منتدى البحوث الاقتصادية
ECONOMIC RESEARCH FORUM

Panel Models

By Chahir Zaki

Training on Applied Micro-Econometrics and Public Policy Evaluation

Economic Research Forum

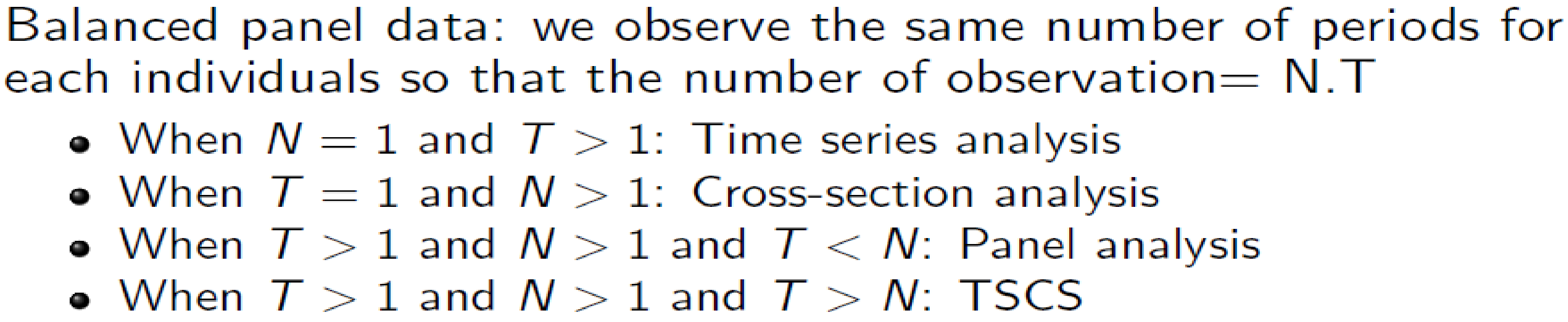
Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

Introduction



Balanced panel data: we observe the same number of periods for each individuals so that the number of observation= N.T

- When $N = 1$ and $T > 1$: Time series analysis
- When $T = 1$ and $N > 1$: Cross-section analysis
- When $T > 1$ and $N > 1$ and $T < N$: Panel analysis
- When $T > 1$ and $N > 1$ and $T > N$: TSCS

The Basic Data Structure

$$
\begin{array}{lccccl}
 & x_{111} & x_{211} & \cdot\cdot & x_{K11} & \longleftarrow \text{Wave 1} \\
\text{Individual 1} & x_{112} & x_{212} & \cdot\cdot & x_{K12} & \\
 & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \\
 & x_{11T} & x_{21T} & \cdot\cdot & x_{K1T} & \longleftarrow \text{Wave T} \\
 & x_{121} & x_{221} & \cdot\cdot & x_{K21} & \longleftarrow \text{Wave 1} \\
\text{Individual 2} & x_{122} & x_{222} & \cdot\cdot & x_{K22} & \\
 & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \\
 & x_{12T} & x_{22T} & \cdot\cdot & x_{K2T} & \longleftarrow \text{Wave T} \\
 & & & & & \\
 & x_{1N1} & x_{2N1} & \cdot\cdot & x_{KN1} & \longleftarrow \text{Wave 1} \\
\text{Individual N} & x_{1N2} & x_{2N2} & \cdot\cdot & x_{KN2} & \\
 & \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdot\cdot & \\
 & x_{1NT} & x_{2NT} & \cdot\cdot & x_{KNT} & \longleftarrow \text{Wave T}
\end{array}
$$

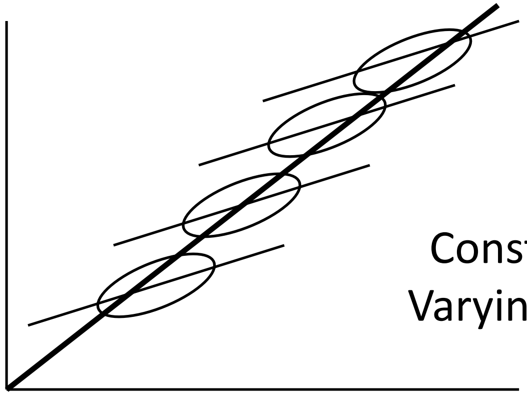
Panel Data Models

- These types of models attempt to account for correlation between observable variables and unobservable variables (Arellano 2003).
- Such heterogeneity can be caused by multiple factors, such as simultaneity (when the independent variables are correlated with the dependent variable), measurement error (which results in the independent variables being correlated with the error term, and the unobserved heterogeneity, which results in both the independent variables being correlated with the error term and bias in the coefficients (Arellano 2003).

Panel Data Models

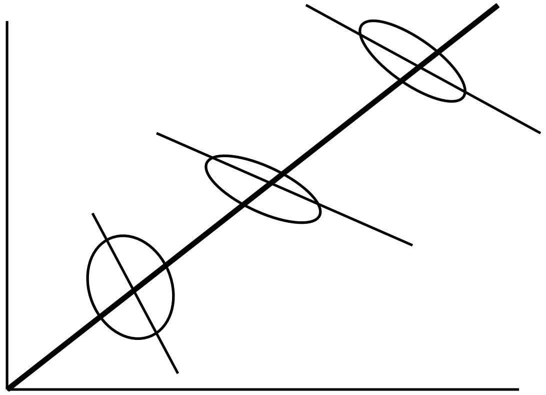
- Panel data regression models are based on panel data, which are observations on the same cross-sectional, or individual, units over several time periods.
- A balanced panel has the same number of time observations for each cross-sectional unit.
- Panel data have several advantages over purely cross-sectional or purely time series data. These include:
    - Increase in the sample size
    - Study of dynamic changes in cross-sectional units over time
    - Study of more complicated behavioral models, including study of time-invariant variables

Possible Combinations of Slopes and Intercepts



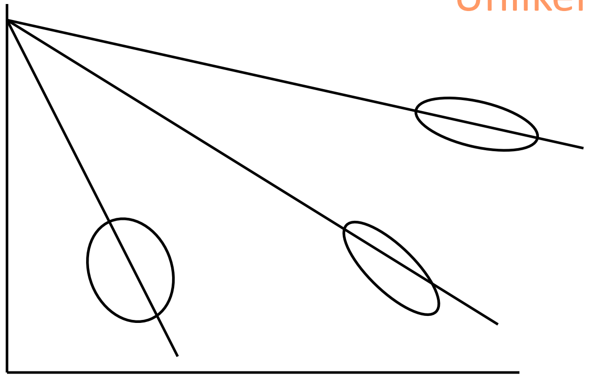
The fixed effects model

Constant slopes
Varying intercepts



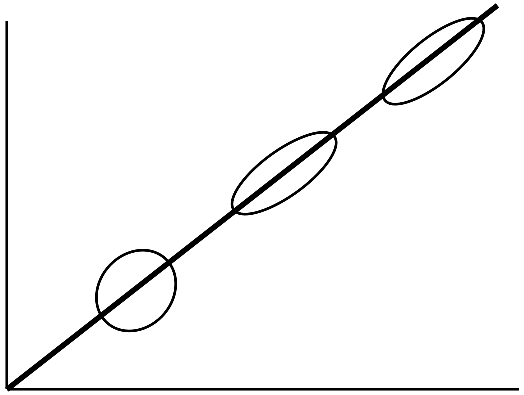
Separate regression for each individual

Varying slopes
Varying intercepts



Unlikely to occur

Varying slopes
Constant intercept



The assumptions required for this model are unlikely to hold

Constant slopes
Constant intercept

Unobserved Heterogeniety

- Omitted variables bias
- Many individual characteristics are not observed
    - e.g. enthusiasm, willingness to take risks
- These vary across individuals – described as unobserved heterogeneity
- If these influence the variable of interest, and are correlated with observed variates, then the estimated effects of these variables will be biased

Problems of Panel

- Some specific problems with panel data model need to be kept in mind:
    - The most serious problem is the problem of attrition, whereby for one reason or another, members of the panel drop out over time so that in the subsequent surveys (i.e., cross-sections) fewer original subjects remain in the panel.
    - Also, over time subjects may refuse or be unwilling to answer some questions.

Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

General Form

$$y_{it} = \alpha + X_{it}\beta + \mu_i + \lambda_t + v_{it}$$
$$u_{it} = \mu_i + \lambda_t + v_{it}$$
$$i = 1,\ldots,N, t = 1,\ldots,T$$

where

y_{it} is the dependent variable,

α is the intercept,

X_{it} is the matrix of explanatory variables with coefficients β,

u_{it} is the disturbance term,

μ_i represents unobserved cross-sectional (individual) effects for N cross sections,

λ_t represents unobserved time-series effects for T time periods, and

v_{it} represents random or idiosyncratic disturbances.

Pooled OLS and Its limitations

- An OLS estimation of panel data would look as follows

$$y_{it} = \alpha + \beta X_{it} + \theta T_i + \gamma T_i t + \delta t + \varepsilon\ _{it}$$

- To get consistent estimates of the parameters α, β,θ and γ using this model, the following conditions must be satisfied:

1. Linearity with respect to independent variables X_{it} T_{it} and e_{it}
2. Exogeneity. Expected value of disturbances ε_{it} is zero and the are not correlated to any regressors (i.e. omitted variables are not correlated with included variables)
3. Disturbances e_{it} are independent and identically distributed, have the same variance (homoscedasticity) and not related to each other (non-auto-correlated)
4. Non-stochastic independent variables
5. No exact multi-collinearity among independent variables (full-rank)

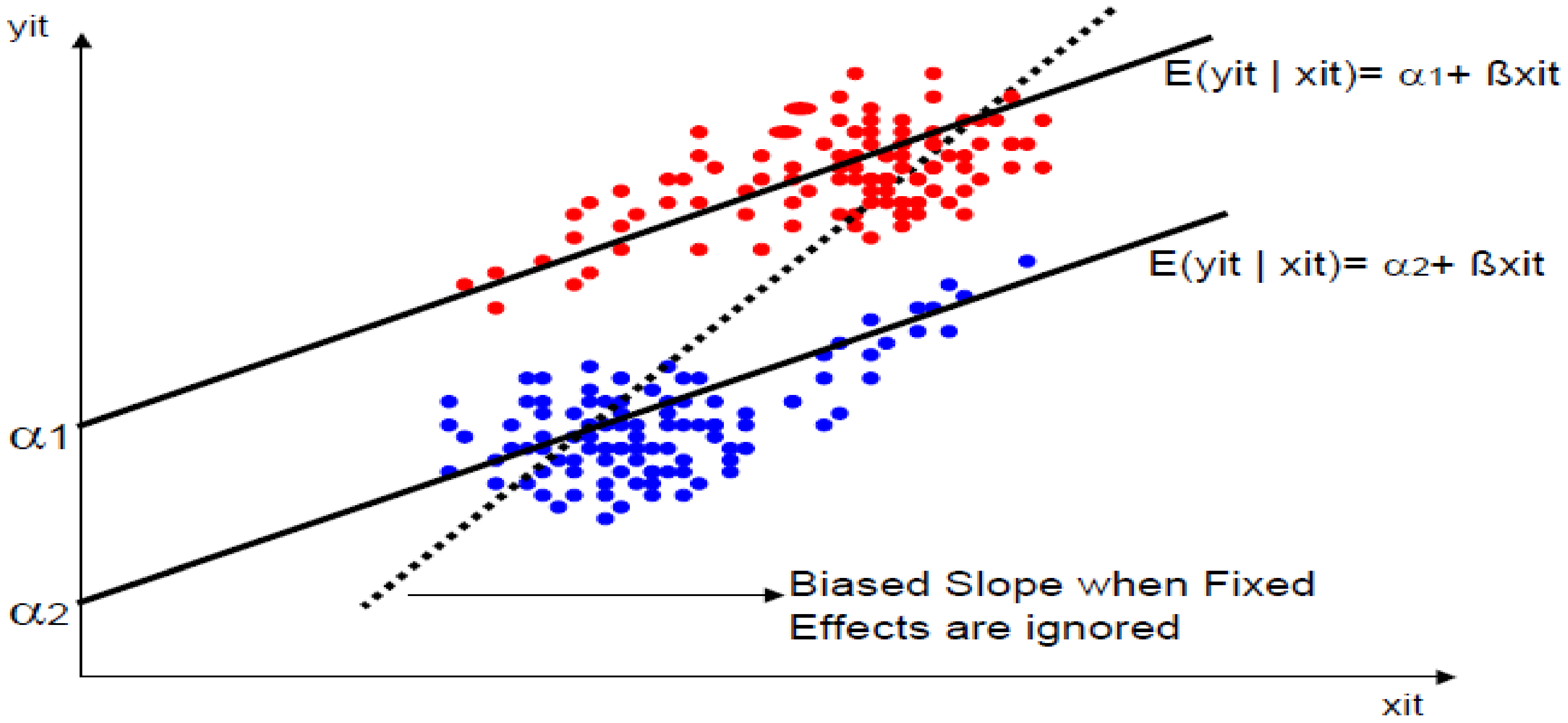
If there are time-invariant individual effects $u_i \neq 0$, this might violate assumptions 2 and 3.

Fixed and Random Effects Models

- Now we will allow for time-invariant individual effects.
- The model can be re-written in two ways
- Fixed effects: $y_{it} = (\alpha + u_i) + \beta X_{it} + \gamma T_i t + \delta t + \nu_{it}$
- Random effects: $y_{it} = \alpha + \beta X_{it} + \gamma T_i t + \delta t + (u_i + \nu_{it})$
- u_i is either a fixed effect specific to each individual (which now includes θT_i)
  - Each individual has a different intercept, but all individual have the same slopes
  - Error terms have constant variance and satisfy assumptions 2 and 3.
  - u_i can be correlated with the other regressors without causing bias
- u_i is a random effect, i.e. part of an individual-specific random component of the error term (error component model).
  - Intercepts and slopes are constant across individuals
  - However, in this case u_i cannot be correlated with X_{it} or $T_i t$ if estimates of β and γ are to remain unbiased (this would violate assumption 2)
  - Disturbances do not have constant variance, but are randomly distributed across individuals

Fixed Effects Models

- In FEM, the intercept in the regression model is allowed to differ among individuals to reflect the unique feature of individual units.
  - This is done by using dummy variables, provided we take care of the dummy variable trap.
  - The FEM using dummy variables is known as the least-squares dummy variable model (LSDV).
- FEM is appropriate in situations where the individual-specific intercept may be correlated with one or more regressors, but consumes a lot of degrees of freedom when N (the number of cross-sectional units) is very large.


yit
E(yit | xit)= α1+ ßxit
E(yit | xit)= α2+ ßxit
α1
α2
Biased Slope when Fixed Effects are ignored
xit

Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

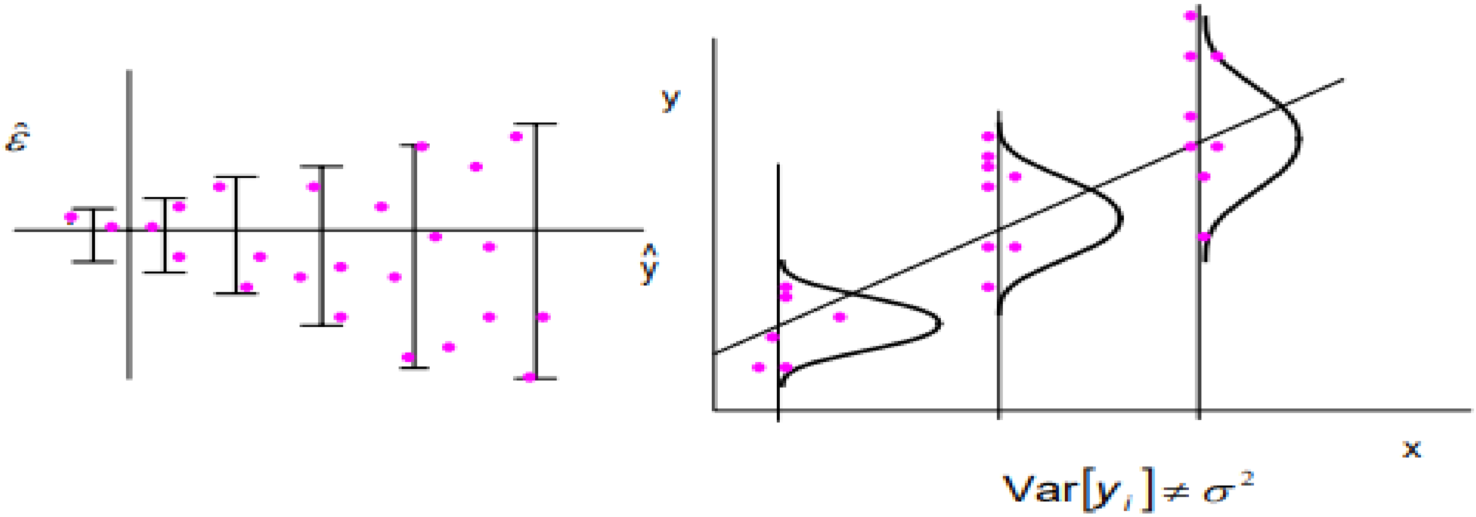
Random Effect Models

- The fixed effects model assumes that each group (firm) has a non-stochastic group-specific component to y. Including dummy variables is a way of controlling for unobservable effects on y.
- But these unobservable effects may be stochastic (i.e. random). The Random Effects Model attempts to deal with this.
- In REM we assume that the intercept value of an individual unit is a random drawing from a much larger population with a constant mean.
- The individual intercept is then expressed as a deviation from the constant mean value.
- REM is more economical than FEM in terms of the number of parameters estimated.
- REM is appropriate in situations where the (random) intercept of each cross-sectional unit is uncorrelated with the regressors.
- Unlike in FEM, time-invariant regressors can be used in REM.

Estimating Random Effects (RE) Models

- The RE model has the following composite errors $w_{it} = u_i + v_{it}$.
- Both components of the error term are assumed independent of the included variables
$$y_{it} = \alpha + \beta X_{it} + \gamma T_i t + \delta t + (u_i + \nu_{it})$$
- Because the model has two parts for the error, we obtain two variance estimates σ_u^2 and σ_v^2
- The variances differs across individuals, making the model heteroskedastic
- We therefore estimate the model using Generalized Least Squares (GLS) instead of OLS

Reminder for GLS



The measurement system used might be a source of variability, and the size of the measurement error is proportional to the measured quantity. In many cases, the variance is a function of the mean

Generalized Least Squares

Heteroskedasticity is known up to a Multiplicative Constant

- Example:

$$sav_i = \beta_0 + \beta_1 inc_i + u_i$$

$$Var(u_i \mid inc_i) = \sigma^2 inc_i$$

$$h_i = inc_i$$

$$sav_i / \sqrt{inc_i} = \beta_0 1/\sqrt{inc_i} + \beta_1 inc_i * 1/\sqrt{inc_i} + k_i$$

- Transformed equation satisfies all G-M assumptions.

Generalized Least Squares

General Least Squares Estimator

- Estimating the transformed equation by OLS is called generalized least squares (GLS)…class of estimators
- GLS will be BLUE in this case
  - Provides more efficient estimates than if used OLS in untransformed analysis.
- Can uses s.e. for t-statistics, p-values, CI, and resulting R^2 is used for F-statistics

Generalized Least Squares

General Least Squares Estimator

- GLS estimator for correcting heteroskedasticity is called WLS estimator.
    - minimize the weighted sum of squared residuals (weighted by $1/h_i$), which is inverse of the variance.
    - Less weight is given to observations with a higher error variance; in contrast, OLS gives same weight to all observations because it is best when error variance is identical for all partitions of the population

Generalized Least Squares

General Least Squares Estimator

□ WLS is great if we know what Var(u_i/$\boldsymbol{x}_i$) looks like, but in most cases won't know form of heteroskedasticity. More typically, we don't know the form of the heteroskedasticity.

□ In this case, you need to estimate $h(\boldsymbol{x}_\mathrm{i})$

- □ Since we are *estimating* h($\boldsymbol{x}_\mathrm{i}$) and using the estimate to transform the equation, call it feasible GLS.

□ Typically, we start with the assumption of a fairly flexible model, such as

- ◘ Var(u/$\boldsymbol{x}$) = $\sigma^2\exp(\delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k)$, where $h(\boldsymbol{x}_\mathrm{i}) = \exp(\delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k)$
- ◘ Since we don't know the δ, it be must estimated

Generalized Least Squares

General Least Squares Estimator

- Can transform above as
  - $u^2 = \sigma^2 \exp(\delta_0 + \delta_1 x_1 + \ldots + \delta_k x_k) v$
  - v is error term
  - assume $E(v/\boldsymbol{x}) = 1$ and $E(v) = 1$
- Taking natural logs of both sides:
  - $\ln(u^2) = \alpha_0 + \delta_1 x_1 + \ldots + \delta_k x_k + e$… α_0 now contains original intercept and $\log(\sigma^2)$
  - Assume $E(e) = 0$ and e is independent of $\boldsymbol{x}$
- We don't have u, so replace with its estimate, $\hat{u}$. Now can estimate this by OLS to get estimate of $h(\boldsymbol{x}_i)$

Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

Testing the Fixed Effects Against the Pooled Cross Section model

- Since the FE model is simply an OLS model with individual dummies added, it can be tested against the pooled-cross section model using an F–test of the joint significance of the dummies

$$F(n-1, nT-n-k) = \frac{(e'e_{pooled} - e'e_{LSDV})/(n-1)}{(e'e_{LSDV})/(nT-n-k)} = \frac{(R^2_{LSDV} - R^2_{pooled})/(n-1)}{(1-R^2_{LSDV})/(nT-n-k)}$$

Testing the Random Effects Against the Pooled Cross Section model

- Testing the RE model against the pooled cross-section model involves the Breusch-Pagan Lagrange Multiplier Test: this is essentially a χ^2 test with one degree of freedom that $\sigma_u^2=0$ and therefore the composite errors can be reduced to regular IID distributed errors
- To determine whether classical OLS (with only one constant term) should be used instead of a fixed or random effects specification, a Lagrange Multiplier test for correlation of error terms is used:

$$H_0 : E\left[u_{it}u_{is}\right] = 0$$
$$H_1 : E\left[u_{it}u_{is}\right] \neq 0$$

- The test statistic has a Chi-squared distribution with one degree of freedom under the null hypothesis and is calculated as follows:

$$LM = \frac{NT}{2(T-1)}\left[\frac{\sum_{i=1}^{n}(T\tilde{\varepsilon}_{i,})^2}{\sum_{i=1}^{n}(T\tilde{\varepsilon}_{i,})^2\sum_{t=1}^{T}(e_{it}{}^2)} - 1\right]^2$$

- If the test statistic exceeds the critical value, OLS should not be used.

Testing FE vs. RE effects

- The null hypothesis of the test is that the individual unobserved effects (the fixed effects and the random effects) are uncorrelated with the included variables X_{it}
  - If that hypothesis is rejected, FE is the preferred model because it is consistent and RE is not
  - It that null hypothesis is not rejected both models are consistent, but RE is more *efficient* and is therefore preferred
- The test compares the coefficient estimates from both models and if they are jointly not significantly different from each other, i.e., there is no detectable bias in RE (Null hypothesis is accepted and RE is preferred.

A test that does that is the **Hausman test**

$$LM = \left(b_{LSDV} - b_{random}\right)' \hat{W}^{-1} \left(b_{LSDV} - b_{random}\right) \sim \chi^2(k),$$

- Where

$$\hat{W} = Var[b_{LSDV} - b_{random}] = Var(b_{LSDV}) - Var(b_{random})$$

Fixed or Random Effects?

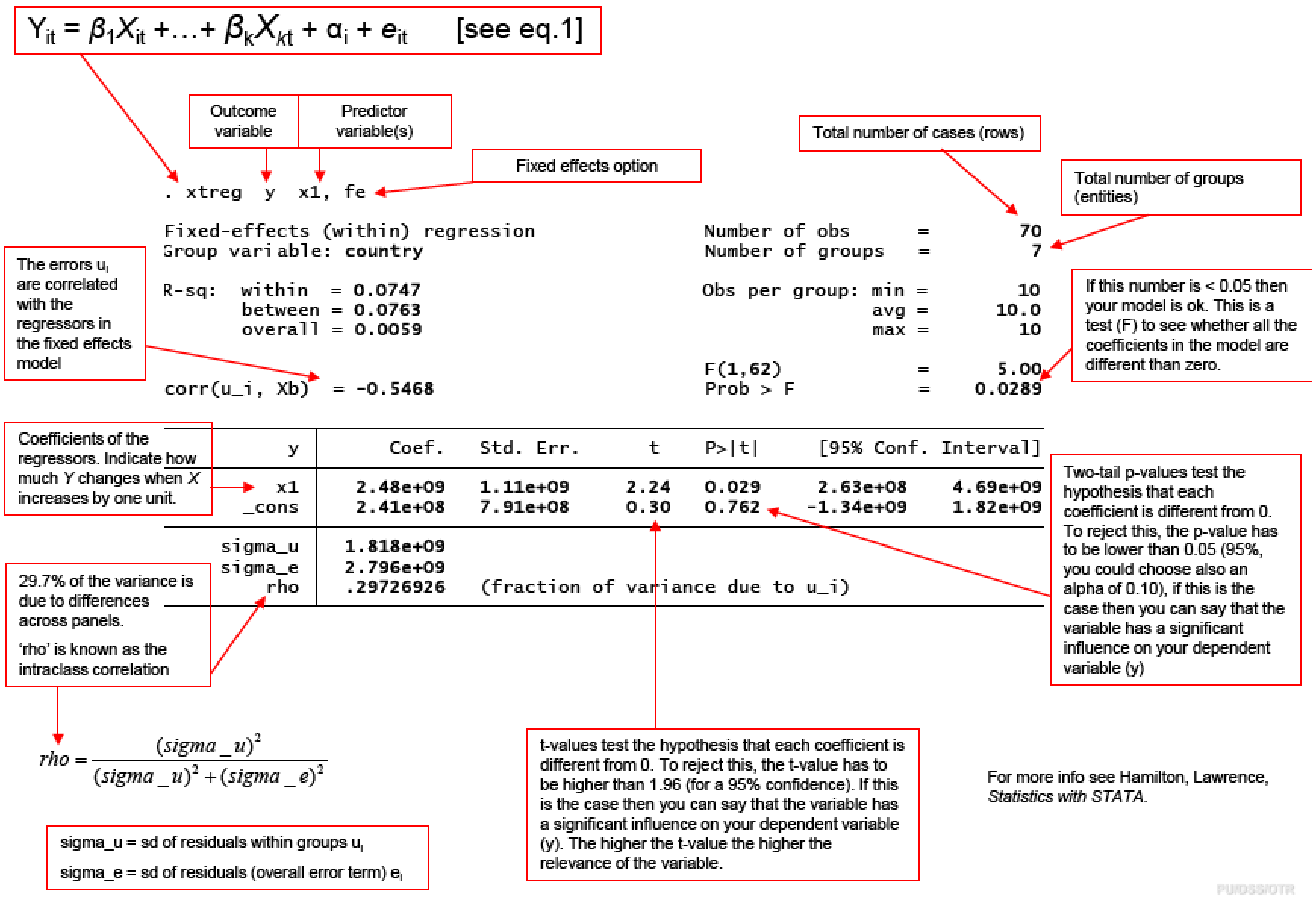
- For random effects:
  - Random effects are efficient
  - Why should we assume one set of unobservables fixed and the other random?
  - Sample information more common than that from the entire population?
  - Can deal with regressors that are fixed across individuals

- Against random effects:
  - Likely to be correlation between the unobserved effects and the explanatory variables. These are assumed to be zero in the random effects model, but in many cases we might expect them to be non-zero.
  - This implies inconsistency due to omitted-variables in the RE model.
  - In this situation, fixed effects is inefficient, but still consistent.

Outline

- Introduction
- Fixed Effects
- Random Effects
- Choosing Between Them
- Stata Application

Fixed effects: *n* entity-specific intercepts (using `xtreg`)



$Y_{it} = \beta_1 X_{it} + \ldots + \beta_k X_{kt} + \alpha_i + e_{it}$ [see eq.1]

Outcome variable

Predictor variable(s)

Fixed effects option

Total number of cases (rows)

Total number of groups (entities)

```
. xtreg  y  x1, fe

Fixed-effects (within) regression               Number of obs      =        70
Group variable: country                         Number of groups   =         7

R-sq:  within  = 0.0747                         Obs per group: min =        10
       between = 0.0763                                        avg =      10.0
       overall = 0.0059                                        max =        10

                                                F(1,62)            =      5.00
corr(u_i, Xb)  = -0.5468                        Prob > F           =    0.0289

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   2.48e+09   1.11e+09     2.24   0.029     2.63e+08    4.69e+09
       _cons |   2.41e+08   7.91e+08     0.30   0.762    -1.34e+09    1.82e+09
-------------+----------------------------------------------------------------
     sigma_u |  1.818e+09
     sigma_e |  2.796e+09
         rho |  .29726926   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

The errors u_i are correlated with the regressors in the fixed effects model

If this number is < 0.05 then your model is ok. This is a test (F) to see whether all the coefficients in the model are different than zero.

Coefficients of the regressors. Indicate how much *Y* changes when *X* increases by one unit.

Two-tail p-values test the hypothesis that each coefficient is different from 0. To reject this, the p-value has to be lower than 0.05 (95%, you could choose also an alpha of 0.10), if this is the case then you can say that the variable has a significant influence on your dependent variable (y)

29.7% of the variance is due to differences across panels.

'rho' is known as the intraclass correlation

$$rho = \frac{(sigma\_u)^2}{(sigma\_u)^2 + (sigma\_e)^2}$$

t-values test the hypothesis that each coefficient is different from 0. To reject this, the t-value has to be higher than 1.96 (for a 95% confidence). If this is the case then you can say that the variable has a significant influence on your dependent variable (y). The higher the t-value the higher the relevance of the variable.

sigma_u = sd of residuals within groups u_i

sigma_e = sd of residuals (overall error term) e_i

For more info see Hamilton, Lawrence, *Statistics with STATA*.

Source: Oscar Torres-Reyna

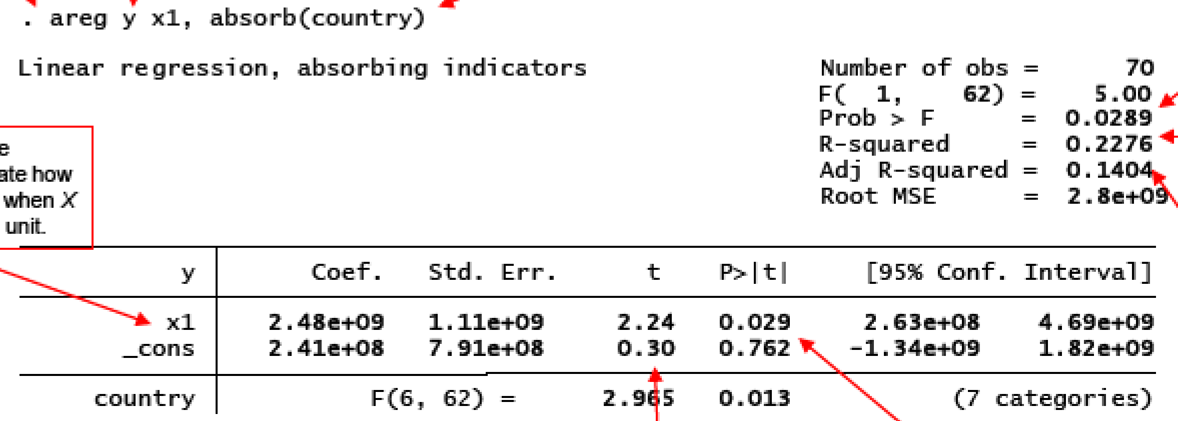
Another way to estimate fixed effects: *n* entity-specific intercepts (using `areg`)

$Y_{it} = \beta_1 X_{it} + \ldots + \beta_k X_{kt} + \alpha_i + e_{it}$ [see eq.1]

Outcome variable

Predictor variable(s)

Hide the binary variables for each entity



```
. areg y x1, absorb(country)

Linear regression, absorbing indicators               Number of obs =      70
                                                      F(  1,    62) =    5.00
                                                      Prob > F      =  0.0289
                                                      R-squared     =  0.2276
                                                      Adj R-squared =  0.1404
                                                      Root MSE      = 2.8e+09

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   2.48e+09   1.11e+09     2.24   0.029     2.63e+08    4.69e+09
       _cons |   2.41e+08   7.91e+08     0.30   0.762    -1.34e+09    1.82e+09
-------------+----------------------------------------------------------------
     country |          F(6, 62) =      2.965   0.013           (7 categories)
```

If this number is < 0.05 then your model is ok. This is a test (F) to see whether all the coefficients in the model are different than zero.

R-square shows the amount of variance of Y explained by X

Adj R-square shows the same as R-sqr but adjusted by the number of cases and number of variables. When the number of variables is small and the number of cases is very large then Adj R-square is closer to R-square.

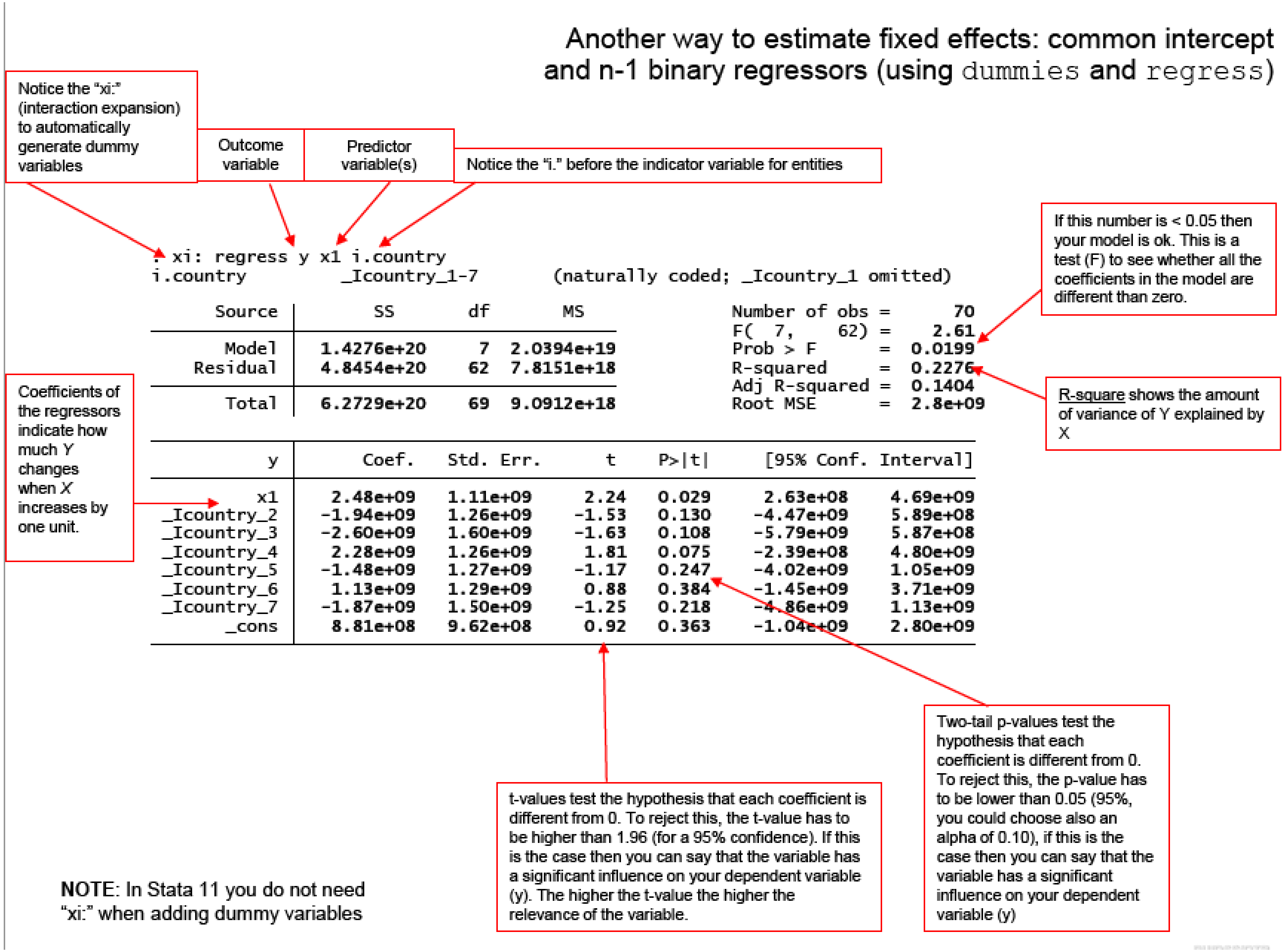
Coefficients of the regressors. Indicate how much *Y* changes when *X* increases by one unit.

"Although its output is less informative than regression with explicit dummy variables, areg does have two advantages. It speeds up exploratory work, providing quick feedback about whether a dummy variable approach is worthwhile. Secondly, when the variable of interest has many values, creating dummies for each of them could lead to too many variables or too large a model …." (Hamilton, 2006, p.180)

t-values test the hypothesis that each coefficient is different from 0. To reject this, the t-value has to be higher than 1.96 (for a 95% confidence). If this is the case then you can say that the variable has a significant influence on your dependent variable (y). The higher the t-value the higher the relevance of the variable.

Two-tail p-values test the hypothesis that each coefficient is different from 0. To reject this, the p-value has to be lower than 0.05 (95%, you could choose also an alpha of 0.10), if this is the case then you can say that the variable has a significant influence on your dependent variable (y)

Source: Oscar Torres-Reyna



Another way to estimate fixed effects: common intercept and n-1 binary regressors (using `dummies` and `regress`)

Notice the "xi:" (interaction expansion) to automatically generate dummy variables

Outcome variable

Predictor variable(s)

Notice the "i." before the indicator variable for entities

```
. xi: regress y x1 i.country
i.country         _Icountry_1-7       (naturally coded; _Icountry_1 omitted)

      Source |       SS       df       MS              Number of obs =      70
-------------+------------------------------           F(  7,    62) =    2.61
       Model |  1.4276e+20     7  2.0394e+19           Prob > F      =  0.0199
    Residual |  4.8454e+20    62  7.8151e+18           R-squared     =  0.2276
-------------+------------------------------           Adj R-squared =  0.1404
       Total |  6.2729e+20    69  9.0912e+18           Root MSE      =  2.8e+09

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   2.48e+09   1.11e+09     2.24   0.029     2.63e+08    4.69e+09
 _Icountry_2 |  -1.94e+09   1.26e+09    -1.53   0.130    -4.47e+09    5.89e+08
 _Icountry_3 |  -2.60e+09   1.60e+09    -1.63   0.108    -5.79e+09    5.87e+08
 _Icountry_4 |   2.28e+09   1.26e+09     1.81   0.075    -2.39e+08    4.80e+09
 _Icountry_5 |  -1.48e+09   1.27e+09    -1.17   0.247    -4.02e+09    1.05e+09
 _Icountry_6 |   1.13e+09   1.29e+09     0.88   0.384    -1.45e+09    3.71e+09
 _Icountry_7 |  -1.87e+09   1.50e+09    -1.25   0.218    -4.86e+09    1.13e+09
       _cons |   8.81e+08   9.62e+08     0.92   0.363    -1.04e+09    2.80e+09
------------------------------------------------------------------------------
```

If this number is < 0.05 then your model is ok. This is a test (F) to see whether all the coefficients in the model are different than zero.

R-square shows the amount of variance of Y explained by X

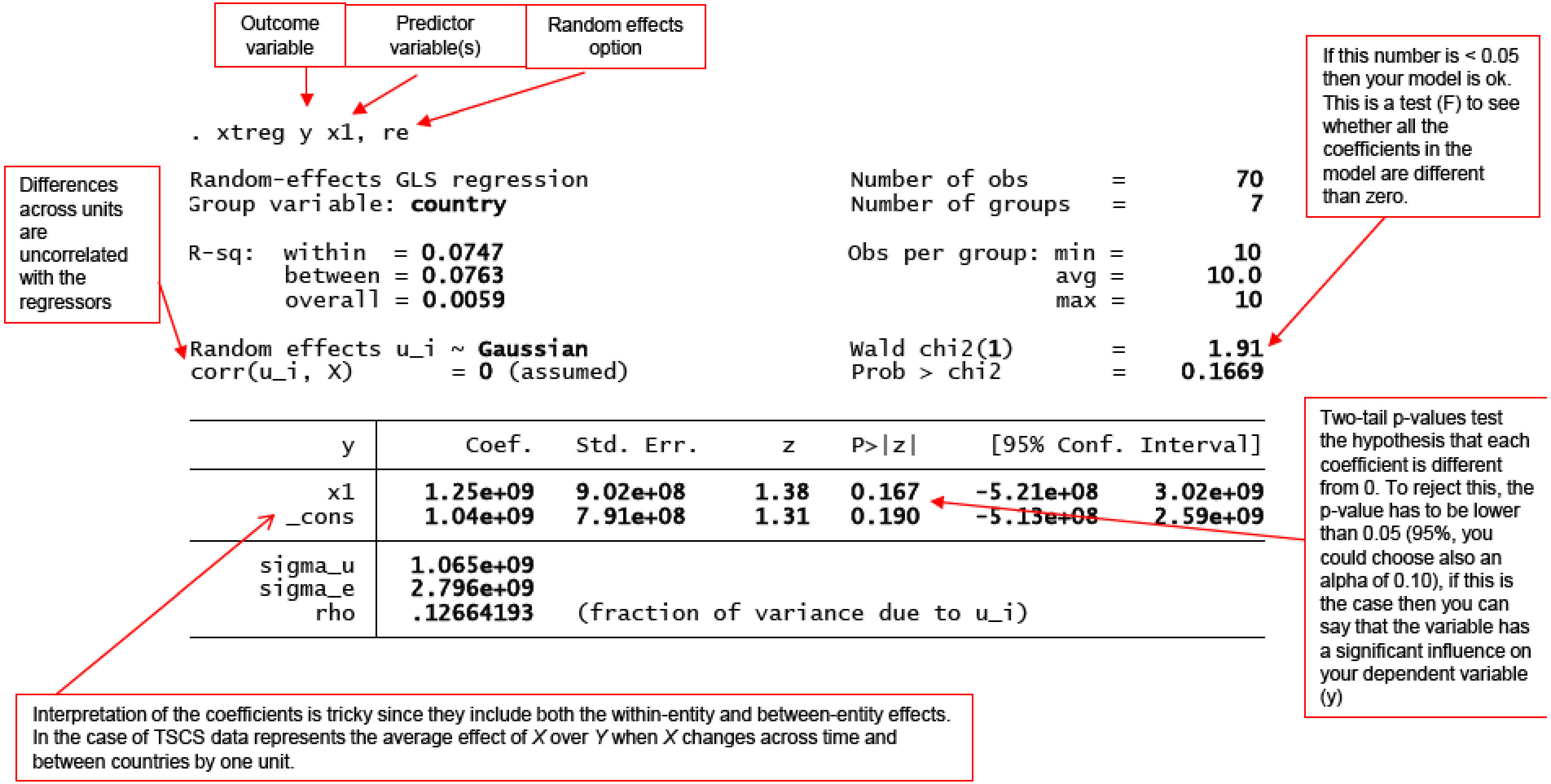
Coefficients of the regressors indicate how much *Y* changes when *X* increases by one unit.

t-values test the hypothesis that each coefficient is different from 0. To reject this, the t-value has to be higher than 1.96 (for a 95% confidence). If this is the case then you can say that the variable has a significant influence on your dependent variable (y). The higher the t-value the higher the relevance of the variable.

Two-tail p-values test the hypothesis that each coefficient is different from 0. To reject this, the p-value has to be lower than 0.05 (95%, you could choose also an alpha of 0.10), if this is the case then you can say that the variable has a significant influence on your dependent variable (y)

NOTE: In Stata 11 you do not need "xi:" when adding dummy variables

Source: Oscar Torres-Reyna



Outcome variable

Predictor variable(s)

Random effects option

```
. xtreg y x1, re

Random-effects GLS regression                   Number of obs      =        70
Group variable: country                         Number of groups   =         7

R-sq:  within  = 0.0747                         Obs per group: min =        10
       between = 0.0763                                        avg =      10.0
       overall = 0.0059                                        max =        10

Random effects u_i ~ Gaussian                   Wald chi2(1)       =      1.91
corr(u_i, X)       = 0 (assumed)                Prob > chi2        =    0.1669

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          x1 |   1.25e+09   9.02e+08     1.38   0.167    -5.21e+08    3.02e+09
       _cons |   1.04e+09   7.91e+08     1.31   0.190    -5.13e+08    2.59e+09
-------------+----------------------------------------------------------------
     sigma_u |  1.065e+09
     sigma_e |  2.796e+09
         rho |  .12664193   (fraction of variance due to u_i)
------------------------------------------------------------------------------
```

Differences across units are uncorrelated with the regressors

If this number is < 0.05 then your model is ok. This is a test (F) to see whether all the coefficients in the model are different than zero.

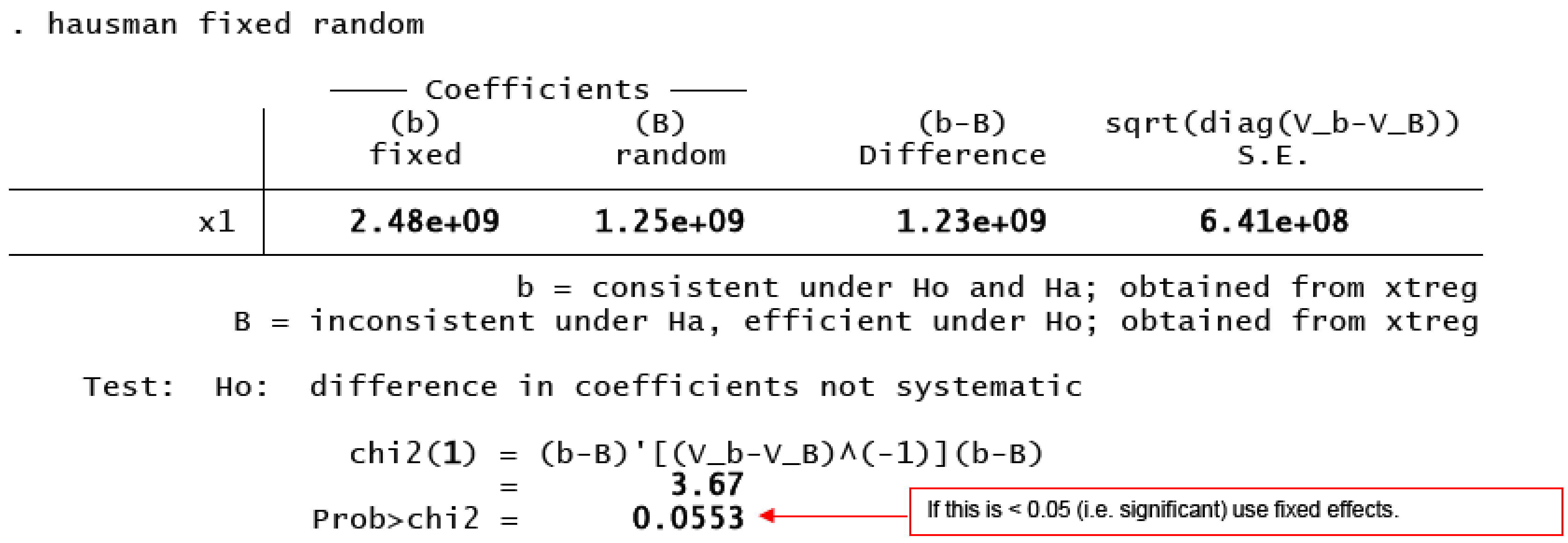
Two-tail p-values test the hypothesis that each coefficient is different from 0. To reject this, the p-value has to be lower than 0.05 (95%, you could choose also an alpha of 0.10), if this is the case then you can say that the variable has a significant influence on your dependent variable (y)

Interpretation of the coefficients is tricky since they include both the within-entity and between-entity effects. In the case of TSCS data represents the average effect of *X* over *Y* when *X* changes across time and between countries by one unit.

Source: Oscar Torres-Reyna

Run a fixed effects model and save the estimates, then run a random model and save the estimates, then perform the test. See below.

```
xtreg  y x1, fe
estimates store fixed
xtreg  y x1, re
estimates store random
hausman fixed random
```



```
. hausman fixed random

                 ---- Coefficients ----
             |      (b)          (B)            (b-B)     sqrt(diag(V_b-V_B))
             |     fixed        random       Difference          S.E.
-------------+----------------------------------------------------------------
          x1 |    2.48e+09     1.25e+09        1.23e+09        6.41e+08
------------------------------------------------------------------------------
                           b = consistent under Ho and Ha; obtained from xtreg
            B = inconsistent under Ha, efficient under Ho; obtained from xtreg

    Test:  Ho:  difference in coefficients not systematic

                 chi2(1) = (b-B)'[(V_b-V_B)^(-1)](b-B)
                         =        3.67
               Prob>chi2 =      0.0553
```

If this is < 0.05 (i.e. significant) use fixed effects.

Source: Oscar Torres-Reyna

References

- Wooldridge Intro, Chapter 14.
- Wooldridge Panel, Chapter 10.
- Oscar Torres-Reyna, Panel Data Analysis Fixed & Random Effects (using Stata 10)
- Slides of Ragui Assaad and Caroline Krafft.

Thank you for your attention