Spatial and Regional Economic Analysis Mini-Course

Author: Vicente Rios

Lecture 3: Regional Development Analysis

- **Issues in European Union Regional Policy**
- **Geographically Weighted Regressions**
- **Application of GWR to the drivers of regional growth**

Issues in European Union Regional Policy

- Nowadays, the **widespread belief among academics and policy makers is that composite indexes provide a good characterization of the multidimensional nature of societal progress** (Stiglitz et al., 2010).

- Thus, recent years have seen a series of **international initiat**ives to **meet the demand for accurate social development indicators**, incorporating more than purely economic perspectives.

- Examples: Human Development Index, Happy Planet Index, etc

- Most successful of all composite indexes: the **Human Development Index** (HDI) (Annand and Sen, 1994).

Issues in European Union Regional Policy

- Problem with the HDI in Europe: not useful to analyze development

- This failure was recognized by Anand and Sen (1994):

"Yet once we take of the high and similar levels of achievement in basic capabilities, it becomes relevant to asses performance using more refined capabilities''.

This is primarily because GDP per head; literacy, enrollment and life expectancy are all high in Europe

Issues in European Union Regional Policy

- Dijkstra (2010) proposes new composite index based on Lisbon Strategy Goals, the RLI.

What is the RLI?

- Original purpose: improve Lisbon Monitoring Platform /ESPON 3 indicators.
- Specific purpose: serve as a **measure of the achievement of objectives** set out in the Lisbon Agenda or Lisbon Strategy.
- **LS goals trace an action and development plan for the EU re**gions, where the emphasis is laid on advancing towards a *knowledge society*
- **RLI does not include GDP per capita** or GDP per person employed. Correlation with GDP is 0.45. New information that cannot be learned from GDP data: employment inclusiveness, education, R\&D.
- Helps to **perform an assessment of the LS and serves as a benchmark in KBE evolutions** (EC, Central Bank).

Issues in European Union Regional Policy

- **Lisbon Strategy (LS)**

The European Council, convened in Lisbon in March of 2000, announced its intention to:

"turn the European Union into the most competitive knowledge-based economy in the world" (European Council, 2000)

LS goals trace an **action and development plan** for the EU regions. Emphasis is laid on advancing towards a **knowledge society** by developing three dimensions:

(i) labor market inclusiveness

(ii) education and

(iii) research and development.

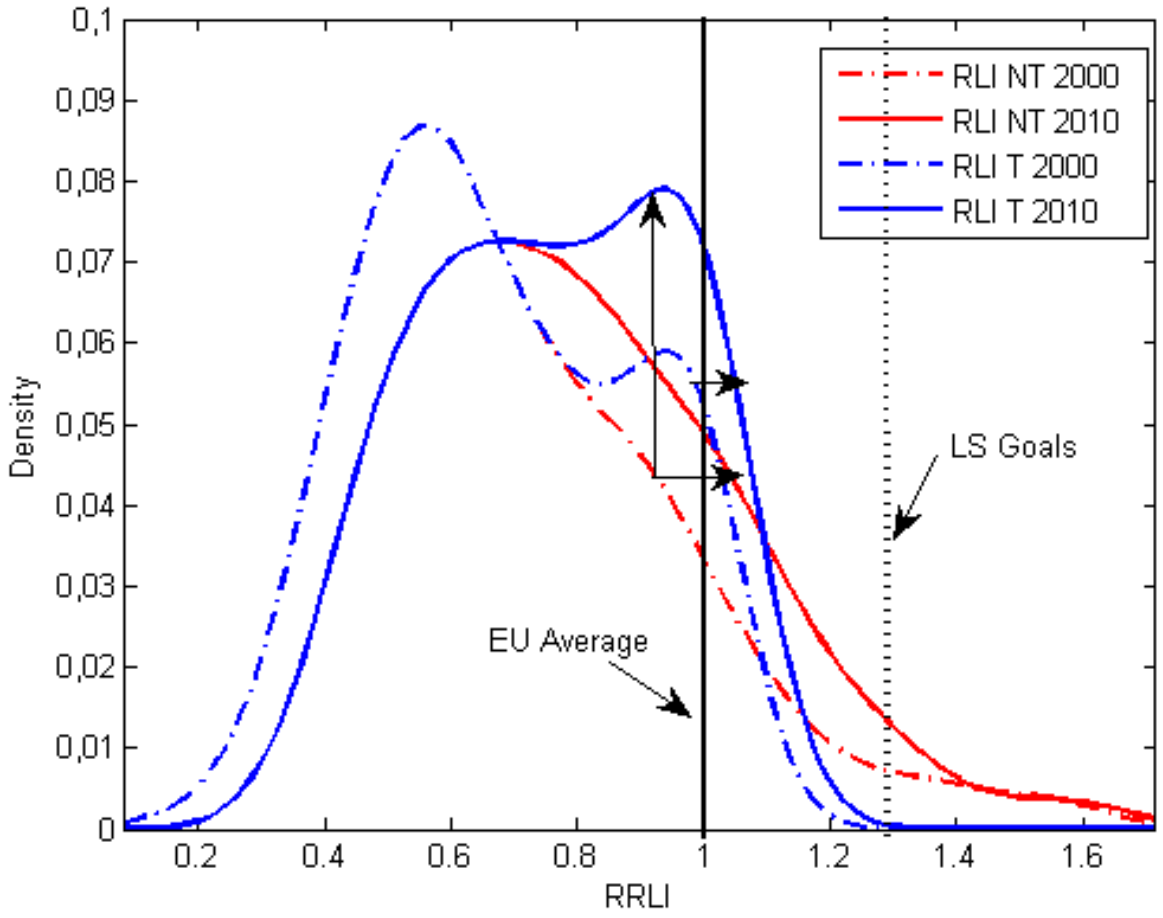
Issues in European Union Regional Policy

- Indicators

Lisbon Indicator	Target	ARLI Score 2000	ARLI Score 2010	MRLI Score 2000	MRLI Score 2010
Emp, men aged 15-54	85	0.70	0.66	0.89	0.86
Emp, women aged 15-54	64	0.86	0.98	0.92	0.99
Emp, people aged 55-64	50	0.63	0.87	0.74	0.91
Early school leavers	10	0.83	0.90	0.58	0.67
Second Educ attaintment	85	0.79	0.86	0.89	0.93
Lifelong learning	12	0.55	0.79	0.55	0.79
Private RD as % of GDP	2	0.60	0.60	0.60	0.60
Government RD as % of GDP	1	0.70	0.85	0.70	0.85
Regional Lisbon Index	100	70.4	80.5	72.1	81.5

Issues in European Union Regional Policy

- Improvement, but failure



Issues in European Union Regional Policy

- Levels

	Multiplicative RLI Ranking	2010	2000	Aditive RLI Ranking	2010	2000	GDP Per Capita Ranking	2010	2000
Top 10 Regions	Hovedstaden	1.31	1.18	Etela-Suomi	0.98	0.96	Inner London	84.24	72.17
	Vstsverige	1.29	0.95	Gloucestershire	0.97	0.96	Luxemburg	58.70	50.74
	Sydsverige	1.26	1.23	Sydsverige	0.97	0.94	Rgion de Bruxelles-Capitale	51.44	49.97
	Stockholm	1.23	0.98	Lansi-Suomi	0.97	0.93	Stockholm	46.50	41,47
	stra Mellansverige	1.20	1.22	Steiermark	0.96	0.84	Hamburg	45.99	42.56
	East Anglia	1.18	1.14	Hovedstaden	0.96	0.97	Hovedstaden	43.12	40.52
	Steiermark	1.13	0.84	Vatsverige	0.96	0.98	North Eastern Scotland	40.27	38.27
	Braunschweig	1.12	1.02	Berkshire	0.96	0.83	Oberbayern	39.00	36.88
	Prov. Vlaams Brabant	1.12	0.95	Vlaams Brabant	0.96	0.88	Berkshire	38.98	38.05
	Ita-Suomi	1.12	1.03	Hampshire	0.95	0.96	Wien	38.29	36.65
Bottom 10 Regions	Yuzhen tsentralen	0.34	0.29	Severen tsentralen	0.44	0.41	Yugoiztochen	2.80	2.58
	Vest	0.33	0.30	Sud-Est	0.43	0.49	Centru	2.78	2.00
	Severoiztochen	0.33	0.29	Severozapaden	0.43	0.41	Nord-Vest	2.58	1.73
	Yugoiztochen	0.32	0.28	Eszak-Magyarorszag	0.43	0.36	Severen tsentralen	2.45	1.21
	Sud - Muntenia	0.30	0.28	Campania	0.42	0.34	Yuzhen tsentralen	2.41	1.03
	Centru	0.30	0.26	Calabria	0.42	0.32	Severozapaden	2.37	2.85
	Sud-Est	0.29	0.29	Malta	0.41	0.27	Sud - Muntenia	2.32	1.44
	Sud-Vest Oltenia	0.28	0.31	Centru	0.41	0.44	Sud-Est	2.22	1.61
	Severozapaden	0.28	0.26	Puglia	0.40	0.34	Sud-Vest Oltenia	2.11	1.60
	Severen tsentralen	0.28	0.27	Sicilia	0.38	0.31	Nord-Est	1.69	1.29

Note: GDP pc data is computed in 2000 constant prices and expressed in thousands.

Issues in European Union Regional Policy

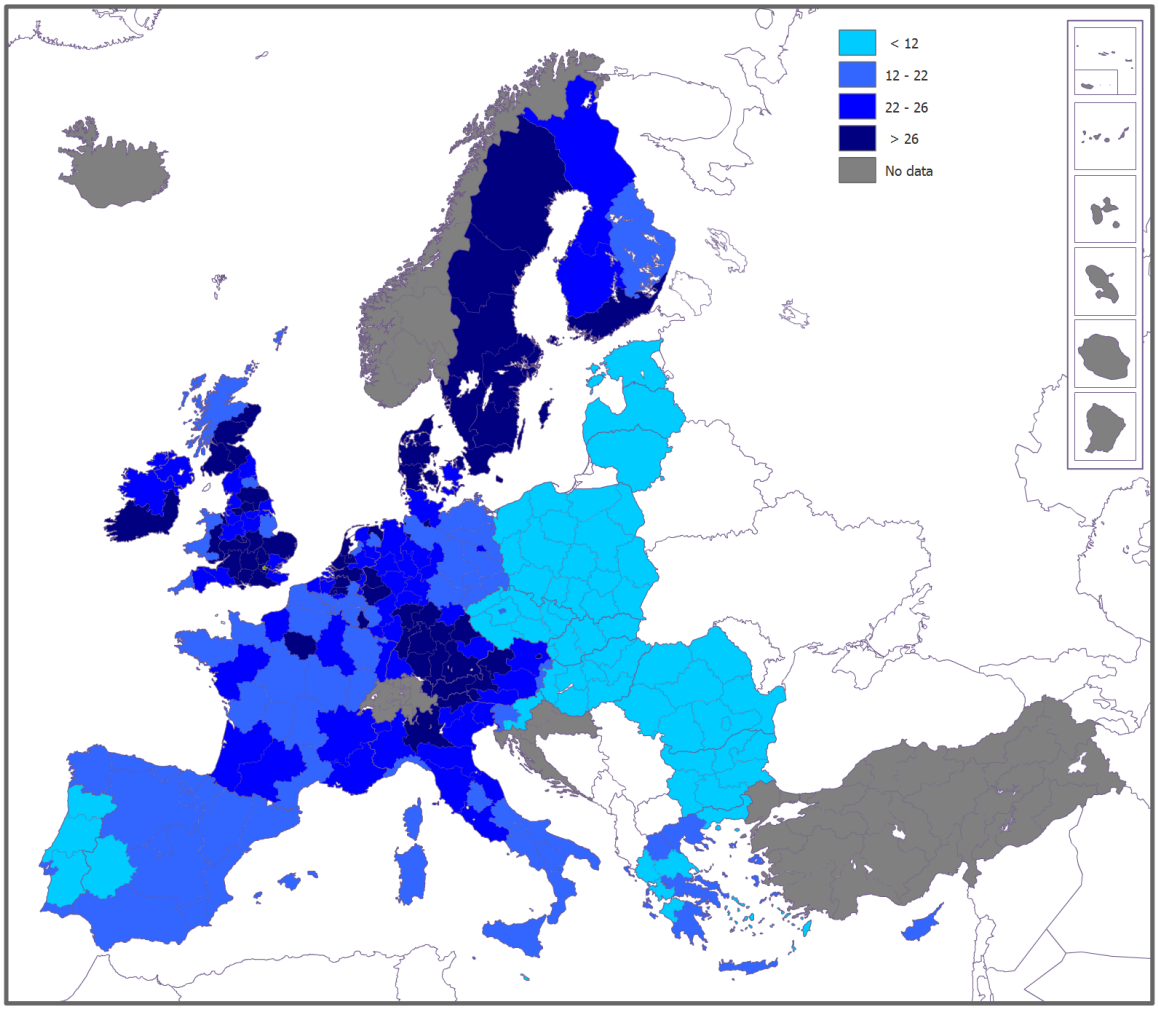
- Dynamics

	Multiplicative RLI		Aditive RLI		GDP Per Capita	
Fast Moving	Lincolnshire	7.36	Corse	13.57	Inner London	1.21
Regions	Norte	6.95	Extremadura	6.30	Luxemburg	0.80
	Swietokrzyskie	5.54	Andalucia	4.37	Ovre Norrland	0.72
	Cantabria	4.97	Sardegna	4.21	Souther and Eastern	0.67
	Nyugat-Dunantul	4.75	Cantabria	4.13	Norra Mellansverige	0.62
	La Rioja	4.68	Asturias	4.09	Highlands and Islands	0.61
	Estonia	4.47	Malta	4.09	Praha	0.61
	Prov. Antwerpen	4.25	Slaskie	4.09	Groningen	0.59
	Luxemburg	4.06	La Rioja	4.06	Attiki	0.58
	Corse	3.85	Champagne-Ardenne	3.86	Aland	0.58
Lagging Behind	Essex	-0.29	Centru	-0.42	Alsace	-0.15
Regions	Friesland	-0.33	Leicester shire	-0.46	Prov Aut Bolzano-Bozen	-0.15
	Greater Manchester	-0.39	Bedfordshire	-0.48	Abruzzo	-0.16
	Dorset and Somerset	-0.47	Muntenia	-0.49	La Rioja	-0.17
	Bedfordshire	-0.47	Yorkshire	-0.59	Prov Aut Trento	-0.18
	Utrecht	-0.72	Greater Manchester	-0.59	Comunidad de Madrid	-0.19
	Ovre Norrland	-0.78	Ostra Mellansverige	-0.62	Catalunya	-0.20
	Opolskie	-1.27	Bratislavsky	-0.64	Comunidad Valenciana	-0.20
	Herefordshire	-1.59	Kent	-0.81	Canarias	-0.23
	Aland	-1.70	Sud-Est	-0.91	Baleares	-0.50

Note: GDP pc data is computed in 2000 constant prices and expressed in thousands.

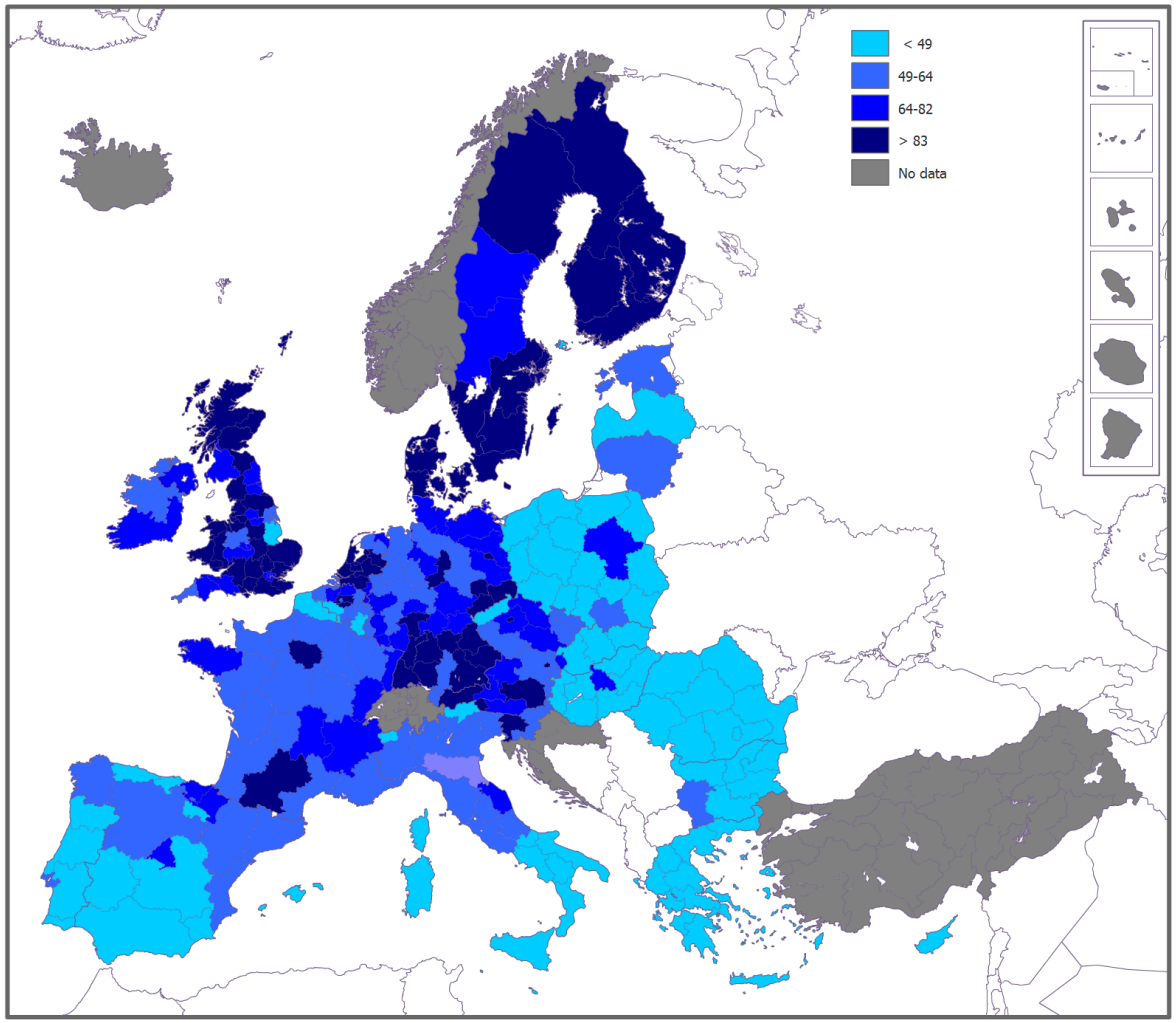
Issues in European Union Regional Policy

- GDP per capita 2000-2010



Issues in European Union Regional Policy

- RLI 2000-2010



Issues in European Union Regional Policy

- Nowadays the planning strategy is **EU2020**
  - Many RLI indicators remain as part of the agenda
  - Addded new targets linked to environmental issues and competitiveness
  - Heterogeneous country and regional targets
  - Reflects increasing awareness on the need of place-based policies (Barca et al., 2011)

- A **'place-based'** narrative has animated debate and shaped ideas about the policy's rationale, governance model and relationship with the EU's 2020 development agenda.

- Key idea: development **intervention should** increasingly **focus on efficiency and social inclusion at the expense of an emphasis on territorial convergence** and how **strategies should consider** economic, social, political, and institutional **diversity** in order **to maximize** both the local and the aggregate potential **for economic development**.
- **What works well in i may not work in j!!**

Issues in European Union Regional Policy

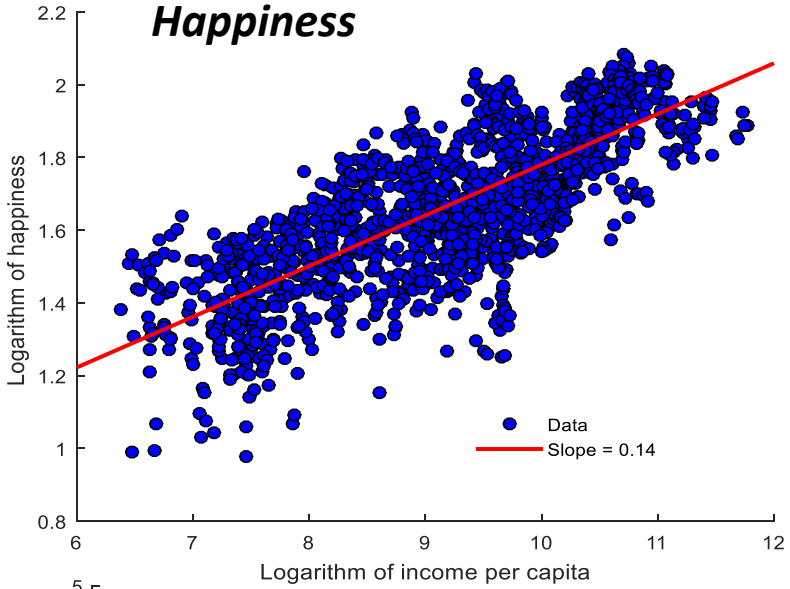
- **“Place-based” intervention** has been argued by many relevant geographers and policy-makers to be **a flip/radical change in the approach to carry out policy** leaving outdated traditional economic wisdom from Solow, Myrdal, Hirschman, Rostow, etc.

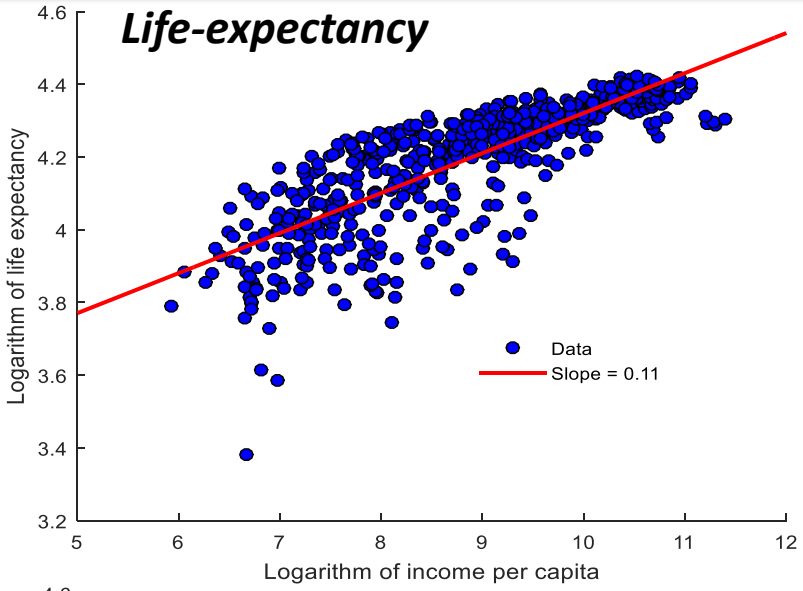
- We have seen that there is an argument for multidimensional indicators…

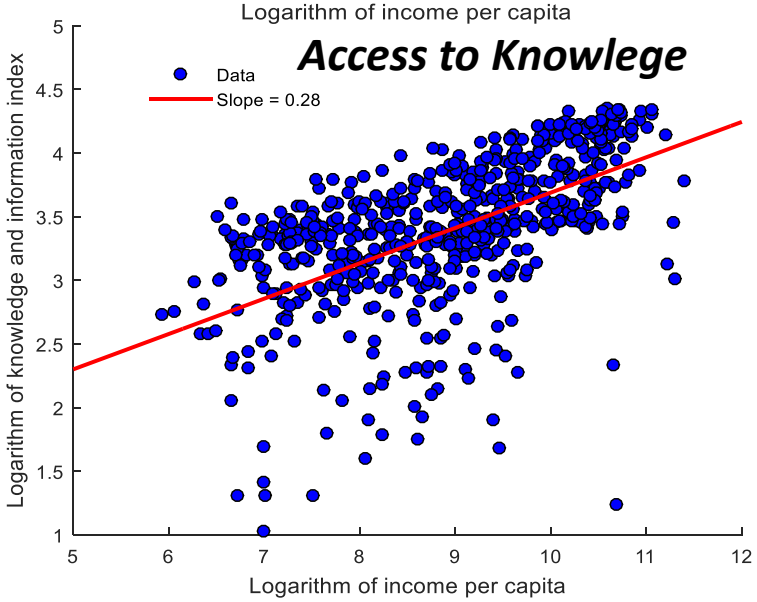
… but growth remains key in all policy discussions.

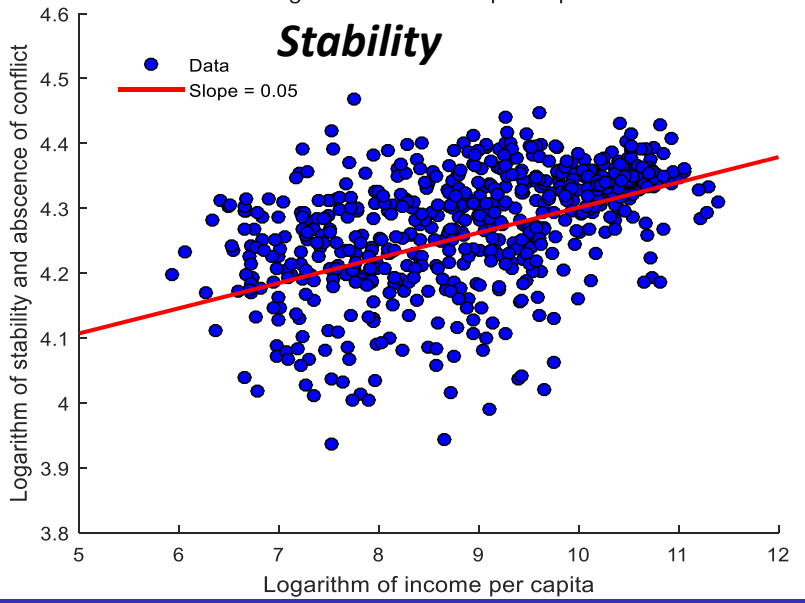
Why this?

The importance of economic growth





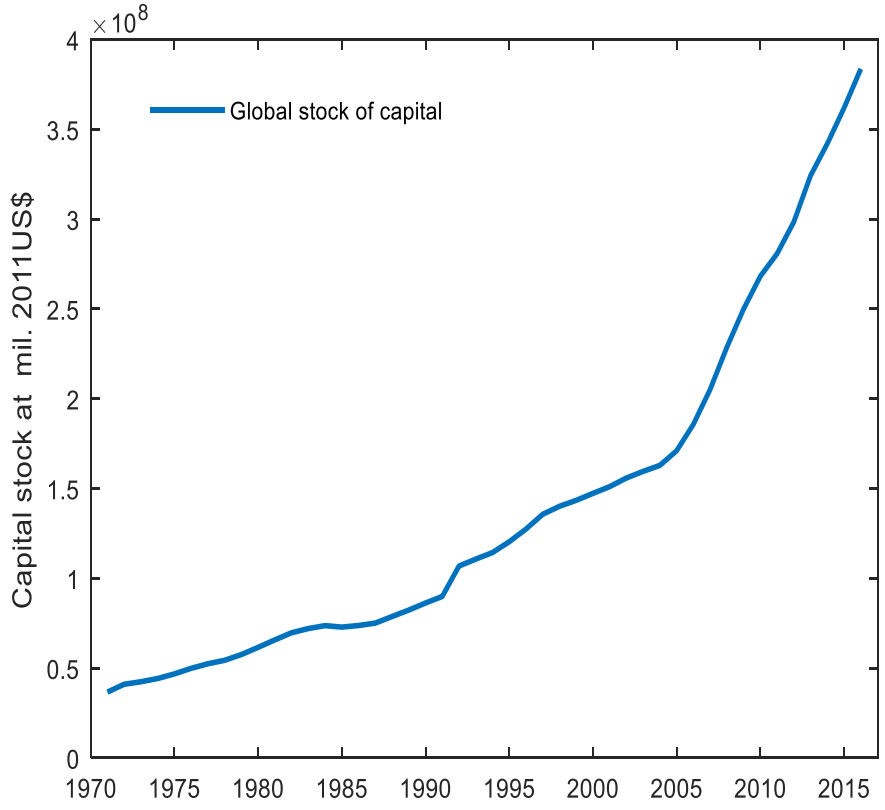




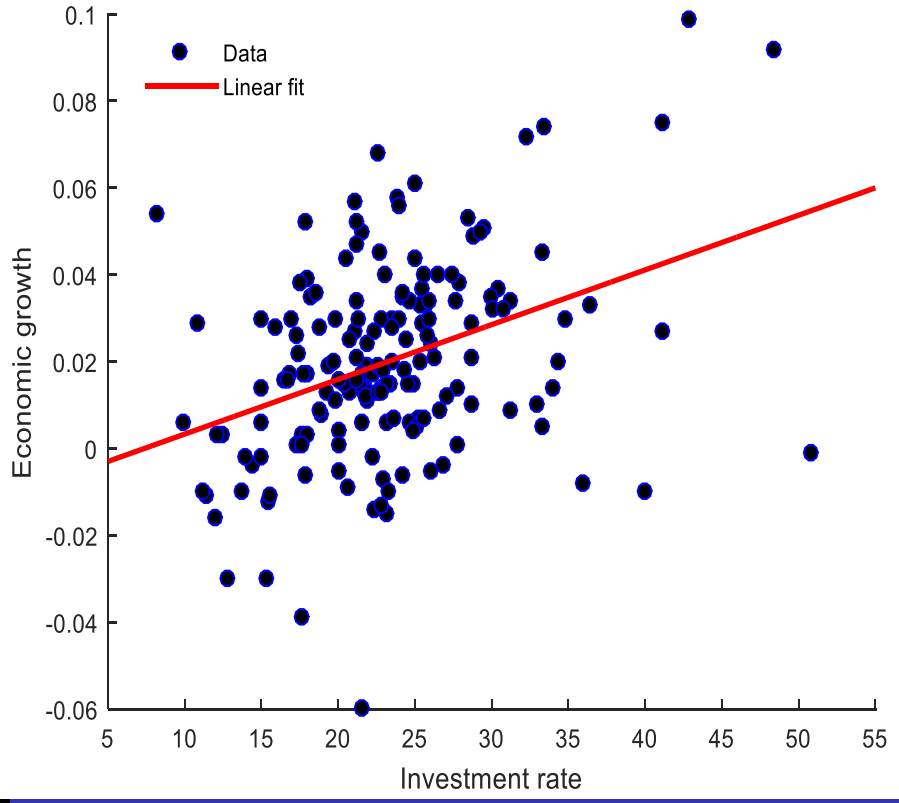
Growth drivers in the Neoclassical growth model

(1) Physical capital and investment

Global capital stock



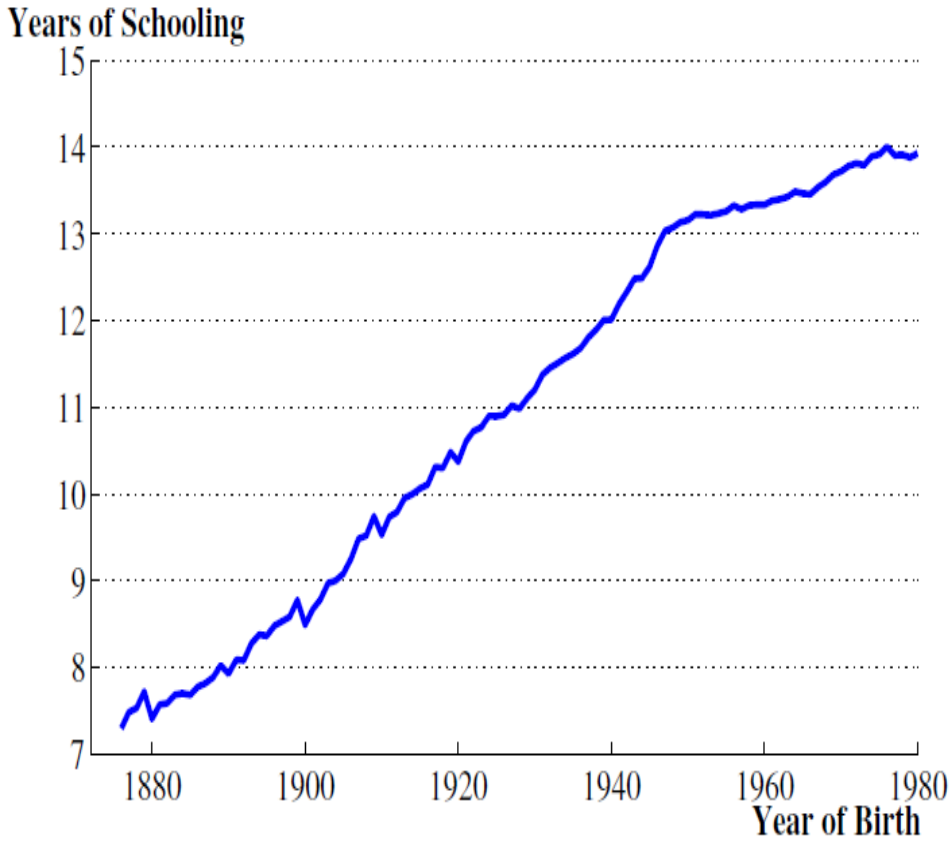
Investment rates and growth



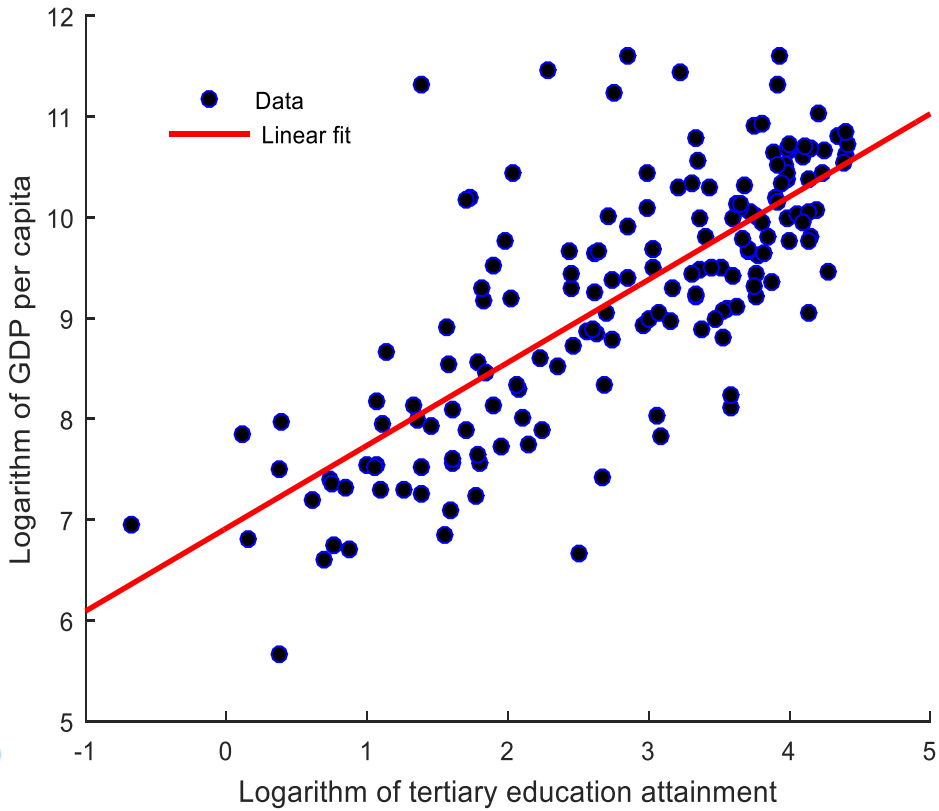
Growth drivers in the Neoclassical growth model

- (2) Human capital and investment in education

Years of schooling



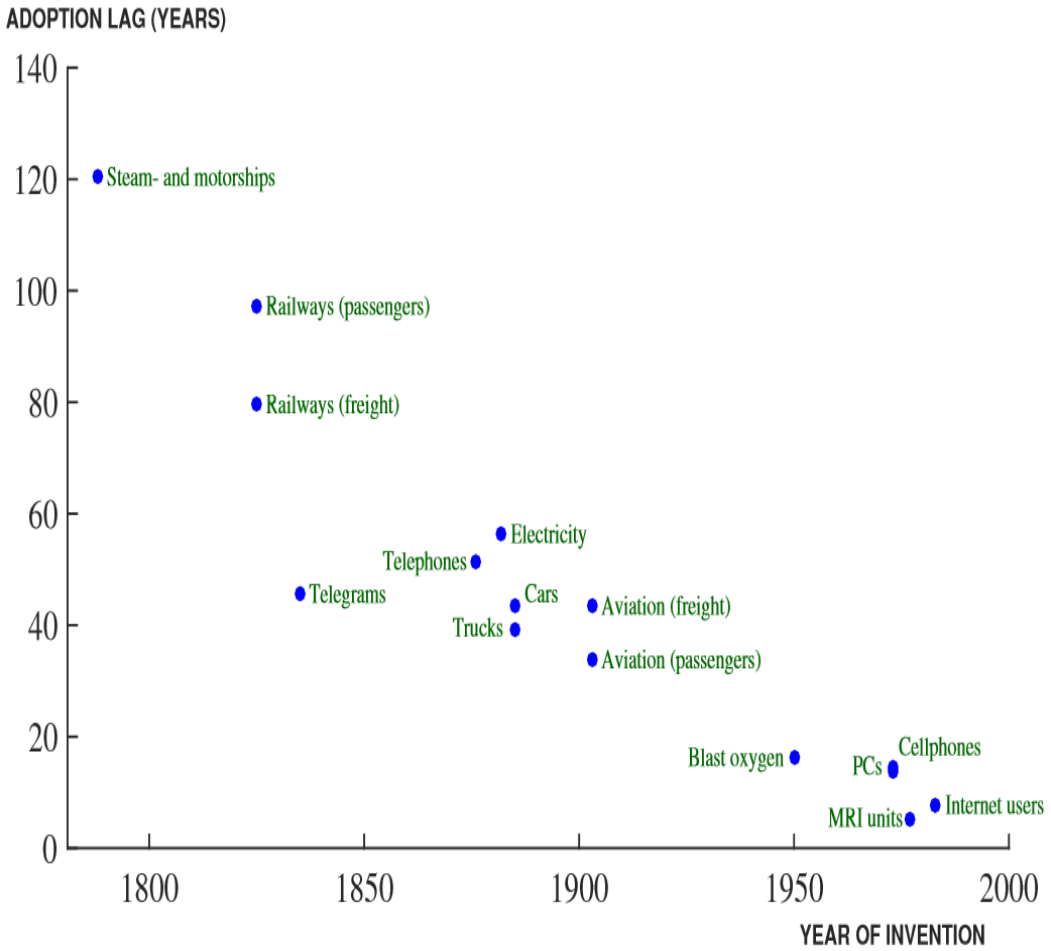
Education and income



Growth drivers in the Neoclassical growth model

- (3) Technological progress

Diffusion and Acceleration of technological progress



The Model

Production: $Y_t = K_t^{\alpha} H_t^{\beta} (A_t L_t)^{1-\alpha-\beta}$

where:

Technology: $A_t = A_0(1+g)^t$

Labor (population): $L_t = L_0(1+n)^t$

} Exogenous forces

Physical capital: $K_{t+1} = K_t + I_t^K - \delta_k K_t$

Human capital: $H_{t+1} = H_t + I_t^H - \delta_H H_t$

} Endogenous processes

Physical Investment: $I_t^K \equiv s_k\, Y_t$

Human capital Investment: $I_t^H \equiv s_H\, Y_t$

Consumption: $C_t \equiv 1 - (s_k + s_H)\, Y_t$

} Accounting identities

Parameter values: $\alpha, \beta, s_H, \delta_H, \delta_k \in (0,1)$ and $g, n > 0$

Initial values: $K_0, H_0, L_0, A_0 > 0$

The Model

If Europe has j=1, …, N economies:

For region j, the production function can be writen as:

$$Y_{jt} = K_{jt}^{\alpha} \left(A_{jt} L_{jt}\right)^{1-\alpha-\beta} H_{jt}^{\beta}$$

After some algebraic manipultations and using the equilibrium values of K and H the evolution of per capita income of region j can be expressed as:

$$y_{jt} = A_{jt} \left(\left(\frac{s_{k,j}}{g_j + \delta_k + n_j} \right)^{\frac{\alpha}{1-\alpha-\beta}} \left(\frac{s_{h,j}}{g_j + \delta_h + n_j} \right)^{\frac{\beta}{1-\alpha-\beta}} \right)$$

Taking logarithms:

$$\ln y_{jt} = \ln A_{j0} + gt + \frac{\alpha}{1-\alpha-\beta} \ln \left(\frac{s_{k,j}}{g_j+\delta_k+n_j} \right) + \frac{\beta}{1-\alpha-\beta} \ln \left(\frac{s_{h,j}}{g_j+\delta_h+n_j} \right)$$

which is equivalent to:

$$\ln y_{jt} = \ln A_{j0} + gt + \frac{\alpha}{1-\alpha-\beta} \ln(s_{k,j}) + \frac{\beta}{1-\alpha-\beta} \ln(s_{h,j}) - \frac{\alpha+\beta}{1-\alpha-\beta} \ln(n_j + 0.05)$$

once we asume $\delta_k = \delta_H$ and that $g + \delta = 0.05$.

We can test the if the expected effects hold when looking at the data!!

Issues with Neoclassical growth model

Estimates of the Augmented Solow Model

	MRW	Updated data	
	1985	1985	2000
$\ln(s_k)$	.69	.65	.96
	(.13)	(.11)	(.13)
$\ln(n+g+\delta)$	-1.73	-1.02	-1.06
	(.41)	(.45)	(.33)
$\ln(s_h)$	.66	.47	.70
	(.07)	(.07)	(.13)
Adj R^2	.78	.65	.60
Implied α	.30	.31	.36
Implied β	.28	.22	.26
No. of observations	98	98	107

→ Positive effect of investment/savings rate

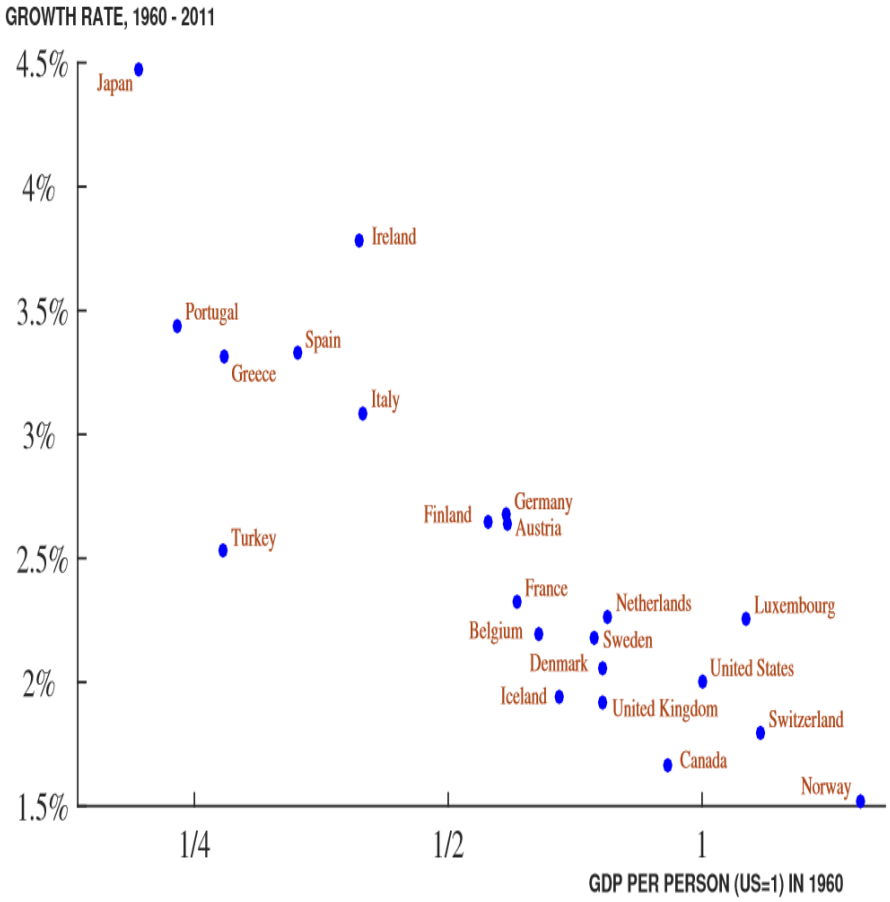
→ Negative effect of population growth and effective depreciation

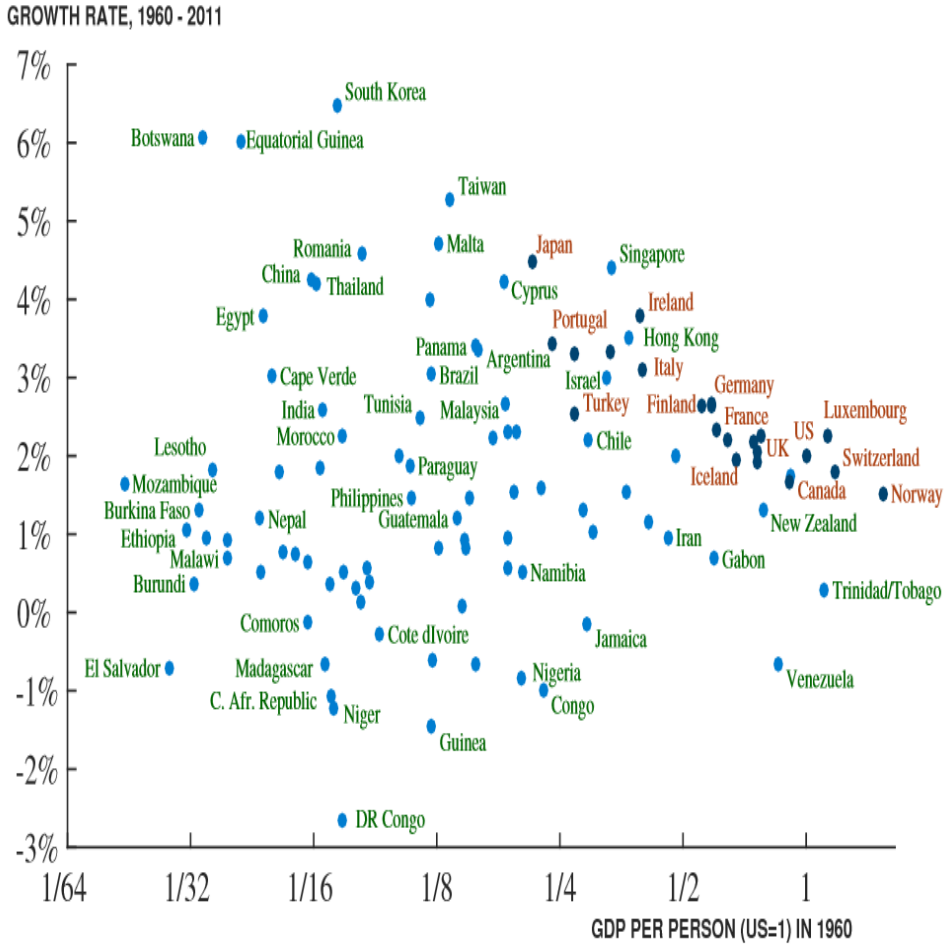
→ Positive effect of human capital investment

→ We can explain 60-75% of income per capita differences based on these three drivers

Issues with the Neoclassical growth model

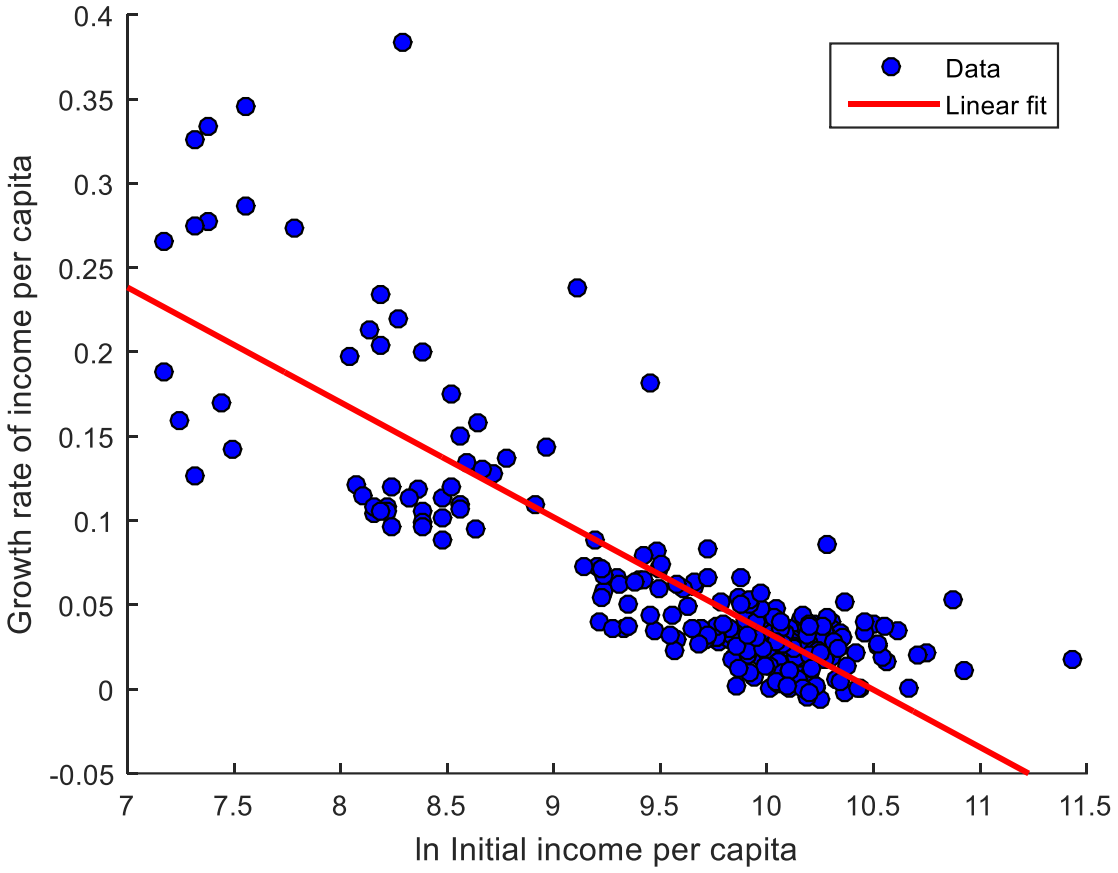
- Key prediction of the model → convergence. This is not in line with global data





Issues with the Neoclassical growth model

- but the prediction of **regional convergence holds in European regions** (scatter is for period 2000-2008)



Issues with the Neoclassical growth model

- Application of neoclassical economic framework to guide development policies too simplistic and can be reduced to:

"if you want more economic growth per capita either (i) invest more on physical capital, human capital or (ii) reduce your population growth"

Corolary: to foster investment/savings create good institutional framework, etc…

But in line with a **place-based philosophy we should question the homogeneity** of the effects of the drivers of growth.

Example: What if you do not need to invest more in a given capital because of increasing investment there has no returns? That is to say, what if running:

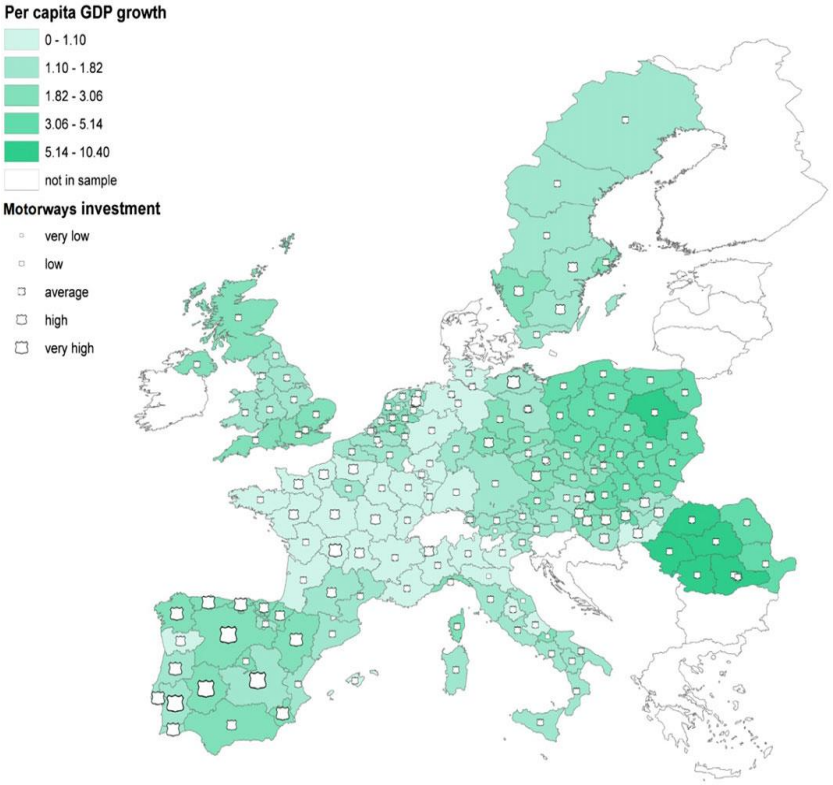
$$gy_i = \beta_{0i} + \beta_{1i} lnY_i + \beta_{2i} I_i + \beta_{3i} H_i + \beta_{3i} n_i + \varepsilon_i$$

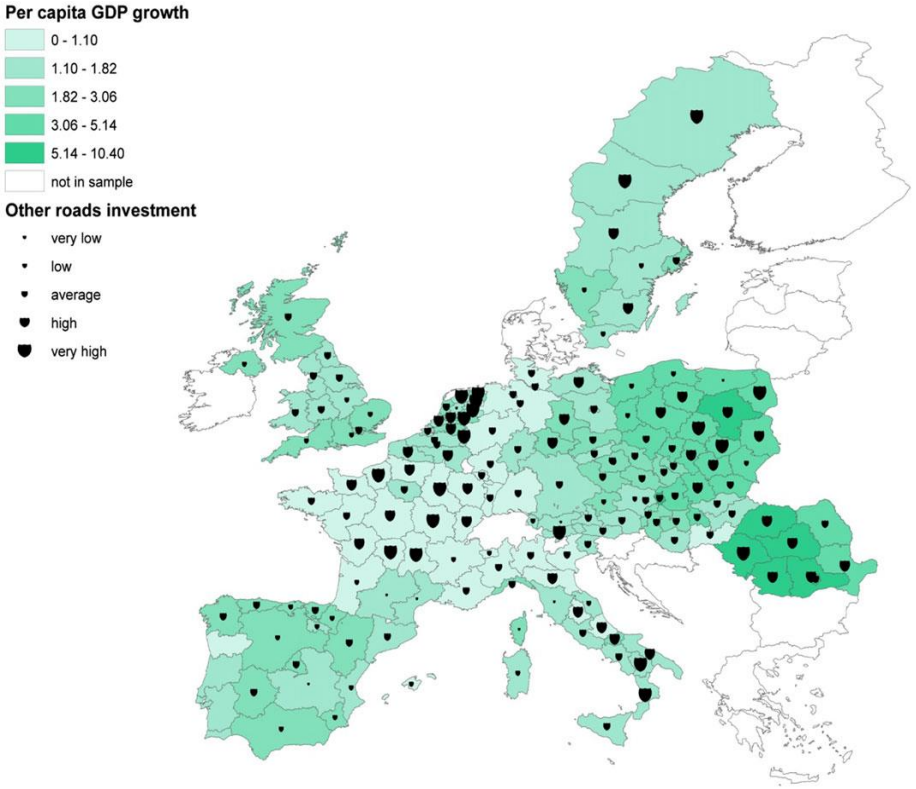
tells you $\beta_{2i} > 0$ for some regions and $\beta_{2i} < 0$ for some others?

Should you always recommend increasing physical capital investment?

Issues with the Neoclassical growth model

- Recent evidence shows that investment returns might be dependent on the quality of the government (i.e, low corruption)





Issues with the Neoclassical growth model

Crescenzi et al (2016)

	Full Sample		Less Developed Regions	
Dep. variable: Change of Log GDP	(1)	(2)	(3)	(4)
Lagged GDP	−0.0252**	−0.0901***	−0.0473***	−0.129***
	(0.0101)	(0.0140)	(0.0138)	(0.0218)
Investment in other roads	0.00102**	0.000607	0.00136	0.000401
	(0.000487)	(0.000476)	(0.00768)	(0.000497)
Quality of Government (QoG)	0.0235***	0.0246***	0.0628***	0.0595***
	(0.00484)	(0.00436)	(0.0109)	(0.00801)
Investment in other roads × QoG	0.00157*	0.00234***	0.00268**	0.00352***
	(0.000829)	(0.000873)	(0.0128)	(0.00118)
Spatial weight of other roads investment		0.00366**		0.00299
		(0.00155)		(0.00204)
Agricultural employment		−0.00352***		−0.00339***
		(0.000626)		(0.000834)
Patent applications		0.00534***		0.00753***
		(0.00180)		(0.00276)
Human capital		0.0136***		0.0420***
		(0.00512)		(0.0134)
Regional population	−4.46e−05***	−1.53e−05**	−3.56e−05*	5.04e−06
	(1.21e−05)	(7.52e−06)	(1.78e−05)	(8.77e−06)
Region dummies	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes
Observations	2,158	2,134	889	876
R^2	0.387	0.472	0.383	0.472
NUTS regions	161	161	66	66

Issues with the Neoclassical growth model

Rodriguez-Pose and Garcilazo (2015)

Dependent variable GDP pc growth	> €80 GMM-sys	> €100 GMM-sys	> €120 GMM-sys	> €150 GMM-sys
Cohesion expenditure pc	3.14e–05** (1.54e–05)	3.82e–05* (2.21e–05)	4.06e–05* (2.27e–05)	3.47e–05* (1.94e–05)
Quality of government	0.0151*** (0.00477)	0.0147*** (0.00463)	0.0142*** (0.00471)	0.0114** (0.00451)
CohesionExp × *QualityGov*	−5.38e–05** (2.44e–05)	−5.23e–05** (2.25e–05)	−5.11e–05** (2.25e–05)	−3.87e–05** (1.95e–05)
ln National Growth	0.352** (0.150)	0.430** (0.171)	0.273 (0.219)	0.309 (0.229)
Constant	0.0299*** (0.00876)	0.0240** (0.0112)	0.0276* (0.0152)	0.0274* (0.0155)
Time controls	Yes	Yes	Yes	Yes
Number of observations	463	410	361	307
Number of countries	11	10	10	10
p-value of AR(4) test	0.189	0.145	0.198	0.231
p-value of Hansen test	0.372	0.427	0.602	0.769
Number of instruments	55	55	55	55

Notes: Different thresholds (GMM-sys). Robust standard errors are given in parentheses.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

These two studies show that variation in a third factor "QOG", changes the effect of investing in infrastructure or that of the cohesion funds investment.

Therefore, in many circumstances we could have **heterogeneous effects**.

Geographically Weighted Regressions

A key assumption on spatial modelling is whether or not the parameters are homogenous o heterogeneous (across space, time, etc)

If we are interested in accounting for potential **spatial heterogeneity in parameters** we can use **GWR** (Fotheringham et al., 2002) as this modeling technique allows local variation int he parameters.

GWR has been **used** primarily **for exploratory data analysis**, rather than hypothesis testing

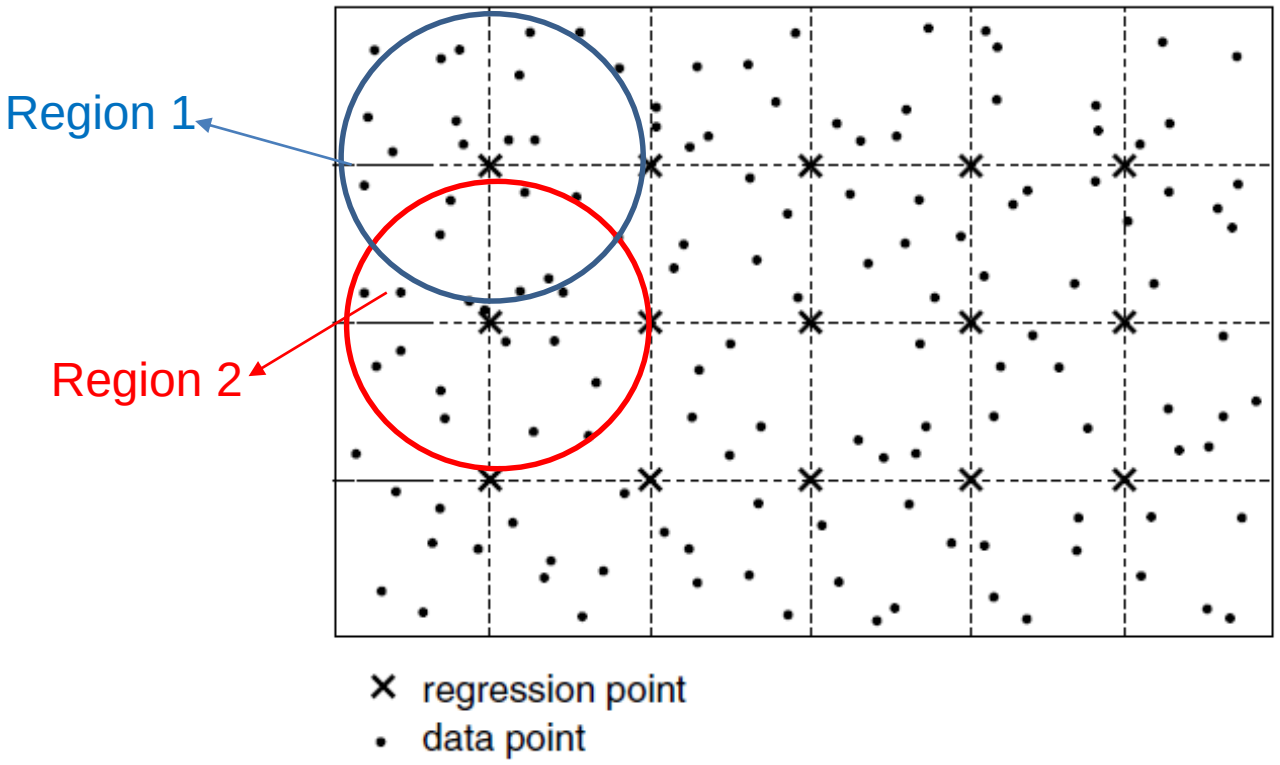
The **basic form of the GW** regression model if we have *m* explanatory variables:

$$y_i = \beta_{0i}(u_i, v_i) + \sum_{k=1}^{m} \beta_{ik}(u_i, v_i)x_{ik} + e_i$$

Where y_i is the dependent variable at location i, x_{ik} is the value of the regressor *k* at location *i*, β_{i0} is the intercept parameter at location *i*, β_{ik} is the local regression coefficient for the kth independent variable at location *i* and e_i is the random error term at location *i*. (u_i,, v_i) denotes the coordinates of the *i-th* point in space

Geographically Weighted Regressions

- You can think about **GWR as a *"spatially moving window regression"***
- **A region can be defined as the four cells around each regression point**
- The **regression model is then calibrated on all data that lie within the region** described around a regression point and the process is repeated for all regression points



Geographically Weighted Regressions

- In the spatial **moving window example** presented above:
  - **a region was described around a regression point and** all the data **points within this region** or window were then **used to calibrate** a model.
  - This process was repeated for all regression points

- GWR works in the same way except that:
  - **Each data point is weighted by its distance from the regression point**
  - Hence, **data points closer to the regression point are weighted more** heavily in the local regression than are data points farther away
  - For a given regression point, **the weight of a data point is at a maximum when it shares the same location as the regression point**. This weight decreases continuously as the distance between the two points increases

Geographically Weighted Regressions

- As data are geographically weighted, **nearer observations have more influence in estimating the local set of regression coefficients** than observations farther away.

- **The model measures the inherent relationships around each regression point** *i*, where each set of regressors is estimated by a weighted least squares approach. The matrix expression for this estimation is:

$$\beta_i = (X'W(u_i, v_i)X)^{-1}\, X'W(u_i, v_i)\mathrm{y}$$

Notice that in this modeling approach $\beta_i = [\beta_{i,0}\ \beta_{i,1}\ \ldots, \beta_{i,m}]'$ is a 1 x k vector, so β is a *n x k* matrix and $W(u_i, v_i)$ is a diagonal *n x n* matrix denoting the geographic observed data for regression point at location (u_i, v_i).

This **W** matrix is determined by some kernel function.

Geographically Weighted Regressions

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0(u_1,v_1) & \beta_1(u_1,v_1) & \dots & \beta_k(u_1,v_1) \\ \beta_0(u_2,v_2) & \beta_1(u_2,v_2) & \dots & \beta_k(u_2,v_2) \\ \dots & \dots & \dots & \dots \\ \beta_0(u_n,v_n) & \beta_1(u_n,v_n) & \dots & \beta_k(u_n,v_n) \end{bmatrix}$$

The parameters in each row of the above matrix are estimated by

$$\hat{\boldsymbol{\beta}}(i) = (X^{\mathrm{T}} W(i) X)^{-1} X^{\mathrm{T}} W(i) Y$$

where i represents a row of the matrix in (2.10) and $W(i)$ is an n by n spatial weighting matrix of the form

$$W(i) = \begin{bmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ . & . & & . \\ . & . & & . \\ 0 & 0 & \dots & w_{in} \end{bmatrix}$$

Geographically Weighted Regressions

Different kernels for determining weights

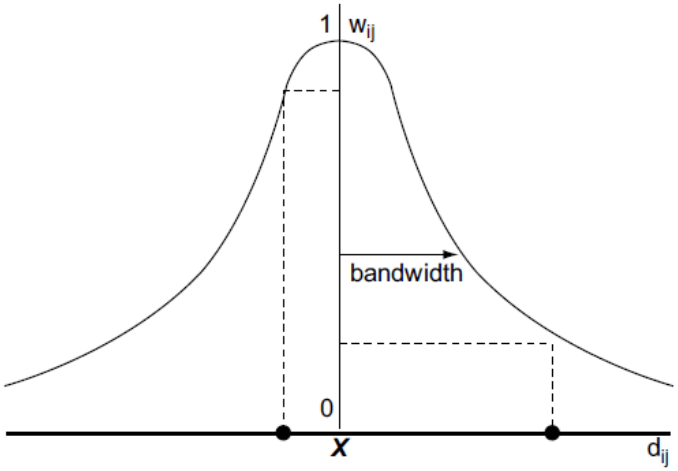
Global Model	$w_{ij} = 1$
Gaussian	$w_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{b}\right)^2\right)$
Exponential	$w_{ij} = \exp\left(-\frac{\lvert d_{ij}\rvert}{b}\right)$
Box-car	$w_{ij} = \begin{cases} 1 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$
Bi-square	$w_{ij} = \begin{cases} (1-(d_{ij}/b)^2)^2 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$
Tri-cube	$w_{ij} = \begin{cases} (1-(\lvert d_{ij}\rvert/b)^3)^3 & \text{if } \lvert d_{ij}\rvert < b, \\ 0 & \text{otherwise} \end{cases}$

"While the specific weighting function does not alter substantially the results the bandwitdh *"b"* matters a lot"

Goodness of fit metrics or selection procedures should be applied to be sure on the quality of the results obtained

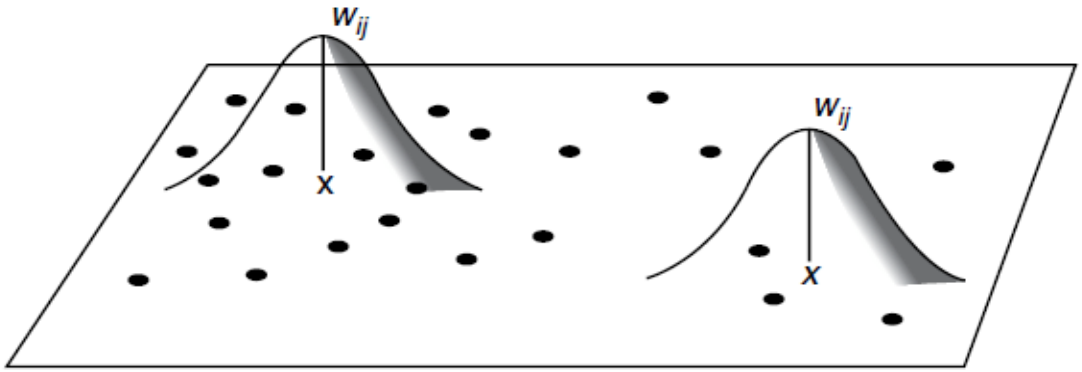
Geographically Weighted Regressions

- Example of kernel function



X regression point w_{ij} is the weight of data point *j* at regression point *i*

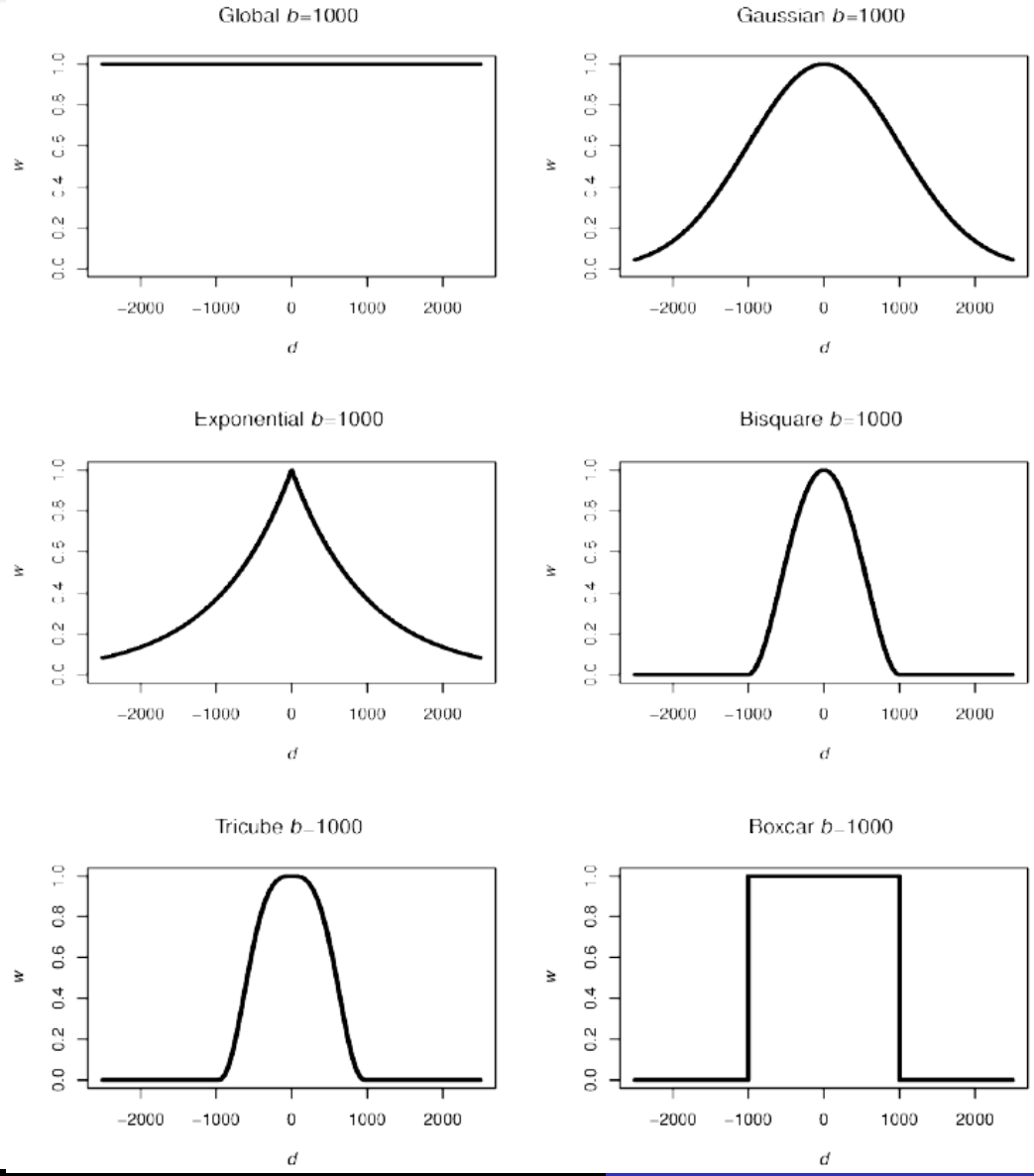
● data point d_{ij} is the distance between regression point *i* and data point *j*



x regression point

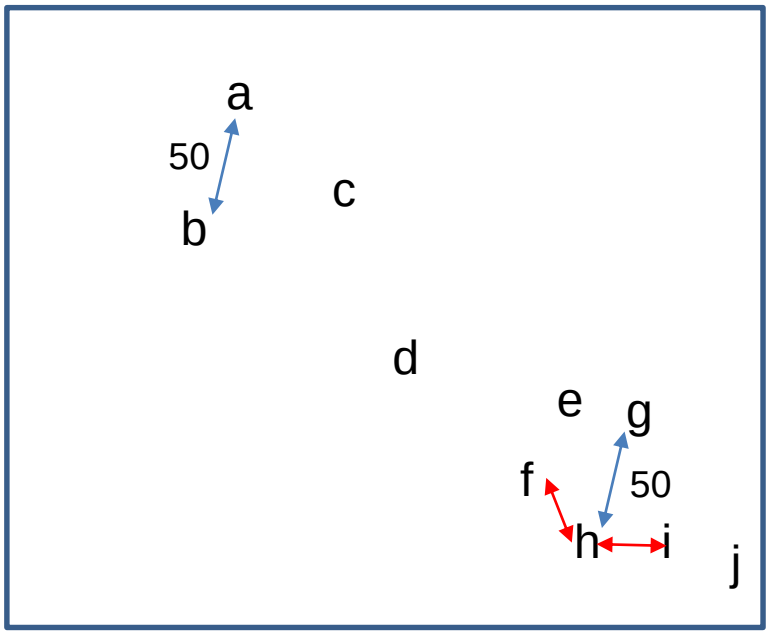
● data point

Geographically Weighted Regressions



Geographically Weighted Regressions

- **The issue of fixed vs adaptative kernels**
  - **Fixed kernels in low-density data point areas** give **less weight to nearby observations** to what they should be attributing **ideal**ly
  - **Fixed kernels in high-density data point areas** give **excessive weight to nearby** observations to what they should be attributing ideally



Fixed kernel: the influence of point b on point a is equal to the impact of h on g

Adaptative kernel: the influence of b on a is higher than that of g on h as far as a has les information surrounding it

Geographically Weighted Regressions

- The distance between points *a* and *b* is the same to that of *g* and *h* but *a* is in a low-density area with very few data points to perform inference whereas *g* is a high-density area

- Let bw denote the bandwidth for notational convenience. For the previous example, a fixed exponential kernel will attribute the following weights:

$$w_{a,b} = e^{-\frac{d_{a,b}}{bw}} = e^{-\frac{50}{bw}} \leftrightarrow w_{h,g} = e^{-\frac{d_{h,g}}{bw}} = e^{-\frac{50}{bw}}$$

Ideally, we would like the weights to be dependent and distributed on the basis of the data availability.

Some approaches to generate **spatially varying kernels**:

1) Replace d in the numerator by the rank in the distribution of distances holding bw fixed
2) Couple it with a kernel with a fixed common number of nearest neighbors
3) Couple a kernel with metrics of data density based on spheres of influence

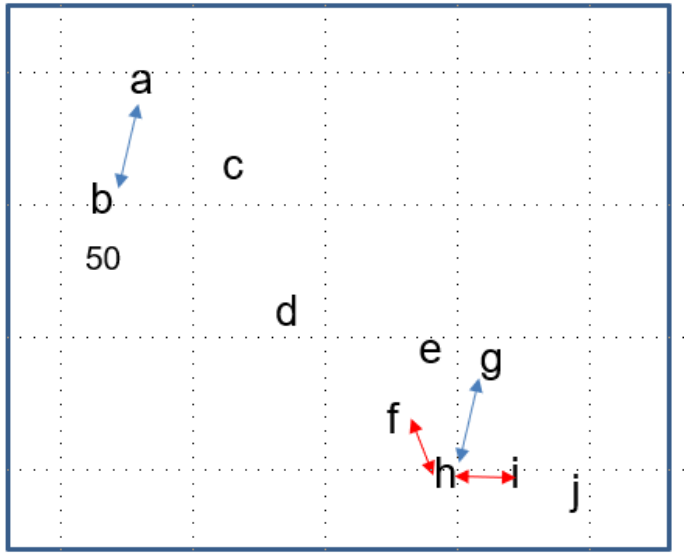
Geographically Weighted Regressions

- Example of Ranks

For point a: $d_{a,b}(1) < d_{a,c}(2) < d_{a,c}(3) < d_{a,e}(4) \ldots < d_{a,j}(9)$

For point h: $d_{h,i}(1) < d_{h,f}(2) < d_{h,g}(3) < d_{h,e}(4) \ldots < d_{h,a}(9)$

$$w_{a,b} = e^{-\frac{d_{a,b}}{bw}} = e^{-\frac{1}{bw}} > w_{h,g} = e^{-\frac{d_{h,g}}{bw}} = e^{-\frac{3}{bw}}$$



Geographically Weighted Regressions

- Hint (1): As the bandwidth has an effect on the results you should perform a model selection using some goodness of fit criteria (i.e, MSE, MAPE, etc)

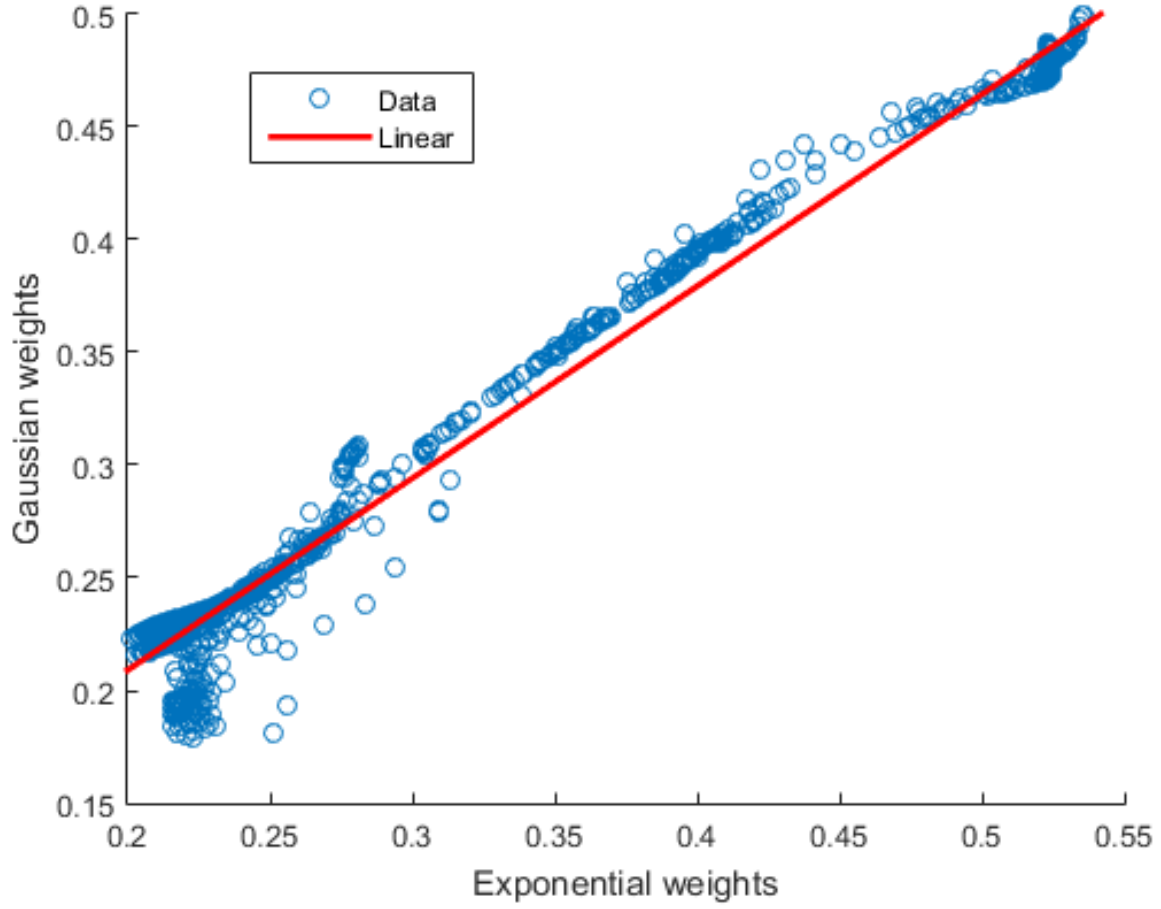
Table 1: GWR Model Selection.

	Exponential Weights		Gaussian Weights	
Kernel	RMSE	R-squared	RMSE	R-squared
Optimal CV bandwidth	0.05110	0.30100	0.05150	0.28770
Adaptive knn =20	0.05942	0.21776	0.05614	0.17968
Adaptive knn =30	0.05836	0.21850	0.06405	0.18344
Adaptive knn =40	0.05457	0.21698	0.05594	0.18428
Adaptive knn =50	0.05433	0.21423	0.05366	0.18297
Adaptive knn =60	0.05450	0.21226	0.05379	0.18188
Adaptive knn =70	0.05313	0.21000	0.05369	0.18013
Adaptive knn =80	0.05317	0.20925	0.05372	0.17930
Adaptive knn =90	0.05317	0.20963	0.05372	0.17927
Adaptive knn =100	0.05319	0.20883	0.05376	0.17801

Notes: These results correspond to averages on 1,000 GWR model draws with expected model size equal to 10.

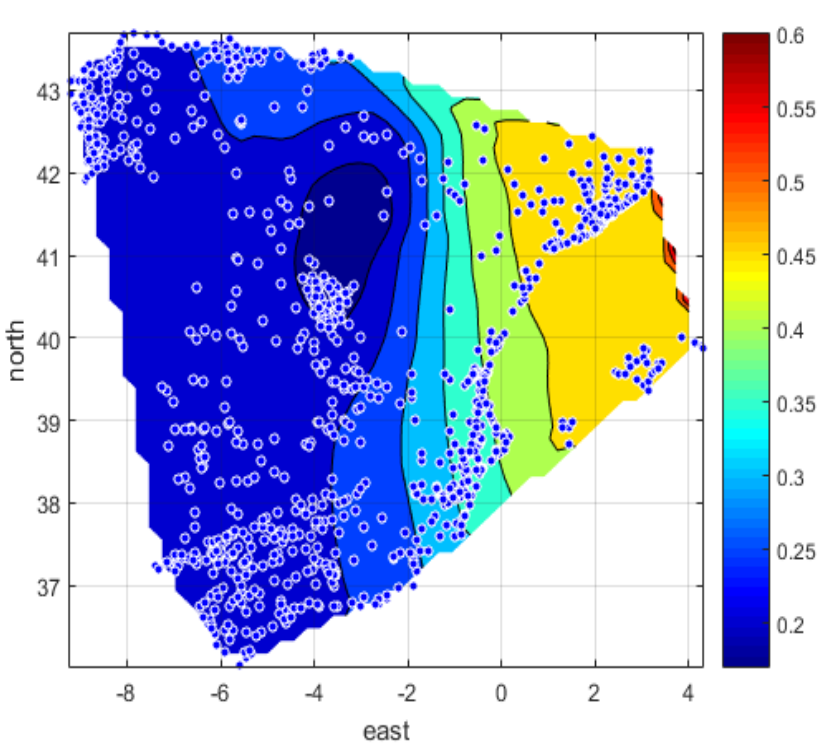
Geographically Weighted Regressions

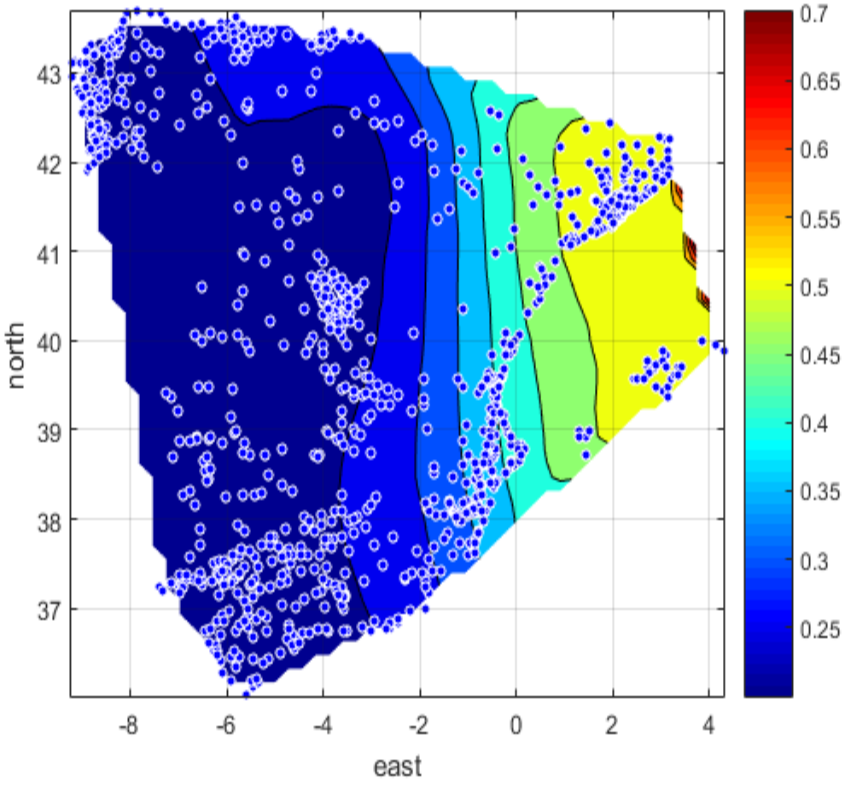
- Hint (2): Check if estimates differ substantially when using different kernel weighting functions



Geographically Weighted Regressions

Example: creative class income inequality effects on Spanish cities using Gaussian and Exponential weights





Practice GWR

- **Goals**
  - Estimate the following GWR with the data of regional economic growth with, the logarithm of initial income per capita (lnY_i), the share of physical capital investment in GDP (I_i) the share of population with tertiary education and (H) and the population growth rate in the dataset *GWR_data.mat*

$$gY_i = \beta_{0i}(u_i, v_i) + \beta_{i1}(u_i, v_i) lnY_i + \beta_{i2}(u_i, v_i) I_i + \beta_{i3}(u_i, v_i) H_i + \beta_{i4}(u_i, v_i) POPG_i + e_i$$

  - Use the function *spatial_stationarity.m* to analyze the significance of spatial instability in the parameters
  - Plot the estimated parameters, also against latittude and longitude, and then with respect both in one shot with a 3d scatter.
  - Compare the fit of the GWR with respect a traditional OLS regression
  - Check the robustness of the results using different kernels: Gaussian, exponential, etc

- Switch to Matlab “Tutorial4_GWR_Growth.m”

Practice GWR

A **first issue when using GWR** is to check if the **model** describes the data significantly better than an ordinary linear regression model. As can be observed, in this example, GWR provides a better **fit** to the regional growth data than non-spatial analysis by looking at the R-squared

```
% run gwr and compare model fit with that of ols
y=GY;
x=[INIT INVEST EDUC POPG];
[n,k]=size(x);
res = gwr(y,[ones(n,1) x],xc,yc,info);
stats=spatial_stationariy_gwr(y,x,xc,yc);
res_ols=OLS_demo(y,[ones(n,1) x]);
res.rsqr
res_ols.rsqr
```

```
res.rsqr
res_ols.rsqr

ans =

    0.9215


ans =

    0.7289
```

Practice GWR

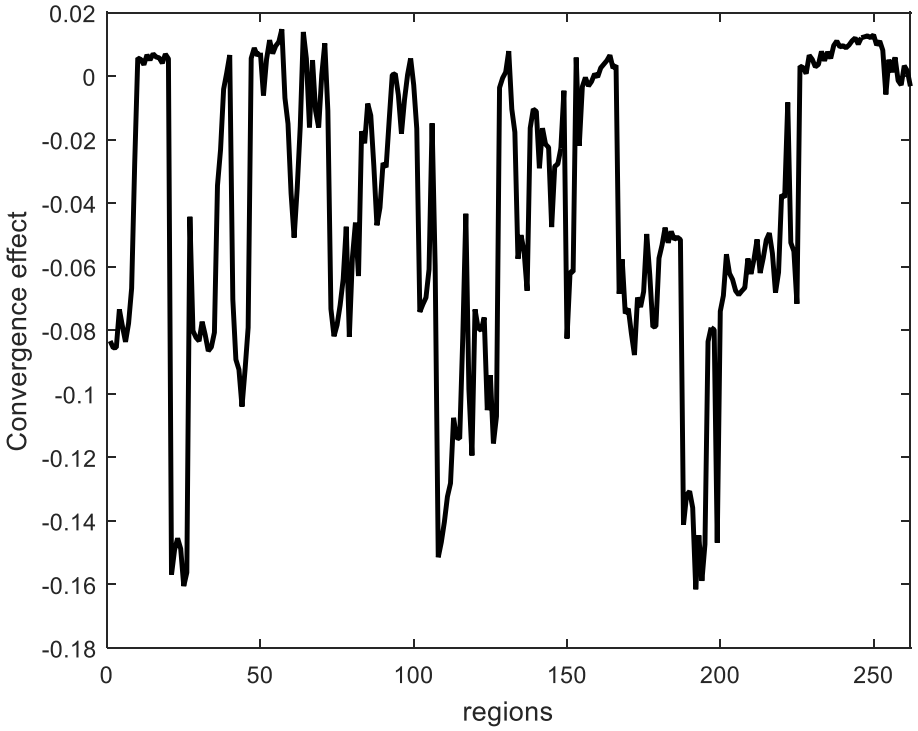
A **second important issue in GWR** is that of the **stationarity** of the parameters. Does the set of parameters β_{ik} exhibit significant variation over the study region?

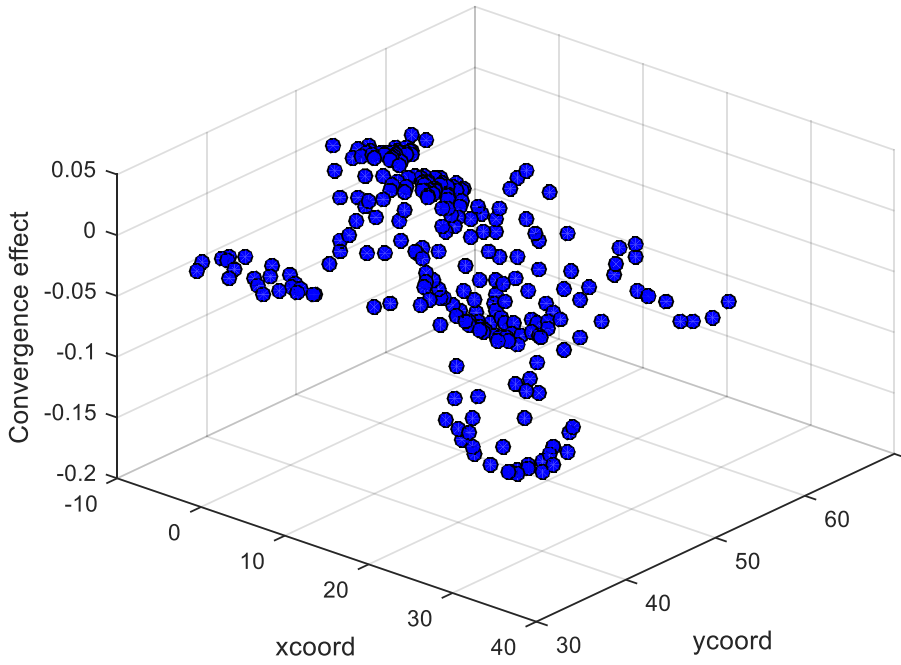
```
>> % check spatial stationarity
east=xc;
north=yc;
stats = spatial_stationariy_gwr(y,x,east,north);
>> stats

stats =

         F: [4x1 double]
    Kansfo: [4x1 double]

>> stats.Kansfo

ans =

    0.0000
    0.0328
    0.2566
    0.0068
```

We find evidence of non-stationarity in (1) initial income, (2) investment and (3) population growth effects. According to this, using a homogeneous parameter LRM would mask true "heterogeneous and place-specific effects".

Practice GWR

- **The effect of initial income**

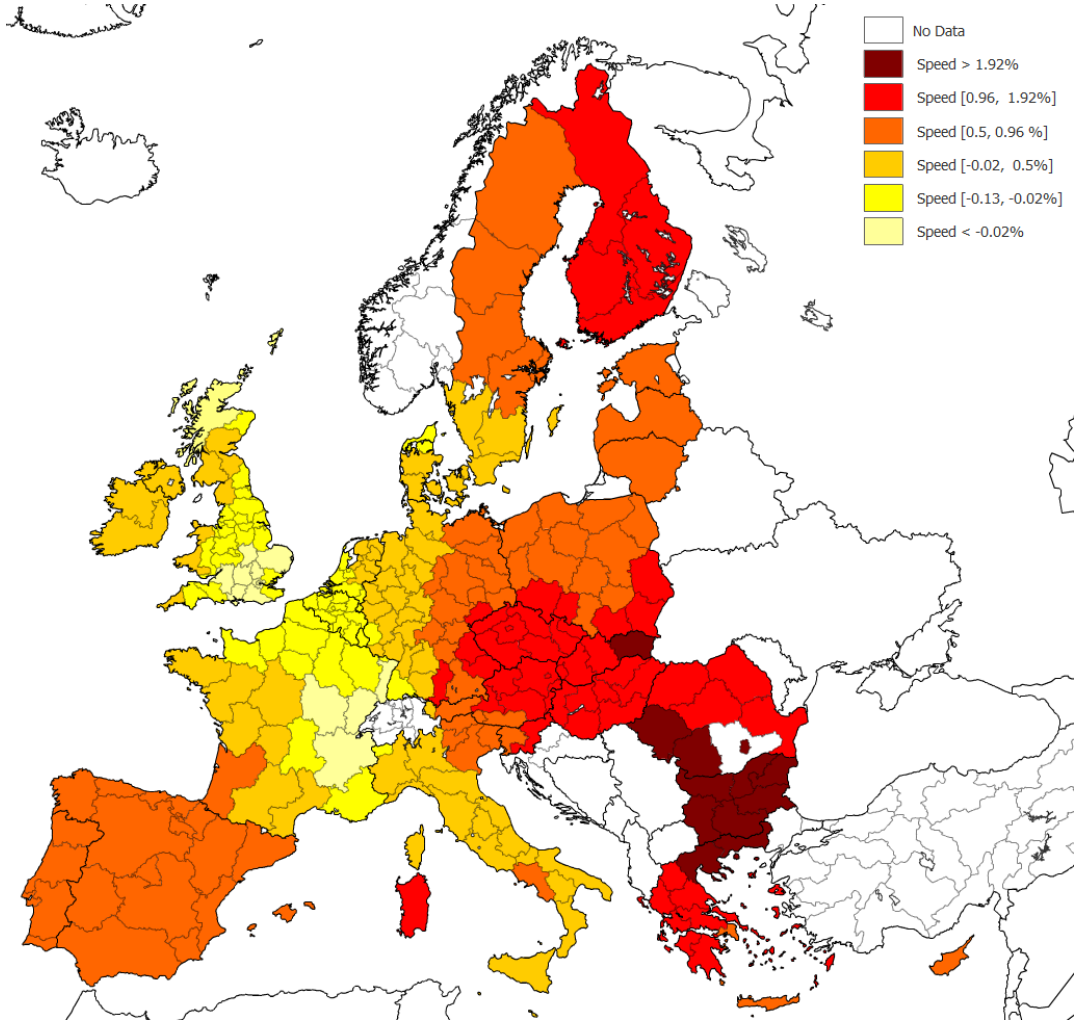




As observed the effect is negative which suggests convergence

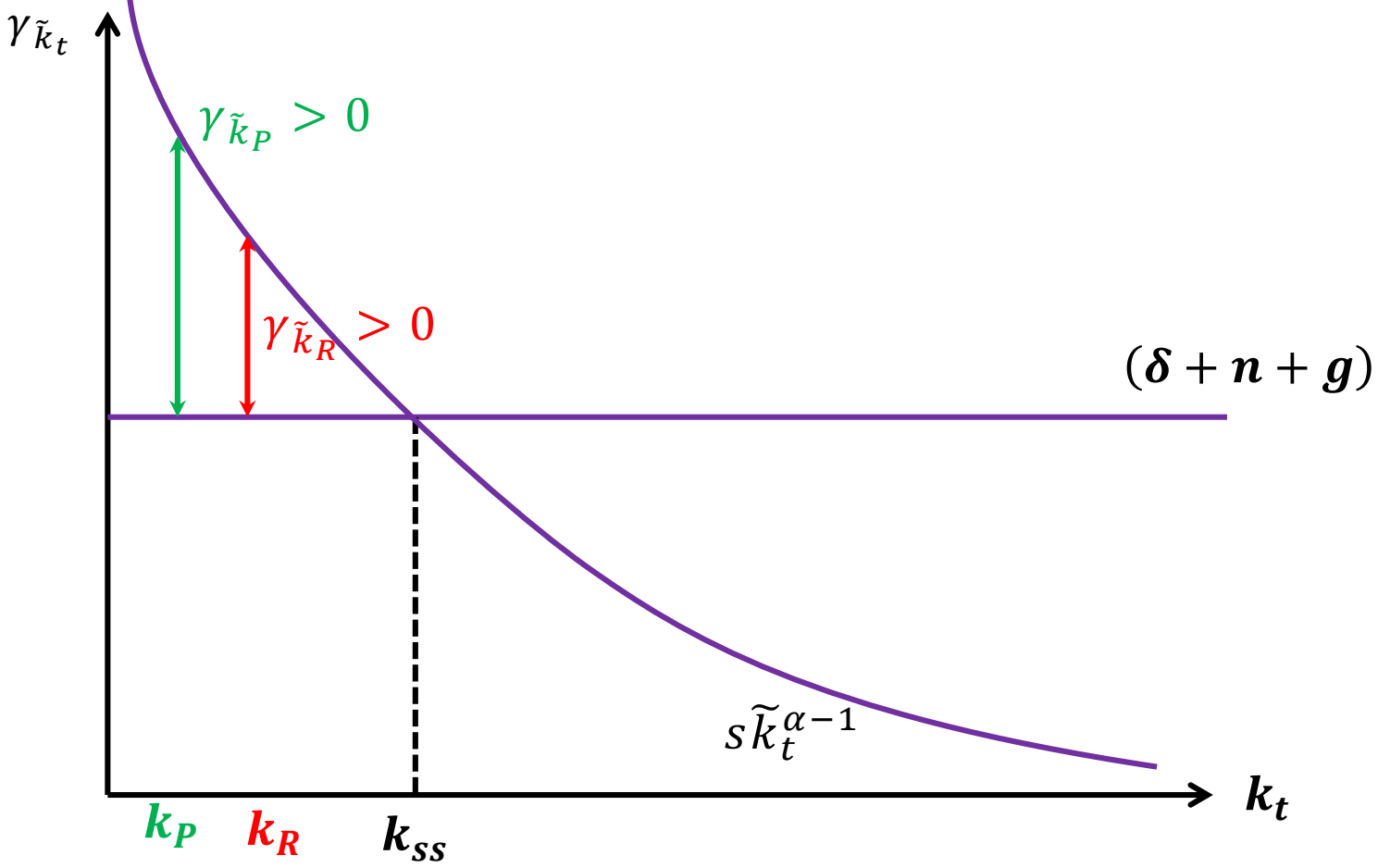
Practice GWR

- **The effect of initial income**



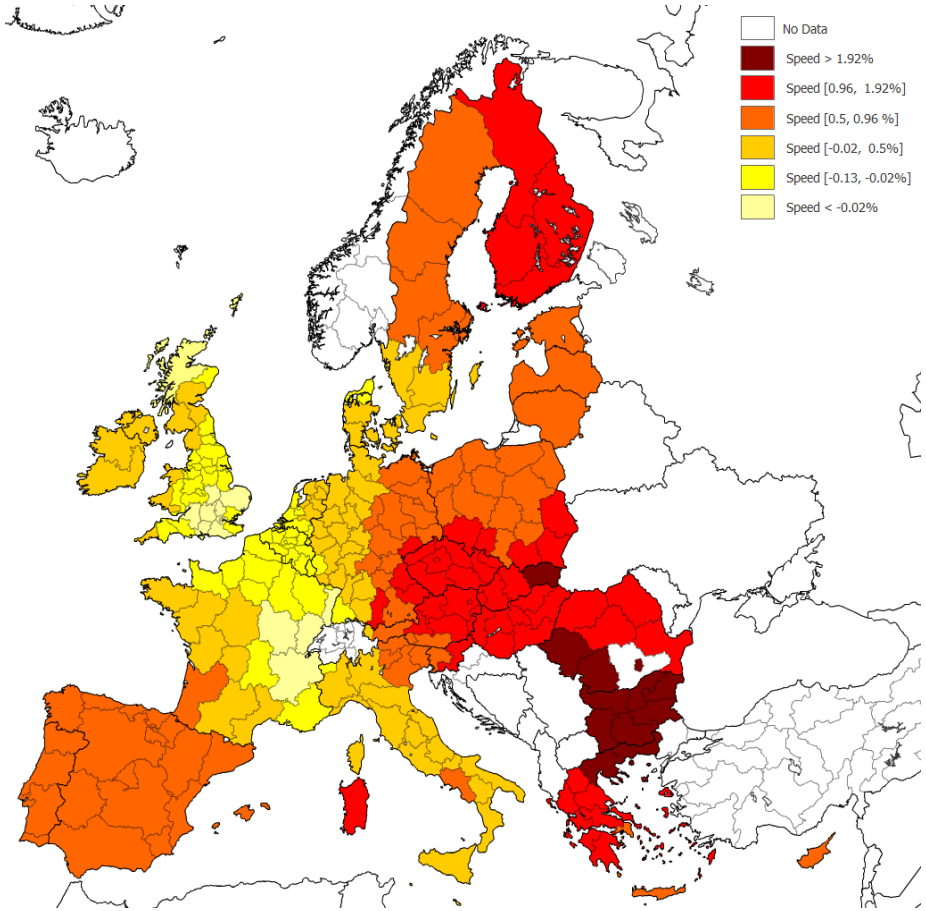
Practice GWR

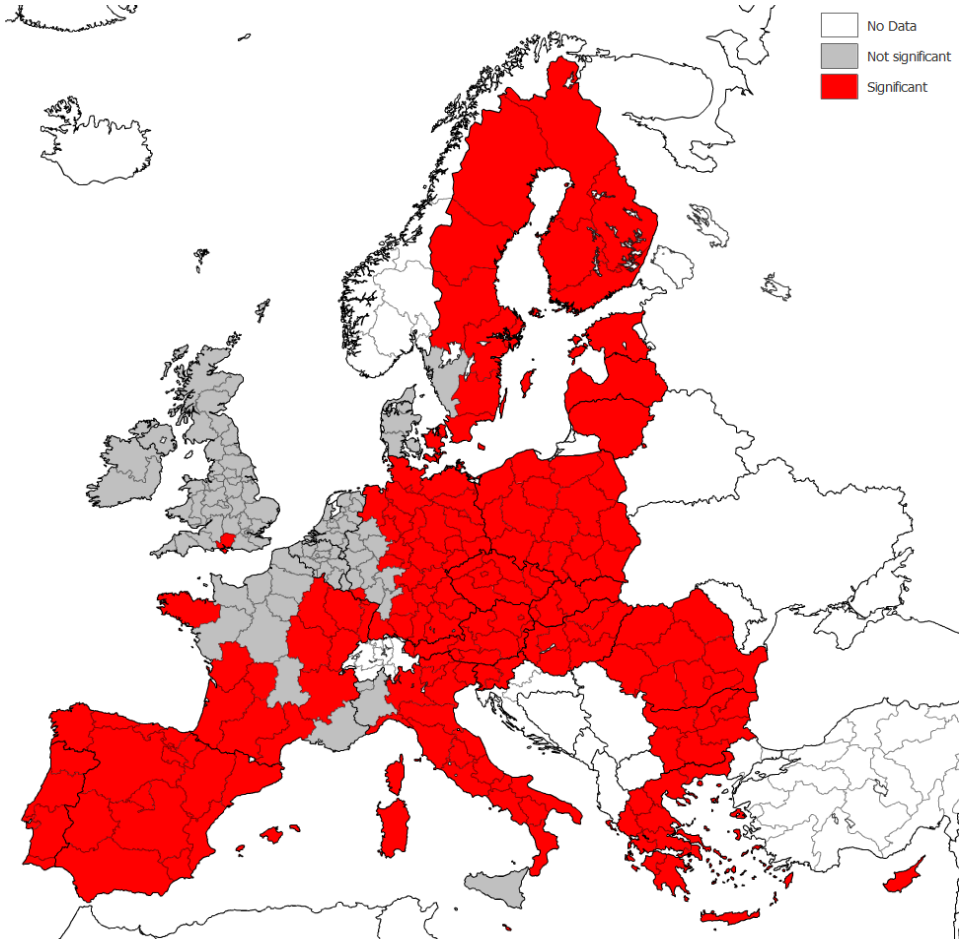
In the Neoclassical growth model poorer economies grow faster since they are far from their steady state (which produces convergence of income)



Practice GWR

- **The effect of initial income**





Practice GWR

- **The effect of initial income**

Eastern European regions from Bulgary, Romania, Czech Republic, Hungary and other peripheric economies from Greece experienced a **fast convergence** rate (i.e, a strong negative effect of income and subsequent growth)

Spain, Portuguese and Polish regions also converged towards their steady state but more slowly than Eastern ones.

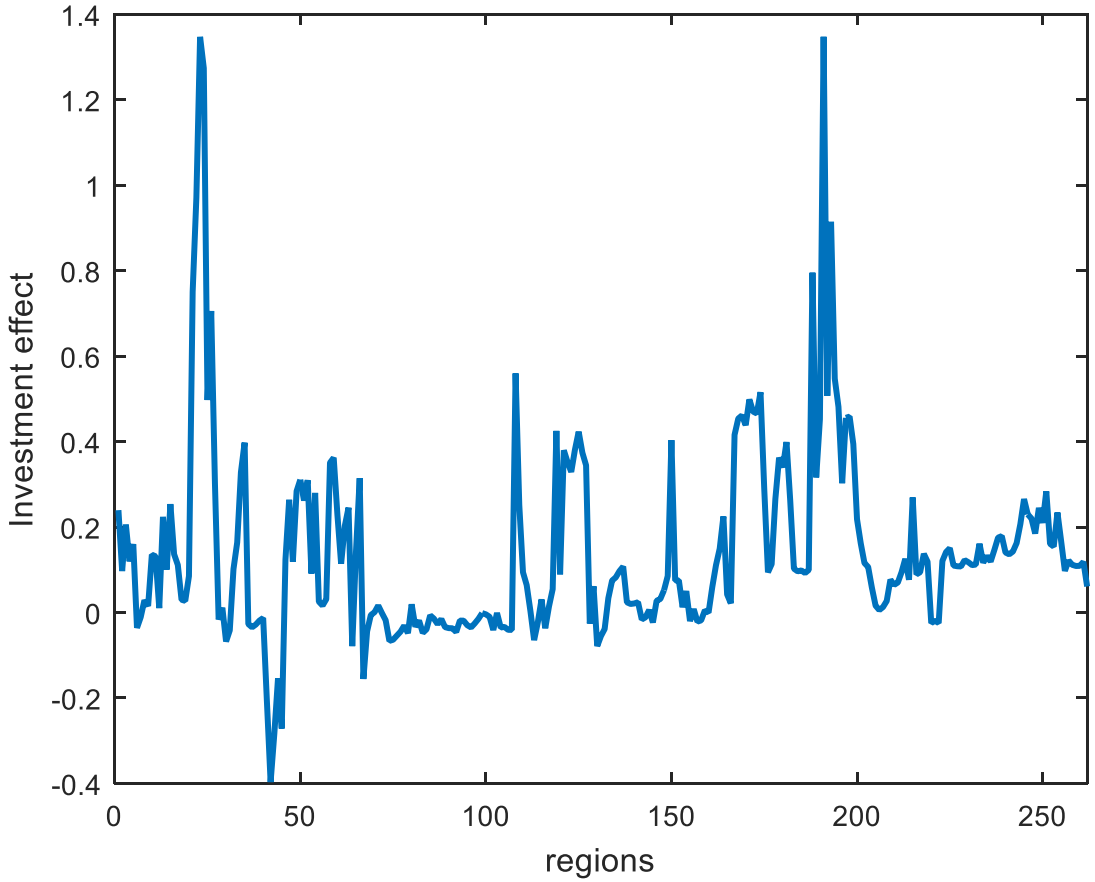
Other regions from countries such as Sweden or Finland also experienced a medium convergence process (given their high initial levels of income, should we question the uniqueness of the SS?)

Divergence dynamics are observed in Eeast Germany, France or Uk, but they're not statistically significant. It is difficult to say if initial income has played a role in this set of regions.

In Italy, significant divergence dynamics are observed for most of their regions. Income levels correlate positively with subsequent growth rates.

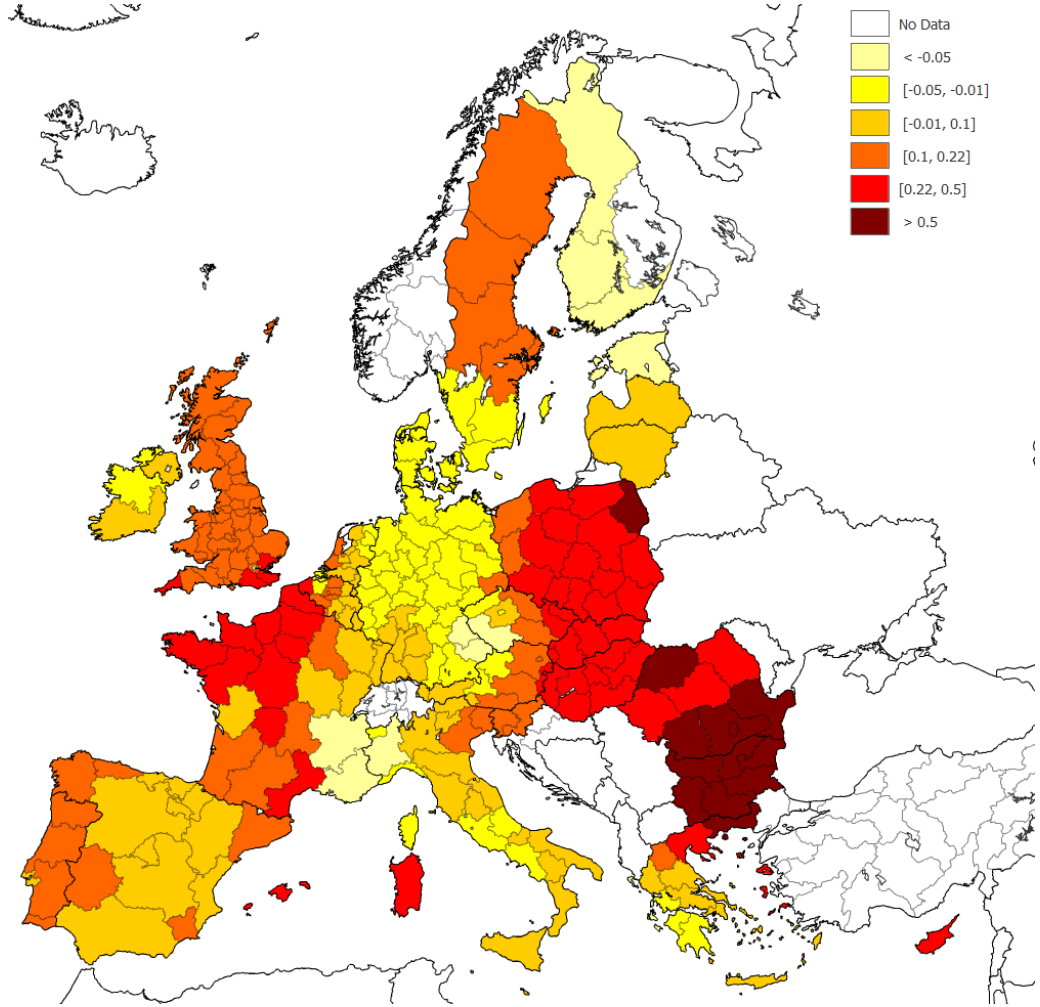
Practice GWR

- **The effect of physical capital investment**



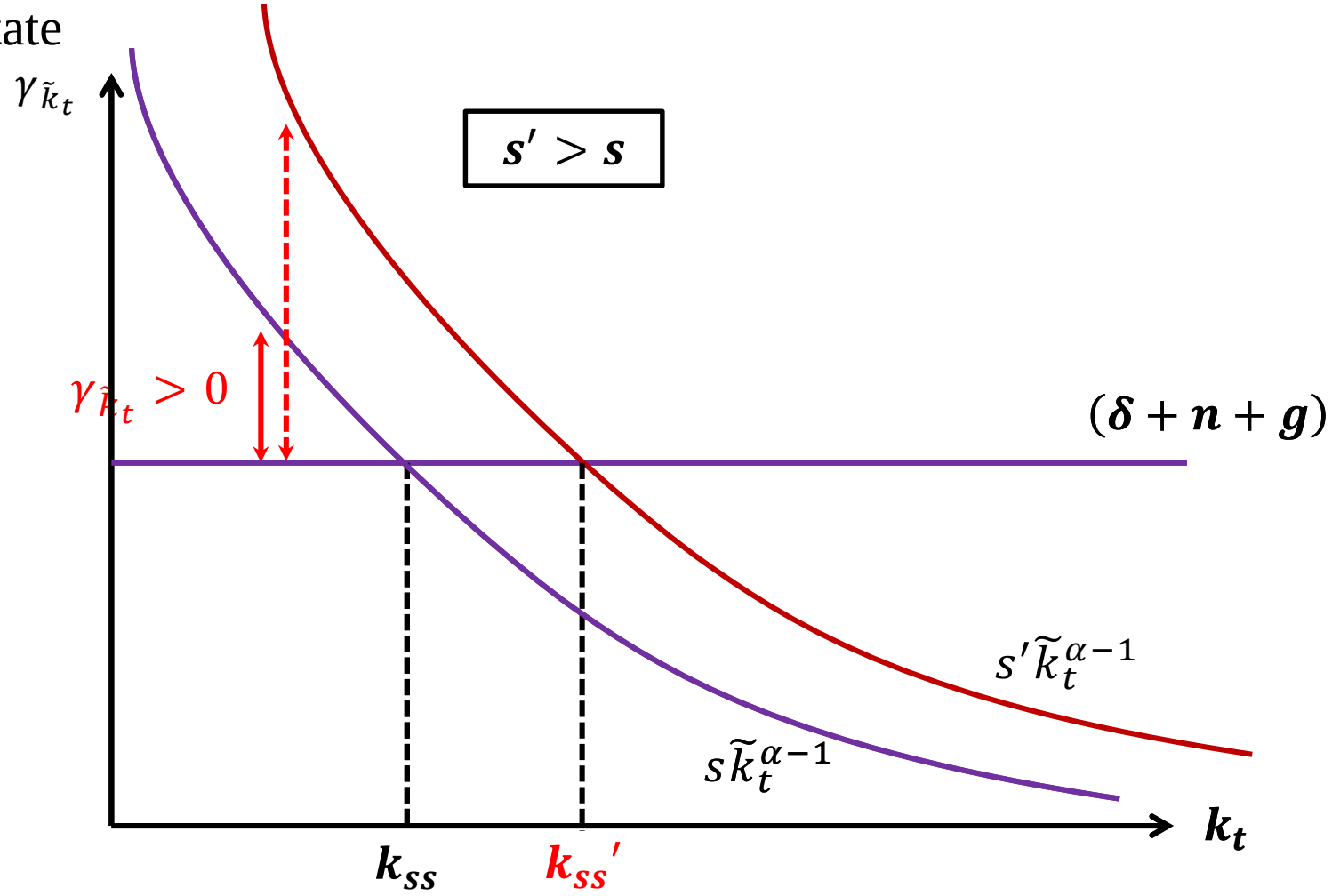
Practice GWR

- **The effect of physical capital investment**



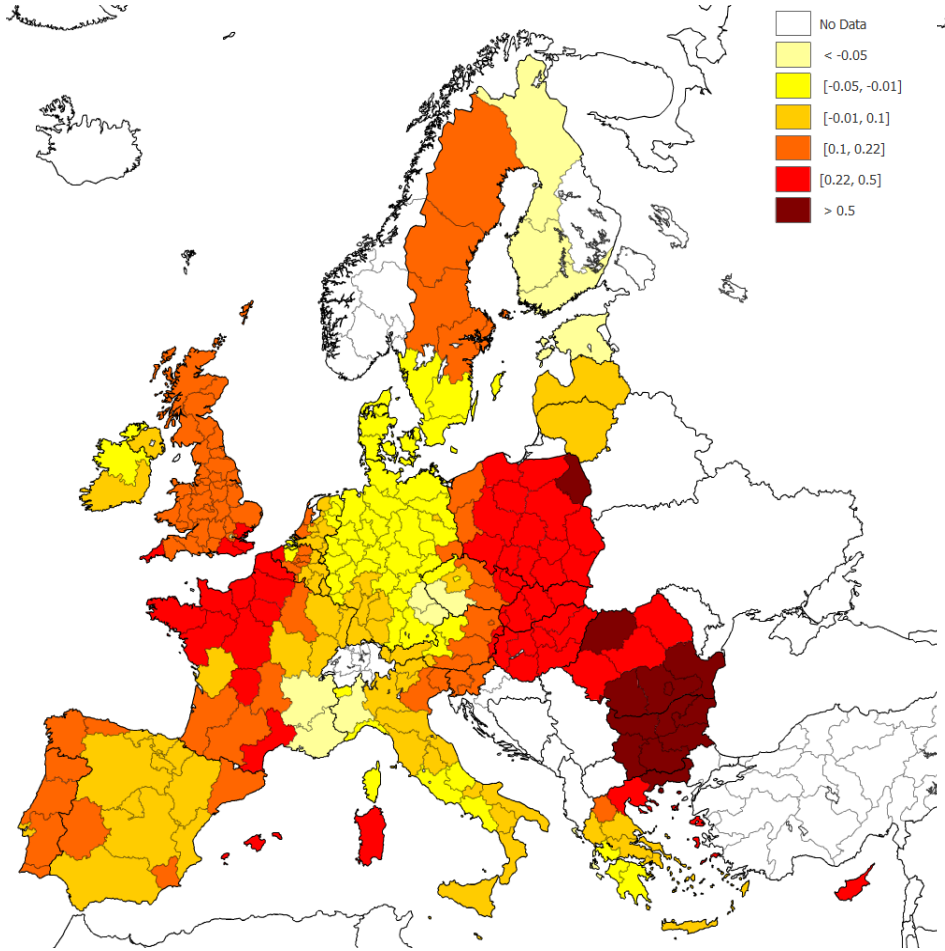
Practice GWR

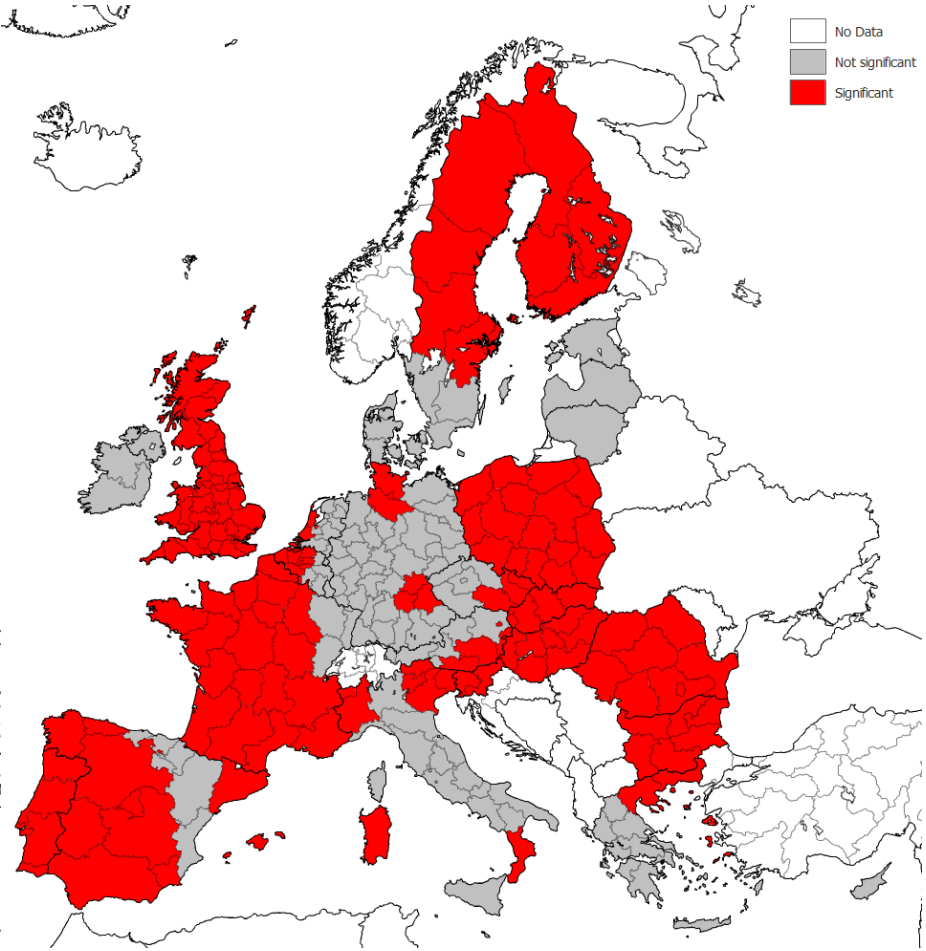
In the Neoclassical growth an increase in the investment rate accelerates growth in the transition towards the steady state and increases the steady state



Practice GWR

- **The effect of physical capital investment**



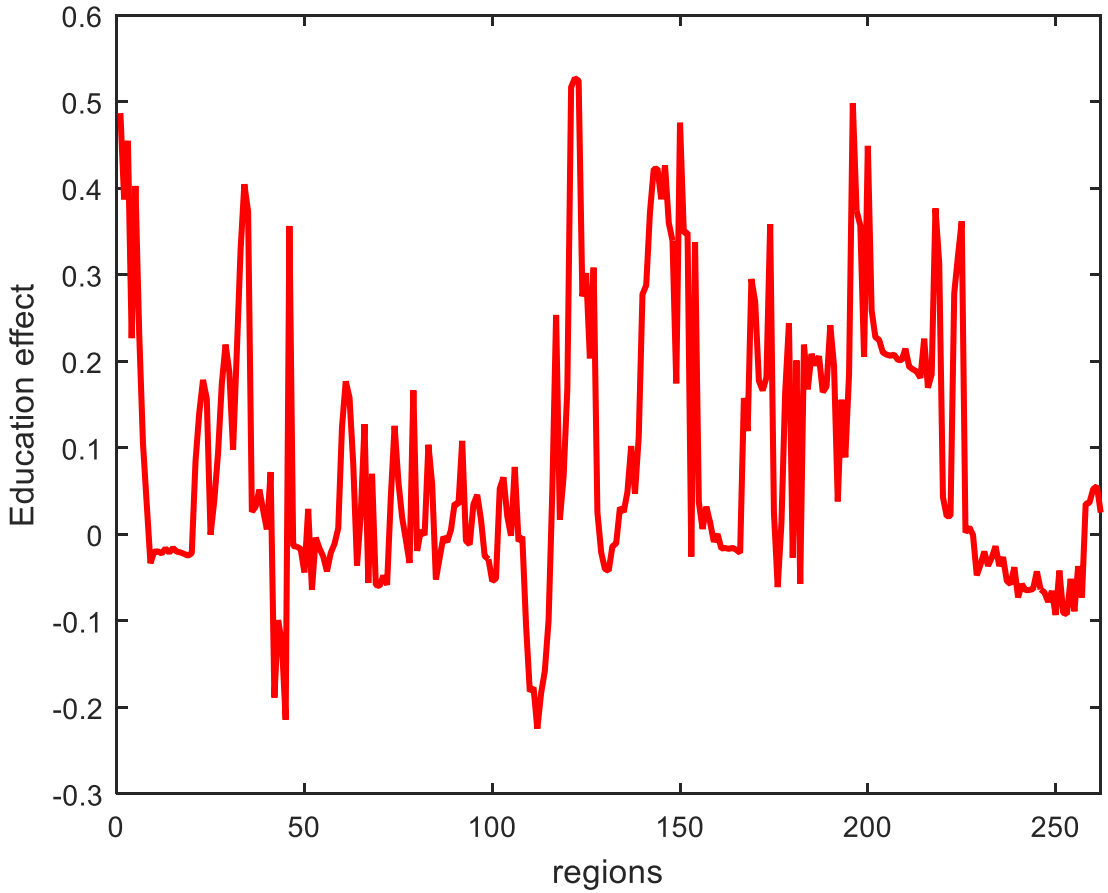


Practice GWR

- **The effect of physical capital investment**
  - For **most of the European regions the observed positive effects are in line with theory predictions**
  - In **Eastern European regions** from Poland, Czech Republic, Bulgary, Romania **very strong positive effects** which suggests that physical capital accumulation is a relevant driver of growth
  - High heterogeneity in Greece (as effects range from - to +)
  - In the west of Spain and Portugal (less developed regions) investment also exerts a positive effect on growth
  - In Italy, Germany, Netherlands, Austria or Ireland physical capital investment does not seem to be contributing to economic growth (not significant)
  - **Small negative** but significant **effects of physical capital investment in northern Economies** are counter-intuitive with theoretical predictions

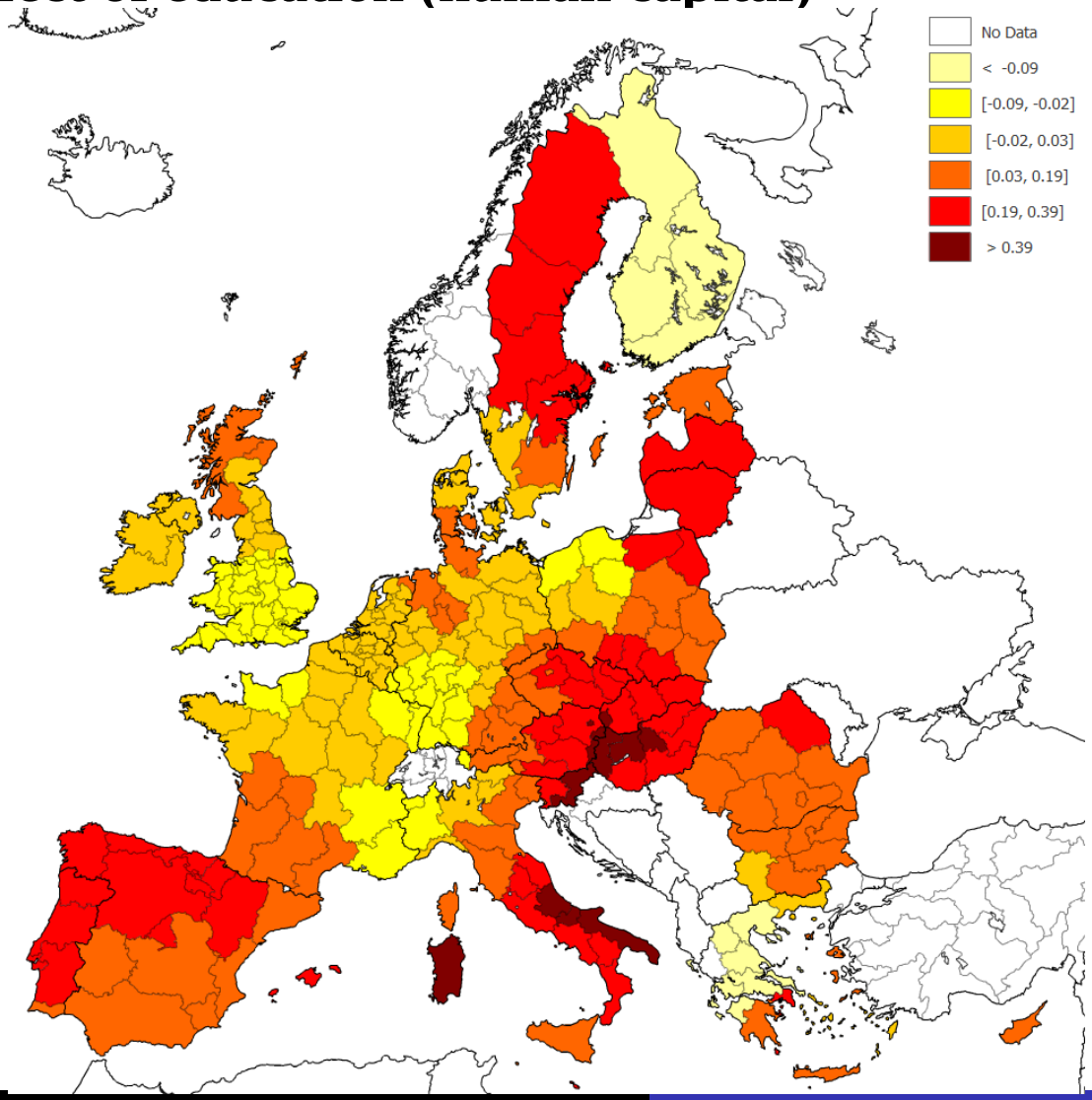
Practice GWR

- **The effect of education (human capital)**



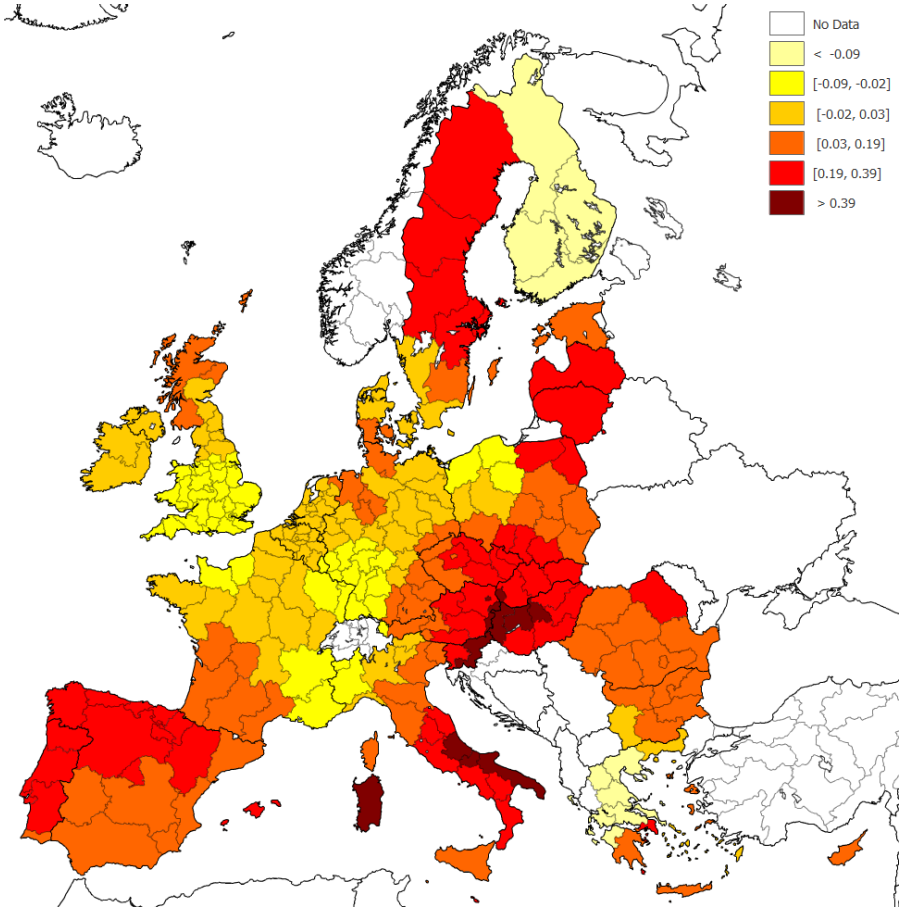
Practice GWR

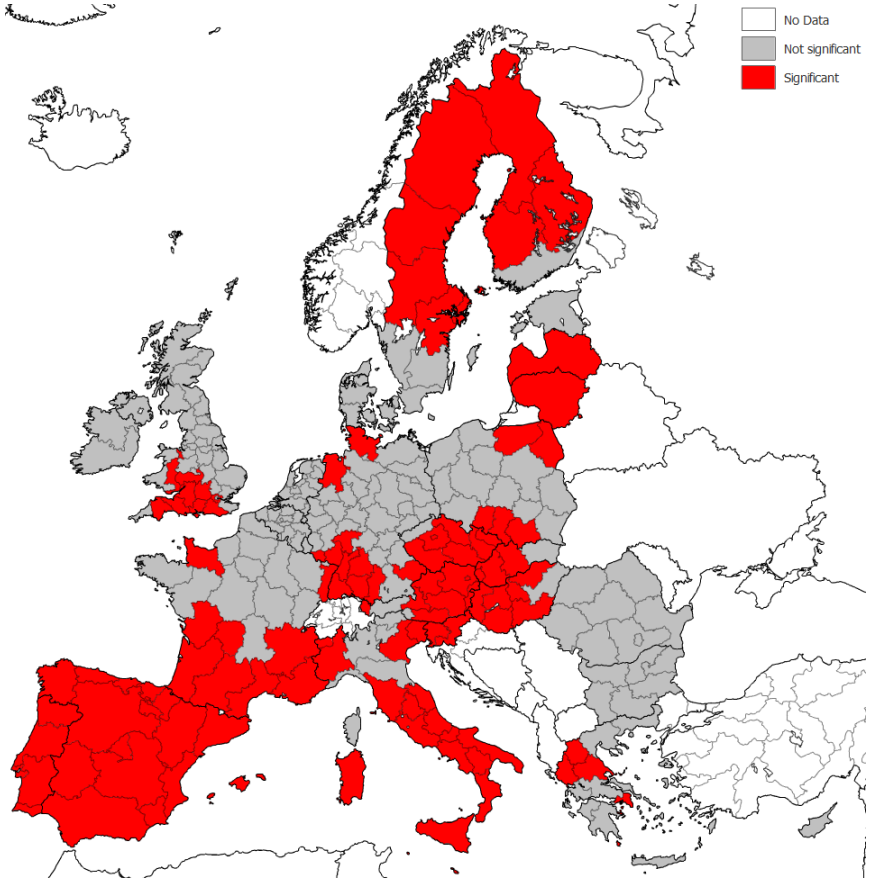
- **The effect of education (human capital)**



Practice GWR

- **The effect of human capital**





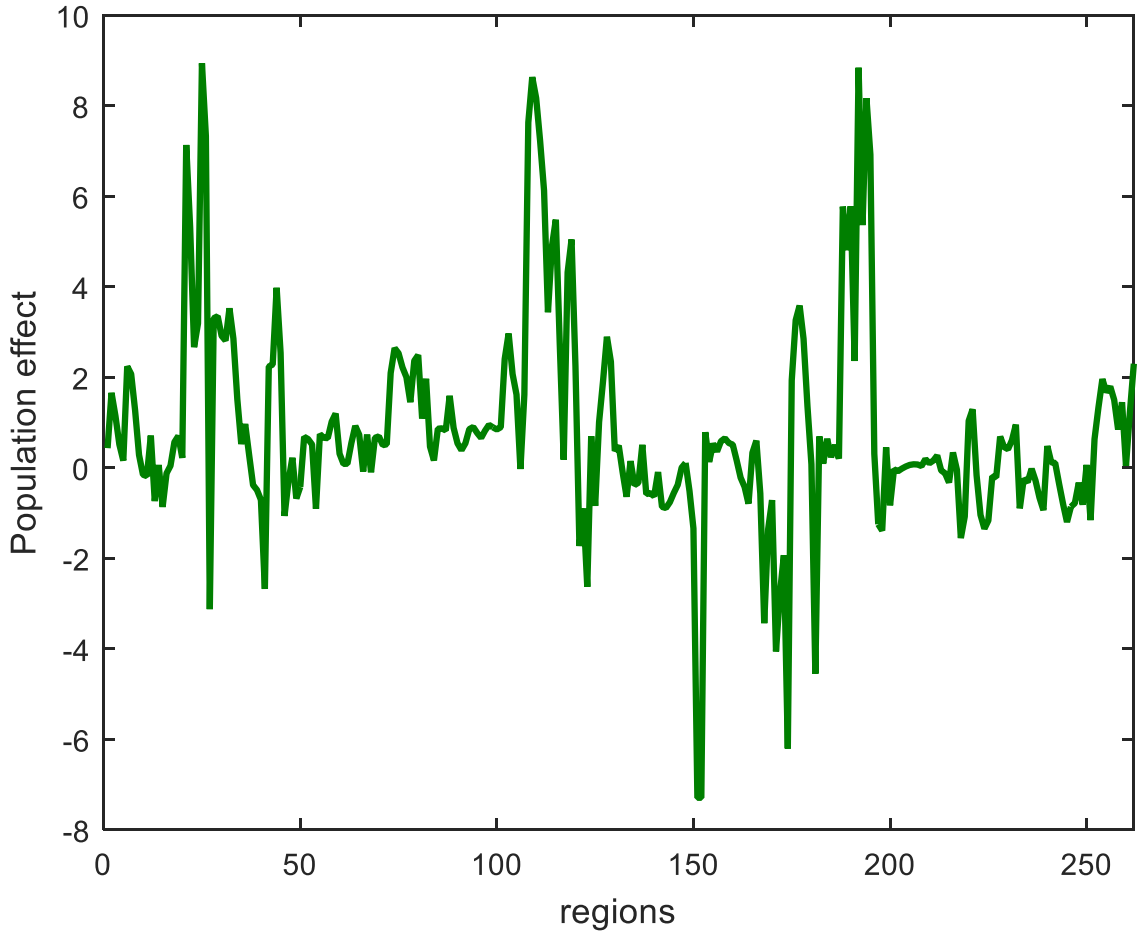
Practice GWR

- **The effect of human capital**
  - Stronger positive effects are observed in less developed regions in the South of Italy, north-west of Spain, South of France, Hungary and even in Austria
  - Negative growth returns to human capital in Greece and Finland (not significant in all regions)
  - North-east of France, South of UK and South of Germany also experience negative returns to education investment.
  - The strong positive effects of human capital accumulation in Eastern regions are not significant. Perhaps there is too much variability in the ability to absorbe knowledge.

Complicated picture for policy-makers regarding the effect of education!

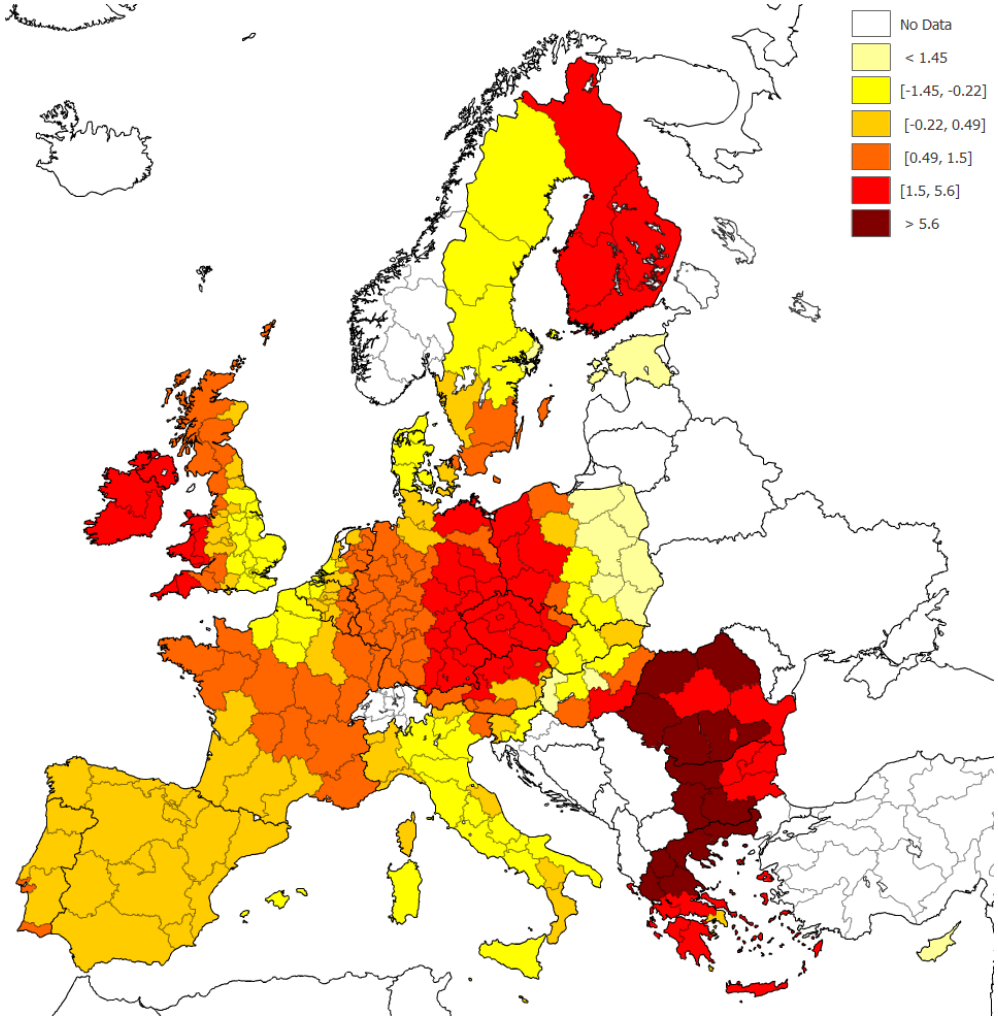
Practice GWR

- **The effect of population growth**



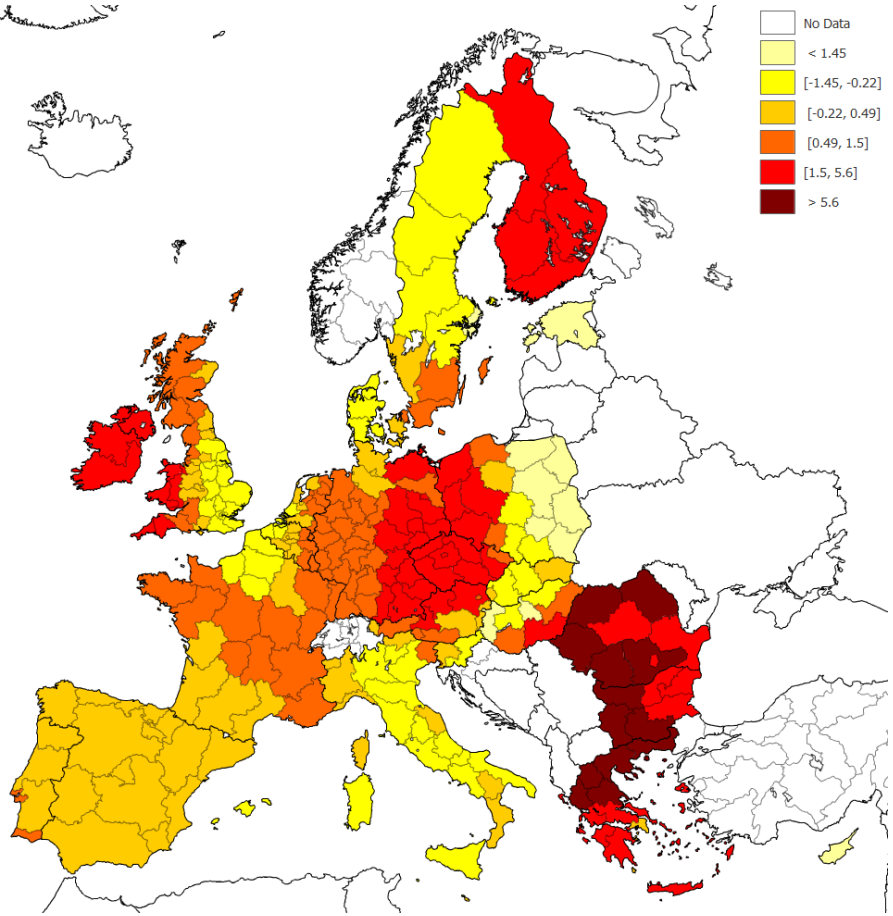
Practice GWR

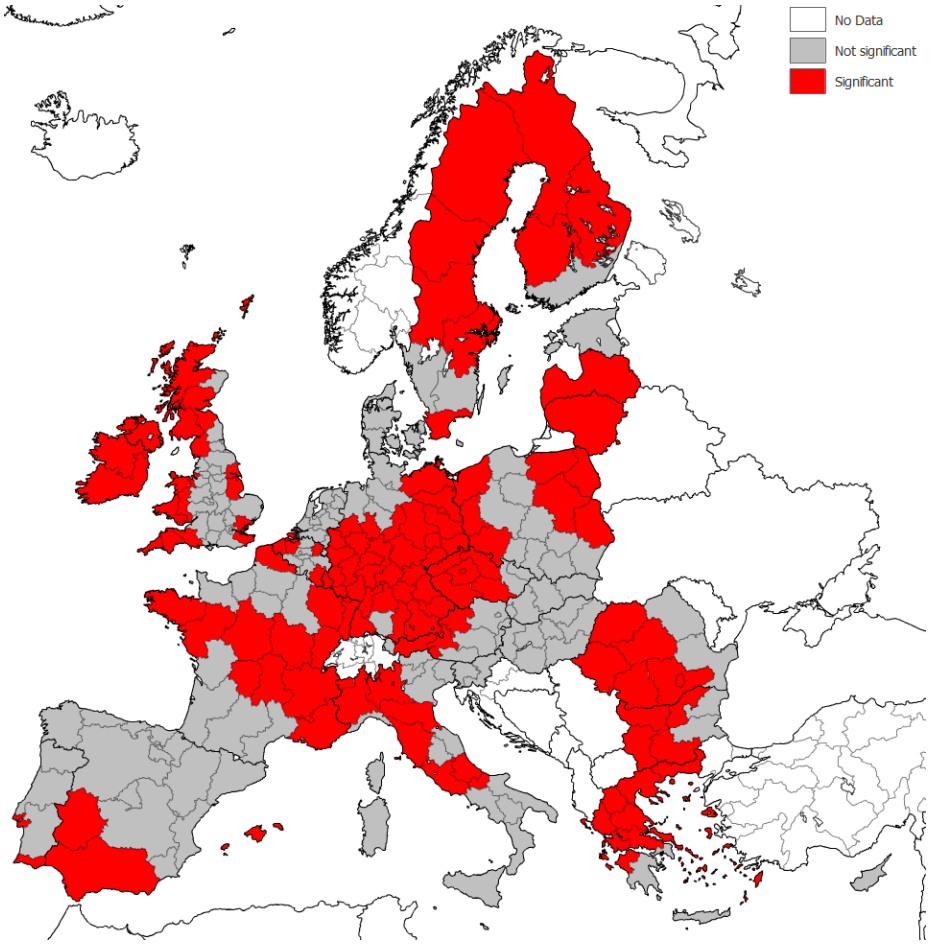
- **The effect of population growth**



Practice GWR

The effect of population growth





Practice GWR

- **Population growth effects**

Italy, Sweden and the South of Spain behave as Neoclassical economies as population growth rates decrease per capita income

Not the same story for Germany, France or Eastern European regions where population growth increases income per capita growth. This suggests some short of increasing returns/or agglomeration process that feedbacks positively

Practice GWR

- As observed GWR produces heterogeneous parameters linked to specific locations

- This might be **useful for policy-making purposes but a causal interpretation on the why there exists such variability is lacking**
- One explanation is that we are missing factors so the variation at some point reflects the omission of variables that are geographically correlated
- If this is a concern, increase the number of X, randomize X and try different combinations

- A possible direction that has not been taken so far in empirical analysis is to explain the variability of the estimated parameters with regressions on third factors, kernels, etc.

- GWR analysis should be a descriptive 1st step that needs a causal *"add in"* for meaningful analysis. Otherwise interpretation becomes too speculative

Practice GWR

- GWR estimates of growth-investment effects decrease with the QOG (similar to the Cohesion Funds investments result of Pose-Garzilazo (2015) paper). Are they substitutes? It looks like once regions arrive to a QOG of 0.75-0.8 additional physical capital investment does not increase growth anymore…