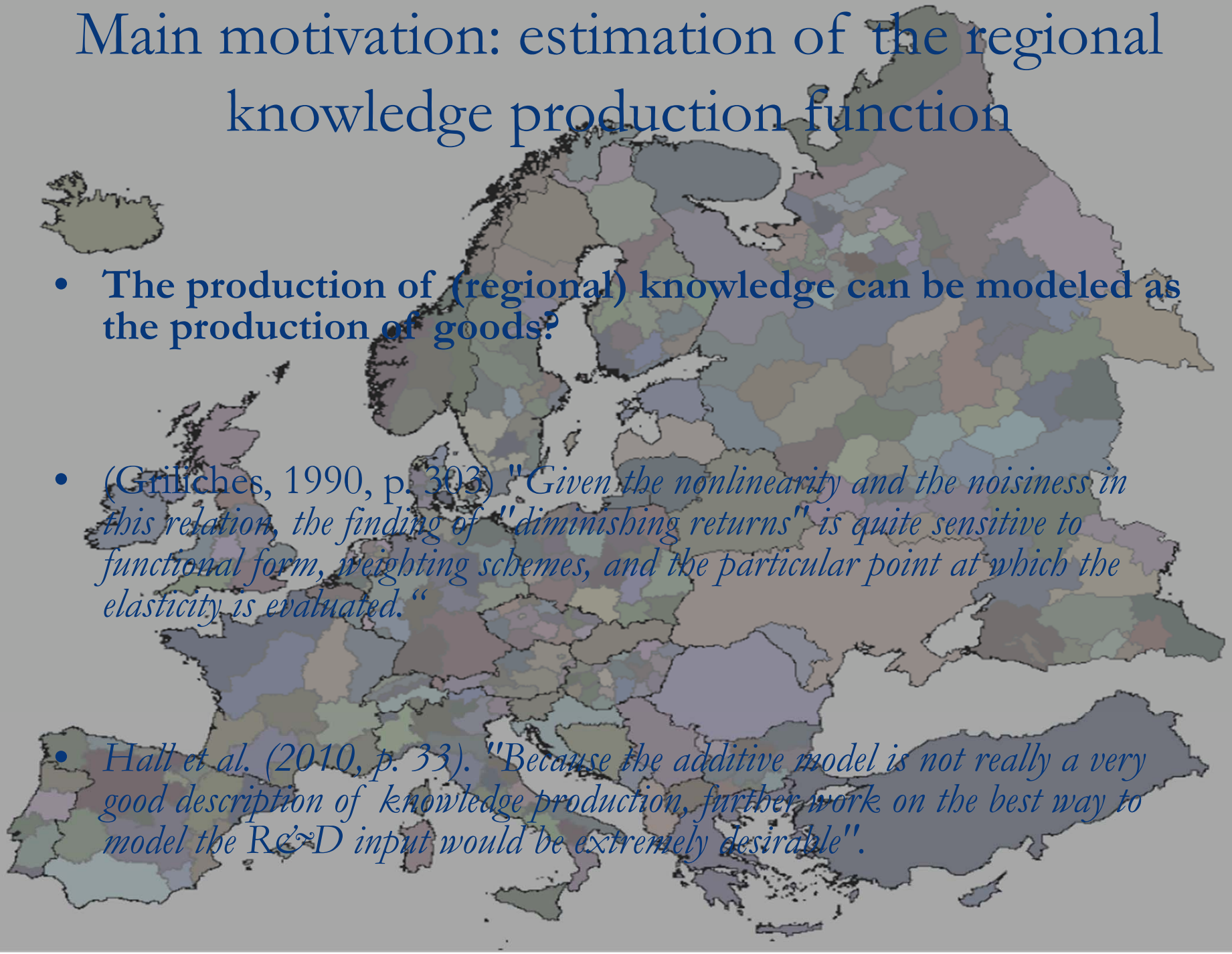
Augmented and unconstrained: revisiting the Regional Knowledge Production Function

Sylvie Charlot (GAEL INRA, Grenoble)
Riccardo Crescenzi (SERC LSE, London)
Antonio Musolesi (University of Ferrara & SEEDS)

20th International Panel Data conference, Tokyo, 2014

Main motivation: estimation of the regional knowledge production function

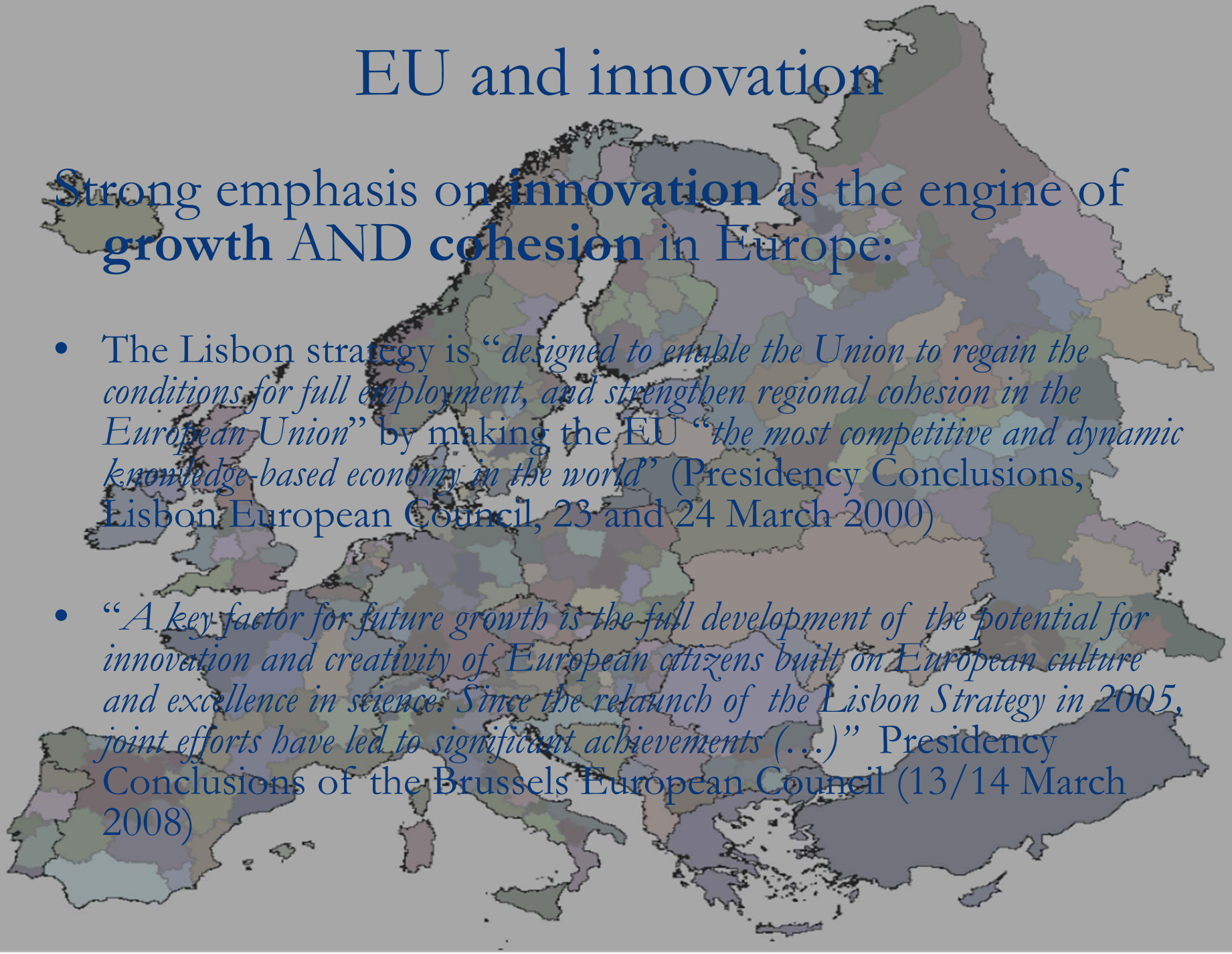


- **The production of (regional) knowledge can be modeled as the production of goods?**

- (Griliches, 1990, p. 303) "*Given the nonlinearity and the noisiness in this relation, the finding of "diminishing returns" is quite sensitive to functional form, weighting schemes, and the particular point at which the elasticity is evaluated.*"

- *Hall et al. (2010, p. 33). "Because the additive model is not really a very good description of knowledge production, further work on the best way to model the R&D input would be extremely desirable".*

EU and innovation



Strong emphasis on **innovation** as the engine of **growth** AND **cohesion** in Europe:

- The Lisbon strategy is "*designed to enable the Union to regain the conditions for full employment, and strengthen regional cohesion in the European Union*" by making the EU "*the most competitive and dynamic knowledge-based economy in the world*" (Presidency Conclusions, Lisbon European Council, 23 and 24 March 2000)

- "*A key factor for future growth is the full development of the potential for innovation and creativity of European citizens built on European culture and excellence in science. Since the relaunch of the Lisbon Strategy in 2005, joint efforts have led to significant achievements (…)*" Presidency Conclusions of the Brussels European Council (13/14 March 2008)

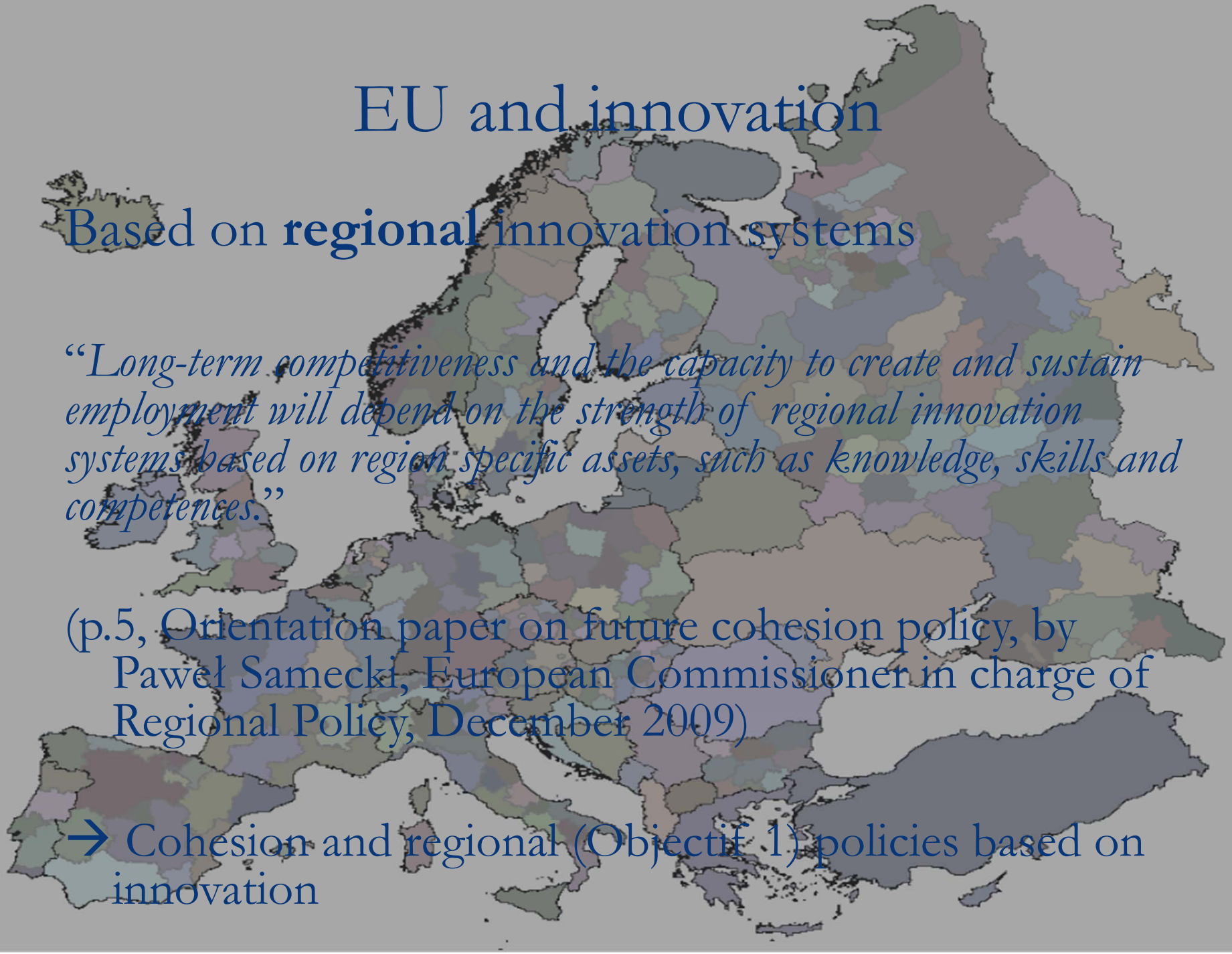
EU and innovation

Based on **regional** innovation systems

"*Long-term competitiveness and the capacity to create and sustain employment will depend on the strength of regional innovation systems based on region specific assets, such as knowledge, skills and competences.*"

(p.5, Orientation paper on future cohesion policy, by Paweł Samecki, European Commissioner in charge of Regional Policy, December 2009)

→ Cohesion and regional (Objectif 1) policies based on innovation

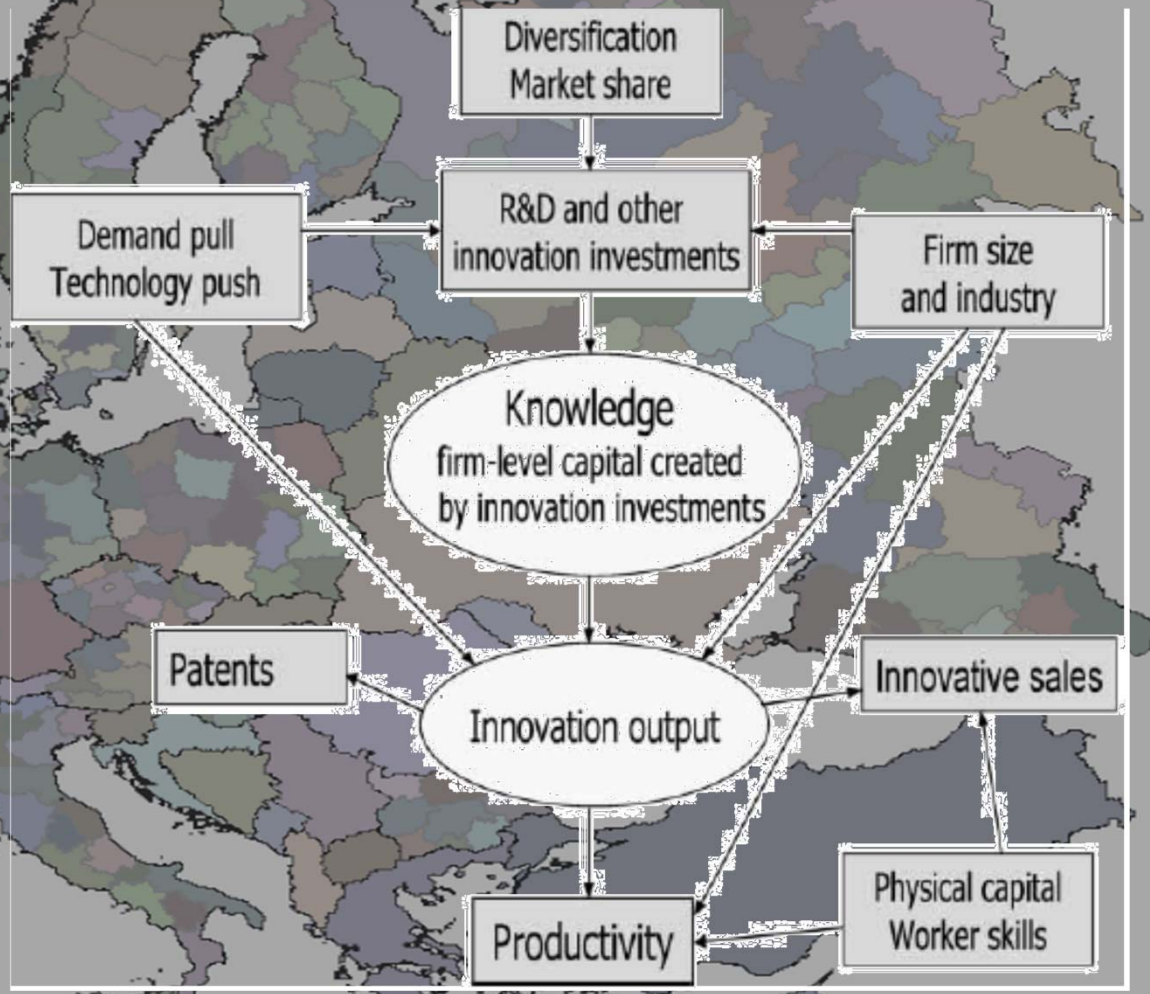


Knowledge production function

Knowledge production function is not a theoretical model

"..is at best a very crude reduced-form type relation whose theoretical underpinnings have still to be worked out"

(Griliches, 1990, p.1672)



Firm level studies

Crépon-Duguet-Mairesse (1998) model:

-RD >patents> productivity

Dealing with many econometric problems: the nature of RD and patents (RD has many zeros, patents is a count), the endogeneity of RD, ...)

More Recent studies uses CIS data:

-RD > innovation (yes/no) > productivity

Other econometric problems to deal with (the filter of CIS, treatment effect)

Regional level studies: economic geography and innovation:

Concentration of economic activities and much more of innovation ← Agglomeration economies (Marshall)

3 key sources of agglomeration economies:

- pecuniary externalities linked to the proximity of customers and suppliers
- labor market thickness conducive to a better matching between employers and employees and
- most relevant for innovation: pure technological externalities

← face to face communication or labor market

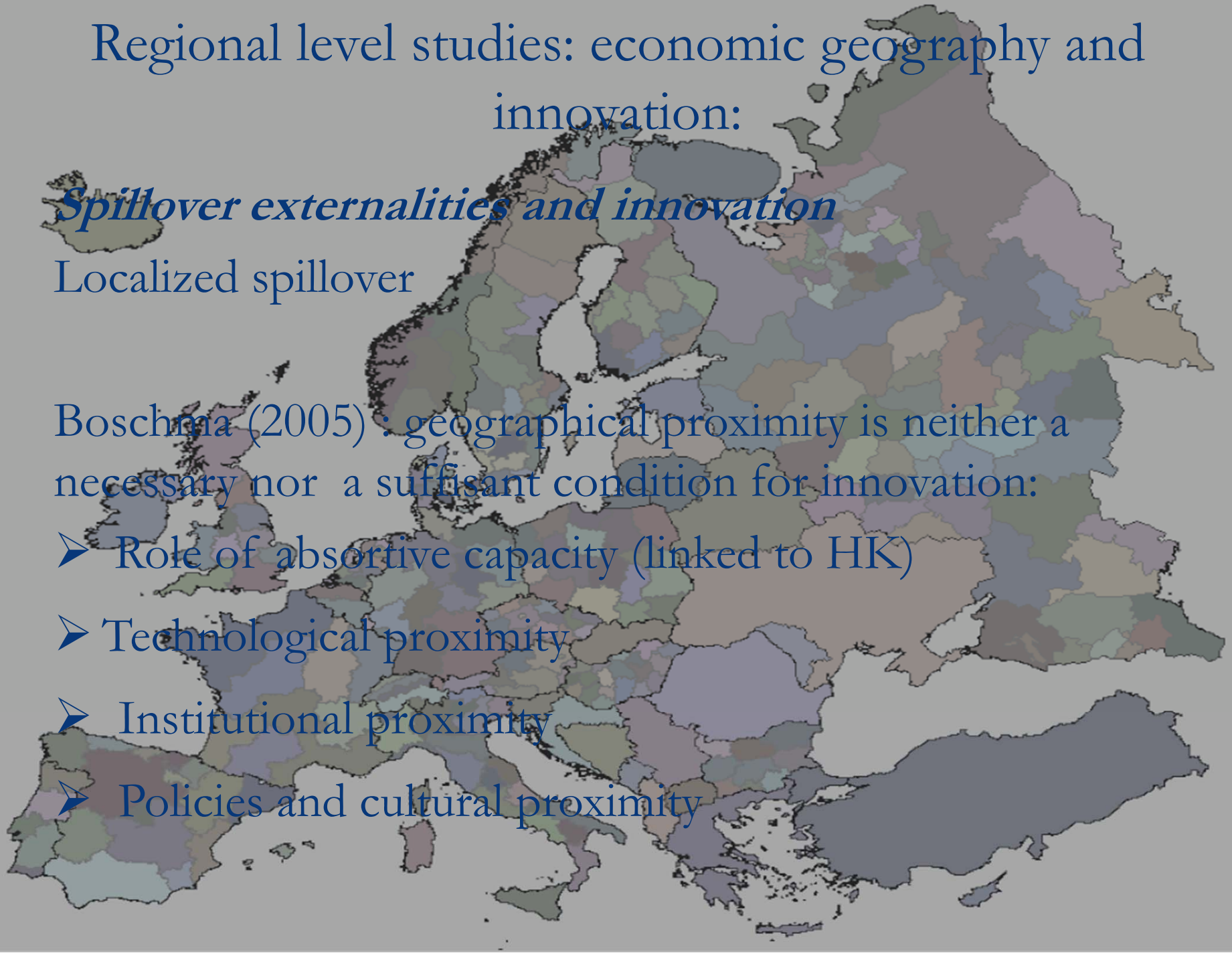
Regional level studies: economic geography and innovation:

Spillover externalities and innovation

Localized spillover

Boschma (2005) : geographical proximity is neither a necessary nor a suffisant condition for innovation:

- Role of absortive capacity (linked to HK)
- Technological proximity
- Institutional proximity
- Policies and cultural proximity



Regional level studies: econometric analyses

Aggregate uni-equation regional knowledge production function (Audretsch (2003), Crescenzi et al. (2007) ….). Linear or log-log specification of the form:

$$K_{r,t} = \beta_1 RD_{r,t} + \beta_2 HK_{r,t} + \beta_3 WRD_{r,t} + \beta_4 WHK_{r,t} + \alpha_r\ (+\lambda_t) + u_{r,t}.$$

"W" indicates external variables (geographic or alternative spillover mechanism)

One- or two-way fixed effects to account for endogenous correlated factors

Econometric biases

$$K_{r,t} = \beta_1 RD_{r,t} + \beta_2 HK_{r,t} + \beta_3 WRD_{r,t} + \beta_4 WHK_{r,t} + \alpha_r\ (+\lambda_t) + u_{r,t}.$$

- ✓ "unobservable factors bias (selection bias)">> one- or two-way specification does nor control for heterogeneous common factors or time varying unobservable variable linked both to patents and the RD and HK

quality of Regional Systems of Innovation, agglomeration economies…

- ✓ Functional form bias

linear or log-log may be to restrictive (e.g. Grilliches, 1990)

Additivity of RD and HK is also a likely restrictive assumption

- ✓ Heterogeneity bias

The effect of the main inputs may differ across regions

Proposed econometric model

Semiparametric variant of the random growth model via Generalised additive model (GAM) framework (Hastie and Tibshirani 1990; Wood, 2004, 2008):

$$K_{r,t} = f_1(RD_{r,t}) + f_2(HK_{r,t}) + f_3(WRD_{r,t}) + f_4(WHK_{r,t}) + \alpha_r + \lambda_t + \gamma_r t + u_{r,t}$$

- $\alpha_r + \lambda_t + \gamma_r t + u_{r,t}$ Random growth (Heckman and Hotz, 1989; Wooldridge, 2005): better account of endogeneity dues. "selection on unobservables"
- GAM: flexibility of nonparametric without the curse of dimensionality

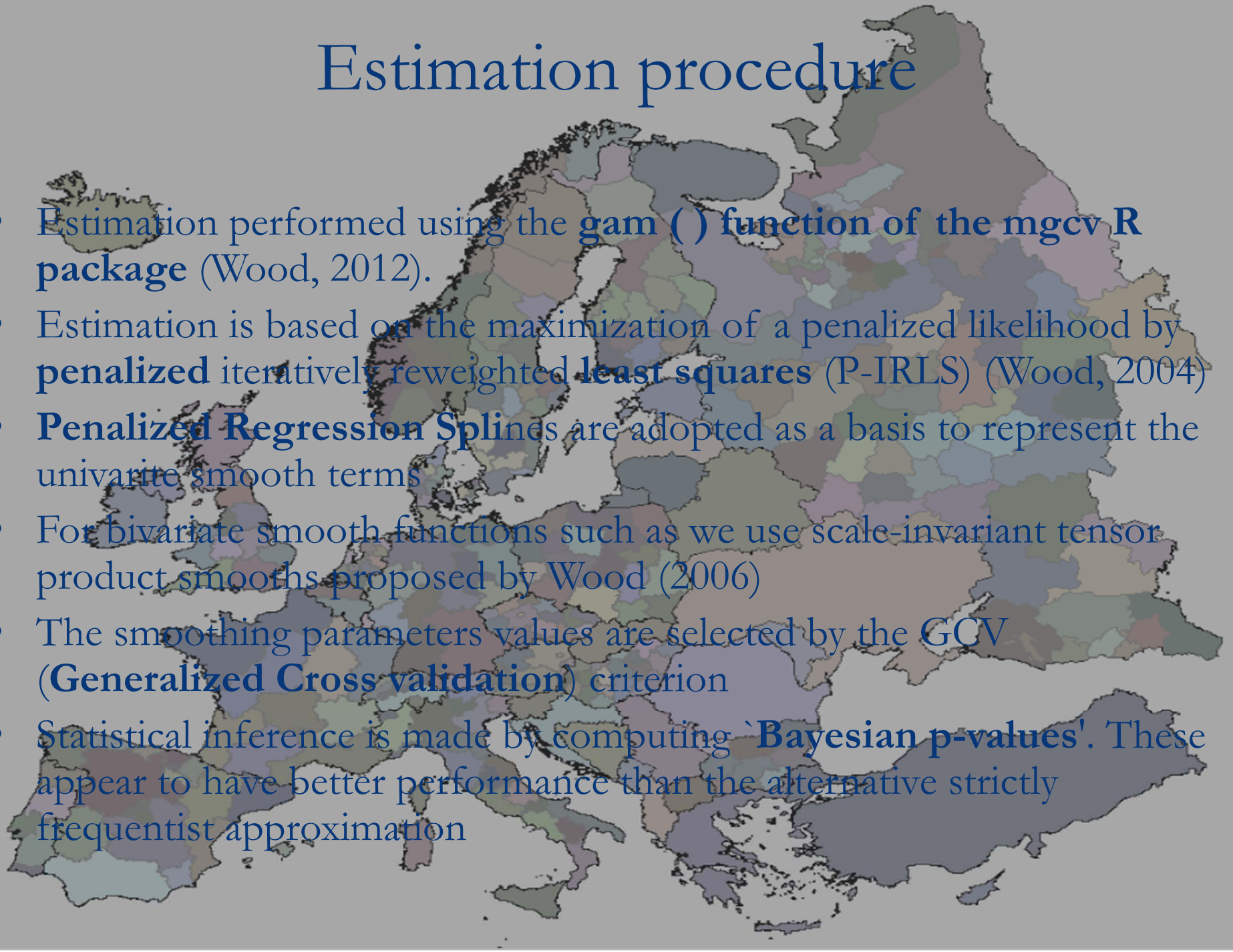
Variants of such an equation (partially relaxing additivity, allowing for heterogeneous relations):

>> $K_{r,t} = Z_t\theta_r + f(RD_{r,t}, HK_{r,t}) + u_{r,t}$

>> $K_{r,t} = Z_t\theta_r + f_{1O1_r}(RD_{r,t}) + f_{2O1_r}(HK_{r,t}) + f_{3O1_r}(WRD_{r,t}) + f_{4O1_r}(WHK_{r,t}) + u_{r,t}.$

$$O1_r = \begin{cases} 1 \text{ if region } \in \text{ objective 1 group} \\ 0 \text{ if region } \notin \text{ objective 1 group} \end{cases}$$

Estimation procedure



- Estimation performed using the **gam ( ) function of the mgcv R package** (Wood, 2012).
- Estimation is based on the maximization of a penalized likelihood by **penalized** iteratively reweighted **least squares** (P-IRLS) (Wood, 2004)
- **Penalized Regression Splines** are adopted as a basis to represent the univarite smooth terms
- For bivariate smooth functions such as we use scale-invariant tensor product smooths proposed by Wood (2006)
- The smoothing parameters values are selected by the GCV (**Generalized Cross validation**) criterion
- Statistical inference is made by computing `**Bayesian p-values**'. These appear to have better performance than the alternative strictly frequentist approximation

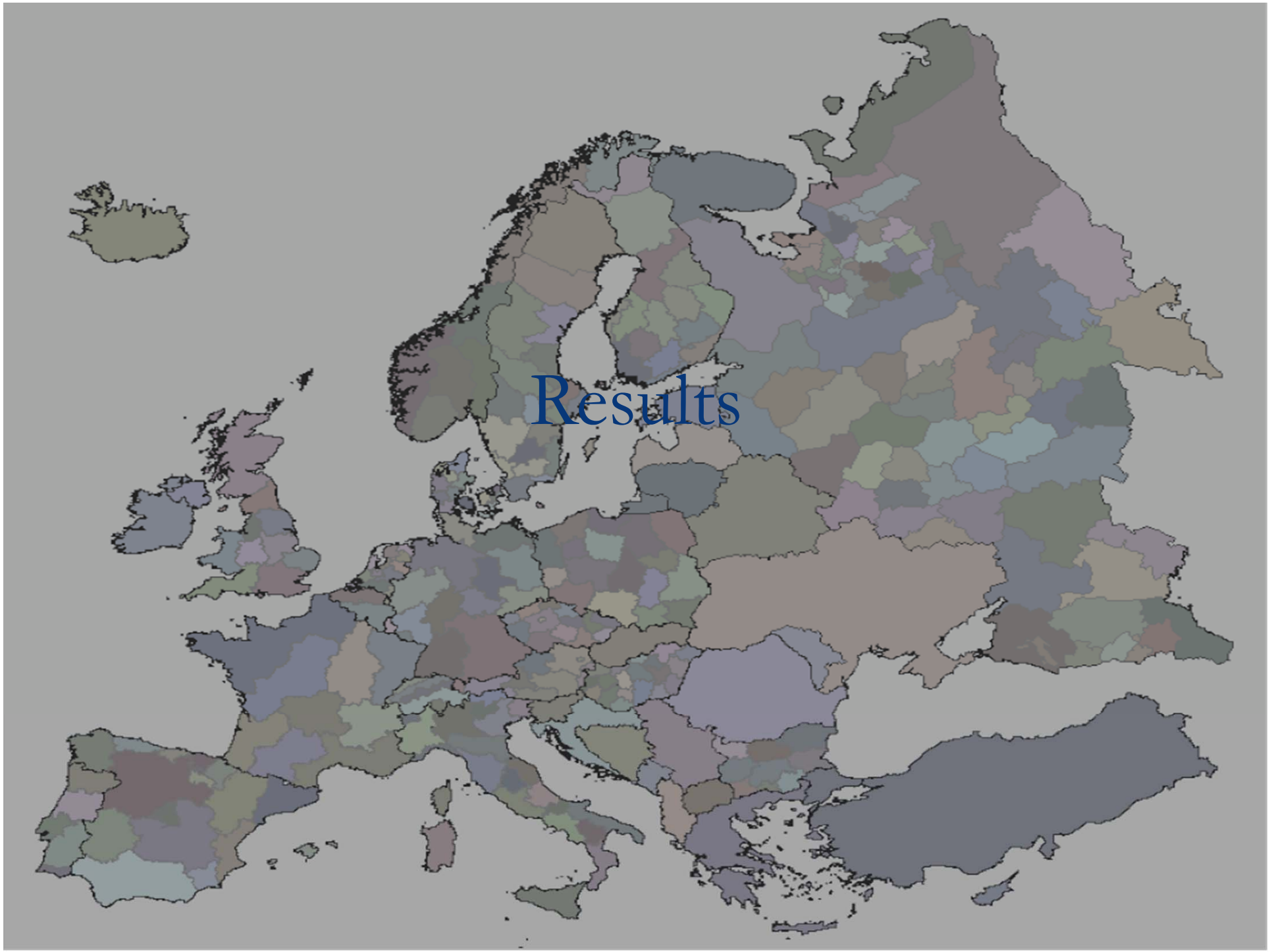
Dataset

- EUROSTAT New Cronos-Regio data
  - Innovation output (K): number of patents per million inhabitants
  - R&D: the share of regional GDP spent on R&D (public and private)
  - HK: the regional share of workers with tertiary education or higher

  All variables arre real numbers

- Regional division: 169 regions
  - NUTS1 regions for Germany, Belgium and the UK
  - NUTS2 for all other countries (Spain, France, Italy, the Netherlands, Greece, Austria, Portugal..).

  Time span: 1995-2004


Results

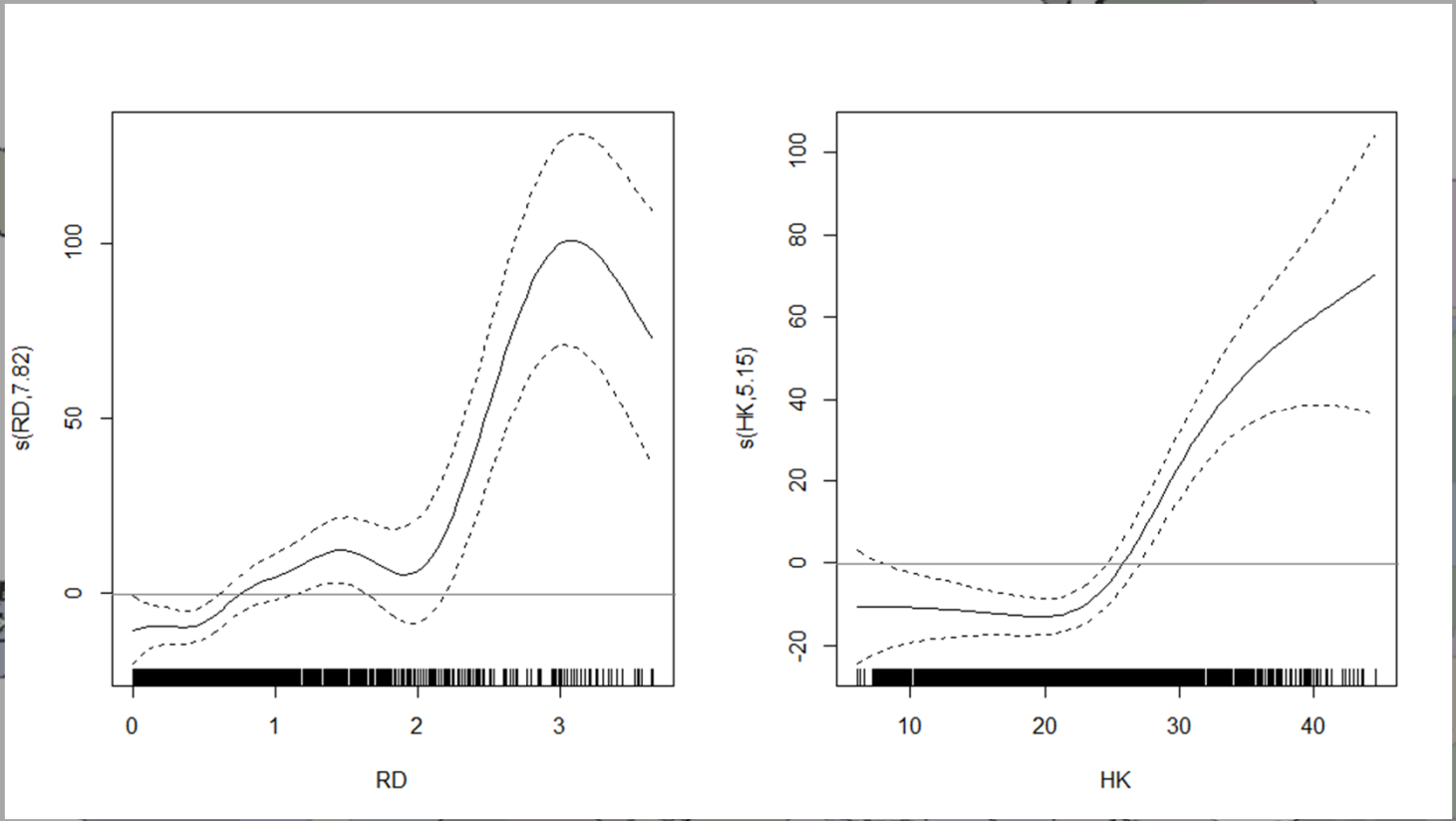
Results (preliminary): the role of unobserved heterogeneity

Parametric estimation of the RKPF (log-log)

	(1)	(2)	(3)
RD	0.260	0.00481	0.600***
	(0.183)	(0.206)	(0.208)
HK	1.397***	0.0382	0.508***
	(0.195)	(0.204)	(0.162)
One way	X	X	X
Two way		X	X
Random trend			X

>>>Empirical Relevance of adopting the random growth specification
>>>F test clearly chooses the random growth
>>> random growth estimates close to firm level ones. Griliches (1990): elasticity of patents with respect to R&D is between 0.3 and 0.6. Blundell et al. (2002) report a preferred estimate of 0.5.

Results (1): RD and HK



R&D expenditure (share of regional GDP)

Significant threshold effects

Positive when > 0.5% but limited

Decreases after the second one > 1.5%

Maximes its effect > 2%

Human capital

More linear impact

Becomes positive after a large level

> 20% with a high slope

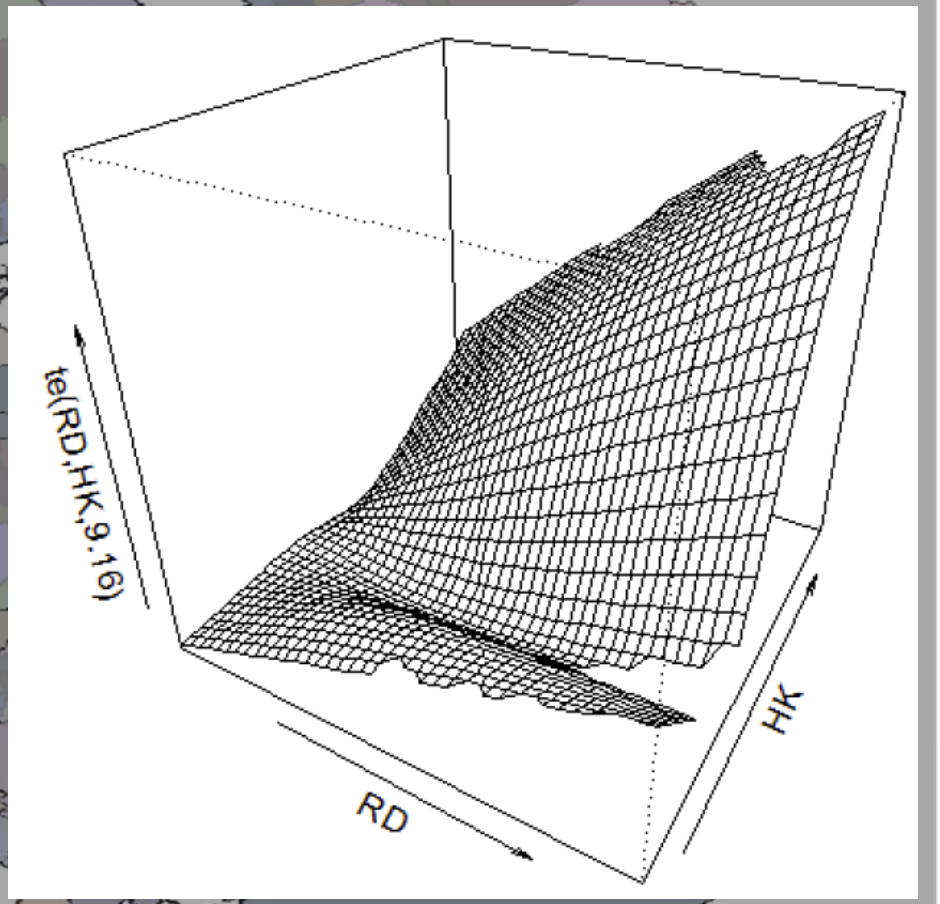
Less sloping and significant > 35%

Results (2): joint effect of R&D and HK

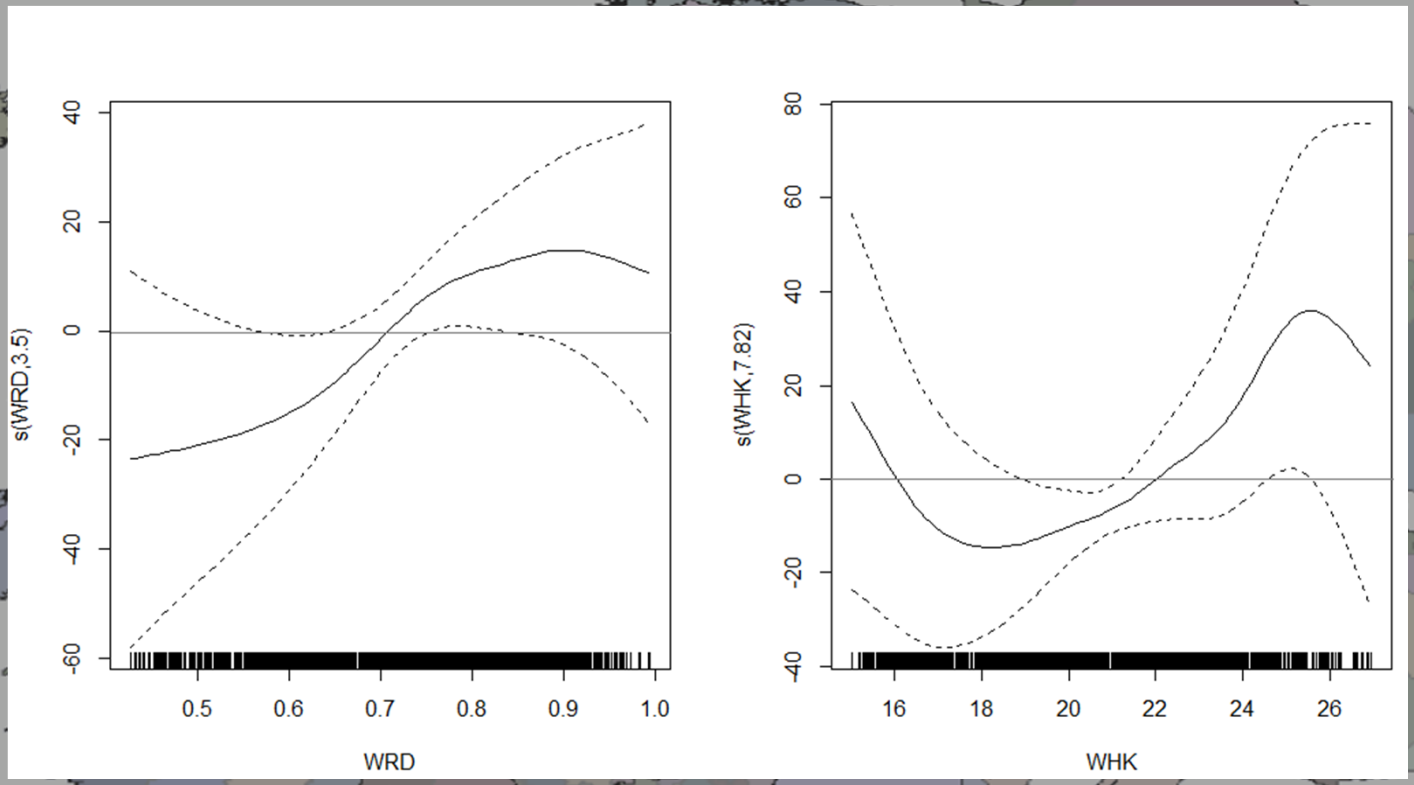
For low level of HK, no effect of R&D on K

For high level of HK, R&D has a positive and increasing effect on K

The effect of HK on innovation increases with RD

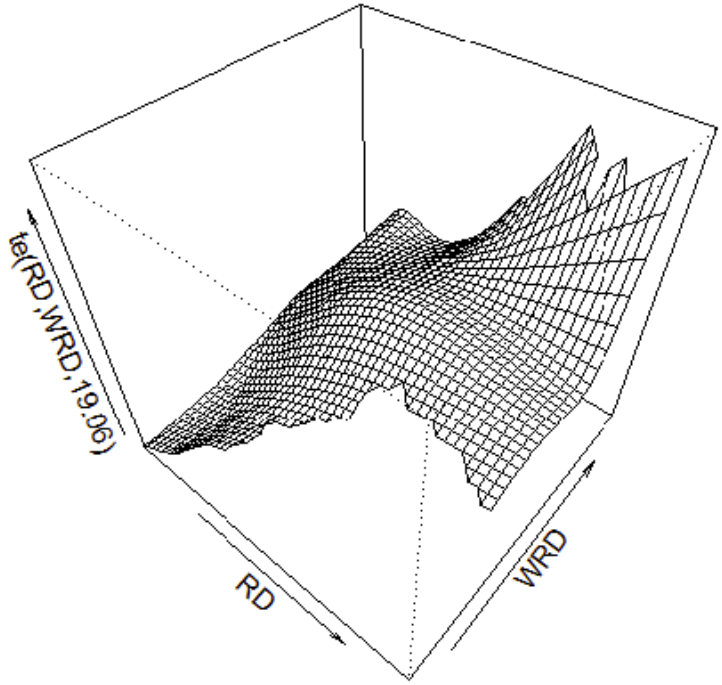


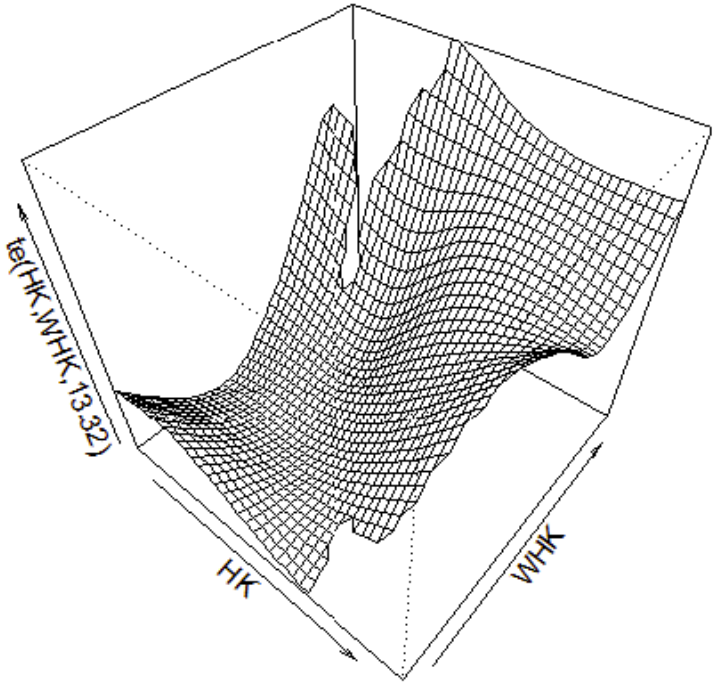
Results (3): spatial spillovers alone



- The effect of foreign inputs appears to be significant locally above a certain threshold (source of structural disavantage of lagging regions)
- This effect tends to wane with proximity to the main centres of HK and RD accumulation (shadow effect)

Results (4): spillover interaction effects





For very low value of local R&D no effect of WRD on knowledge

For intermediate levels of local R&D, WRD increases K until a threshold after which negative impact of WRD

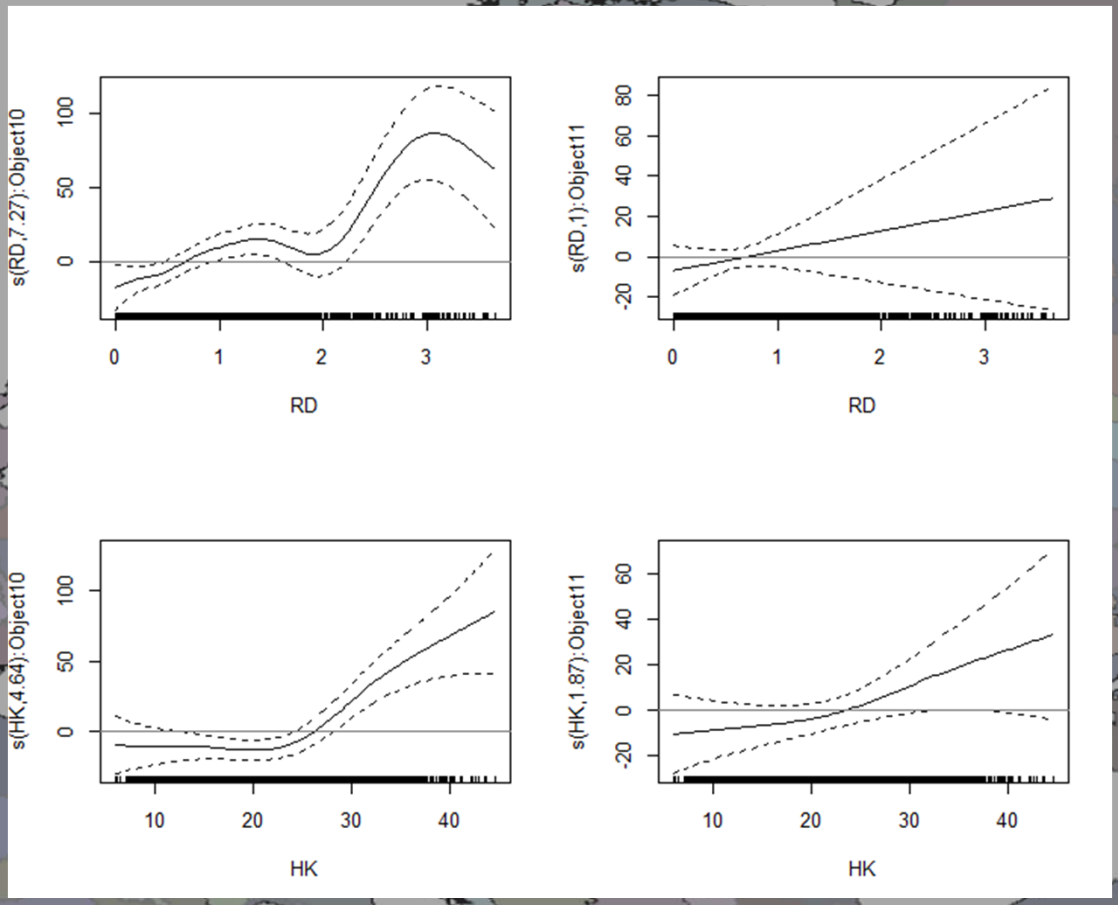
For large levels of local R&D, WRD sharply increases K

No effect of local HK for very low level of WHK

Increasing slithly WHK can be detrimental for K,

Then, local and neigbhorood HK are strong complements

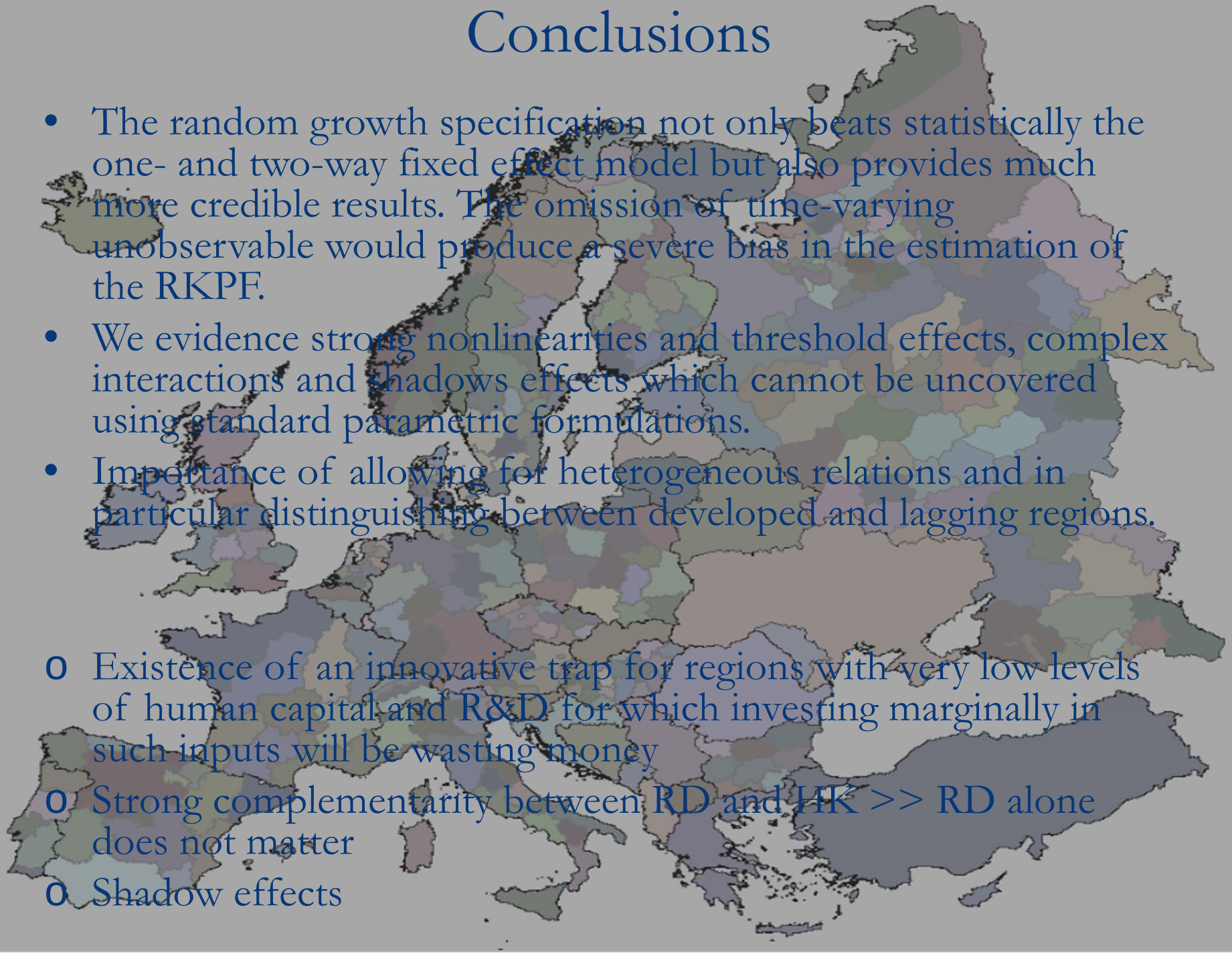
Results (5): developed versus lagging regions



always very significant for developed regions
Not for lagging regions
Previous picture mainly for richer European regions

Conclusions

- The random growth specification not only beats statistically the one- and two-way fixed effect model but also provides much more credible results. The omission of time-varying unobservable would produce a severe bias in the estimation of the RKPF.
- We evidence strong nonlinearities and threshold effects, complex interactions and shadows effects which cannot be uncovered using standard parametric formulations.
- Importance of allowing for heterogeneous relations and in particular distinguishing between developed and lagging regions.



- Existence of an innovative trap for regions with very low levels of human capital and R&D for which investing marginally in such inputs will be wasting money
- Strong complementarity between RD and HK >> RD alone does not matter
- Shadow effects