Growth regression

Lorenzo Rocco

Department of Economics, Padova

March, 2012

Galton Fallacy (Quah, 1993)

- Galton fallacy refers to a misleading interpretation of growth regression results.
- Galton, a British aristocrat, noticed that the sons of taller fathers where taller than average but shorter than their fathers (on average) and conversely the sons of shorter fathers were shorter than average but taller than their fathers. Therefore he concluded that a "regression (= going back) to the mean" was at work and he feared of a future of homogenous mediocrity (no dispersion). However this outcome would have been incompatible with the fact that dispersion of men height was stable overtime.
- Galton conclusion was wrong because the sons of "outlier" parents are unlikely to be outliers themselves but this not implies that the new generation does not include as many "outliers" as the previous one. Therefore a less than one-to-one transmission of father height to the children does not imply convergence in children height (i.e. dispersion reduction).

Galton Fallacy

Let's consider how this applies to growth regression.
For a given set of countries, take two points in time $t1$ and $t2$ and consider the model

$$Y_i(t2) = \alpha + \beta Y_i(t1) + \varepsilon_i \tag{1}$$

where Y_i is income per capita. (This looks like regressing children height on father height).
Model (1) is equivalent to a standard growth regression as the latter can be obtained by subtracting $Y_i(t1)$ from both sides of (1)

$$Y_i(t2) - Y_i(t1) = \alpha + (\beta - 1) Y_i(t1) + \varepsilon_i \tag{2}$$

and the parameter $(\beta - 1)$ is the convergence parameter, which is negative (i.e. evidence of convergence) if $\beta < 1$

Galton Fallacy

Now suppose that the dispersion of $Y(t1)$ is equal to that of $Y(t2)$, i.e. $Var(Y(t1)) = Var(Y(t2))$. This means that between time t1 and t2 there is no convergence, i.e. there is no tendency of reducing differences in income across countries.

Suppose to estimate model (1) by OLS. By using standard OLS formula we get

$$\beta = \frac{cov(Y(t2), Y(t1))}{VarY(t1)}$$

Since $Var(Y(t1)) = Var(Y(t2))$ we can write
$Var(Y(t1)) = Var(Y(t1))^{1/2} Var(Y(t2))^{1/2} = \sigma(Y(t1))\sigma(Y(t2))$
This implies

$$\beta = \frac{cov(Y(t2), Y(t1))}{\sigma(Y(t1))\sigma(Y(t2))} \leq 1$$

by the Cauchy-Schwartz inequality (recall: $\frac{cov(Y(t2),Y(t1))}{\sigma(Y(t1))\sigma(Y(t2))}$ is the correlation index which is bounded between -1 and 1).

Galton Fallacy

Suppose now $Var(Y(t1)) < Var(Y(t2))$, i.e. divergence.

$$\beta = \frac{cov(Y(t2), Y(t1))}{VarY(t1)} \frac{Var(Y(t2))}{Var(Y(t2))} = \frac{cov(Y(t2), Y(t1))}{\sigma(Y(t1))\sigma(Y(t2))} \frac{\sigma(Y(t2))}{\sigma(Y(t1))}$$

It might occur that $\beta < 1$ even in case of divergence if the correlation index is sufficiently small.

Therefore the simple growth regression could provide evidence of convergence also in the absence of convergence.

Galton Fallacy

- Note: the same argument applies also to the case of conditional convergence: imagine to regress the residual of $Y(t2)$ on the residual of $Y(t1)$ after having removed the partial correlation with the controls X.
- An entire new field of research developed since Quah (1993), that applied time series methods (e.g. distribution dynamics) to test convergence. Generally, results reject the hypothesis of convergence (and therefore the Solow model) reinforcing the need for sensible endogenous growth model.
- Nevertheless, the fact that growth regression have little power in testing convergence does not mean that growth regressions are useless. They can still be used to investigate the determinants of growth.

Czernich, Falk, Kretschmer, Woessmann (2011)

- This recent paper aims at estimating the causal effect of broadband networks on economic growth in the OECD countries.
- This is the only paper we shall consider that does not look at developing countries. Nonetheless, it might have important policy implication for developing countries.
- Estimating the causal effect is difficult because
  - 1) a more dynamic economy could stimulate the demand for more broadband network (reverse causation)
  - 2) liberal governments could go for a more competitive telecom market that quickly adopt broadband, and at the same time they might liberalize other sectors with beneficial effects on growth.
- A less-than-standard IV technique is adopted. The exogenous source of variation is the pre-existing voice telephony and cable TV diffusion, which facilitates investments in broadband
  - for instance, the same copper wires and cable-ducts can be used

Internet and growth

High-speed internet via broadband infrastructure

- accelerates the distribution of ideas and information
- fosters competition for and development of new products and processes
- facilitates the introduction of new work practices, entrepreneurial activities and improved job matching

Data

- Annual panel of 25 OECD countries in 1996–2007.
- Broadband penetration is measured as the number of broadband subscribers per 100 inhabitants, provided by the OECD Broadband Portal.
- A broadband line is defined as a line (DSL, cable, fibre or other) that offers download speeds of at least 256 kbit / s.
- Canada was the first country in the sample that started rolling out broadband in 1997.
    - To start the analysis in the year before any broadband introduction, the first observation year is 1996.

Data

- All data on economic performance are taken from the OECD Economic Outlook and are available until 2007.
  - Real GDP is expressed in 2000 purchasing power parity (PPP) and normalised by working-age population
  - The propensity to accumulate physical capital is proxied by the ratio of real private non-residential fixed capital formation to real private GDP.
  - Human capital accumulation of the workforce is proxied by the average number of years of schooling of the population aged 15–64 years.
- Complete data are available for a panel of 20 OECD countries on an annual basis in 1996–2007. In models that do not control for investment and education, five more countries with missing data on these controls can be included.

The model

The production function evaluated at the steady state of a Solow model augmented by human capital is

$$logy_{it} = logA_i + \beta_1 log\, s_{it} + \beta_2 log\, h_{it} + \beta_3 n_{it} \qquad (3)$$

where s_{it} is the saving rate (i.e. accumulation of physical capital) h_{it} is human capital per capita, n_{it} is population (workforce) growth rate and A_i is the technology parameter in country i.
Assume that technology evolves over time according to

$$A_{it} = A_{i0} e^{\lambda_i t} \qquad (4)$$

Broadband affects the speed of technological progress λ_i according to

$$\lambda_i t = \alpha_t + \alpha_1 D_{it} \qquad (5)$$

where D_{it} is a dummy variable which takes 1 for countries and years where a broadband network is in place.
This specification assumes a one-shot shift in the rate of technological progress equal in all countries.

This specification generates a simple DiD model:

$$log\ y_{it} = \alpha_i + \alpha_t + \alpha_1 D_{it} + \beta_1 log\ s_{it} + \beta_2 log\ h_{it} + \beta_3 n_{it} + \varepsilon_{it} \quad (6)$$

An alternative specification of (5) is

$$\lambda_i = \alpha + \alpha_1 B_{it} \quad (7)$$

where B_{it} is the broadband penetration rate (proportion of subscribers) in country i at time t (intensity of broadband diffusion).

Taking first differences of (3) and using (7) we end up with

$$\Delta log\ y_{it} = \alpha + \alpha_1 B_{it} + \beta_1 \Delta log\ s_{it} + \beta_2 \Delta log\ h_{it} + \beta_3 \Delta n_{it} + \Delta\varepsilon_{it} \quad (8)$$

The model

To equation (8) two controls are added

- the initial level of GDP, one year before the first appearance of broadband (to be into a growth regression set up)
- the years since broadband introduction in the country (this ensures that we only compare rates of broadband penetration at the same point in the diffusion process, which may differ across countries)

$$\Delta log\ y_{it} = \alpha + \alpha_1 B_{it} + \beta_1 \Delta log\ s_{it} + \beta_2 \Delta log\ h_{it} + \beta_3 \Delta n_{it} + \tag{9}$$

$$+\beta_4 log\ y_{i0} + \beta_5 T_{it}^{B} + \mu_{it}$$

Note: first differences remove country fixed effects.

Results - model (6) by OLS

Table 2

Broadband Introduction and Gross Domestic Product (GDP) per Capita

Dependent variable: log of GDP per capita	Model 1	Model 2	Model 3
Dummy broadband introduction	0.023**	0.025**	0.019**
	(2.59)	(2.62)	(2.09)
Log of capital formation/GDP	0.110**		
	(2.35)		
Log of years of education	0.083		
	(0.54)		
Growth of working-age population	−0.006		
	(0.01)		
Country dummies	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes
Constant	3.644***	3.563***	3.492***
	(13.03)	(255.64)	(304.32)
R^2 (within)	0.86	0.85	0.84
Observations	240	240	300
Countries	20	20	25

Notes. Ordinary least squares estimation for 1996–2007. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Robust t statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels.

Results - model (9) by OLS

Table 3

Broadband Diffusion and Growth of Gross Domestic Product (GDP) per Capita

Dependent variable: growth of GDP per capita	Model 1	Model 2	Model 3
Broadband penetration rate	0.065***	0.091***	0.083***
	(3.08)	(4.31)	(3.03)
Years since broadband introduction	−0.003***	−0.004***	−0.003**
	(2.95)	(3.89)	(2.59)
Growth of capital formation/GDP	0.069***		
	(6.76)		
Growth of years of education	−0.007		
	(0.30)		
ΔGrowth of working-age population	−0.231		
	(1.06)		
GDP per capita in 1996	−0.001***	−0.001***	−0.001***
	(3.90)	(3.97)	(3.68)
Constant	0.049***	0.053***	0.046***
	(6.98)	(7.68)	(8.60)
R^2	0.30	0.19	0.11
Observations	240	240	300
Countries	20	20	25

Notes. Ordinary least squares estimation for 1996–2007. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Broadband penetration rate measured as broadband subscribers per 100 inhabitants; broadband line refers download speeds of at least 256 kbit/s. Robust t statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels.

Remark

Although model (6) is a DiD setting, the interpretation of the results cannot be causal

- The treatment "is chosen" and might depend on time-varying country preferences (omitted variables).
- Both the year of broadband introduction and its subsequent penetration may be endogenous and depend on economic development (reverse causation).
- Another source of endogeneity may be state intervention in the telecommunication sector. This issue has become evident in the response to the economic crisis in 2008 and 2009, when the governments of many countries initiated economic stimulus packages that emphasised investments in high-speed internet (again reversed causation).

IV strategy is based on the fact that most commonly used broadband standards (e.g. ADSL) rely on the copper wire. In countries where fibre is rolled out the existing ducts of traditional networks are used.

The IV strategy is non standard and rely on a non-linear first stage.

- It is assumed that the maximum reach of broadband γ_i (i.e. the ultimate ceiling of broadband penetration) is determined by the spread of the voice telephony (number of telecommunication access lines) and cable TV networks (number of cable TV subscribers) that existed before broadband infrastructure (1996)

$$\gamma_i = \gamma_0 + \delta_1 telnet_{i0} + \delta_2 cablenet_{i0} \tag{10}$$

[Data from the International Telecommunications Union (ITU) World Telecommunication / ICT Indicators Database]

Based on the work of Griliches (1957), the diffusion of a new technology in country i at time t is best described through a logistic curve of the following form

$$B_{it} = \frac{\gamma_i}{1 + exp(-\theta(t - \tau))} + \varepsilon_{it} \quad (11)$$

The diffusion curve can be characterised by the parameters γ, θ and τ that determine the maximum broadband penetration level (saturation level), the diffusion speed and the inflexion point of the diffusion process, respectively. At the inflexion point τ, the diffusion curve has its maximum growth rate $\theta/2$.
Substituting (10) in (11) we have a non-linear first stage which relates the penetration rates only to exogenous/predetermined variables.
The predicted values from (11) are used in (9).

IV - remarks

- (11) is non linear - it is estimated by non-linear least squares
- (11) does not include the exogenous controls in (9)
- Using a non-linear first stage to generate fitted values that are subsequently used in the second-stage equation is not innocuous because it generates consistent estimates only if the non-linear first-stage model is correctly specified (Angrist and Krueger, 2001).
- This model also generates time-variant predicted values of broadband penetration even though the instruments are time invariant.
- The year of broadband introduction in country i is estimated from the same model by defining it as the year when the penetration rate reached 1 percent

First stage

Table 4

The Diffusion Curve: First Stage of the Instrumental Variable Model

Dependent variable: broadband penetration rate	Model 1	Model 2	Model 3
Voice telephony penetration rate	0.585***	0.487***	0.653***
	(6.16)	(4.59)	(7.24)
Cable TV penetration rate	0.279**	0.301**	0.334**
	(2.11)	(2.54)	(2.47)
Diffusion speed (β)	0.647***	0.623***	0.613***
	(15.60)	(13.36)	(9.15)
Inflexion point (τ)	2,004.531***	2,004.532***	2,004.485***
	(10,423.78)	(10,508.86)	(7,939.33)
Log of capital formation/GDP			0.035**
			(2.24)
Log of years of education			0.034*
			(2.08)
Growth of working-age population			0.347
			(0.86)
Constant (γ_0)	−0.057	0.002	−0.104**
	(1.25)	(0.04)	(2.25)
R^2	0.96	0.93	0.95
F-test (cable TV penetration rate = voice telephony penetration rate = 0)	26.69	32.18	31.92
Observations	260	325	240
Countries	20	25	20

Notes. Non-linear least squares estimation for 1996–2008. Diffusion speed and inflexion point do not vary across countries. In contrast, the saturation level is country-specific and is a linear function of the voice telephony penetration rate and the cable TV penetration rate in the year before the first emergence of broadband. Model 3 contains additive control variables and is for 1996–2007. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Robust t-statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels. GDP, gross domestic product.

Second Stage - model (6)

Table 5

The Effect of Broadband Introduction on Gross Domestic Product (GDP) per Capita: Second Stage of the Instrumental Variable Model

Dependent variable: log of GDP per capita	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Predicted broadband introduction	0.027***	0.031***	0.039*** (4.56)	0.028***	0.031***	0.039***
	(2.84)	(3.18)		(2.79)	(3.10)	(4.60)
Log of capital formation/GDP	0.107**			0.108**		
	(2.11)			(2.13)		
Log of years of education	0.088			0.088		
	(0.61)			(0.60)		
Growth of working-age population	0.043			0.028		
	(0.06)			(0.04)		
Dummy 2 years before broadband introduction				0.006	0.007	−0.001
				(1.14)	(1.18)	(0.11)
Dummy 3 years before broadband introduction				0.010*	0.009	0.001
				(1.93)	(1.63)	(0.11)
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Constant	3.627***	3.563***	3.492***	3.627***	3.561***	3.492***
	(12.10)	(65.67)	(61.41)	(12.05)	(65.11)	(61.22)
R^2 (within)	0.86	0.85	0.84	0.86	0.85	0.84
Observations	240	240	300	240	240	300
Countries	20	20	25	20	20	25

Notes. Second-stage estimation for 1996–2007. The year of broadband introduction is predicted from the first-stage diffusion curve reported in Table 4, referring to the year in which the predicted broadband penetration rate was larger than 1%. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Bootstrapped z-statistics in parentheses. Significance at * 10%, ** 5%, and *** 1% levels.

Remarks

- To account for the fact that broadband penetration is itself the result of a (non-linear) first-stage estimation, standard errors are bootstrapped (500 repetitions) in this and subsequent models.
- To test the assumption of common trend in the DiD model (6), columns 4–6 in Table 5 additionally include placebo introduction dummies for points in time before the predicted introduction of broadband (dummies for two and three years before broadband introduction). The placebo introduction dummies do not enter the model significantly.
- Similarly, placebo introduction dummies for two and three years after actual broadband introduction (not shown) do not enter the model significantly

Table 6

The Effect of Broadband Diffusion on Growth of Gross Domestic Product (GDP) per Capita: Second Stage of the Instrumental Variable Model

Dependent variable: growth of GDP per capita	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Predicted broadband penetration rate	0.092***	0.135***	0.145***			
	(3.57)	(5.27)	(4.10)			
Lagged predicted broadband penetration rate				0.116***	0.156***	0.159***
				(4.05)	(5.20)	(4.49)
Years since predicted broadband introduction	−0.003***	−0.005***	−0.005***	−0.004***	−0.005***	−0.004***
	(3.92)	(5.40)	(4.03)	(4.66)	(5.53)	(4.24)
Growth of capital formation/GDP	0.069***			0.071***		
	(5.73)			(5.28)		
Growth of years of education	−0.004			−0.018		
	(0.19)			(1.18)		
ΔGrowth of working-age population	−0.18			−0.227		
	(0.77)			(0.98)		
GDP per capita in 1996	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***	−0.001***
	(6.06)	(6.57)	(4.38)	(7.01)	(7.24)	(4.18)
Constant	0.052***	0.056***	0.049***	0.057***	0.060***	0.050***
	(10.29)	(11.59)	(8.36)	(11.72)	(12.33)	(7.74)
R^2	0.29	0.19	0.12	0.35	0.25	0.14
Observations	240	240	300	220	220	275
Countries	20	20	25	20	20	25

Notes. Second-stage estimation for 1996–2007. Broadband penetration rates and year of broadband introduction are predicted from the first-stage diffusion curve reported in Table 4. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Bootstrapped z-statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels.

Remark

- Columns 4–6 in Table 6 add time lags in the growth-enhancing effect of broadband. This is based on the idea that the effect of dissemination and processing of information enabled by new broadband infrastructure may not show immediately but may take some time.
- In Table 7, country fixed effects are included in the growth regression.
  - In this specification, identification comes only from variation in growth rates within countries over time.
  - Columns 4–6 additionally include time dummies for three-year intervals (1999–2001, 2002–4 and 2005–7, with 1996–8 as the reference time period) to capture different phases of global economic growth.

Further results

Table 7

The Effect of Broadband Diffusion on Growth of Gross Domestic Product (GDP) per Capita: Instrumental Variable Results with Fixed Effects

Dependent variable: growth of GDP per capita	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Predicted broadband penetration rate	0.125***	0.170***	0.133***	0.135***	0.162***	0.118*
	(2.97)	(4.23)	(2.79)	(2.63)	(3.62)	(1.85)
Years since predicted broadband introduction	−0.004***	−0.006***	−0.004***	−0.004**	−0.005***	−0.005**
	(2.92)	(3.98)	(2.61)	(2.13)	(2.65)	(2.55)
Growth of capital formation/GDP	0.061***			0.068***		
	(4.04)			(4.56)		
Growth of years of education	0.003			0.002		
	(0.14)			(0.11)		
ΔGrowth of working-age population	−0.248			−0.242		
	(1.03)			(1.10)		
Country dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year controls	No	No	No	Yes	Yes	Yes
Constant	0.027***	0.029***	0.028***	0.024***	0.028***	0.026***
	(8.37)	(9.56)	(10.94)	(7.84)	(8.86)	(9.05)
R^2 (within)	0.27	0.15	0.06	0.30	0.17	0.10
Observations	240	240	300	240	240	300
Countries	20	20	25	20	20	25

Notes. Second-stage estimation for 1996–2007. Broadband penetration rates and year of broadband introduction are predicted from the first-stage diffusion curve reported in Table 4. Models 4–6 include dummies for three-year time intervals (1999–2001, 2002–4 and 2005–7; 1996–8 is the reference interval). Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Bootstrapped z-statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels.

Instruments validity

Given that we measure the extent of the voice telephony and cable networks in 1996, they are predetermined.
However, predetermination does not necessarily constitute exogeneity in an econometric sense.
Are there other channels through which instruments operate? are instruments excludable?

- the spread of the traditional telephone and cable networks may not only affect the deployment of the broadband network but also the diffusion of other growth-enhancing technologies, such as mobile telephony and computers.
    - We estimate diffusion curves for mobile telephony and computers where the ceiling γ_i is again determined by the penetration of the traditional fixed-line telephone network and the cable TV network.
    - No significant effects

Instruments validity

- Fixed-line voice telephony could also have a direct effect on growth. However the direct effect of fixed-line voice telephony is likely to be minor in the late 1990s, being fixed phone a mature technology. Moreover, all models control for the initial level of GDP per capita in 1996, which would capture any potential earlier indirect effects
    - Still, the non-linear form of our first stage allows us to test for any direct effects of fixed-line voice telephony and cable TV networks might have on growth.
    - Identification relies on the different functional forms of the first and second stage.

Instruments validity

Table 9

Traditional Networks and the Diffusion of Mobile Telephony and Computers

Dependent variable	Model 1 Mobile subscribers	Model 2 Computer users
Voice telephony penetration rate	0.257	3.049*
	(0.79)	(2.03)
Cable TV penetration rate	−0.186	0.700
	(0.74)	(0.94)
Diffusion speed (β)	0.548***	0.149***
	(12.38)	(5.53)
Inflexion point (τ)	1.999.347***	2.005.767***
	(11,636.54)	(367.34)
Constant (γ_0)	0.827***	−0.497
	(5.11)	(1.25)
R^2	0.96	0.95
F-test (cable TV penetration rate = voice telephony penetration rate = 0)	0.67	2.06
Observations	200	183
Countries	20	20

Notes. Non-linear least squares estimation for 1996–2005. Diffusion speed and inflexion point do not vary across countries. In contrast, the saturation level is country-specific and is a linear function of telephone access lines per 100 inhabitants and cable TV subscribers per 100 inhabitants in the year before the first emergence of broadband. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Robust t statistics in parentheses. Significance at * 10%, ** 5% and *** 1% levels.

Instruments validity

Table 10

Additional Robustness Specifications

Dependent variable: growth of GDP per capita	Model 1	Model 2†	Model 3	Model 4
Predicted broadband penetration rate	0.090***	0.104***	0.084***	0.097***
	(3.26)	(3.57)	(3.29)	(3.64)
Years since predicted broadband introduction	−0.003***	−0.004***	−0.003***	−0.004***
	(3.75)	(3.78)	(3.99)	(3.98)
Voice telephony penetration rate	−0.009			
	(0.95)			
Cable TV penetration rate	−0.003			
	(0.37)			
Trade openness			0.007***	
			(2.83)	
Log of years of education				−0.009
				(1.22)
Growth of capital formation/GDP	0.069***	0.070***	0.073***	0.069***
	(5.47)	(5.07)	(6.14)	(5.54)
Growth of years of education	−0.002	−0.001	−0.008	0.0004
	(0.09)	(0.03)	(0.35)	(0.02)
ΔGrowth of working-age population	−0.174	−0.196	−0.195	−0.179
	(0.75)	(0.78)	(0.87)	(0.76)
GDP per capita in 1996	−0.001***	−0.001***	−0.001***	−0.001***
	(4.94)	(5.44)	(4.59)	(5.84)
Constant	0.054***	0.053***	0.042***	0.067***
	(8.80)	(9.58)	(6.96)	(4.56)
R^2	0.23	0.29	0.32	0.30
Observations	240	216	240	240
Countries	20	18	20	20

Notes. Second-stage estimation for 1996–2007. Broadband penetration rates and year of broadband introduction are predicted from the first-stage diffusion curve reported in Table 4. Sample of OECD countries. Five countries drop from the full model because of lack of data on control variables. Bootstrapped z-statistics

Conclusions

The introduction and diffusion of broadband had an important impact on growth in GDP per capita.

- After a country had introduced broadband, GDP per capita was 2.7–3.9% higher on average than before its introduction
- An increase in the broadband penetration rate by 10 percentage points raised annual growth in per capita GDP by 0.9–1.5 percentage points.