' `"GRC"' `"ITA"'

. reshape wide csh_c csh_g csh_i, i(year) j(countrycode) string
(note: j = ESP GRC ITA)

Data                               long   ->   wide
-----------------------------------------------------------------------------

Number of obs.                      165   ->     55
Number of variables                   5   ->     10
j variable (3 values)       countrycode   ->   (dropped)
xij variables:
                                  csh_c   ->   csh_cESP csh_cGRC csh_cITA
                                  csh_g   ->   csh_gESP csh_gGRC csh_gITA
                                  csh_i   ->   csh_iESP csh_iGRC csh_iITA
-----------------------------------------------------------------------------

. tsset year, yearly
        time variable:  year, 1960 to 2014
                delta:  1 year
```

We build up a list of equations for `sureg` using the list of country codes created by `levelsof`:

```
. loc eqns
. foreach c of local ctylist {
  2.         loc eqns "`eqns´ (csh_c`c´ L.csh_c`c´ csh_i`c´ csh_g`c´)"
  3. }
. display "`eqns´"
 (csh_cESP L.csh_cESP csh_iESP csh_gESP) (csh_cGRC L.csh_cGRC csh_iGRC csh_gGRC) (csh_cITA L.csh_cITA csh_
> iITA csh_gITA)
```

```
. sureg "`eqns´", corr
Seemingly unrelated regression
```

Equation	Obs	Parms	RMSE	"R-sq"	chi2	P
csh_cESP	54	3	.0082906	0.9644	1594.02	0.0000
csh_cGRC	54	3	.0176729	0.9160	606.11	0.0000
csh_cITA	54	3	.0093013	0.5868	87.74	0.0000

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
csh_cESP						
csh_cESP						
L1.	.7429981	.0803418	9.25	0.000	.5855311	.9004651
csh_iESP	-.1297547	.0366487	-3.54	0.000	-.2015848	-.0579245
csh_gESP	-.2479771	.1023932	-2.42	0.015	-.4486642	-.0472901
_cons	.2298568	.0693135	3.32	0.001	.0940048	.3657088
csh_cGRC						
csh_cGRC						
L1.	.6566215	.0779777	8.42	0.000	.5037881	.8094549
csh_iGRC	-.1672854	.0702747	-2.38	0.017	-.3050213	-.0295495
csh_gGRC	.2234098	.1396253	1.60	0.110	-.0502508	.4970704
_cons	.2305564	.0647088	3.56	0.000	.1037294	.3573833

(continued)

```
------------------------------------------------------------------------------
csh_cITA     |
    csh_cITA |
         L1. |   .6442806   .1024233     6.29   0.000     .4435346    .8450267
             |
    csh_iITA |  -.1245503   .0658817    -1.89   0.059     -.253676    .0045753
    csh_gITA |   .0637873   .1391879     0.46   0.647     -.209016    .3365906
       _cons |   .2338423   .0622189     3.76   0.000     .1118955     .355789
------------------------------------------------------------------------------

Correlation matrix of residuals:

          csh_cESP  csh_cGRC  csh_cITA
csh_cESP    1.0000
csh_cGRC    0.2292    1.0000
csh_cITA    0.3216    0.1134    1.0000

Breusch-Pagan test of independence: chi2(3) =     9.115, Pr = 0.0278
```

Note from the displayed correlation matrix of residuals and the Breusch–Pagan test of independence that there is evidence of cross-equation correlation of the residuals.

Given our systems estimates, we may test hypotheses on coefficients in different equations: for instance, that the coefficients on `csh_g` are equal across equations. Note that in the `test` command we must specify in which equation each coefficient appears.

```
. test [csh_cESP]csh_gESP = [csh_cGRC]csh_gGRC = [csh_cITA]csh_gITA

 ( 1)  [csh_cESP]csh_gESP - [csh_cGRC]csh_gGRC = 0
 ( 2)  [csh_cESP]csh_gESP - [csh_cITA]csh_gITA = 0

           chi2(  2) =    8.12
         Prob > chi2 =    0.0173
```

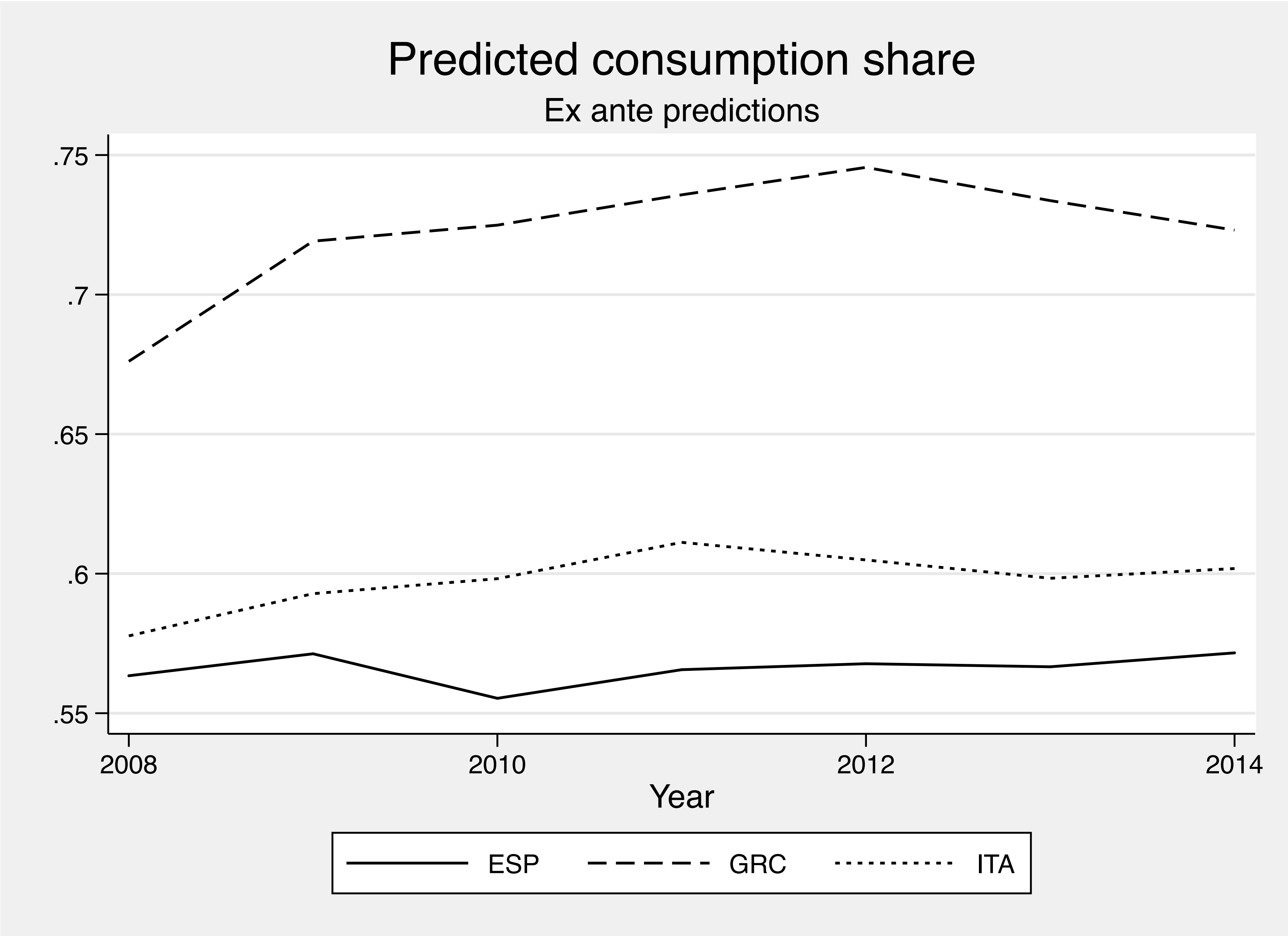
We can produce *ex post* or *ex ante* forecasts from `sureg` with `predict`, specifying a different variable name for each equation's predictions:

```
. sureg "`eqns´" if year<=2007, notable
Seemingly unrelated regression

Equation             Obs   Parms        RMSE    "R-sq"       chi2        P

csh_cESP              47       3    .0080839    0.9518     978.96   0.0000
csh_cGRC              47       3    .0163868    0.9053     453.21   0.0000
csh_cITA              47       3    .0089431    0.6135      81.72   0.0000

. foreach c of local ctylist {
  2.          qui predict double `c´hat if year>2007, xb equation(csh_c`c´)
  3.          label var `c´hat "`c´"
  4. }
. su *hat
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
      ESPhat |          7    .5659447    .0055302   .5553351   .5716213
      GRChat |          7    .7225971    .0223778   .6760783   .7455726
      ITAhat |          7    .5978607    .0106106   .5777059   .6112091
. tsline *hat if year>2007, legend(rows(1)) ti("Predicted consumption share") ///
> t2("Ex ante predictions") ylab(,angle(0) labs(small)) xlab(,labs(small)) ///
> legend(size(small)) scheme(s2mono)
```


Predicted consumption share
Ex ante predictions
.75
.7
.65
.6
.55
2008
2010
2012
2014
Year
ESP
GRC
ITA