Quantitative Economics for the Evaluation of the European Policy

Dipartimento di Economia e Management

Co-funded by the Erasmus+ Programme of the European Union

Project funded by
European Commission Erasmus + Programme –Jean Monnet Action
Project number 553280-EPP-1-2015-1-IT-EPPJMO-MODULE

Irene Brunetti   Davide Fiaschi   Angela Parenti[1]

14/10/2016

[1] ireneb@ec.unipi.it, davide.fiaschi@unipi.it, and aparenti@ec.unipi.it.

From Theory to Empirics in Solow model

From the standard Solow model we have that:

$$\dot{k} = sf(k,h) - (\delta + g_A + n)\,k, \tag{1}$$

where

$$k \equiv \frac{K}{AL}\ ,\ f \equiv F\left(\frac{K}{AL}, h\right) \text{ and } f_k > 0, f_{kk} < 0 \tag{2}$$

and s is the exogenous saving/investment rate, h the level of human capital, δ the depreciation rate of physical capital, g_A the growth rate of technological change, and n the growth rate of employment.

Equilibrium level

Assuming a Cobb-Douglas function that $F(K, AH) = K^{\alpha}(AH)^{1-\alpha}$ then:

$$\dot{k} = sk^{\alpha}h^{1-\alpha} - (\delta + g_A + n)\,k; \tag{3}$$

$$k_t^{EQ} = h\left(\frac{s}{\delta + g_A + n}\right)^{1/(1-\alpha)}; \tag{4}$$

and:

$$y_t^{EQ} \equiv \frac{Y}{AL} = h\left(\frac{s}{\delta + g_A + n}\right)^{\alpha/(1-\alpha)}. \tag{5}$$

Log-linearization around the euqilibrium

The generic one-sector growth model implies, to the first-order approximation, that:

$$\log(y_t) = (1 - e^{-\lambda t})\log(y^{EQ}) + e^{-\lambda t}\log(y_0), \tag{6}$$

where y_t is the level of GDP per worker in efficiency units at time t, and the parameter λ measures the **rate of convergence**. Eq. (6) is expressed in terms of the unobservable y_t.

Suppose that technological progress is evolving as:

$$A_t = A_0 e^{g_A t} \tag{7}$$

In order to describe the dynamics in terms of the observable GDP per worker, $\frac{Y_t}{L_t}$, we get:

$$\log\left(\frac{Y_t}{L_t}\right) - g_A t - \log(A_0) = (1 - e^{-\lambda t})\log(y^{EQ}) + e^{-\lambda} t \left[\log\left(\frac{Y_0}{L_0}\right) - \log(A_0)\right] \tag{8}$$

Log-linearization around the euqilibrium (2)

Rewriting Eq. (8) we get:

$$\log\left(\frac{Y_t}{L_t}\right) = g_A t + e^{-\lambda t}\log\left(\frac{Y_0}{L_0}\right) + (1-e^{-\lambda t})\log(y^{EQ}) + (1-e^{-\lambda t})\log(A_0) \quad (9)$$

or, equivalently:

$$\begin{aligned}\log\left(\frac{Y_t}{L_t}\right) - \log\left(\frac{Y_0}{L_0}\right) &= g_A t + (1-e^{-\lambda t})\log\left(\frac{Y_0}{L_0}\right) + (1-e^{-\lambda t})\log(y^{EQ}) + \\ &+ (1-e^{-\lambda t})\log(A_0) \quad (10)\end{aligned}$$

Defining the average growth rate of GDP per worker, $\overline{g_{Y/L}}$, as:

$$\overline{g_{Y/L}} \approx \frac{\log\left(\frac{Y_t}{L_t}\right) - \log\left(\frac{Y_0}{L_0}\right)}{t} \quad (11)$$

and:

$$\beta \equiv -\frac{(1-e^{-\lambda t})}{t} \quad (12)$$

Log-linearization around the euqilibrium (3)

Therefore we get:

$$\overline{g_{Y/L}} = g_A - \beta \log\left(\frac{Y_0}{L_0}\right) - \beta \log(y^{EQ}) - \beta \log(A_0) \tag{13}$$

Substituting the level of equilibrium y^{EQ} we get:

$$\begin{aligned} \overline{g_{Y/L}} &= g_A + \beta \log\left(\frac{Y_0}{L_0}\right) + \beta \frac{\alpha}{1-\alpha} \log(n + g_A + \delta) - \beta \frac{\alpha}{1-\alpha} \log(s) \\ &- \beta \log(h) - \beta \log A_0 \end{aligned} \tag{14}$$

Issues for the estimation

- Initial condition on TFP not observable (A_0) $\Rightarrow$ we need proxies!
- Constrained on parameters
- Intercept is an estimate of the growth rate of exogenous technological progress
- Conditional convergence occurs when $\beta < 0$ and $\beta > -1$ and depends on t $\Rightarrow$ speed of convergence
- All the determinants are exogenous and no relevant variable are omitted $\Rightarrow$ OLS are unbiased estimators!

Cross-region estimation

$$\overline{g_{Y/L}}_i = intercept + \beta_0 \log\left(Y/L_{i,1991}\right) + \beta_1 \bar{s}_i + \beta_2 \bar{n}_i + \beta_3 \bar{h}_i + \epsilon_i \tag{15}$$

	Estimate	Std. Error	t-Stat.	P-value
(Intercept)	-0.0929	0.0123	-7.53	0.0000
β_0	-0.0154	0.0011	-14.57	0.0000
β_1	0.0027	0.0029	0.93	0.3532
β_2	-0.0146	0.0034	-4.31	0.0000
β_3	0.0204	0.0024	8.57	0.0000
	Res.se=0.008956 (255) DF			
	R-squared=0.6411, Adj.R-squared=0.6354			
	F-stat.=112.6 (1,255) DF, p-value=$< 2e^{-16}$			

Endogeneity in cross-region regression

Simultaneity Problem

The fact that the right-hand-side variables are not exogenous, but are **jointly determined with the growth rate** (for example the level of investment is highly correlated with growth).

- *Estimation issue*: estimates can be biased.
- *Identification issue*: the value of β can fail to illustrate how initial conditions affect expected future income differences if the saving rate is itself function of income. Hence, $\beta \geq 0$ may be compatible with at least partial convergence, while $\beta < 0$ with economic divergence if physical and human capital accumulation for rich and poor are diverging across time.

Endogeneity in cross-region regression

Measurement Error

In this case we would like to measure the (partial) effect of a variable but we can **observe only an imperfect measure** $\Rightarrow$ we introduce measurement error.

Omitted Variables

Omitted variables appear when we would like to control for one or more additional variables but, usually because of data unavailability, we cannot include them in a regression model. $\Rightarrow$ one way to represent this situation

is to write the regression equation considering the omitted variable as part of the error term.

Instrumental Variables and Two-Stage Least Squares

Consider the linear model:

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_K x_K + u \quad (16)$$

$$E(u) = 0, \ Cov(x_j, u) = 0, \ j = 1, 2, ..., K - 1 \quad (17)$$

therefore x_K might be correlated with u. In other words, $x_1, ..., x_{K-1}$ are exogenous while x_K is potentially endogenous

$\Rightarrow$ OLS estimation generally results in **inconsistent** estimators of all the β_j if $Cov(x_K, u) \neq 0$

Instrumental Variables and Two-Stage Least Squares (2)

The method of instrumental variables (IV) provides a general solution to the problem of an endogenous explanatory variable. To use the IV approach with x_K endogenous, we need an observable variable, z_1, not in equation (17) that satisfies two conditions:

- z_1 must be uncorrelated with u: $Cov(z_1, u) = 0 \Rightarrow z_1$ is exogenous
- The second requirement involves the relationship between z_1 and the endogenous variable, x_K. Consider the regression of x_K on *all* the exogenous variables:

$$x_K = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + ... + \delta_{K-1} x_{K-1} + \theta_1 z_1 + e_K \tag{18}$$

  where $E(e_K) = 0$ and e_K is uncorrelated with $x_1, ..., x_{K-1}$ and z_1 $\Rightarrow \theta_1 \neq 0$

z_1 is an **instrumental variable** candidate for x_K!

Two-stage least squares (2SLS) estimator

Under certain assumptions, the two-stage least squares (2SLS) estimator is the most efficient IV estimator:

1. Obtain the fitted values $\hat{x}_K$ from the regression:

$$x_K = \delta_0 + \delta_1 x_1 + \delta_2 x_2 + ... + \delta_{K-1} x_{K-1} + \theta_1 z_1 + e_K \tag{19}$$

   This is called **first-stage regression**.

2. Run the OLS regression

$$y = \beta_0 + \beta_1 x_1 + ... + \beta_K \hat{x}_K + u \tag{20}$$

   This is called the **second-stage regression**, and it produces the $\hat{\beta}_j$